Topologia e Análise no Espaço $\mathbb{R}^n$

Ronaldo Freire de Lima

Para Alice, minha maravilha.

A prolixidade tem sua dose de obscuridade, não menor que a da concisão.

Johann Kepler

Prefácio

A Análise é a teoria matemática que culminou das diversas transformações do Cálculo — estabelecido no século XVII, independentemente, por Isaac Newton (1642–1727) e Gottfried Leibniz (1646–1716) —, as quais foram protagonizadas por eminentes matemáticos europeus ao longo dos séculos XVIII e XIX e início do século XX.

Os conceitos primordiais da Análise, a saber, os de derivada e integral, expressam-se através das noções de limite e convergência, que, presentemente, são consideradas em toda a sua generalidade por uma outra teoria, a Topologia, surgida em meados do século XIX. Desta forma, do ponto de vista da Análise, uma abordagem atual às funções passa, necessariamente, pelo estudo dos aspectos topológicos — isto é, da topologia — dos conjuntos onde estas são definidas.

Este livro se propõe, então, a estudar a topologia dos espaços euclidianos multidimensionais (capítulos de 1 a 3) e a diferenciabilidade das funções definidas nestes espaços (capítulos 4 e 5). Ele destina-se, principalmente, a estudantes de matemática em fim de graduação ou início de mestrado, e o pré-requisito à sua leitura é o conhecimento de fatos elementares da Álgebra Linear e da Análise na reta — muitos dos quais são relembrados nos capítulos 0 e 1.

O conteúdo, a ser coberto num curso semestral de quatro horas semanais, foi escolhido de modo a apresentar os conceitos e resultados mais fundamentais da Topologia e da Análise — de funções diferenciáveis — em espaços euclidianos, bem como contemplar o material necessário a estudos posteriores onde estas teorias são aplicadas ou generalizadas, tais como Geometria Diferencial, Análise Complexa, Equações Diferenciais e Topologia Geral.

Com o propósito de finalizar os capítulos de forma "poética", em cada um deles, a seção que antecede a dos exercícios — grifada $(*)$ — introduz um resultado elegante e, ao mesmo tempo, significativo, o qual está intimamente relacionado com os assuntos ali discutidos. Este espírito é estendido ao livro como um todo, que inclui um apêndice onde, a partir da teoria apresentada nos capítulos precedentes, discute-se sobre o importante conceito de variedade diferenciável. Um segundo apêndice fornece soluções de todos os exercícios propostos e encerra, desta forma, o texto.

Em conclusão, gostaria de deixar consignados os meus agradecimentos a Roberto Sá, que me sugeriu escrever este livro e muito me apoiou ao longo do processo; aos amigos Rubens Leão de Andrade, Victor Giraldo, Cassio Neri, Edmundo Pereira, Airton von Sohsten e Tatiana Roque, por todos os ensinamentos; a todos aqueles que leram versões preliminares do texto e contribuíram para a sua melhoria, quer apontando erros ou obscuridades, quer sugerindo diferentes abordagens — em especial, aos estudantes do curso de bacharelado em matemática da UFRN,

que, enquanto alunos da disciplina Análise II, por mim ministrada, tiveram a oportunidade e a bondade de fazê-lo.

Ronaldo F. de Lima

Natal RN
Julho 2013

Conteúdo

CAPíTULO 0

Funções e Números Reais

Neste capítulo preliminar, faremos uma breve digressão sobre funções e números reais. Nosso intuito é o de fixar notação, relembrar os principais conceitos associados a estes objetos — com os quais o leitor, supostamente, está familiarizado — e destacar suas propriedades mais fundamentais. Assim é que, na primeira seção, dedicada às funções, abordamo-las de forma abstrata e omitimos, em favor da objetividade, não só a apresentação de exemplos, mas também as demonstrações das proposições ali contidas. Já na segunda, consideramos o conjunto dos números reais a partir de sua estrutura de corpo ordenado e completo, bem como discutimos o conceito de convergência de sequências de números reais.

1. Funções

Dentre os conceitos mais elementares e profundos da Matemática, encontra-se o de *função*. Para introduzi-lo, consideramos dois conjuntos A e B, ditos, respectivamente, o *domínio* e o *contradomínio* da função (a qual indicaremos por f) e uma lei de correspondência entre estes conjuntos, que associa a cada elemento x de A um único elemento $f(x)$ (lê-se "f de x") de B. Dizemos, então, que f é uma função de A em B ou, equivalentemente, definida em A e que toma valores em B, e a representamos da seguinte forma:

$$\begin{array}{cccc} f: & A & \to & B \\ & x & \mapsto & f(x). \end{array}$$

Um elemento $y = f(x) \in B$ é dito a *imagem* de x pela função f, e o subconjunto de B formado por todos os seus elementos que são imagem de algum elemento $x \in A$ é chamado de *conjunto-imagem* de f, o qual denota-se por $f(A)$. Assim,

$$f(A) = \{y \in B;\, y = f(x),\, x \in A\}.$$

Uma função $f : A \to B$ é dita

- *injetiva* se cumpre a condição: $x \neq x' \Rightarrow f(x) \neq f(x'),\ x, x' \in A$;
- *sobrejetiva* se $f(A) = B$;
- *bijetiva* se é injetiva e sobrejetiva.

Diz-se que dois conjuntos A e B têm *mesma cardinalidade* quando existe uma função bijetiva de A em B. Um conjunto é dito *enumerável* quando tem a mesma cardinalidade do conjunto dos números naturais, $\mathbb{N} = \{1, 2, 3, \dots\}$.

Dados uma função $f : A \to B$ e um subconjunto de A, X, a *restrição* de f a X é a função

$$\begin{array}{cccc} f|_X: & X & \to & B \\ & x & \mapsto & f(x). \end{array}$$

Neste caso, diz-se também que $f : A \to B$ é uma *extensão* de $f|_X$ a A.

1.1. Inversão – Composição. Dada uma função bijetiva, $f : A \to B$, tem-se que cada elemento $y \in B$ está associado a um único $x \in A$ através da igualdade $y = f(x)$. Assim, esta correspondência define uma função $g : B \to A$, dita a *inversa* de f, em que $g(y) = x$. Nestas condições, diz-se que a função f é *invertível* e denota-se sua inversa, g, por f^{-1}. Logo, quando $f : A \to B$ é invertível, para quaisquer $x \in A$ e $y \in B$, valem as igualdades:

$$f^{-1}(f(x)) = x \qquad \text{e} \qquad f(f^{-1}(y)) = y. \tag{1}$$

Consideremos agora funções $f : A \to B$ e $g : B \to C$ e observemos que a função

$$\begin{array}{rccc} g \circ f : & A & \to & C \\ & x & \mapsto & g(f(x)), \end{array}$$

dita a *composta* de g com f, está bem definida (Fig. 1).

$$\begin{tikzcd} B \arrow[r, "g"] & C \\ A \arrow[u, "f"] \arrow[ur, "g \circ f"'] & \end{tikzcd}$$

Figura 1

$$\begin{tikzcd} B \arrow[r, "g"] & C \arrow[d, "h"] \\ A \arrow[u, "f"] \arrow[r, "h \circ g \circ f"'] & D \end{tikzcd}$$

Figura 2

Deve-se notar que se f, g, h são funções tais que as compostas $g \circ f$ e $h \circ g$ estão bem definidas, então estão bem definidas as compostas $(h \circ g) \circ f$ e $h \circ (g \circ f)$. Além disso, vale a igualdade $(h \circ g) \circ f = h \circ (g \circ f)$, isto é, a composição de funções é uma operação associativa (Fig. 2).

Segue-se das igualdades (1) que, quando $f : A \to B$ é bijetiva, tem-se

$$f^{-1} \circ f = i_A \qquad \text{e} \qquad f \circ f^{-1} = i_B\,,$$

em que $i_A : A \to A$ e $i_B : B \to B$ são as funções *identidade* de A e B, respectivamente, definidas por $i_A(x) = x$ e $i_B(y) = y$.

1.2. Imagem – Imagem Inversa. Consideremos uma função $f : A \to B$ e subconjuntos $X \subset A$, $Z \subset B$. Definem-se a *imagem* de X por f, $f(X)$, e a *imagem inversa* de Z por f, $f^{-1}(Z)$, por

$$f(X) = \{y \in B;\, y = f(x),\, x \in X\} \quad \text{e} \quad f^{-1}(Z) = \{x \in A;\, f(x) \in Z\}.$$

Note que $f(X)$ é o conjunto-imagem da restrição de f a X e que $f^{-1}(Z)$ está bem definido, mesmo que f não seja invertível.

Dados conjuntos $X, Y \subset A$, $Z, W \subset B$, e uma função $f : A \to B$, verificam-se as propriedades da imagem e da imagem inversa listadas na tabela seguinte. Note que a primeira inclusão na segunda linha reduz-se a uma igualdade quando f é sobrejetiva. O mesmo ocorre com a segunda inclusão (na mesma linha) quando f é injetiva. Observem-se também as propriedades referentes à interseção e à diferença de conjuntos, que são preservadas pela imagem inversa, mas não necessariamente

IMAGEM	IMAGEM INVERSA
$f(\emptyset)=\emptyset$	$f^{-1}(\emptyset)=\emptyset$
$f(f^{-1}(Z))\subset Z$	$f^{-1}(f(X))\supset X$
$X\subset Y\Rightarrow f(X)\subset f(Y)$	$Z\subset W\Rightarrow f^{-1}(Z)\subset f^{-1}(W)$
$f(X\cup Y)=f(X)\cup f(Y)$	$f^{-1}(Z\cup W)=f^{-1}(Z)\cup f^{-1}(W)$
$f(X\cap Y)\subset f(X)\cap f(Y)$	$f^{-1}(Z\cap W)=f^{-1}(Z)\cap f^{-1}(W)$
$f(X-Y)\supset f(X)-f(Y)$	$f^{-1}(Z-W)=f^{-1}(Z)-f^{-1}(W)$

pela imagem. No entanto, as inclusões nas duas últimas linhas reduzem-se a igualdades quando a função f é injetiva.

Dadas funções $f:A\to B$ e $g:B\to C$, valem, para quaisquer conjuntos $X\subset A$ e $Y\subset C$, as igualdades:

$$(g\circ f)(X)=g(f(X)) \qquad \text{e} \qquad (g\circ f)^{-1}(Y)=f^{-1}(g^{-1}(Y)).$$

1.3. Gráficos – Projeções. O *gráfico* de uma função $f:A\to B$ é o conjunto

$$\operatorname{graf}(f)=\{(x,y)\in A\times B\,;\ y=f(x)\}\subset A\times B,$$

em que $A\times B$ é o *produto cartesiano* de A por B (Fig. 3),

$$A\times B=\{(a,b); a\in A,\ b\in B\}.$$

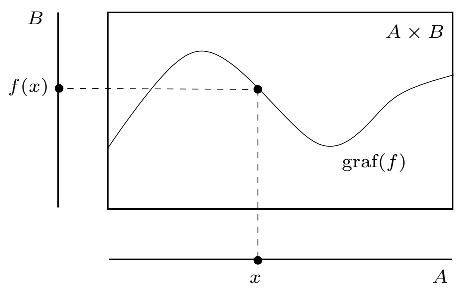

Figura 3

As funções

$$\begin{array}{rccc} P_A: & A\times B & \to & A \\ & (a,b) & \mapsto & a \end{array} \qquad \text{e} \qquad \begin{array}{rccc} P_B: & A\times B & \to & B \\ & (a,b) & \mapsto & b \end{array}$$

são denominadas, respectivamente, a *projeção* sobre A e a *projeção* sobre B.

Considerando-se, então, subconjuntos $X\subset A$ e $Y\subset B$, vale a igualdade

$$X\times Y=P_A^{-1}(X)\cap P_B^{-1}(Y).$$

Além disso, dada uma função $f:A\to B$, tem-se que a restrição de P_A ao gráfico de f,

$$\begin{array}{rccc} P_A|_{\operatorname{graf}(f)}: & \operatorname{graf}(f) & \to & A \\ & (x,f(x)) & \mapsto & x, \end{array}$$

é, claramente, bijetiva.

1.4. Sequências e Famílias. Uma função cujo domínio é o conjunto dos números naturais, $\mathbb{N} = \{1, 2, 3, \dots\}$, é chamada de sequência. Mais precisamente, uma *sequência* num conjunto A é uma função $x : \mathbb{N} \to A$. A imagem de um natural $k \in \mathbb{N}$ pela função x é denotada por x_k (ao invés de $x(k)$), sendo este denominado, então, o k-ésimo *termo* da sequência. Fixando-se o conjunto A, denota-se uma sequência $x : \mathbb{N} \to A$ por $(x_k)_{k\in\mathbb{N}}$ ou, simplesmente, por (x_k). Quando os termos x_k são definidos por uma expressão, esta é dita o *termo geral* da sequência (x_k).

Uma *subsequência* de uma sequência $x : \mathbb{N} \to A$ é uma restrição de x a um subconjunto infinito de $\mathbb{N}$, $\mathbb{N}_0 = \{k_1 < k_2 < k_3 < \dots\}$, a qual denotamos por $(x_k)_{k\in\mathbb{N}_0}$ ou $(x_{k_i})_{i\in\mathbb{N}}$. Note que toda subsequência é também uma sequência.

Uma função arbitrária de um conjunto Λ num conjunto A pode ser vista como uma "sequência generalizada", na qual o conjunto $\mathbb{N}$ é substituído por Λ, dito, neste caso, um *conjunto de índices*. Deste ponto de vista, a função é chamada de *família* (de elementos de A) e denotada por $(x_\lambda)_{\lambda\in\Lambda}$, isto é, para cada $\lambda \in \Lambda$, x_λ é um elemento de A, dito um *membro* da família $(x_\lambda)_{\lambda\in\Lambda}$. Quando $\Lambda_0 \subset \Lambda$, a família $(x_\lambda)_{\lambda\in\Lambda_0}$ é dita uma *subfamília* de $(x_\lambda)_{\lambda\in\Lambda}$.

No caso particular em que $(X_\lambda)_{\lambda\in\Lambda}$ é uma família de conjuntos, definem-se a *união* e a *interseção* dos membros desta família, respectivamente, por

$$\bigcup_{\lambda\in\Lambda} X_\lambda = \{x\,;\, x \in X_\lambda \text{ para algum } \lambda \in \Lambda\} \quad \text{e} \quad \bigcap_{\lambda\in\Lambda} X_\lambda = \{x\,;\, x \in X_\lambda\ \forall \lambda \in \Lambda\}.$$

Quando não há ambiguidade com respeito ao conjunto de índices, Λ, escrevem-se

$$\bigcup_{\lambda\in\Lambda} X_\lambda = \bigcup X_\lambda \qquad \text{e} \qquad \bigcap_{\lambda\in\Lambda} X_\lambda = \bigcap X_\lambda\,.$$

Dados um conjunto A e uma família de subconjuntos de A, $(X_\lambda)_{\lambda\in\Lambda}$, valem as igualdades

- $A - \bigcup X_\lambda = \bigcap (A - X_\lambda)$;
- $A - \bigcap X_\lambda = \bigcup (A - X_\lambda)$.

Além disso, se $f : A \to B$ é uma função e $(Y_\gamma)_{\gamma\in\Gamma}$ é uma família de subconjuntos de B, tem-se:

- $f(\bigcup X_\lambda) = \bigcup f(X_\lambda)$ e $f^{-1}(\bigcup Y_\gamma) = \bigcup f^{-1}(Y_\gamma)$;
- $f(\bigcap X_\lambda) \subset \bigcap f(X_\lambda)$ e $f^{-1}(\bigcap Y_\gamma) = \bigcap f^{-1}(Y_\gamma)$.

Devido ao fato do conceito de função permear, praticamente, todas as partes da Matemática, o termo "função" admite diversos sinônimos, os quais variam de acordo com a natureza dos conjuntos envolvidos. Nos contextos da Álgebra Linear, da Topologia e da Análise, por exemplo, as funções são chamadas, geralmente, de *transformações* ou *aplicações*, reservando-se a designação "função" àquelas cujo contradomínio é um conjunto numérico.

2. Números Reais

Neste texto, consideraremos estabelecido o conjunto dos números reais, denotado por $\mathbb{R}$, como o único corpo que é ordenado e completo, sendo a estrutura de corpo determinada pelas operações de *adição* e *multiplicação*[(i)],

$$(x, y) \mapsto x + y \quad \text{e} \quad (x, y) \mapsto xy, \;\; x, y \in \mathbb{R},$$

a ordem pela relação *menor que*[(ii)] ($<$) e a completude pela propriedade de que todo conjunto limitado superiormente possui um supremo[(iii)].

Convém-nos mencionar que, em $\mathbb{R}$, a compatibilidade da relação $<$ com as operações de adição e multiplicação exprime-se através das propriedades:

- $x < y \Rightarrow x + z < y + z$;
- $x < y$ e $0 < z \Rightarrow xz < yz$.

Daí, conclui-se facilmente que:

- $x < y$ e $z < 0 \Rightarrow yz < xz$;
- $x < y$ e $z < w \Rightarrow x + z < y + w$;
- $0 < x < y$ e $0 < z < w \Rightarrow xz < yw$;
- $0 < x < y \Rightarrow 0 < \frac{1}{y} < \frac{1}{x}\cdot$

Cabe-nos lembrar também que um conjunto $X \subset \mathbb{R}$ é dito *limitado superiormente* se existe $b \in \mathbb{R}$, tal que, para todo $x \in X$, tem-se $x < b$ ou $x = b$, isto é, $x \leq b$. Neste caso, b é dito uma *cota superior* de X. O *supremo* de um tal X, denotado por $\sup X$, é, por definição, a menor de todas as suas cotas superiores.

Analogamente, diz-se que um conjunto $Y \subset \mathbb{R}$ é *limitado inferiormente* se existe $a \in \mathbb{R}$, dito, então, uma *cota inferior* de Y, tal que $a \leq y \, \forall y \in Y$. Neste caso, o *ínfimo* de Y, $\inf Y$, é a maior de suas cotas inferiores.

Em suma,

$$\sup X = \min\{b \in \mathbb{R};\, x \leq b \;\forall x \in X\} \quad \text{e} \quad \inf Y = \max\{a \in \mathbb{R};\, a \leq y \;\forall y \in Y\}.$$

Dentre os subconjuntos de $\mathbb{R}$, destacam-se o dos números *naturais*, o dos *inteiros* e o dos *racionais*, dados, respectivamente, por

$$\mathbb{N} = \{1, 2, 3, \dots\}, \quad \mathbb{Z} = \{\dots, -2, -1, 0, 1, 2, \dots\} \quad \text{e} \quad \mathbb{Q} = \{p/q\,; p, q \in \mathbb{Z},\, q \neq 0\}.$$

Ressaltamos que valem as inclusões $\mathbb{N} \subset \mathbb{Z} \subset \mathbb{Q} \subset \mathbb{R}$. Lembramos ainda que $\mathbb{Q}$ é enumerável e que o conjunto $\mathbb{R} - \mathbb{Q}$, dito dos números *irracionais*, é não-enumerável.

Há também os *intervalos*, que são subconjuntos especiais de $\mathbb{R}$ determinados por dois números reais distintos, $a < b$. São eles:

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

ii) $[a, b] = \{x \in \mathbb{R}; a \leq x \leq b\}$;

iii) $(a, b] = \{x \in \mathbb{R}; a < x \leq b\}$;

iv) $[a, b) = \{x \in \mathbb{R}; a \leq x < b\}$.

[(i)]Juntamente com suas propriedades de associatividade, comutatividade, existência de elemento neutro e inverso e distributividade da multiplicação com respeito à adição.

[(ii)]Juntamente com suas propriedades de transitividade, tricotomia e compatibilidade com respeito à adição e à multiplicação.

[(iii)]Equivalentemente: Todo conjunto limitado inferiormente possui um ínfimo.

De cada um dos quatro intervalos acima, diz-se que é *limitado* e tem *extremos* a e b, sendo o intervalo (a,b) chamado de *aberto* e $[a,b]$ de *fechado*. O intervalo $[a,a]=\{a\}$ é dito *degenerado*. Escrevem-se também

v) $(a,+\infty)=\{x\in\mathbb{R}; a<x\}$;

vi) $[a,+\infty)=\{x\in\mathbb{R}; a\leq x\}$;

vii) $(-\infty,b]=\{x\in\mathbb{R}; x\leq b\}$;

viii) $(-\infty,b)=\{x\in\mathbb{R}; x<b\}$;

e denomina-se cada um dos conjuntos de (v) a (viii) de *intervalo ilimitado*.

Neste momento, merece menção o fato de que o conjunto dos números racionais é *denso* em $\mathbb{R}$, significando que em qualquer intervalo não degenerado de $\mathbb{R}$ existe pelo menos um número racional. O mesmo vale para o conjunto dos números irracionais $\mathbb{R}-\mathbb{Q}$. Evidentemente, $\mathbb{N}$ e $\mathbb{Z}$ não têm esta propriedade.

Em consideração à completude de $\mathbb{R}$, devemos mencionar, igualmente, duas de suas muitas consequências. Primeiramente, que ela permite que se estenda a $\mathbb{R}$ as operações de adição, multiplicação e potenciação definidas em $\mathbb{Q}$. Além disso, ela implica na *propriedade arquimediana* de $\mathbb{R}$, qual seja:

> Dados números reais $a,b>0$, existe $k\in\mathbb{N}$, tal que $a<kb$.

A fim de ilustrar a sua efetividade, apliquemos a propriedade arquimediana de $\mathbb{R}$ para constatar que, dados $x>0$ e $0<y<1$, fazendo-se

$$X=\left\{\frac{x}{k}\,;k\in\mathbb{N}\right\} \qquad \text{e} \qquad Y=\{y^k\,;k\in\mathbb{N}\},$$

tem-se $\inf X=\inf Y=0$.

Claramente, 0 é uma cota inferior de ambos estes conjuntos. Verifiquemos, então, que nenhum real $r>0$ é cota inferior de X ou Y. De fato, pela propriedade arquimediana, existe $k>0$, tal que $x<rk$, donde $X\ni x/k<r$. Logo, r não é cota inferior de X.

No caso do conjunto Y, observemos inicialmente que, para todo $a>0$ e todo $k\in\mathbb{N}$, vale a *desigualdade de Bernoulli*, $(1+a)^k\geq 1+ka$, a qual se prova facilmente por indução. Dado, então, $y\in(0,1)$, para algum $a>0$, $1<1/y=1+a$, e, pela propriedade arquimediana, existe $k\in\mathbb{N}$, tal que $ka>1/r$. Assim,

$$\frac{1}{y^k}=(1+a)^k\geq 1+ka>ka>\frac{1}{r}\,,$$

isto é, $Y\ni y^k<r$, donde r não é cota inferior de Y.

Analogamente, verifica-se que, para todo $x'<0$, tem-se $\sup\{\frac{x'}{k}\,;k\in\mathbb{N}\}=0$.

O *módulo* ou *valor absoluto* de um número real x é o número $|x|$, definido da seguinte forma:

$$|x|=\begin{cases} x & \text{se} \quad x\geq 0\\ -x & \text{se} \quad x<0.\end{cases}$$

Note que, para todo $x\in\mathbb{R}$, valem as igualdades

$$|x|=\max\{x,-x\}=\sqrt{x^2}\,.$$

Além disso, verificam-se, para quaisquer $x,y\in\mathbb{R}$, as seguintes propriedades:

- $|x| \geq 0$ e $|x| = 0 \Leftrightarrow x = 0$;
- $|xy| = |x||y|$;
- $|x + y| \leq |x| + |y|$;
- $||x| - |y|| \leq |x - y|$.

Observemos ainda que, dados $a, x \in \mathbb{R}$ e $r > 0$, tem-se

- $|x - a| < r \Leftrightarrow x \in (a - r, a + r)$;
- $|x - a| \leq r \Leftrightarrow x \in [a - r, a + r]$.

Geometricamente, o conjunto $\mathbb{R}$ representa-se por uma reta ℓ, em que cada ponto de ℓ é associado a um único número real e vice-versa. Esta associação é feita através da seguinte regra: Se um número real x é menor que y, então o ponto de ℓ associado a x está à esquerda daquele associado a y. Assim, cada intervalo não degenerado e limitado de $\mathbb{R}$ corresponde a um segmento de reta de ℓ cujos extremos correspondem àqueles do intervalo (daí a terminologia).

Define-se, então, a *distância* entre dois pontos $x, y \in \mathbb{R}$ por $d(x, y) = |x - y|$, o que corresponde ao comprimento do segmento de ℓ determinado por qualquer intervalo cujos extremos sejam x e y (Fig. 4).

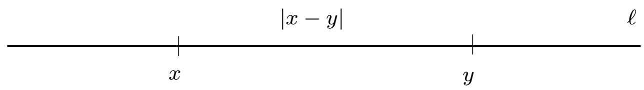

FIGURA 4

Devido a esta representação, é comum referir-se aos números reais como pontos e designar-se ℓ a *reta real*.

2.1. Sequências em $\mathbb{R}$. Diz-se que uma sequência $(x_k)_{k \in \mathbb{N}}$ em $\mathbb{R}$ é *limitada inferiormente* (respectivamente *superiormente*) se o conjunto dos seus termos, $\{x_k\,;\, k \in \mathbb{N}\}$, é limitado inferiormente (respectivamente superiormente). Uma sequência é dita *limitada* se é limitada inferiormente e superiormente. Assim, uma sequência (x_k) em $\mathbb{R}$ é limitada se, e somente se, existem $a, b \in \mathbb{R}$, tais que $x_k \in (a, b)\, \forall k \in \mathbb{N}$. Neste caso, tomando-se $\mu > 0$, tal que $\mu > b$ e $-\mu < a$, tem-se

$$|x_k| < \mu \;\forall k \in \mathbb{N}.$$

Um número real $a \in \mathbb{R}$ é dito um *limite* de uma sequência (x_k) se cumpre a seguinte condição:

> Para todo $\epsilon > 0$, existe $k_0 \in \mathbb{N}$, tal que $k \geq k_0 \Rightarrow |x_k - a| < \epsilon$.

No caso afirmativo, diz-se que (x_k) *converge* para a e que (x_k) é *convergente*, caso contrário, diz-se que (x_k) é *divergente*.

Note que $|x_k - a| < \epsilon \Leftrightarrow x_k \in (a - \epsilon, a + \epsilon)$. A partir daí, verifica-se facilmente a unicidade do limite de uma sequência convergente. De fato, se $a, b \in \mathbb{R}$ fossem limites distintos de uma sequência (x_k), poderíamos tomar $\epsilon = |a-b|/2 > 0$. Neste caso, tomando-se $k \in \mathbb{N}$ suficientemente grande, teríamos

$$x_k \in (a-\epsilon, a+\epsilon) \cap (b-\epsilon, b+\epsilon) = \emptyset.$$

Logo, devemos ter $a = b$.

Quando $a \in \mathbb{R}$ é o limite de uma sequência convergente (x_k), escrevemos

$$x_k \to a \qquad \text{ou} \qquad \lim x_k = a.$$

Observemos que se $(x_k), (y_k)$ e (z_k) são sequências em $\mathbb{R}$ que satisfazem

i) $x_k \leq y_k \leq z_k \ \forall k \in \mathbb{N}$;

ii) $\lim x_k = \lim z_k = a$;

então (y_k) é convergente(iv) e $\lim y_k = a$.

Com efeito, pela condição (ii), dado $\epsilon > 0$, existe $k_0 \in \mathbb{N}$, tal que, para todo $k \geq k_0$, tem-se $|x_k - a| < \epsilon$ e $|z_k - a| < \epsilon$. Segue-se daí e de (i) que, para tais valores de k, $-\epsilon < x_k - a \leq y_k - a \leq z_k - a < \epsilon$, donde $|y_k - a| < \epsilon \ \forall k \geq k_0$, isto é, $y_k \to a$.

Quando um conjunto $X \subset \mathbb{R}$ é limitado superiormente, tem-se que $a = \sup X$ é limite de uma sequência (x_k), tal que $x_k \in X \ \forall k \in \mathbb{N}$. De fato, dado $k \in \mathbb{N}$, existe pelo menos um elemento de X, o qual designamos por x_k, o qual satisfaz $a - \frac{1}{k} < x_k \leq a$. Caso contrário $a - \frac{1}{k}$ seria uma cota superior de X menor que a, o que iria de encontro ao fato de a ser o supremo de X. A sequência (x_k), assim definida, cumpre $|x_k - a| < 1/k$. No entanto, pela propriedade arquimediana de $\mathbb{R}$, dado $\epsilon > 0$, existe $k_0 \in \mathbb{N}$, tal que $1 < k_0\epsilon$. Logo, para todo $k \geq k_0$, tem-se

$$|x_k - a| < \frac{1}{k} \leq \frac{1}{k_0} < \epsilon,$$

donde $x_k \to a$.

De modo análogo, verifica-se que se $Y \subset \mathbb{R}$ é limitado inferiormente, então o ínfimo de Y é limite de uma sequência de pontos de Y.

Diz-se que uma sequência (x_k) em $\mathbb{R}$ é *monótona* quando cumpre uma das seguintes condições:

- $x_{k+1} \geq x_k \ \forall k \in \mathbb{N}$;
- $x_{k+1} \leq x_k \ \forall k \in \mathbb{N}$.

No primeiro caso, diz-se que (x_k) é *monótona crescente* e, no segundo, *monótona decrescente*.

Verifiquemos agora que *toda sequência monótona e limitada é convergente.* Mais especificamente, se (x_k) e (y_k) são sequências limitadas em $\mathbb{R}$, tais que (x_k) é monótona crescente e (y_k) é monótona decrescente, então

$$x_k \to a = \sup\{x_k\,;k \in \mathbb{N}\} \quad \text{e} \quad y_k \to b = \inf\{y_k\,;k \in \mathbb{N}\}.$$

De fato, dado $\epsilon > 0$, existe $k_0 \in \mathbb{N}$, tal que $a - \epsilon < x_{k_0} \leq a$, donde, para todo $k \geq k_0$, $a - \epsilon \leq x_{k_0} \leq x_k \leq a$, pois (x_k) é crescente e a é cota superior do conjunto $\{x_k\,;k \in \mathbb{N}\}$. Assim, $k \geq k_0 \Rightarrow |x_k - a| < \epsilon$, isto é, $x_k \to a$. De forma análoga, verifica-se a convergência de (y_k).

(iv) Este resultado é conhecido como *Teorema do Confronto*.

Dados $x' < 0 < x$ e $0 \leq y < 1$, vimos que

$$\sup\left\{\frac{x'}{k}\,; k \in \mathbb{N}\right\} = \inf\left\{\frac{x}{k}\,; k \in \mathbb{N}\right\} = \inf\{y^k\,; k \in \mathbb{N}\} = 0.$$

Logo, pelo exposto no parágrafo anterior, tem-se

- $\frac{x}{k} \to 0\,\forall x \in \mathbb{R}$;
- $x^k \to 0\,\forall x \in [0,1)$;

pois as sequências cujos termos gerais são $\frac{x}{k}$ e x^k são, ambas, monótonas.

Vejamos agora um resultado clássico da Análise Real relacionado ao conceito de convergência de sequências.

Teorema de Bolzano-Weierstrass. *Toda sequência limitada em $\mathbb{R}$ possui uma subsequência convergente.*

Demonstração. Seja (x_k) uma sequência limitada em $\mathbb{R}$. Então, existe um intervalo aberto $I = (\lambda, \mu)$, tal que $x_k \in I$ para todo $k \in \mathbb{N}$. Definindo-se

$$\Omega = \{t \in \mathbb{R};\, t \leq x_k \text{ para infinitos valores de } k\},$$

vemos que este conjunto é não-vazio (pois $\lambda \in \Omega$) e limitado superiormente (pois μ é cota superior de Ω). Logo, existe $a = \sup \Omega$. Provaremos que a é limite de uma subsequência de (x_k).

Pela escolha de a, tem-se, para cada $i \in \mathbb{N}$, que:

i) $\left(a - \frac{1}{i}, a\right] \cap \Omega \neq \emptyset$;

ii) $a + \frac{1}{i}$ não pertence a Ω.

Segue-se de (i) que existem infinitos termos de (x_k) que são maiores que $a - \frac{1}{i}$, isto é, $a - \frac{1}{i} \in \Omega$, e de (ii) que existe, no máximo, um número finito de termos de (x_k) que são maiores que $a + \frac{1}{i}$. Assim, para todo $i \in \mathbb{N}$, existem infinitos termos de (x_k) no intervalo $I_i = \left(a - \frac{1}{i}\,,\, a + \frac{1}{i}\right)$. Podemos, então, tomar um deles no intervalo I_1 e denotá-lo por x_{k_1}. Uma vez que existem infinitos termos de (x_k) em I_2, existe $k_2 \in \mathbb{N}$, tal que $k_1 < k_2$ e $x_{k_2} \in I_2$. Procedendo-se indutivamente, obtém-se uma subsequência (x_{k_i}), de (x_k), tal que, para cada $i \in \mathbb{N}$, $x_{k_i} \in I_i$, isto é, $|x_{k_i} - a| < \frac{1}{i}\cdot$ Logo, $x_{k_i} \to a$. $\square$

Em conclusão às nossas considerações sobre funções e números reais, devemos mencionar que, além do conjunto $\mathbb{R}$, consideraremos estabelecidas, juntamente com suas propriedades de continuidade e diferenciabilidade, as funções reais elementares de uma variável real, isto é, as funções polinomiais, as funções exponencial e logarítmica e as funções trigonométricas.

CAPíTULO 1

O Espaço Vetorial Normado $\mathbb{R}^n$

A Análise, enquanto teoria das funções reais de uma variável, está subjugada a uma forte estrutura algébrica dos números reais, qual seja, a de corpo — ordenado e completo, inclusive. Neste contexto, a noção de valor absoluto, em $\mathbb{R}$, é também essencial, uma vez que, a partir dela, se estabelece um conceito natural de distância que, por sua vez, conduz ao de limite de funções.

Em contrapartida, os espaços euclidianos multidimensionais, em geral, não dispõem da estrutura de corpo. Desta forma, a fim de estender às funções definidas nesses espaços os conceitos e resultados da Análise em $\mathbb{R}$, é necessário que se faça uso da estrutura algébrica que lhes é comum — a de espaço vetorial — e que a ela se incorpore um análogo do valor absoluto — o que se chama de norma. Neste caso, o espaço vetorial em questão é dito normado.

Sendo assim, com o propósito de estabelecer o ambiente adequado às nossas considerações, introduziremos neste capítulo o espaço euclidiano n-dimensional, $\mathbb{R}^n$, como um espaço vetorial normado.

Para tanto, iniciaremos discutindo o conceito de norma e, logo após, relembraremos algumas noções elementares da Álgebra Linear, como produtos internos, formas bilineares e determinantes, as quais surgirão naturalmente em diversos contextos da teoria que desenvolveremos nos capítulos subsequentes.

Consideraremos, em seguida, o espaço das transformações lineares de $\mathbb{R}^n$ em $\mathbb{R}^m$ munido de uma norma com propriedades especiais, dita espectral. Abordaremos, então, o conceito fundamental de convergência de sequências, estabelecendo, inclusive, alguns resultados essenciais que o envolvem, como o Teorema de Bolzano-Weierstrass. Feito isto, discutiremos brevemente sobre o conceito de espaço métrico, que generaliza aquele de espaço vetorial normado.

Por fim, com o intuito de ilustrar a aplicabilidade dos assuntos discutidos, apresentaremos um caso particular de um importante resultado da Teoria dos Sistemas Dinâmicos, conhecido como Teorema Ergódico da Média.

1. Normas em $\mathbb{R}^n$

Dado um número natural n, o *espaço euclidiano n-dimensional*, denotado por $\mathbb{R}^n$, é o produto cartesiano de n cópias de $\mathbb{R}$, isto é,

$$\mathbb{R}^n = \underbrace{\mathbb{R} \times \mathbb{R} \times \cdots \times \mathbb{R} \times \mathbb{R}}_{n \text{ vezes}}.$$

Um elemento $x \in \mathbb{R}^n$ é dito um *ponto* de $\mathbb{R}^n$ e é, então, denotado por uma n-upla de números reais, isto é, $x = (x_1, x_2, \ldots, x_n)$, em que, para cada $i \in \{1, \ldots, n\}$, x_i é um número real, dito a i-ésima *coordenada* de x.

Dados $x = (x_1, x_2, \ldots, x_n)$, $y = (y_1, y_2, \ldots, y_n)$ em $\mathbb{R}^n$ e $\lambda \in \mathbb{R}$, definimos a *soma*, $x + y$, e o *produto*, λx, por

$$x + y = (x_1 + y_1, x_2 + y_2, \ldots, x_n + y_n) \qquad \text{e} \qquad \lambda x = (\lambda x_1, \lambda x_2, \ldots, \lambda x_n).$$

Estas operações, juntamente com suas propriedades(i), concedem a $\mathbb{R}^n$ uma estrutura de *espaço vetorial real*(ii). Neste caso, os elementos de $\mathbb{R}^n$ são também chamados de *vetores.* Note que

$$x = (x_1, x_2, \ldots, x_n) = x_1 e_1 + x_2 e_2 + \cdots + x_n e_n ,$$

em que $\{e_1, e_2, \ldots, e_n\}$ é a *base canônica* de $\mathbb{R}^n$, isto é, cada vetor e_i tem todas as suas coordenadas nulas, exceto pela i-ésima, que é igual a 1. Os reais x_i são ditos, então, mais especificamente, as *coordenadas* de x *com respeito à base canônica de* $\mathbb{R}^n$.

A *norma* (*euclidiana*) de um vetor $x = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$ é, por definição, o número real $\|x\|$, dado por

$$\|x\| = \sqrt{x_1^2 + x_2^2 + \cdots + x_n^2}\,.$$

A *distância* (*euclidiana*) entre $x, y \in \mathbb{R}^n$, $d(x, y)$, é, então, definida por

$$d(x, y) = \|x - y\|.$$

Considerando-se em $\mathbb{R}^n$ um sistema cartesiano de coordenadas, tem-se que, geometricamente, a distância entre dois pontos corresponde ao comprimento do segmento de reta que os une. Em particular, para todo $x \in \mathbb{R}^n$, $\|x\| = d(x, 0)$ é o comprimento do segmento cujos extremos são a *origem*, 0, e o ponto x (Fig. 1).

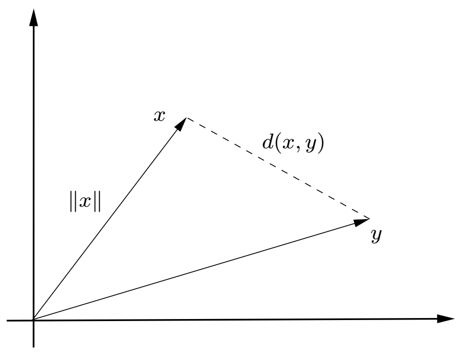

FIGURA 1

Dados $x, y \in \mathbb{R}^n$ e $\lambda \in \mathbb{R}$, verificam-se as seguintes propriedades:

i) $\|x\| \geq 0$ e $\|x\| = 0 \Leftrightarrow x = 0$;

ii) $\|\lambda x\| = |\lambda|\, \|x\|$;

iii) $\|x + y\| \leq \|x\| + \|y\|$.

(i) Comutatividade, distributividade, existência de elemento neutro e inverso.

(ii) Ao longo do texto, todos os espaços vetoriais serão considerados, implicitamente, reais, isto é, terão $\mathbb{R}$ como corpo de escalares.

A validez de (i) e (ii) é imediata. Já a propriedade (iii), chamada de *desigualdade triangular*, é uma consequência de uma outra desigualdade, dita *de Cauchy-Schwarz*, que provaremos adiante.

De modo geral, dado um espaço vetorial $\mathbb{V}$, uma *norma* em $\mathbb{V}$ é uma função $\| \ \| : \mathbb{V} \to \mathbb{R}$ que, para quaisquer $x, y \in \mathbb{V}$ e $\lambda \in \mathbb{R}$, cumpre as condições (i), (ii) e (iii) listadas acima. Um espaço vetorial $\mathbb{V}$ munido de uma norma $\| \ \|$ é dito *normado* e é denotado por $(\mathbb{V}, \| \ \|)$ quando há necessidade de se especificar a norma.

Neste contexto, a norma euclidiana em $\mathbb{R}^n$ é apenas um caso particular. Há, por exemplo, duas outras normas em $\mathbb{R}^n$ que, eventualmente, fazem-se mais convenientes que a norma euclidiana. São elas a *norma do máximo* e a *norma da soma*, as quais definem-se, respectivamente, por

$$\|x\|_{\max} = \max\{|x_1|, \dots, |x_n|\} \quad \text{e} \quad \|x\|_{\mathrm{s}} = |x_1| + \cdots + |x_n|, \quad x = (x_1, \dots, x_n) \in \mathbb{R}^n.$$

Como no caso da norma euclidiana, as propriedades (i) e (ii) são facilmente verificadas. Provemos, então, que $\| \ \|_{\max}$ satisfaz a desigualdade triangular. De fato, dados $x, y \in \mathbb{R}^n$, para algum $i \in \{1, 2, \dots, n\}$, tem-se

$$\begin{aligned} \|x + y\|_{\max} &= \max\{|x_1 + y_1|, \dots, |x_n + y_n|\} \\ &= |x_i + y_i| \leq |x_i| + |y_i| \\ &\leq \max\{|x_1|, \dots, |x_n|\} + \max\{|y_1|, \dots, |y_n|\} \\ &= \|x\|_{\max} + \|y\|_{\max}. \end{aligned}$$

Deixamos como exercício a verificação de que $\| \ \|_{\mathrm{s}}$, igualmente, satisfaz a desigualdade triangular.

Adotaremos a seguinte notação: Dado $x \in \mathbb{R}^n$, salvo menção em contrário, $\|x\|$ indicará a norma euclidiana de x. As normas do máximo e da soma serão denotadas como acima.

Vejamos agora que, dado $x = (x_1, \dots, x_n) \in \mathbb{R}^n$, para algum $i \in \{1, 2, \dots, n\}$, tem-se

$$\begin{aligned} \|x\|_{\max} &= |x_i| = \sqrt{x_i^2} \leq \sqrt{x_1^2 + \cdots + x_n^2} \quad (= \|x\|) \\ &\leq \sqrt{x_1^2} + \cdots + \sqrt{x_n^2} = |x_1| + \cdots + |x_n| \quad (= \|x\|_{\mathrm{s}}) \\ &\leq n|x_i| = n\|x\|_{\max}, \end{aligned}$$

isto é,

$$\|x\|_{\max} \leq \|x\| \leq \|x\|_{\mathrm{s}} \leq n\|x\|_{\max} \quad \forall x \in \mathbb{R}^n. \tag{2}$$

Na Seção 4, introduziremos o importante conceito de sequência convergente, o qual, por sua vez, relaciona-se com o de norma. Veremos então que, em decorrência das desigualdades (2), se uma sequência for convergente relativamente a uma destas normas, então o será com relação às outras duas, isto é, do ponto de vista da convergência de sequências, as normas euclidiana, do máximo e da soma são equivalentes. Isto motiva a definição seguinte.

Definição 1 (equivalência de normas). Duas normas $\| \ \|_1$ e $\| \ \|_2$ num espaço vetorial $\mathbb{V}$ são ditas *equivalentes* se existem constantes positivas $\lambda, \mu \in \mathbb{R}$, tais que, para todo $x \in \mathbb{V}$,

$$\|x\|_1 \leq \lambda \|x\|_2 \qquad \text{e} \qquad \|x\|_2 \leq \mu \|x\|_1 .$$

Segue-se diretamente da definição que a equivalência entre normas é, de fato, uma relação de equivalência, isto é, ela é reflexiva (toda norma é equivalente a ela mesma), simétrica (se uma norma é equivalente à outra, esta é equivalente à primeira) e transitiva (se uma norma é equivalente a uma segunda que, por sua vez, é equivalente a uma terceira, então a primeira é equivalente à última).

Note que, pelas desigualdades (2), as normas euclidiana, do máximo e da soma são, duas a duas, equivalentes. Na verdade, vale um resultado bem mais forte, que provaremos na Seção 4: *Duas normas quaisquer em* $\mathbb{R}^n$ *são equivalentes.*

Sejam $\mathbb{V}$ um espaço vetorial, $(\mathbb{W}, \| \ \|)$ um espaço vetorial normado e $T : \mathbb{V} \to \mathbb{W}$ uma aplicação linear injetiva. Fazendo-se $\|x\|_0 = \|Tx\|$, $x \in \mathbb{V}$, verifica-se facilmente que $\| \ \|_0$ define uma norma em $\mathbb{V}$, dita *induzida* por $\| \ \|$ e T. Note que a injetividade de T é necessária para que se tenha $\|x\|_0 = 0 \Leftrightarrow x = 0$.

Em particular, todo espaço vetorial de dimensão finita, por ser isomorfo a algum espaço $\mathbb{R}^n$, admite uma norma.

Exemplo 1. Designando-se por $\mathrm{M}(2)$ o espaço vetorial formado pelas matrizes quadradas de ordem 2, temos que a correspondência que associa a cada matriz

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

de $\mathrm{M}(2)$ o vetor $v = (a, b, c, d)$ de $\mathbb{R}^4$ é um isomorfismo linear. Este isomorfismo, juntamente com a norma euclidiana de $\mathbb{R}^4$, induz a norma $\| \ \|_0$ em $\mathrm{M}(2)$, em que

$$\|A\|_0 = \sqrt{a^2 + b^2 + c^2 + d^2}.$$

2. Produtos Internos – Determinantes

Relembremos que uma *forma bilinear* num espaço vetorial (real) $\mathbb{V}$ é uma função $f : \mathbb{V} \times \mathbb{V} \to \mathbb{R}$ que é linear em cada uma de suas variáveis, isto é, dados $x, y, z \in \mathbb{V}$ e $\lambda \in \mathbb{R}$, tem-se

- $f(\lambda x + y, z) = \lambda f(x, z) + f(y, z)$ e
- $f(x, \lambda y + z) = \lambda f(x, y) + f(x, z)$.

Um *produto interno* em $\mathbb{V}$ é uma forma bilinear f que é simétrica (isto é, $f(x, y) = f(y, x) \, \forall x, y \in \mathbb{V}$) e positiva definida (isto é, $f(x, x) > 0 \, \forall x \in \mathbb{V} - \{0\}$).

Dados $x = (x_1, \dots, x_n)$ e $y = (y_1, \dots, y_n)$ em $\mathbb{R}^n$, o *produto escalar* $\langle x, y \rangle$ é definido por

$$\langle x, y \rangle = x_1 y_1 + \cdots + x_n y_n \, . \tag{3}$$

Verifica-se facilmente que o produto escalar é um produto interno em $\mathbb{R}^n$ (dito *canônico*) e que a norma euclidiana $\| \ \|$ satisfaz

$$\|x\|^2 = \langle x, x \rangle \ \ \forall \, x \in \mathbb{R}^n.$$

De modo geral, um produto interno $\langle \, , \, \rangle$ num espaço vetorial $\mathbb{V}$ determina uma norma $\| \ \|$ através da relação $\|x\| = \sqrt{\langle x, x \rangle}$, $x \in \mathbb{V}$. Neste caso, diz-se que a norma $\| \ \|$ *provém* do produto interno $\langle \, , \, \rangle$. Temos, então, que a norma euclidiana em $\mathbb{R}^n$ provém do produto escalar.

Vale salientar que nem toda norma em $\mathbb{R}^n$ provém de um produto interno. Este é o caso, por exemplo, das normas da soma e do máximo (vide Exercício 6).

Dados vetores v e w de um espaço vetorial $\mathbb{V}$ munido de um produto interno $\langle\ ,\ \rangle$, vale a implicação seguinte:

$$\langle x, v\rangle = \langle x, w\rangle \ \forall x \in \mathbb{V} \ \Rightarrow \ v = w.$$

Com efeito, fazendo-se $x = v - w$, obtém-se $\langle v-w, v-w\rangle = 0$, donde $\|v-w\| = 0$ e, portanto, $v = w$.

Observação 1. Dado um produto interno qualquer em $\mathbb{R}^n$, prova-se que existe uma base relativamente a qual sua expressão assume a forma (3). Por este motivo, e por simplicidade, *o produto escalar será o único produto interno de* $\mathbb{R}^n$ *que consideraremos.*

No teorema seguinte, estabeleceremos a desigualdade fundamental da álgebra linear dos espaços vetoriais com produto interno.

Teorema 1 (desigualdade de cauchy-schwarz). *Dados* $x, y \in \mathbb{R}^n$, *tem-se*

$$|\langle x, y\rangle| \le \|x\|\,\|y\|,$$

valendo a igualdade se, e somente se, um dos vetores x, y *é múltiplo do outro.*

Demonstração. O resultado é imediato para $x = 0$ ou $y = 0$. Supondo-se, então, $x, y \neq 0$, definimos a função $f(t) = \langle x - ty, x - ty\rangle$, $t \in \mathbb{R}$. Note que f é não-negativa e que f possui um zero se, e somente se, x é múltiplo de y. Além disso,

$$f(t) = \langle x - ty, x - ty\rangle = \|y\|^2t^2 - 2\langle x, y\rangle t + \|x\|^2,$$

isto é, f é uma função quadrática não-negativa de t. Logo, seu discriminante deve ser não-positivo, isto é,

$$4\langle x, y\rangle^2 - 4\|x\|^2\|y\|^2 \le 0,$$

em que vale a igualdade se, e só se, f possui um único zero. □

Observação 2. A desigualdade de Cauchy-Schwarz é válida em qualquer espaço vetorial normado $(\mathbb{V}, \|\ \|)$ cuja norma é advinda de um produto interno $\langle\ ,\ \rangle$. A demonstração é a mesma. Devido a isto, pode-se introduzir num tal espaço o conceito de ângulo entre vetores: Dados $v, w \in \mathbb{V} - \{0\}$, o *ângulo* entre v e w, denotado por $\measuredangle(v, w)$, é o único $\theta \in [0, \pi]$, tal que[(iii)]

$$\cos\theta = \frac{\langle v, w\rangle}{\|v\|\,\|w\|}\cdot$$

Como aplicação da desigualdade de Cauchy-Schwarz, verifiquemos agora que a norma euclidiana satisfaz a desigualdade triangular. De fato, dados $x, y \in \mathbb{R}^n$, temos

$$\|x+y\|^2 = \|x\|^2 + 2\langle x, y\rangle + \|y\|^2 \le \|x\|^2 + 2\|x\|\,\|y\| + \|y\|^2 = (\|x\| + \|y\|)^2,$$

donde se obtém

$$\|x+y\| \ \le \ \|x\| + \|y\|.$$

A partir da desigualdade triangular, verifica-se facilmente que, para quaisquer $x, y \in \mathbb{R}^3$ (vide Exercício 2), tem-se

$$|\,\|x\| - \|y\|\,| \le \|x - y\|. \tag{4}$$

[(iii)]Note que, pela desigualdade de Cauchy-Schwarz, $-1 \le \frac{\langle v,w\rangle}{\|v\|\,\|w\|} \le 1$, o que garante a existência e unicidade de θ.

Além disso, usando-se indução, prova-se que se $x_1, x_2, \ldots, x_n$ são vetores de $\mathbb{R}^n$, então

$$\|x_1 + x_2 + \cdots + x_n\| \leq \|x_1\| + \|x_2\| + \cdots + \|x_n\|.$$

2.1. Ortogonalidade. Dois vetores x e y de $\mathbb{R}^n$ são ditos *ortogonais* se $\langle x, y\rangle = 0$. Um subconjunto de $\mathbb{R}^n$ é dito *ortogonal* se seus elementos são, dois a dois, ortogonais. Um conjunto ortogonal cujos elementos têm, todos, norma euclidiana igual a 1 é dito *ortonormal*.

Dados vetores $x, y \in \mathbb{R}^n$, tem-se

$$\|x + y\|^2 = \langle x + y, x + y\rangle = \|x\|^2 + 2\langle x, y\rangle + \|y\|^2,$$

donde se obtém o resultado seguinte.

PROPOSIÇÃO 1. *Dois vetores $x, y \in \mathbb{R}^n$ são ortogonais se, e somente se, cumprem a igualdade $\|x + y\|^2 = \|x\|^2 + \|y\|^2$.*

Vale salientar que a parte "somente se" da Proposição 1 constitui o célebre *Teorema de Pitágoras.*

Dada uma base $\mathfrak{B} = \{v_1, \ldots, v_m\} \subset \mathbb{V}$ de um subespaço vetorial m-dimensional $\mathbb{V}$ de $\mathbb{R}^n$, obtém-se uma base ortogonal $\mathfrak{B}_0 = \{w_1, \ldots, w_m\}$ deste subespaço através das seguintes relações de recorrência:

$$w_1 = v_1\,, \quad w_{i+1} = v_{i+1} - \sum_{j=1}^{i} \frac{\langle v_{i+1}, w_j\rangle}{\langle w_j, w_j\rangle} w_j\,, \quad i = 1, \ldots, m-1.$$

Este método é conhecido como o *processo de ortogonalização de Gram-Schmidt*. Note que, a partir da base $\mathfrak{B}_0 = \{w_1, \ldots, w_m\}$, obtém-se uma base ortonormal de $\mathbb{V}$, $\mathfrak{B}_1 = \{u_1, \ldots, u_m\}$, fazendo-se, para cada $i = 1, \cdots, m$, $u_i = w_i/\|w_i\|$.

EXEMPLO 2 (PROJEÇÃO ORTOGONAL). Seja $\mathbb{V} \subset \mathbb{R}^n$ um subespaço vetorial m-dimensional de $\mathbb{R}^n$. Dado um vetor $x \in \mathbb{R}^n$, existe um único vetor $v \in \mathbb{V}$, tal que $x - v$ é ortogonal a todos os vetores de $\mathbb{V}$. O vetor v é chamado de *projeção ortogonal* de x sobre $\mathbb{V}$ (Fig. 2). Para determiná-lo, fixemos uma base ortonormal $\mathfrak{B} = \{u_1, \ldots, u_m\}$ de $\mathbb{V}$ e observemos que um vetor de $\mathbb{R}^n$ é ortogonal a todos os vetores de $\mathbb{V}$ se, e somente se, é ortogonal a cada um dos vetores da base $\mathfrak{B}$. Fazendo-se, então, $v = \lambda_1 u_1 + \cdots + \lambda_m u_m$ e impondo-se a condição $\langle x - v, u_i\rangle = 0\ \forall\, i \in \{1, \ldots, m\}$, conclui-se que, para cada i, λ_i é determinado pela igualdade $\lambda_i = \langle x, u_i\rangle$, resultando na existência e unicidade do vetor v.

Segue-se desta unicidade e da bilinearidade do produto interno que, designando-se v por $P_{\mathbb{V}}(x)$, a aplicação *projeção ortogonal* de $\mathbb{R}^n$ sobre $\mathbb{V}$,

$$\begin{array}{rccl} P_{\mathbb{V}}: & \mathbb{R}^n & \to & \mathbb{V} \\ & x & \mapsto & P_{\mathbb{V}}(x) = \displaystyle\sum_{i=1}^{m} \langle x, u_i\rangle u_i\,, \end{array}$$

está bem definida e é linear.

Note que, além da linearidade, toda projeção ortogonal $P_{\mathbb{V}} : \mathbb{R}^n \to \mathbb{V}$ sobre um subespaço vetorial $\mathbb{V}$ de $\mathbb{R}^n$ tem as seguintes propriedades:

- $P_{\mathbb{V}} \circ P_{\mathbb{V}} = P_{\mathbb{V}}$ (idempotência);
- $\langle P_{\mathbb{V}}(x), y\rangle = \langle x, P_{\mathbb{V}}(y)\rangle\ \forall x, y \in \mathbb{R}^n$ (simetria);
- $\langle P_{\mathbb{V}}(x), x\rangle \geq 0\ \forall x \in \mathbb{R}^n$ (positividade);

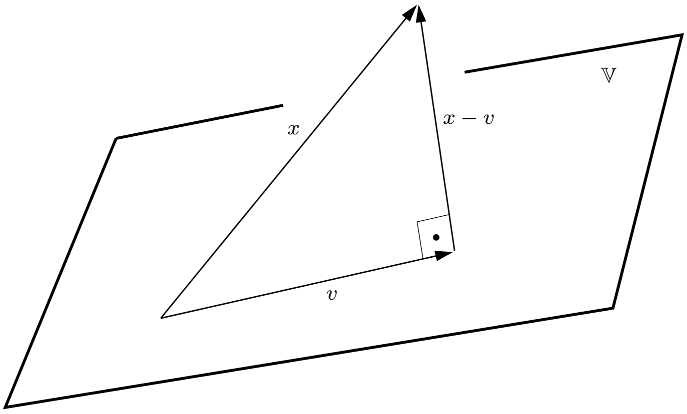

FIGURA 2

- $\|P_{\mathbb{V}}(x)\| \leq \|x\| \, \forall x \in \mathbb{R}^n$ e $\|P_{\mathbb{V}}(x)\| = \|x\| \Leftrightarrow x \in \mathbb{V}$ (semi-contratilidade).

Vale mencionar também a seguinte propriedade da projeção ortogonal: Dentre todos os pontos de um subespaço vetorial $\mathbb{V}$ de $\mathbb{R}^n$, o que está mais próximo de um dado $x \in \mathbb{R}^n$ é a projeção ortogonal de x sobre $\mathbb{V}$, $P_{\mathbb{V}}(x)$. Com efeito, dado $w \in \mathbb{V}$, $w \neq P_{\mathbb{V}}(x)$, temos que $P_{\mathbb{V}}(x) - w \in \mathbb{V}$. Logo, $x - P_{\mathbb{V}}(x)$ é ortogonal a $P_{\mathbb{V}}(x) - w$. Assim, pelo Teorema de Pitágoras, $\|x - P_{\mathbb{V}}(x)\|^2 + \|P_{\mathbb{V}}(x) - w\|^2 = \|x - w\|^2$. Uma vez que $\|P_{\mathbb{V}}(x) - w\|^2 > 0$, tem-se, então,

$$\|x - P_{\mathbb{V}}(x)\| < \|x - w\| \; \forall x \in \mathbb{R}^n, \, w \in \mathbb{V} - \{P_{\mathbb{V}}(x)\}. \tag{5}$$

Dados subespaços $\mathbb{V}_1, \mathbb{V}_2, \mathbb{W}$ de $\mathbb{R}^n$, diz-se que $\mathbb{W}$ é a *soma direta* de $\mathbb{V}_1$ e $\mathbb{V}_2$, e escreve-se $\mathbb{W} = \mathbb{V}_1 \oplus \mathbb{V}_2$, se, para todo $w \in \mathbb{W}$, existem únicos $v_1 \in \mathbb{V}_1$ e $v_2 \in \mathbb{V}_2$, tais que $w = v_1 + v_2$. Neste caso, verifica-se facilmente que $\mathbb{V}_1 \cap \mathbb{V}_2 = \{0\}$ e que $\dim \mathbb{W} = \dim \mathbb{V}_1 + \dim \mathbb{V}_2$.

O *complemento ortogonal* de um subespaço $\mathbb{V}$ de $\mathbb{R}^n$ é o conjunto

$$\mathbb{V}^{\perp} = \{w \in \mathbb{R}^n; \, \langle v, w \rangle = 0 \, \forall v \in \mathbb{V}\}.$$

Suas principais propriedades são:

i) $\mathbb{V}^{\perp}$ é um subespaço de $\mathbb{R}^n$;

ii) $\mathbb{V}^{\perp\perp} = \mathbb{V}$;

iii) $\mathbb{R}^n = \mathbb{V} \oplus \mathbb{V}^{\perp}$.

Dado um subespaço $\mathbb{V} \subset \mathbb{R}^n$, devido à propriedade (iii), todo vetor $x \in \mathbb{R}^n$ se escreve de forma única como $x = v + w$, em que $v \in \mathbb{V}$ e $w \in \mathbb{V}^{\perp}$. Observando-se que, nesta igualdade, $v = P_{\mathbb{V}}(x)$, conclui-se facilmente que

$$x \in \mathbb{V}^{\perp} \;\Leftrightarrow\; P_{\mathbb{V}}(x) = 0.$$

2.2. Aplicações n-lineares – Determinantes. Sejam $\mathbb{V}_1, \mathbb{V}_2, \ldots, \mathbb{V}_n$ e $\mathbb{W}$ espaços vetoriais. Uma aplicação $f : \mathbb{V}_1 \times \cdots \times \mathbb{V}_n \to \mathbb{W}$ é dita *n-linear* se, para todo $i \in \{1, \ldots, n\}$, tem-se

$$f(v_1, \ldots, \lambda v_i + v_i', \ldots, v_n) = \lambda f(v_1, \ldots, v_i, \ldots, v_n) + f(v_1, \ldots, v_i', \ldots, v_n)$$

quaisquer que sejam $\lambda \in \mathbb{R}$, $v_j \in \mathbb{V}_j$ $(j = 1, 2, \ldots, i-1, i+1, \ldots, n)$ e $v_i, v_i' \in \mathbb{V}_i$. Quando $\mathbb{V}_1 = \cdots = \mathbb{V}_n = \mathbb{V}$ e $\mathbb{W} = \mathbb{R}$, a aplicação f é dita uma *forma* n-linear ou n-*forma* em $\mathbb{V}$. Aplicações 2-lineares, assim como as 2-formas, são chamadas *bilineares*.

Um exemplo de n-forma em $\mathbb{R}^n$ que convém ser considerado é o *determinante*. Dado $n \geq 2$, o determinante é uma função que associa um número real a cada matriz (real) quadrada de ordem n, $A = (a_{ij})_{n\times n}$, cuja expressão é um polinômio de grau n e variáveis a_{ij}, $i, j \in \{1, \ldots, n\}$. Quando $n = 3$, por exemplo, tem-se

$$\det A = a_{11}a_{22}a_{33} - a_{11}a_{32}a_{23} + a_{21}a_{32}a_{13} - a_{21}a_{12}a_{33} + a_{31}a_{12}a_{23} - a_{31}a_{22}a_{13}\,.$$

Cada matriz quadrada $A = (a_{ij})_{n\times n}$ pode ser identificada com uma n-upla de vetores de $\mathbb{R}^n$, $(x_1, x_2, \ldots, x_n) \in \mathbb{R}^n \times \cdots \times \mathbb{R}^n$, de tal modo que, para cada $j \in \{1, \ldots, n\}$, $x_j \in \mathbb{R}^n$ é o j-ésimo vetor-coluna de A, isto é, $x_j = (a_{1j}, \ldots, a_{nj})$. Feita esta identificação, o determinante passa a ser uma função definida no produto cartesiano de n cópias de $\mathbb{R}^n$, $\mathbb{R}^n \times \cdots \times \mathbb{R}^n$.

Verifica-se, então, que a função determinante é uma n-forma *anti-simétrica* em $\mathbb{R}^n$, isto é, det é n-linear e, para quaisquer $x_1, \ldots, x_i, \ldots, x_j, \ldots, x_n \in \mathbb{R}^n$, tem-se

$$\det(x_1, \ldots, x_i, \ldots, x_j, \ldots, x_n) = -\det(x_1, \ldots, x_j, \ldots, x_i, \ldots, x_n).$$

Em particular, det é uma n-forma *alternada*, isto é,

$$\det(x_1, \ldots, x_i, \ldots, x_j, \ldots, x_n) = 0 \quad \text{sempre que} \quad x_i = x_j\,.$$

Deste último fato e da n-linearidade do determinante, segue-se que toda matriz cujos vetores-coluna são linearmente dependentes tem determinante nulo. Consequentemente, toda matriz cujo determinante é não-nulo é invertível.

Vale mencionar que a função determinante é caracterizada pelo fato de ser a única n-forma anti-simétrica em $\mathbb{R}^n$ que assume o valor 1 em $(e_1, \ldots, e_n)$. Uma das muitas consequências desta caracterização é a validez, para quaisquer matrizes quadradas de mesma ordem, A e B, da igualdade

$$\det(AB) = (\det A)(\det B),$$

da qual decorre imediatamente que toda matriz invertível A tem determinante não-nulo e que o determinante de A^{-1} é $1/\det A$.

Dados $i, j \in \{1, \ldots, n\}$ e uma matriz quadrada de ordem n, $A = (a_{ij})_{n\times n}$, designemos por A_{ij} a matriz quadrada de ordem n cujos vetores-coluna são os mesmos de A, exceto pelo j-ésimo, que é o i-ésimo vetor da base canônica de $\mathbb{R}^n$, e_i. A matriz $B = (b_{ij})_{n\times n}$, em que $b_{ij} = \det A_{ij}$, é dita a *matriz dos cofatores* de A, e sua transposta é chamada de *adjunta* (clássica) de A, a qual se denota por $\operatorname{adj} A$. Verifica-se, então, que

$$A(\operatorname{adj} A) = (\det A)I,$$

em que I denota a matriz identidade de ordem n. Em particular, se A for invertível, tem-se

$$A^{-1} = \frac{1}{\det A}\operatorname{adj} A.$$

Por fim, consideremos uma transformação linear $T : \mathbb{R}^n \to \mathbb{R}^n$ e tomemos duas bases distintas de $\mathbb{R}^n$, $\mathfrak{B}_0$ e $\mathfrak{B}_1$. Denotando-se por A_0 e A_1 as respectivas matrizes de T com respeito a $\mathfrak{B}_0$ e $\mathfrak{B}_1$, temos que as mesmas são *semelhantes*, isto é, existe uma matriz invertível M (neste caso, a matriz de mudança da base $\mathfrak{B}_0$ para a base $\mathfrak{B}_1$), tal que $A_0 = M^{-1}A_1M$, donde

$$\det A_0 = \det(M^{-1}A_1M) = (\det M^{-1})(\det A_1)(\det M) = \det A_1 \,.$$

Assim, os determinantes das matrizes associadas à transformação linear T assumem o mesmo valor, o qual se define como o *determinante* de T e se denota por $\det T$. Deve-se notar, pelas considerações dos parágrafos anteriores, que uma transformação linear $T : \mathbb{R}^n \to \mathbb{R}^n$ é invertível se, e somente se, $\det T \neq 0$.

3. O Espaço $\mathrm{L}(\mathbb{R}^n, \mathbb{R}^m)$ e a Norma Espectral

O conjunto

$$\mathrm{L}(\mathbb{R}^n, \mathbb{R}^m) = \{T : \mathbb{R}^n \to \mathbb{R}^m;\ T \text{ é linear}\},$$

munido das operações usuais de soma de funções e multiplicação de um número real por uma função, é um espaço vetorial real.

Como se sabe, a cada transformação linear $T \in \mathrm{L}(\mathbb{R}^n, \mathbb{R}^m)$, corresponde uma única matriz A do tipo $m \times n$ que é a representação de T nas bases canônicas de $\mathbb{R}^n$ e $\mathbb{R}^m$, respectivamente. Por sua vez, listando-se os elementos de uma matriz $m \times n$, associamo-la facilmente a um único vetor v de $\mathbb{R}^{mn}$. Designando-se, então, o espaço das matrizes reais $m \times n$ por $\mathrm{M}(m, n)$, temos que as correspondências

$$\begin{array}{ccccc} \mathrm{L}(\mathbb{R}^n, \mathbb{R}^m) & \to & \mathrm{M}(m,n) & \to & \mathbb{R}^{mn} \\ T & \mapsto & A & \mapsto & v \end{array}$$

são, claramente, aplicações lineares bijetivas, isto é, $\mathrm{L}(\mathbb{R}^n, \mathbb{R}^m)$, $\mathrm{M}(m,n)$ e $\mathbb{R}^{mn}$ são espaços vetoriais isomorfos. Em particular, cada norma em $\mathbb{R}^{mn}$ induz, através desses isomorfismos, uma norma em $\mathrm{L}(\mathbb{R}^n, \mathbb{R}^m)$, bem como em $\mathrm{M}(m,n)$.

Denotaremos por $\| \ \|_e$ as normas de $\mathrm{L}(\mathbb{R}^n, \mathbb{R}^m)$ e $\mathrm{M}(m,n)$ induzidas pela norma euclidiana de $\mathbb{R}^{mn}$, e as chamaremos, igualmente, de *norma euclidiana* em $\mathrm{L}(\mathbb{R}^n, \mathbb{R}^m)$ e $\mathrm{M}(m,n)$, respectivamente.

Uma transformação linear $T : \mathbb{R}^n \to \mathbb{R}^n$ é também chamada de *operador linear* em $\mathbb{R}^n$. Escreveremos $\mathrm{L}(\mathbb{R}^n)$ e $\mathrm{M}(n)$ para denotar o espaço dos operadores lineares em $\mathbb{R}^n$ e o conjunto das matrizes reais $n \times n$, respectivamente. O conjunto dos operadores lineares invertíveis de $\mathrm{L}(\mathbb{R}^n)$ e o conjunto das matrizes invertíveis de $\mathrm{M}(n)$ serão denotados, respectivamente, por $\mathrm{I}(\mathbb{R}^n)$ e $\mathrm{I}(n)$.

O fato de $\mathrm{L}(\mathbb{R}^n, \mathbb{R}^m)$ e $\mathrm{M}(m,n)$ serem espaços vetoriais isomorfos, nos leva muitas vezes a identificá-los. Por esta razão, dados $x \in \mathbb{R}^n$, $T \in \mathrm{L}(\mathbb{R}^n, \mathbb{R}^m)$ e $Z \in \mathrm{L}(\mathbb{R}^m, \mathbb{R}^p)$, adotam-se as notações Tx e ZT, que sugerem produtos, em lugar das tradicionais $T(x)$ e $Z \circ T$.

Exemplo 3. Consideremos a transformação linear $T : \mathbb{R}^2 \to \mathbb{R}^3$, dada por $T(x_1, x_2) = (x_1 + x_2, 2x_1 - x_2, x_1 + 3x_2)$. Uma vez que $T(1,0) = (1,2,1)$ e $T(0,1) = (1,-1,3)$, temos que a matriz de T com respeito às bases canônicas de $\mathbb{R}^2$ e $\mathbb{R}^3$, respectivamente, é

$$A = \begin{pmatrix} 1 & 1 \\ 2 & -1 \\ 1 & 3 \end{pmatrix}.$$

Associando-a ao vetor $v = (1,1,2,-1,1,3)$, tem-se

$$\|T\|_e = \|A\|_e = \|v\| = \sqrt{17}.$$

Introduz-se em $\mathrm{L}(\mathbb{R}^n,\mathbb{R}^m)$ uma norma com propriedades especiais (vide Proposição 2, abaixo), chamada de *norma espectral* e definida por

$$\|T\| = \sup\{\,\|Tx\|\,;\, x\in\mathbb{R}^n,\ \|x\|=1\,\}\,,\ \ T\in\mathrm{L}(\mathbb{R}^n,\mathbb{R}^m).$$

Verifiquemos, inicialmente, que a função $T\mapsto\|T\|$ está bem definida em $\mathrm{L}(\mathbb{R}^n,\mathbb{R}^m)$. Para tanto, consideremos $T\in\mathrm{L}(\mathbb{R}^n,\mathbb{R}^m)$ e $x=(x_1,\dots,x_n)\in\mathbb{R}^n$, tal que $\|x\|=1$ (neste caso, dizemos que x é um vetor *unitário*). Fazendo-se, então, $\mu=\max\{\|Te_1\|,\dots,\|Te_n\|\}$, tem-se, pela desigualdade triangular e por (2), que

$$\begin{aligned}\|Tx\| &= \|T(x_1e_1+\cdots+x_ne_n)\| \le |x_1|\,\|Te_1\|+\cdots+|x_n|\,\|Te_n\|\\ &\le \mu(|x_1|+\cdots+|x_n|) = \mu\|x\|_{\mathrm{s}} \le \mu n\|x\| = \mu n.\end{aligned}$$

Desta forma, o conjunto $\{\|Tx\|;\, x\in\mathbb{R}^n,\ \|x\|=1\}$ é limitado superiormente e, portanto, possui um supremo, donde se conclui que a norma espectral, como função, está bem definida.

A verificação de que a norma espectral em $\mathrm{L}(\mathbb{R}^n,\mathbb{R}^m)$ é, de fato, uma norma, segue-se das seguintes considerações: Dados dois subconjuntos limitados e não vazios de $\mathbb{R}$, X e Y, tem-se:

- se $\lambda>0$, então $\sup(\lambda X)=\lambda\sup X$, em que $\lambda X=\{\lambda x\,;\, x\in X\}$;
- se $x\le y$ quaisquer que sejam $x\in X$ e $y\in Y$, então $\sup X\le\sup Y$;
- $\sup(X+Y)=\sup X+\sup Y$, em que $X+Y=\{x+y\,;\, x\in X, y\in Y\}$.

Deixamos os detalhes a cargo do leitor.

Exemplo 4. Consideremos a transformação linear $T:\mathbb{R}^2\to\mathbb{R}^2$, dada por $T(x_1,x_2)=(x_1+4x_2,-2x_1+2x_2)$. Supondo-se $\|(x_1,x_2)\|=1$, por um cálculo direto, obtém-se

$$\|T(x_1,x_2)\|^2 = 5x_1^2+20x_2^2 = 5(4-3x_1^2).$$

Segue-se que, sujeito à condição $\|(x_1,x_2)\|=1$, o valor máximo de $\|T(x_1,x_2)\|$ é $\sqrt{20}$, que é atingido nos pontos $(0,1)$ e $(0,-1)$. Logo, $\|T\|=\sqrt{20}$.

O isomorfismo natural entre $\mathrm{L}(\mathbb{R}^n,\mathbb{R}^m)$ e $\mathrm{M}(m,n)$ induz uma norma espectral em $\mathrm{M}(m,n)$, isto é, se $A\in\mathrm{M}(m,n)$, então a *norma espectral* de A, $\|A\|$, é, por definição, igual à norma espectral da transformação linear $T:\mathbb{R}^n\to\mathbb{R}^m$ cuja matriz, com respeito às bases canônicas de $\mathbb{R}^n$ e $\mathbb{R}^m$, é A.

Proposição 2 (propriedades da norma espectral). *A norma espectral, $\|\ \|:\mathrm{L}(\mathbb{R}^n,\mathbb{R}^m)\to\mathbb{R}$, tem as seguintes propriedades:*

i) $\|Tx\|\le\|T\|\,\|x\|\ \ \forall\, T\in\mathrm{L}(\mathbb{R}^n,\mathbb{R}^m),\ x\in\mathbb{R}^n$;

ii) $\|ZT\|\le\|Z\|\,\|T\|\ \ \forall\, T\in\mathrm{L}(\mathbb{R}^n,\mathbb{R}^m),\ Z\in\mathrm{L}(\mathbb{R}^m,\mathbb{R}^p)$.

Demonstração. A desigualdade (i) é trivial para $x=0$. Dado $x\ne0$, temos que $\|T(x/\|x\|)\|\le\|T\|$, pois o vetor $x/\|x\|$ é unitário. Juntando-se isto à linearidade de T, obtém-se

$$\|Tx\| = \|x\|\left\|T\left(\frac{x}{\|x\|}\right)\right\| \le \|T\|\,\|x\|,$$

o que prova (i).

Quanto à desigualdade (ii), tomando-se um vetor unitário $x\in\mathbb{R}^n$ e considerando-se (i), tem-se

$$\|ZTx\|\le\|Z\|\,\|Tx\|\le\|Z\|\,\|T\|\,\|x\| = \|Z\|\,\|T\|.$$

Logo, $\|Z\|\,\|T\|$ é uma cota superior do conjunto $\{\|ZTx\|;\, x \in \mathbb{R}^n,\ \|x\| = 1\}$, o que implica $\|ZT\| \leq \|Z\|\,\|T\|$. $\square$

Segue-se imediatamente do item (ii) desta proposição que, dadas matrizes $A \in \mathrm{M}(p, m)$ e $B \in \mathrm{M}(m, n)$, tem-se

$$\|AB\| \leq \|A\|\,\|B\|.$$

As normas espectral e euclidiana em $\mathrm{L}(\mathbb{R}^n, \mathbb{R}^m)$ são equivalentes. Mais precisamente, valem as seguintes desigualdades

$$\|T\| \leq \|T\|_e \leq \sqrt{n}\,\|T\| \quad \forall T \in \mathrm{L}(\mathbb{R}^n, \mathbb{R}^m). \tag{6}$$

Antes de verificá-las, vejamos alguns conceitos e resultados relativos à teoria de operadores lineares em $\mathbb{R}^n$.

Relembremos inicialmente que, dada $T \in \mathrm{L}(\mathbb{R}^n, \mathbb{R}^m)$, a *adjunta* de T é a transformação linear $T^* \in \mathrm{L}(\mathbb{R}^m, \mathbb{R}^n)$, definida pela relação

$$\langle Tx, y\rangle = \langle x, T^*y\rangle \quad \forall x \in \mathbb{R}^n,\ y \in \mathbb{R}^m.$$

Se A é a matriz de $T \in \mathrm{L}(\mathbb{R}^n, \mathbb{R}^m)$ relativamente a quaisquer bases ortonormais de $\mathbb{R}^n$ e $\mathbb{R}^m$, respectivamente, verifica-se facilmente que a matriz de T^* (com respeito às mesmas bases) é a transposta de A, a qual denotamos por A^*. Daí, conclui-se que:

i) $(T^*)^* = T \quad \forall T \in \mathrm{L}(\mathbb{R}^n, \mathbb{R}^m)$;

ii) $(\lambda T + Z)^* = \lambda T^* + Z^* \quad \forall T, Z \in \mathrm{L}(\mathbb{R}^n, \mathbb{R}^m),\ \lambda \in \mathbb{R}$;

iii) $(ZT)^* = T^* Z^* \quad \forall T \in \mathrm{L}(\mathbb{R}^n, \mathbb{R}^m),\ Z \in \mathrm{L}(\mathbb{R}^m, \mathbb{R}^p)$.

Convém observar que, pela propriedade (ii), acima, a aplicação

$$\begin{array}{ccc} \mathrm{L}(\mathbb{R}^n, \mathbb{R}^m) & \to & \mathrm{L}(\mathbb{R}^m, \mathbb{R}^n) \\ T & \mapsto & T^* \end{array}$$

é linear.

Diz-se que um operador $T \in \mathrm{L}(\mathbb{R}^n)$ é *auto-adjunto* quando $T = T^*$.

Os operadores auto-adjuntos em $\mathbb{R}^n$, também pela propriedade (ii), constituem um subespaço vetorial de $\mathrm{L}(\mathbb{R}^n)$. Sua propriedade mais importante traduz-se no célebre *Teorema Espectral*[iv], o qual assegura, para cada tal T, a existência de uma base ortonormal de $\mathbb{R}^n$, $\mathfrak{B} = \{u_1, \dots, u_n\}$, em que cada vetor u_i é um *autovetor* de T, isto é, $Tu_i = \lambda_i u_i$, $\lambda_i \in \mathbb{R}$. Neste caso, o real λ_i é dito um *autovalor* de T associado ao autovetor u_i.

Vejamos agora que se $T \in \mathrm{L}(\mathbb{R}^n)$ é um operador auto-adjunto cujos autovalores são $\lambda_1, \dots, \lambda_n$, então

$$\|T\| = \max\{|\lambda_1|, \dots, |\lambda_n|\}. \tag{7}$$

Com efeito, seja $\mathfrak{B} = \{u_1, \dots, u_n\}$ uma base ortonormal de $\mathbb{R}^n$ formada por autovetores de T. Dado $x \in \mathbb{R}^n$, $\|x\| = 1$, sejam $t_1, \dots, t_n$ as coordenadas de x com respeito à base $\mathfrak{B}$. Então,

$$Tx = T(t_1u_1 + \cdots + t_nu_n) = t_1Tu_1 + \cdots + t_nTu_n = t_1\lambda_1u_1 + \cdots + t_n\lambda_nu_n\,,$$

[iv]No Apêndice A, forneceremos uma demonstração deste teorema como aplicação da teoria das variedades diferenciáveis.

donde, fazendo-se $\mu = \max\{|\lambda_1|, \dots, |\lambda_n|\}$, obtém-se

$$\|Tx\|^2 = t_1^2\lambda_1^2 + \cdots + t_n^2\lambda_n^2 \le \mu^2(t_1^2 + \cdots + t_n^2) = \mu^2,$$

isto é, $\|Tx\| \le \mu$.

Desta forma, μ é uma cota superior do conjunto $C = \{\|Tx\|; \|x\| = 1\}$. No entanto, $\mu = |\lambda_i| = \|Tu_i\|$ para algum $i = 1, \dots, n$, isto é, $\mu \in C$. Logo, $\|T\| = \sup C = \mu$, como desejávamos provar.

Segue-se das considerações do Exemplo 2 que, para todo subespaço vetorial $\mathbb{V}$ de $\mathbb{R}^n$, a projeção ortogonal sobre $\mathbb{V}$, $P_{\mathbb{V}} : \mathbb{R}^n \to \mathbb{V} \subset \mathbb{R}^n$, é um operador auto-adjunto que cumpre $\|P_{\mathbb{V}}\| = 1$.

Consideremos agora um outro tipo especial de operador em $\mathbb{R}^n$, dito ortogonal, que pode ser introduzido a partir do conceito de operador adjunto. Mais precisamente, um operador linear $T \in \mathrm{L}(\mathbb{R}^n)$ é chamado de *ortogonal* se $TT^* = I$, em que I é o operador identidade de $\mathbb{R}^n$. Em particular, todo operador ortogonal T é invertível e satisfaz $T^{-1} = T^*$.

Observemos que se $T \in \mathrm{L}(\mathbb{R}^n)$ é ortogonal, $\langle x, y\rangle = \langle T^*Tx, y\rangle = \langle Tx, Ty\rangle$ quaisquer que sejam $x, y \in \mathbb{R}^n$, isto é, T preserva produto interno (e, portanto, preserva norma). Reciprocamente, suponhamos que T preserva produto interno. Neste caso, dado $x \in \mathbb{R}^n$, tem-se, para todo $y \in \mathbb{R}^n$, $\langle x, y\rangle = \langle Tx, Ty\rangle = \langle T^*Tx, y\rangle$, donde $T^*Tx = x$, isto é, T^* é ortogonal. Assim, um operador em $\mathrm{L}(\mathbb{R}^n)$ é ortogonal se, e somente se, preserva produto interno.

Considerando-se agora a *identidade de polarização*

$$\langle v_1, v_2\rangle = \frac{1}{2}(\|v_1\|^2 + \|v_2\|^2 - \|v_1 - v_2\|^2)\ \forall v_1, v_2 \in \mathbb{R}^n, \tag{8}$$

verifica-se facilmente que se um operador linear em $\mathbb{R}^n$ preserva norma, então ele preserva produto interno.

Em suma, são equivalentes as seguintes afirmações a respeito de um operador linear $T \in \mathrm{L}(\mathbb{R}^n)$:

- T é ortogonal;
- T preserva produto interno;
- T preserva norma.

É imediato que a norma espectral de qualquer operador ortogonal é 1, razão pela qual estes operadores são também chamados de *unitários*. Deve-se observar também que, diferentemente do conjunto dos operadores auto-adjuntos, aquele formado pelos operadores ortogonais de $\mathbb{R}^n$ não é um subespaço vetorial de $\mathrm{L}(\mathbb{R}^n)$, pois o operador nulo não é ortogonal.

Provemos agora as desigualdades (6). Para isto, dadas $T, U \in \mathrm{L}(\mathbb{R}^n, \mathbb{R}^m)$, façamos[v]

$$\langle T, U\rangle = \text{traço}(T^*U)$$

e verifiquemos que esta igualdade define em $\mathrm{L}(\mathbb{R}^n, \mathbb{R}^m)$ um produto interno que é aquele induzido[vi] pelo isomorfismo natural entre este espaço e $\mathbb{R}^{mn}$.

---

[v] O traço da matriz de um operador linear T em $\mathbb{R}^n$ independe da base escolhida e é definido como o *traço* de T.

[vi] Dado um isomorfismo linear entre espaços vetoriais, $T : \mathbb{V} \to \mathbb{W}$, se $\langle\, ,\, \rangle$ é um produto interno em $\mathbb{W}$, então a função $\langle\, ,\, \rangle_0$, dada por $\langle v_1, v_2\rangle_0 = \langle Tv_1, Tv_2\rangle$, $v_1, v_2 \in \mathbb{V}$, define em $\mathbb{V}$ um produto interno, dito *induzido* por T e $\langle\, ,\, \rangle$.

De fato, denotando-se por $A = (a_{ij})$ e $B = (b_{ij})$, $1 \le i \le m$, $1 \le j \le n$, as respectivas matrizes de T e U com respeito às bases canônicas de $\mathbb{R}^n$ e $\mathbb{R}^m$, temos que o traço de T^*U é, por definição, o traço da matriz $A^*B = (c_{ij})$, $1 \le i, j \le n$, sendo $c_{ij} = \sum_{k=1}^m a_{ki}b_{kj}$. Logo,

$$\text{traço}(A^*B) = \sum_{i=1}^{n} c_{ii} = \sum_{i,k=1}^{n,m} a_{ki}b_{ki} = \langle v, w\rangle,$$

em que v e w são os vetores de $\mathbb{R}^{mn}$ obtidos listando-se os elementos de A e B, respectivamente. Segue-se que $\langle T, U\rangle$, como definido, é o produto interno mencionado. Em particular,

$$\|T\|_e = \sqrt{\text{traço}(T^*T)} \quad \forall T \in \mathrm{L}(\mathbb{R}^n, \mathbb{R}^m). \tag{9}$$

Dada $T \in \mathrm{L}(\mathbb{R}^n, \mathbb{R}^m)$, temos que $Z = T^*T \in \mathrm{L}(\mathbb{R}^n)$ é um operador auto-adjunto, pois $Z^* = (T^*T)^* = T^*(T^*)^* = T^*T = Z$. Podemos, desta forma, tomar uma base ortonormal de $\mathbb{R}^n$, $\mathfrak{B} = \{u_1, \dots, u_n\}$, formada por autovetores de Z. Note que cada um dos autovalores correspondentes é não-negativo, já que, para cada i, $\lambda_i = \langle Zu_i, u_i\rangle = \langle T^*Tu_i, u_i\rangle = \langle Tu_i, Tu_i\rangle \ge 0$. Além disso, dado um vetor unitário $x \in \mathbb{R}^n$, tem-se

$$\|Tx\|^2 = \langle Tx, Tx\rangle = \langle T^*Tx, x\rangle = \langle Zx, x\rangle \le \|Zx\| \le \|Z\|,$$

donde $\sqrt{\|Z\|}$ é uma cota superior do conjunto $\{\|Tx\|; \|x\| = 1\}$. No entanto, por (7), $\|Z\| = |\lambda_i| = \lambda_i$ para algum $i = 1, \dots, n$. Além disso, $\|Tu_i\|^2 = \langle Zu_i, u_i\rangle = \lambda_i$, isto é, $\|Tu_i\| = \sqrt{\|Z\|}$. Logo[(vii)],

$$\|T\| = \sqrt{\|T^*T\|} = \max\{\sqrt{\lambda_1}, \dots, \sqrt{\lambda_n}\}. \tag{10}$$

Considerando-se, então, a igualdade (9), tem-se $\|T\|_e^2 = \lambda_1 + \cdots + \lambda_n$. Desta forma,

$$\|T\|^2 = \lambda_i \le \|T\|_e^2 \le n\lambda_i = n\,\|T\|^2,$$

donde se obtém as desigualdades (6).

4. Sequências em $\mathbb{R}^n$

Neste momento, em que temos posse do conceito de norma em $\mathbb{R}^n$, faz-se apropriado abordar as sequências neste espaço e, na medida do possível, estender a estas os principais conceitos e resultados das sequências em $\mathbb{R}$.

Dada, então, uma sequência $(x_k)_{k\in\mathbb{N}}$ em $\mathbb{R}^n$, denotaremos o seu k-ésimo termo, x_k, por

$$x_k = (x_{k1}, x_{k2}, \dots, x_{kn}), \;\; x_{ki} \in \mathbb{R} \quad (i = 1, 2 \dots, n).$$

Assim, cada sequência (x_k) em $\mathbb{R}^n$ determina n sequências $(x_{ki})_{k\in\mathbb{N}}$ em $\mathbb{R}$, podendo, desta forma, ser vista como uma "matriz" com infinitas linhas e n colunas, em que a k-ésima linha é formada pelas coordenadas do k-ésimo termo da sequência (x_k) e a i-ésima coluna, que é infinita, é formada pelos termos da sequência $(x_{ki})_{k\in\mathbb{N}}$, conforme indicado abaixo.

[(vii)] A igualdade (10) justifica a nomenclatura adotada para a norma espectral, uma vez que o conjunto dos autovalores de um operador é chamado de *espectro* do mesmo.

$$
\begin{array}{cccc}
x_{11} & x_{12} & \dots & x_{1n} \\
x_{21} & x_{22} & \dots & x_{2n} \\
\vdots & \vdots & & \vdots \\
x_{k1} & x_{k2} & \dots & x_{kn} \\
\vdots & \vdots & & \vdots
\end{array}
$$

Para cada i fixo, a sequência $(x_{ki})_{k\in\mathbb{N}}$ é dita uma *sequência-coordenada* de (x_k).

Definição 2 (sequência limitada). Diz-se que uma sequência $(x_k)_{k\in\mathbb{N}}$ em $\mathbb{R}^n$ é *limitada* se existe um real $\mu > 0$, tal que $\|x_k\| \le \mu \ \ \forall k \in \mathbb{N}$.

As sequências limitadas de $\mathbb{R}^n$ devem sua importância, principalmente, ao Teorema de Bolzano-Weierstrass, que estabelece uma importante propriedade das mesmas e sobre o qual discutiremos adiante.

Proposição 3 (coordenadas limitadas). *Uma sequência $(x_k)_{k\in\mathbb{N}}$ em $\mathbb{R}^n$ é limitada se, e somente se, cada uma de suas sequências-coordenada $(x_{ki})_{k\in\mathbb{N}}$ é limitada em $\mathbb{R}$.*

Demonstração. Suponhamos que (x_k) seja limitada. Então, existe $\mu > 0$, tal que $\|x_k\| \le \mu$ para todo $k \in \mathbb{N}$. Porém, para todo $i \in \{1, 2, \dots, n\}$ e todo $k \in \mathbb{N}$, tem-se

$$|x_{ki}| = \sqrt{x_{ki}^2} \le \sqrt{x_{k1}^2 + \dots + x_{ki}^2 + \dots + x_{kn}^2} = \|x_k\| \le \mu,$$

donde se infere que $(x_{ki})_{k\in\mathbb{N}}$ é limitada para todo $i \in \{1, \dots, n\}$.

Reciprocamente, se cada uma das n sequências $(x_{ki})_{k\in\mathbb{N}}$ é limitada, então, para cada $i \in \{1, \dots, n\}$, existe um real $\mu_i > 0$ satisfazendo $|x_{ki}| \le \mu_i \, \forall k \in \mathbb{N}$. Logo, tomando-se $\mu = \max\{\mu_1, \dots, \mu_n\}$ e considerando-se (2), obtém-se

$$\|x_k\| \le n\|x_k\|_{\max} = n \max\{|x_{k1}|, \dots, |x_{kn}|\} \le n\mu,$$

donde (x_k) é limitada. □

Observação 3. A última parte da demonstração acima alude ao fato de que podemos considerar outras normas em $\mathbb{R}^n$, que não a euclidiana, na conceituação de sequência limitada, isto é, podemos dizer que uma sequência (x_k) é *limitada relativamente* a uma norma $\| \ \|_0$ em $\mathbb{R}^n$ se existe $\mu > 0$, tal que $\|x_k\|_0 < \mu \, \forall k \in \mathbb{N}$.

É imediato que se $\| \ \|_0$ é equivalente à norma euclidiana, então uma sequência (x_k) em $\mathbb{R}^n$ é limitada se, e somente se, é limitada relativamente à norma $\| \ \|_0$. No caso afirmativo, então, o enunciado da Proposição 3 permanece válido quando substituímos "limitada" (primeira menção) por "limitada relativamente à norma $\| \ \|_0$".

4.1. Sequências Convergentes. Passemos agora ao nosso conceito mais fundamental.

Definição 3 (sequência convergente). Dado $a \in \mathbb{R}^n$, diz-se que uma sequência (x_k) em $\mathbb{R}^n$ *converge* para a ou, equivalentemente, que a é *limite* da sequência (x_k), se cumpre a seguinte condição:

$$\boxed{\text{Para todo } \epsilon > 0, \quad \text{existe } k_0 \in \mathbb{N}, \quad \text{tal que } k \geq k_0 \;\Rightarrow\; \|x_k - a\| < \epsilon.}$$

No caso afirmativo, a sequência (x_k) é dita *convergente*, caso contrário, ela é dita *divergente*.

Dados $a \in \mathbb{R}^n$ e $r > 0$, o conjunto dos pontos $x \in \mathbb{R}^n$ que cumprem a desigualdade $\|x - a\| < r$, denotado por $B(a, r)$, é chamado de *bola* (*euclidiana*) *aberta* de centro a e raio r. Desta forma, podemos dizer que uma sequência (x_k) em $\mathbb{R}^n$ converge para a se, para *toda* bola aberta $B(a, \epsilon)$, a partir de um certo $k_0 \in \mathbb{N}$, *todos* os seus termos pertencem a $B(a, \epsilon)$ (Fig. 3). Em geral, o quão grande deve ser k_0 depende de quão pequeno é o ϵ dado.

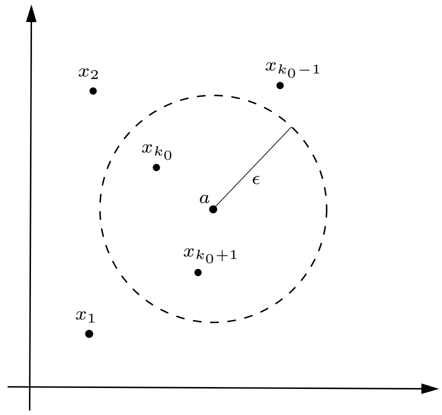

Figura 3

Observação 4 (unicidade do limite). Toda sequência convergente em $\mathbb{R}^n$ possui um único limite. De fato, se (x_k) fosse uma sequência em $\mathbb{R}^n$ com dois limites distintos, a e b, poderíamos tomar $\epsilon = \|b - a\|/2 > 0$. Então, a partir de um certo $k_0 \in \mathbb{N}$, todos os termos da sequência (x_k) estariam em ambas as bolas abertas $B(a, \epsilon)$ e $B(b, \epsilon)$, o que é absurdo, uma vez que estas são disjuntas.

Quando $a \in \mathbb{R}^n$ é o limite de uma sequência (x_k), escrevemos

$$x_k \to a \quad \text{ou} \quad a = \lim x_k \quad \text{ou ainda} \quad a = \lim_{k \to \infty} x_k\,.$$

Proposição 4. *Toda sequência convergente é limitada.*

Demonstração. De fato, se (x_k) é uma sequência em $\mathbb{R}^n$ que converge para a, temos que existe $k_0 \in \mathbb{N}$, tal que $k \geq k_0 \Rightarrow \|x_k - a\| < 1$. Logo,

$$\|x_k\| \;=\; \|(x_k - a) + a\| \;\leq\; \|x_k - a\| + \|a\| < 1 + \|a\| \;\forall k \geq k_0\,.$$

Fazendo-se $\mu = \max\{\|x_1\|, \dots, \|x_{k_0-1}\|, 1 + \|a\|\}$, tem-se $\|x_k\| \leq \mu \;\forall k \in \mathbb{N}$. □

Observação 5. Considerando-se em $\mathbb{R}^n$ uma sequência (x_k) e um ponto a, é imediato, a partir da definição de convergência, que

$$x_k \to a \quad (\text{em } \mathbb{R}^n) \quad \Leftrightarrow \quad \|x_k - a\| \to 0 \quad (\text{em } \mathbb{R}),$$

isto é, a é limite de (x_k) se, e somente se, a distância entre x_k e a converge para zero. Além disso, uma vez que, para todo $k \in \mathbb{N}$, $|\,\|x_k\| - \|a\|\,| \leq \|x_k - a\|$, tem-se

$$(11) \qquad x_k \to a \quad (\text{em } \mathbb{R}^n) \quad \Rightarrow \quad \|x_k\| \to \|a\| \quad (\text{em } \mathbb{R}).$$

Deve-se observar também que, dada uma sequência (x_k) em $\mathbb{R}^n$, se $\mu > 0$ é tal que, para todo $\epsilon > 0$, existe $k_0 \in \mathbb{N}$ satisfazendo

$$k \geq k_0 \Rightarrow \|x_k - a\| < \mu\epsilon,$$

então $x_k \to a$. Com efeito, dado $\epsilon' > 0$, seja $\epsilon = \epsilon'/\mu$. Por hipótese, para este ϵ, existe $k_0 \in \mathbb{N}$, tal que $k \geq k_0 \Rightarrow \|x_k - a\| < \mu\epsilon = \epsilon'$. Uma vez que ϵ' foi tomado arbitrariamente, conclui-se que $x_k \to a$.

Observação 6 (convergência relativa). O conceito de sequência convergente que introduzimos é relativo à norma euclidiana. No entanto, como o fizemos quando discutimos sobre sequências limitadas, pode-se considerá-lo relativamente a outras normas. Neste sentido, é relevante mencionarmos que se $\|\ \|_0$ é uma norma em $\mathbb{R}^n$ equivalente à norma euclidiana, então uma sequência é convergente com relação à norma $\|\ \|_0$ se, e somente se, o é com relação à norma euclidiana. No caso afirmativo, os limites de ambas estas sequências coincidem.

Para constatarmos isto, consideremos uma sequência (x_k) em $\mathbb{R}^n$ que converge para $a \in \mathbb{R}^n$ com relação à norma $\|\ \|_0$. Neste caso, dado $\epsilon > 0$, existe $k_0 \in \mathbb{N}$, tal que $k \geq k_0$ implica $\|x_k - a\|_0 < \epsilon$. Além disso, pela equivalência das normas dadas, existe $\mu > 0$ satisfazendo $\|x\| \leq \mu\|x\|_0 \,\forall x \in \mathbb{R}^n$. Logo, para todo $k \geq k_0$,

$$\|x_k - a\| \;\leq\; \mu\|x_k - a\|_0 \;<\; \mu\epsilon,$$

isto é, (x_k) é convergente (com mesmo limite) com respeito à norma euclidiana. A recíproca segue-se diretamente da simetria da equivalência de normas.

Proposição 5 (convergência das coordenadas). *Uma sequência* $(x_k)_{k\in\mathbb{N}}$ *em* $\mathbb{R}^n$, $x_k = (x_{k1}, x_{k2}, \dots, x_{kn})$, *converge para* $a = (a_1, \dots, a_n) \in \mathbb{R}^n$ *se, e somente se, para cada* $i = 1, 2, \dots, n$, *a sequência-coordenada* $(x_{ki})_{k\in\mathbb{N}}$ *converge para* a_i .

Demonstração. Suponhamos, inicialmente, que $x_k \to a$. Então, dado $\epsilon > 0$, existe $k_0 \in \mathbb{N}$, tal que $k \geq k_0 \Rightarrow \|x_k - a\| < \epsilon$. Porém, para todo $i \in \{1, \dots, n\}$ e todo $k \in \mathbb{N}$, tem-se $|x_{ki} - a_i| \leq \|x_k - a\|$. Assim, para todo $k \geq k_0$, vale $|x_{ki} - a_i| < \epsilon$, isto é, $x_{ki} \to a_i$ qualquer que seja $i \in \{1, \dots, n\}$.

Reciprocamente, suponhamos que $x_{ki} \to a_i$ para todo $i \in \{1, \dots, n\}$. Então, dado $\epsilon > 0$, para cada $i \in \{1, \dots, n\}$, existe $k_i \in \mathbb{N}$, tal que $k \geq k_i \Rightarrow |x_{ki} - a_i| < \epsilon$. Tomando-se, então, $k_0 = \max\{k_1, \dots, k_n\}$, tem-se, para todo $k \geq k_0$, $\|x_k - a\|_{\max} = \max\{\,|x_{k1} - a_1|, \dots, |x_{kn} - a_n|\,\} < \epsilon$, donde $x_k \to a$ (vide Observação 6). □

Observação 7. Dado $p \in \mathbb{N}$, o conceito de sequência-coordenada, bem como o resultado da Proposição 5, generalizam-se facilmente para sequências definidas em $\mathbb{R}^p$ quando se decompõe este espaço como $\mathbb{R}^p = \mathbb{R}^{j_1} \times \mathbb{R}^{j_2} \times \cdots \times \mathbb{R}^{j_n}$, em que, para cada $i \in \{1, \dots, n\}$, $j_i \in \mathbb{N}$ e $j_1 + \cdots + j_n = p$. Mais especificamente, se (x_k) é uma sequência em $\mathbb{R}^p$, cada um dos seus termos se escreve como $x_k = (x_{1k}, \dots, x_{nk})$, em que $x_{ik} \in \mathbb{R}^{j_i} \,\forall i \in \{1, \dots, n\}$. Assim, para cada tal i, $(x_{ik})_{k\in\mathbb{N}}$ é uma sequência em $\mathbb{R}^{j_i}$. Dado, então, $a = (a_1, \dots, a_n) \in \mathbb{R}^p$, tem-se que $x_k \to a$ se, e somente se, $x_{ik} \to a_i \,\forall i \in \{1, \dots, n\}$.

Vejamos, na proposição seguinte, o comportamento das sequências convergentes de $\mathbb{R}^n$ com respeito às suas operações.

PROPOSIÇÃO 6 (PROPRIEDADES OPERATÓRIAS). *Sejam* (x_k), (y_k) *sequências em* $\mathbb{R}^n$ *e* (λ_k) *uma sequência em* $\mathbb{R}$, *tais que* $x_k \to a$, $y_k \to b$ *e* $\lambda_k \to \lambda$. *Então,*

$$(x_k + y_k) \to a + b \qquad e \qquad (\lambda_k x_k) \to \lambda a.$$

DEMONSTRAÇÃO. Dado $\epsilon > 0$, existem $k_1, k_2, k_3 \in \mathbb{N}$, tais que

$$k \geq k_1 \Rightarrow \|x_k - a\| < \epsilon, \quad k \geq k_2 \Rightarrow \|y_k - b\| < \epsilon \quad \text{e} \quad k \geq k_3 \Rightarrow \|\lambda_k - \lambda\| < \epsilon.$$

Então, para todo $k \geq k_0 = \max\{k_1, k_2, k_3\}$, tem-se

$$\|(x_k + y_k) - (a + b)\| \leq \|x_k - a\| + \|y_k - b\| < 2\epsilon,$$

ou seja, $(x_k + y_k) \to a + b$.

Temos também que a sequência (x_k), por ser convergente, é limitada. Assim, existe $\mu > 0$, tal que $\|x_k\| < \mu \ \forall k \in \mathbb{N}$. Desta forma, para todo $k \geq k_0$,

$$\|\lambda_k x_k - \lambda a\| = \|(\lambda_k - \lambda)x_k + \lambda(x_k - a)\| \leq |\lambda_k - \lambda| \, \|x_k\| + |\lambda| \, \|x_k - a\| < (\mu + |\lambda|)\epsilon,$$

donde se infere que $(\lambda_k x_k) \to \lambda a$. $\square$

Consideremos, agora, subsequências de sequências em $\mathbb{R}^n$. Como sabemos, toda subsequência é uma sequência e, portanto, cabe-nos falar sobre subsequências convergentes e indagar sobre as relações de convergência entre uma sequência e as suas subsequências. Neste contexto, é fácil constatar que toda subsequência de uma sequência convergente é convergente e tem o mesmo limite que a sequência. É fácil, também, obter exemplos de sequências divergentes que possuem subsequências convergentes.

Conforme constataremos, em Análise, muitos resultados importantes são obtidos quando se assegura que uma determinada sequência possui, pelo menos, uma subsequência convergente, não se fazendo necessário, portanto, que a sequência, em si, seja convergente. Por esta razão, o Teorema de Bolzano-Weierstrass, que apresentamos a seguir, constitui um dos resultados mais importantes da Análise no espaço $\mathbb{R}^n$.

TEOREMA DE BOLZANO-WEIERSTRASS. *Toda sequência limitada em* $\mathbb{R}^n$ *possui uma subsequência convergente.*

DEMONSTRAÇÃO. Uma vez que, no capítulo anterior, provamos o teorema no caso em que $n = 1$, podemos supor que $n > 1$.

Tomemos, então, uma sequência limitada (x_k), em $\mathbb{R}^n$, e escrevamos, para cada $k \in \mathbb{N}$, $x_k = (x_{k1}, \dots, x_{kn})$. Pela Proposição 3, cada sequência-coordenada $(x_{ki})_{k \in \mathbb{N}}$ é limitada em $\mathbb{R}$. Daí, temos que existe um subconjunto infinito $\mathbb{N}_1$, de $\mathbb{N}$, tal que a subsequência $(x_{k1})_{k \in \mathbb{N}_1}$, de $(x_{k1})_{k \in \mathbb{N}}$, converge para um real a_1.

Analogamente, existe $\mathbb{N}_2 \subset \mathbb{N}_1$, infinito, tal que a subsequência $(x_{k2})_{k \in \mathbb{N}_2}$, de $(x_{k2})_{k \in \mathbb{N}_1}$, converge para um real a_2. Note que a subsequência $(x_{k1})_{k \in \mathbb{N}_2}$, de $(x_{k1})_{k \in \mathbb{N}_1}$, também converge para a_1.

Repetindo-se este procedimento, após n passos, obteremos um subconjunto infinito $\mathbb{N}_n$, de $\mathbb{N}$, tal que, para cada $i = 1, 2, \dots, n$, a subsequência $(x_{ki})_{k \in \mathbb{N}_n}$, de $(x_{ki})_{k \in \mathbb{N}}$, converge para um real a_i. Logo, pela Proposição 5, a subsequência $(x_k)_{k \in \mathbb{N}_n}$, de $(x_k)_{k \in \mathbb{N}}$, converge para $a = (a_1, a_2, \dots, a_n) \in \mathbb{R}^n$. $\square$

Observação 8. Segue-se da demonstração acima e das considerações da Observação 3 que o Teorema de Bolzano-Weierstrass é válido para qualquer norma de $\mathbb{R}^n$ que seja equivalente à norma euclidiana, isto é, se $\| \ \|_0$ é uma tal norma, então toda sequência limitada relativamente à norma $\| \ \|_0$ possui uma subsequência que é convergente relativamente à norma $\| \ \|_0$.

Aplicaremos agora o Teorema de Bolzano-Weierstrass para estabelecer o resultado seguinte, mencionado no fim da Seção 1.

Proposição 7. *Duas normas quaisquer em $\mathbb{R}^n$ são equivalentes.*

Demonstração. Seja $\| \ \|_0$ uma norma arbitrária em $\mathbb{R}^n$. Provaremos que esta é equivalente à norma da soma $\| \ \|_s$. O resultado, então, se seguirá por transitividade.

Seja $\lambda = \max\{\|e_1\|_0, \dots, \|e_n\|_0\}$. Assim, para todo $x = (x_1, \dots, x_n) \in \mathbb{R}^n$, tem-se,

$$\|x\|_0 = \|x_1e_1 + \dots + x_ne_n\|_0 \le |x_1| \, \|e_1\|_0 + \dots + |x_n| \, \|e_n\|_0 \le \lambda\|x\|_s \, .$$

Resta-nos, pois, mostrar que existe $\mu > 0$, tal que $\|x\|_s \le \mu\|x\|_0 \, \forall x \in \mathbb{R}^n$. Suponhamos, por absurdo, que um tal μ não exista. Então, para todo $k \in \mathbb{N}$, existe $x_k \in \mathbb{R}^n$ satisfazendo $\|x_k\|_s > k\|x_k\|_0$. Fazendo-se $u_k = x_k/\|x_k\|_s$, tem-se

$$\|u_k\|_s = 1 \qquad \text{e} \qquad \|u_k\|_0 = \frac{\|x_k\|_0}{\|x_k\|_s} < \frac{1}{k} \cdot$$

Em particular, a sequência (u_k) é limitada relativamente à norma da soma. Uma vez que esta é equivalente à norma euclidiana, pelo Teorema de Bolzano-Weierstrass (vide Observação 8), existe uma subsequência (u_{k_i}) que converge (relativamente à norma da soma) para $u \in \mathbb{R}^n$. Neste caso, devemos ter[(viii)] $\|u\|_s = 1$ e, portanto, $u \ne 0$. Por outro lado, para todo $i \in \mathbb{N}$,

$$\|u\|_0 \le \|u - u_{k_i}\|_0 + \|u_{k_i}\|_0 < \lambda\|u - u_{k_i}\|_s + \frac{1}{k_i} \cdot$$

Logo,

$$\|u\|_0 \le \lim_{i\to\infty} \left(\lambda\|u - u_{k_i}\|_s + \frac{1}{k_i}\right) = 0,$$

donde $\|u\|_0 = 0$, o que contradiz o fato de u ser não-nulo. □

Observação 9. Segue-se da Proposição 7, da Observação 3 e da Observação 6 que, em $\mathbb{R}^n$, os conceitos de sequência limitada e sequência convergente independem da norma adotada.

4.2. Sequências de Cauchy. Introduziremos agora uma condição sobre sequências em $\mathbb{R}^n$ que é equivalente a de ser convergente e que está relacionada com o importante conceito de espaço métrico completo, sobre o qual discutiremos na próxima seção.

Definição 4 (sequência de cauchy). Uma sequência $(x_k)_{k\in\mathbb{N}}$ em $\mathbb{R}^n$ é dita *de Cauchy* se para todo $\epsilon > 0$, existe $k_0 \in \mathbb{N}$, tal que $k, l \ge k_0 \Rightarrow \|x_k - x_l\| < \epsilon$.

[(viii)] Verifica-se facilmente que a implicação (11) é válida para qualquer norma de $\mathbb{R}^n$, em particular, para a norma da soma.

Assim, uma sequência é de Cauchy quando, dado $\epsilon > 0$, a partir de um certo termo, a distância entre dois termos quaisquer da sequência é menor que ϵ, isto é,

$$(x_k)_{k\in\mathbb{N}} \text{ é de Cauchy} \quad \Leftrightarrow \quad \lim_{k\to\infty} \|x_{k+p} - x_k\| = 0 \ \forall p \in \mathbb{N}.$$

Uma vez que esta condição, contrariamente à de convergência, envolve apenas os termos da sequência, o resultado seguinte, dentre outros benefícios, se faz útil quando se necessita estabelecer a convergência de uma sequência da qual se ignora o limite (vide Exemplo 5, abaixo).

Teorema 2 (completude de $\mathbb{R}^n$). *Uma sequência em $\mathbb{R}^n$ é de Cauchy se, e somente se, é convergente.*

Demonstração. Suponhamos que (x_k) seja uma sequência de Cauchy em $\mathbb{R}^n$ e provemos, inicialmente, que ela é limitada. Com efeito, neste caso, existe $k_0 \in \mathbb{N}$, tal que $\|x_k - x_l\| < 1$ quaisquer que sejam $k, l \geq k_0$. Logo, para todo $k \geq k_0$, $\|x_k\| = \|x_k - x_{k_0} + x_{k_0}\| \leq \|x_k - x_{k_0}\| + \|x_{k_0}\| < 1 + \|x_{k_0}\|$. Desta forma, fazendo-se $\mu = \max\{\|x_1\|, \dots, \|x_{k_0-1}\|, 1 + \|x_{k_0}\|\}$, tem-se, para todo $k \in \mathbb{N}$, que $\|x_k\| < \mu$, isto é, (x_k) é limitada.

Segue-se, então, do Teorema de Bolzano-Weierstrass, que existe uma subsequência de $(x_k)_{k\in\mathbb{N}}$, $(x_k)_{k\in\mathbb{N}_0}$, que é convergente. Designando-se por $a \in \mathbb{R}^n$ o limite dessa subsequência, temos que, dado $\epsilon > 0$, existem $l_1 \in \mathbb{N}_0$ e $k_1 \in \mathbb{N}$ satisfazendo

$$\|x_l - a\| < \epsilon \ \ \forall\, l \geq l_1\,,\ l \in \mathbb{N}_0 \qquad \text{e} \qquad \|x_k - x_l\| < \epsilon \ \ \forall\, k, l \geq k_1\,,\ k, l \in \mathbb{N}.$$

Então, fazendo-se $k_2 = \max\{l_1, k_1\}$ e tomando-se $l \in \mathbb{N}_0$, tal que $l \geq k_2$, tem-se, para todo $k \geq k_2$,

$$\|x_k - a\| \leq \|x_k - x_l\| + \|x_l - a\| < 2\epsilon,$$

donde se infere que $x_k \to a$.

Reciprocamente, suponhamos que (x_k) seja uma sequência convergente em $\mathbb{R}^n$ e que seu limite seja a. Então, dado $\epsilon > 0$, existe $k_0 \in \mathbb{N}$, tal que $\|x_k - a\| < \epsilon$ para todo $k \geq k_0$. Logo, se $k, l \geq k_0$, tem-se

$$\|x_k - x_l\| \leq \|x_k - a\| + \|x_l - a\| < 2\epsilon,$$

isto é, (x_k) é de Cauchy. □

Exemplo 5. Suponhamos que (x_k) seja uma sequência em $\mathbb{R}^n$, a qual, para um dado $\mu \in (0,1)$, satisfaz (Fig. 4)

$$\|x_{k+1} - x_k\| \leq \mu \|x_k - x_{k-1}\| \ \ \forall k \in \mathbb{N}.$$

Verifica-se facilmente, por indução sobre k, que uma tal sequência, então, cumpre

$$\|x_{k+1} - x_k\| \leq \mu^k \|x_1 - x_0\| \ \ \forall k \in \mathbb{N}.$$

Desta forma, escrevendo-se

$$x_{k+p} - x_k = \sum_{i=1}^{p} (x_{k+i} - x_{k+i-1})$$

e considerando-se, juntamente com a desigualdade triangular, a desigualdade

$$\sum_{i=1}^{p} \mu^{i-1} < \frac{1}{1-\mu} \ \ \forall p \in \mathbb{N}$$

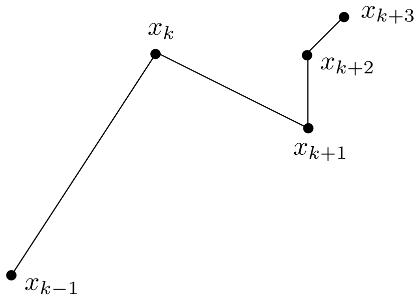

FIGURA 4

(a qual se verifica por indução sobre p), obtém-se, para quaisquer $k, p \in \mathbb{N}$,

$$\|x_{k+p} - x_k\| \le \sum_{i=1}^{p} \|x_{k+i} - x_{k+i-1}\| \le \sum_{i=1}^{p} \mu^{k+i-1}\|x_1 - x_0\| \le \frac{\|x_1 - x_0\|}{1-\mu}\mu^k.$$

Logo, para todo $p \in \mathbb{N}$, tem-se

$$\lim_{k\to\infty} \|x_{k+p} - x_k\| = 0,$$

donde se infere que (x_k) é uma sequência de Cauchy. Segue-se, então, do Teorema 2, que (x_k) é convergente.

4.3. Transformações Lineares e Sequências. Como um primeiro indício da efetividade da norma espectral de $\mathrm{L}(\mathbb{R}^n, \mathbb{R}^m)$, apliquemo-la para verificar que transformações lineares levam sequências convergentes em sequências convergentes.

PROPOSIÇÃO 8 (LINEARIDADE E CONVERGÊNCIA). *Sejam $T \in \mathrm{L}(\mathbb{R}^n, \mathbb{R}^m)$ uma transformação linear e $(x_k)_{k\in\mathbb{N}}$ uma sequência convergente em $\mathbb{R}^n$ cujo limite é $a \in \mathbb{R}^n$. Então, a sequência $(Tx_k)_{k\in\mathbb{N}}$, em $\mathbb{R}^m$, é convergente e seu limite é Ta.*

DEMONSTRAÇÃO. Sendo a o limite de (x_k), temos que, dado $\epsilon > 0$, existe $k_0 \in \mathbb{N}$, tal que $k \ge k_0 \Rightarrow \|x_k - a\| < \epsilon$. Logo, para $k \ge k_0$,

$$\|Tx_k - Ta\| = \|(T(x_k - a)\| \le \|T\|\,\|x_k - a\| < \|T\|\epsilon,$$

donde se conclui que $Tx_k \to Ta$. □

Na seção anterior, valendo-nos do Teorema Espectral, deduzimos a igualdade $\|T\| = \sqrt{\|T^*T\|}$, válida para toda transformação linear $T \in \mathrm{L}(\mathbb{R}^n, \mathbb{R}^m)$, obtendo, para cada tal T, um vetor unitário $u \in \mathbb{R}^n$ que satisfaz $\|Tu\| = \|T\|$.

Através da Proposição 8 e do Teorema de Bolzano-Weierstrass, podemos, entretanto, provar a existência do vetor u de forma independente do Teorema Espectral. Com efeito, lembrando-se que, por definição, $\|T\|$ é o supremo do conjunto $\{\|Tx\|;\, x \in \mathbb{R}^n, \|x\| = 1\}$, obtém-se uma sequência (x_k), de vetores unitários de $\mathbb{R}^n$, tal que $\|Tx_k\| \to \|T\|$. Uma vez que (x_k) é uma sequência limitada, pelo Teorema de Bolzano-Weierstrass, existe uma subsequência convergente de (x_k), $(x_{k_i})_{i\in\mathbb{N}}$. Fazendo-se $u = \lim x_{k_i}$, tem-se $\|u\| = \lim \|x_{k_i}\| = 1$ e, pela Proposição 8, $Tu = \lim Tx_{k_i}$, donde $\|Tu\| = \lim \|Tx_{k_i}\| = \|T\|$, como desejado.

Em particular, podemos escrever

$$\|T\| = \max\{\|Tx\|;\, x \in \mathbb{R}^n, \|x\| = 1\}.$$

O resultado da Proposição 8 generaliza-se para aplicações n-lineares, conforme a proposição seguinte.

Proposição 9 (n-linearidade e convergência). *Considere n números naturais, $j_1, \dots, j_n$, uma aplicação n-linear, $f : \mathbb{R}^{j_1} \times \cdots \times \mathbb{R}^{j_n} \to \mathbb{R}^m$, e uma sequência $(x_k)_{k\in\mathbb{N}}$ em $\mathbb{R}^{j_1} \times \cdots \times \mathbb{R}^{j_n}$. Então, se $x_k \to a$ em $\mathbb{R}^{j_1} \times \cdots \times \mathbb{R}^{j_n}$, tem-se $f(x_k) \to f(a)$ em $\mathbb{R}^m$.*

Demonstração. Escrevamos $a = (a_1, \dots, a_n)$ e, para cada $k \in \mathbb{N}$, $x_k = (x_{1k}, \dots, x_{nk})$, em que $a_i\,, x_{ik} \in \mathbb{R}^{j_i}\ \forall i \in \{1, \dots, n\}$. Uma vez que $x_k \to a$, para cada $i \in \{1, \dots, n\}$, a sequência-coordenada $(x_{ik})_{k\in\mathbb{N}}$, em $\mathbb{R}^{j_i}$, satisfaz $x_{ik} \to a_i$ (vide Observação 7). Em particular, para cada tal i, existe $\lambda_i > 0$, tal que $\|x_{ik}\| \le \lambda_i\, \forall k \in \mathbb{N}$.

Agora, da n-linearidade de f, segue-se que

$$f(x_k) - f(a) = \sum_{i=1}^{n} f(a_1\,, \dots, a_{i-1}, x_{ik} - a_i\,, x_{(i+1)k}, , \dots, x_{nk}).$$

Além disso, existe uma constante $\mu > 0$ (vide Exercício 8), tal que

$$\|f(y_1, \dots, y_n)\| \le \mu\|y_1\| \dots \|y_n\| \ \ \forall y_i \in \mathbb{R}^{j_i}, \ \ i \in \{1, \dots, n\}.$$

Segue-se destas considerações e da desigualdade triangular que, fazendo-se $\lambda = \max\{\lambda_1\,, \dots, \lambda_n\,, \|a_1\|, \dots, \|a_n\|, \mu\}$, tem-se

$$\begin{aligned}
\|f(x_k) - f(a)\| &\le \sum_{i=1}^{n} \|f(a_1\,, \dots, a_{i-1}, x_{ik} - a_i\,, x_{(i+1)k}, , \dots, x_{nk})\| \\
&\le \sum_{i=1}^{n} \mu\|a_1\| \dots \|a_{i-1}\|\, \|x_{ik} - a_i\|\, \|x_{(i+1)k}\| \dots \|x_{nk}\| \\
&\le \sum_{i=1}^{n} \lambda^n \|x_{ik} - a_i\|,
\end{aligned}$$

donde se infere que $\|f(x_k) - f(a)\| \to 0$ e, portanto, que $f(x_k) \to f(a)$. □

Dadas sequências (x_k) e (y_k) em $\mathbb{R}^n$, tem-se, pela Proposição 9 e pela bilinearidade do produto interno, que

$$x_k \to a \quad \text{e} \quad y_k \to b \quad \Rightarrow \quad \langle x_k, y_k\rangle \to \langle a, b\rangle.$$

Exemplo 6 (sequências em subespaços). O limite de toda sequência convergente num subespaço vetorial $\mathbb{V}$ de $\mathbb{R}^n$ é, necessariamente, um ponto de $\mathbb{V}$. Com efeito, dado $x \in \mathbb{R}^n$, temos que $x \in \mathbb{V}$ se, e somente se, $\langle x, w\rangle = 0\, \forall w \in \mathbb{V}^{\perp}$, pois $\mathbb{V}^{\perp\perp} = \mathbb{V}$. Assim, se (x_k) é uma sequência em $\mathbb{V}$ e $a \in \mathbb{R}^n$ é o seu limite, vale, para todo $w \in \mathbb{V}^{\perp}$ e todo $k \in \mathbb{N}$, a igualdade $\langle x_k, w\rangle = 0$. Logo, $\langle a, w\rangle = \langle \lim x_k, w\rangle = \lim\langle x_k, w\rangle = 0$, donde se conclui que $a \in \mathbb{V}$.

4.4. Sequências em $\mathrm{L}(\mathbb{R}^n, \mathbb{R}^m)$ e $\mathrm{M}(m, n)$. Feitas as devidas alterações, os conceitos que introduzimos sobre sequências em $\mathbb{R}^n$, bem como os resultados que obtivemos, aplicam-se a qualquer espaço vetorial normado de dimensão finita, particularmente, $(\mathrm{L}(\mathbb{R}^n, \mathbb{R}^m), \|\ \|)$ e $(\mathrm{M}(m, n), \|\ \|)$. Façamos, então, algumas considerações sobre sequências nesses espaços.

Exemplo 7 (potências de um operador). Dado um operador linear T de $\mathrm{L}(\mathbb{R}^n)$, seja $(T_k)_{k\in\mathbb{N}}$ a sequência em $\mathrm{L}(\mathbb{R}^n)$ definida por $T_k = T^k$. Então, $T_k \to 0$ se $\|T\| < 1$. Com efeito, pela Proposição 2-(ii), $\|T_k\| = \|T^k\| \le \|T\|^k$. Logo, se $\|T\| < 1$, $\|T_k\| \to 0$, o que implica $T_k \to 0$.

Exemplo 8 (produtos de sequências). Segue-se diretamente da Proposição 9, bem como da bilinearidade da aplicação

$$\begin{array}{ccc} \mathrm{L}(\mathbb{R}^n,\mathbb{R}^m)\times\mathrm{L}(\mathbb{R}^m,\mathbb{R}^p) & \to & \mathrm{L}(\mathbb{R}^n,\mathbb{R}^p) \\ (T,Z) & \mapsto & ZT\,, \end{array}$$

que se (T_k) e (Z_k) são sequências em $\mathrm{L}(\mathbb{R}^n,\mathbb{R}^m)$ e $\mathrm{L}(\mathbb{R}^m,\mathbb{R}^p)$, respectivamente, tais que $T_k \to T$ e $Z_k \to Z$, então $Z_kT_k \to ZT$.

Analogamente, se (A_k) é uma sequência em $\mathrm{M}(m,n)$ e (B_k) é uma sequência em $\mathrm{M}(n,p)$, então

$$A_k \to A \quad \text{e} \quad B_k \to B \quad \Rightarrow \quad A_kB_k \to AB.$$

Além disso, pela n-linearidade da função determinante, tem-se

$$A_k \to A \quad \Rightarrow \quad \det A_k \to \det A.$$

Exemplo 9 (sequências de operadores auto-adjuntos). Uma vez que os operadores auto-adjuntos de $\mathrm{L}(\mathbb{R}^n)$ formam um subespaço vetorial deste, segue-se das considerações do Exemplo 6 que o limite de toda sequência convergente de operadores auto-adjuntos é também um operador auto-adjunto.

Em espaços de funções, em particular em $\mathrm{L}(\mathbb{R}^n,\mathbb{R}^m)$, é natural considerar-se o conceito de convergência pontual, conforme a definição seguinte.

Definição 5 (convergência pontual). Diz-se que uma sequência $(T_k)_{k\in\mathbb{N}}$ em $\mathrm{L}(\mathbb{R}^n,\mathbb{R}^m)$ *converge pontualmente* para uma aplicação $T : \mathbb{R}^n \to \mathbb{R}^m$ se, para todo $x \in \mathbb{R}^n$, $T_kx \to T(x)$.

Vejamos, na proposição seguinte, que em $\mathrm{L}(\mathbb{R}^n,\mathbb{R}^m)$, diferentemente de outros espaços de funções, convergência e convergência pontual são conceitos equivalentes.

Proposição 10 (linearidade e convergência pontual). *Uma sequência $(T_k)_{k\in\mathbb{N}}$ em $\mathrm{L}(\mathbb{R}^n,\mathbb{R}^m)$ converge pontualmente para $T : \mathbb{R}^n \to \mathbb{R}^m$ se, e somente se, T é linear e $T_k \to T$.*

Demonstração. Suponhamos que $(T_k)_{k\in\mathbb{N}}$ convirja pontualmente para T e tomemos, arbitrariamente, $x, y \in \mathbb{R}^n$ e $\lambda \in \mathbb{R}$. Teremos, então,

$$T(x+\lambda y) = \lim T_k(x+\lambda y) = \lim(T_kx + \lambda T_ky) = \lim T_kx + \lambda \lim T_ky = Tx + \lambda Ty,$$

donde T é linear. Além disso, fazendo-se $Z_k = T_k - T$, vale, para todo $x \in \mathbb{R}^n$, $\lim Z_kx = 0$. Logo, pela igualdade (9),

$$\|Z_k\|_e^2 = \text{traço}(Z_k^*Z_k) = \sum_{i=1}^n \langle (Z_k^*Z_k)e_i, e_i\rangle = \sum_{i=1}^n \langle Z_ke_i, Z_ke_i\rangle = \sum_{i=1}^n \|Z_ke_i\|^2.$$

Daí, segue-se que

$$\lim \|Z_k\|_e = \sqrt{\sum_{i=1}^n \lim \|Z_ke_i\|^2} = 0,$$

donde $Z_k \to 0$, isto é $T_k \to T$.

Reciprocamente, suponhamos que T seja linear e que $T_k \to T$. Neste caso, dado $x \in \mathbb{R}^n$, tem-se, para todo $k \in \mathbb{N}$, $\|T_k x - Tx\| = \|(T_k - T)x\| \leq \|T_k - T\|\|x\|$, donde $\|T_k x - Tx\| \to 0$, pois $\|T_k - T\| \to 0$. Logo, $T_k x \to Tx$, ou seja, a sequência $(T_k)_{k\in\mathbb{N}}$ converge pontualmente para T. $\square$

5. O Espaço Métrico $\mathbb{R}^n$

5.1. Métricas em $\mathbb{R}^n$. Relembremos que a *distância euclidiana* entre dois pontos $x, y \in \mathbb{R}^n$, $d(x,y)$, é definida por

$$d(x,y) = \|x - y\|.$$

Decorre imediatamente das propriedades da norma que, assim definida, a função d, para quaisquer $x, y, z \in \mathbb{R}^n$, satisfaz:

i) $d(x,y) \geq 0$ e $d(x,y) = 0$ se, e somente se, $x = y$ (positividade);

ii) $d(x,y) = d(y,x)$ (simetria);

iii) $d(x,z) \leq d(x,y) + d(y,z)$ (desigualdade triangular).

Estas são as propriedades essenciais da distância, isto é, intuitivamente, espera-se que a distância entre dois pontos distintos seja positiva e que seja nula entre dois pontos iguais (propriedade (i)), que a distância entre x e y seja igual àquela entre y e x (propriedade (ii)), e que a distância entre dois pontos nunca seja superior à soma das distâncias destes a um terceiro (propriedade (iii)).

Diz-se, então, que uma *distância* ou, equivalentemente, *métrica*, em $\mathbb{R}^n$ é uma função $d : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ que tem estas propriedades. A distância euclidiana, desta forma, constitui um caso particular de métrica em $\mathbb{R}^n$.

De modo geral, dado um conjunto M e uma métrica $d : M \times M \to \mathbb{R}$, chama-se o par (M,d) de *espaço métrico*.

Note que, assim como a norma euclidiana, cada norma em $\mathbb{R}^n$ determina neste uma métrica. Definindo-se, por exemplo, $d_{\max}, d_{\mathrm{s}} : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ por

$$d_{\max}(x,y) = \|x - y\|_{\max} \qquad \text{e} \qquad d_{\mathrm{s}}(x,y) = \|x - y\|_{\mathrm{s}}\,,$$

tem-se que $(\mathbb{R}^n, d)$, $(\mathbb{R}^n, d_{\max})$ e $(\mathbb{R}^n, d_{\mathrm{s}})$ são espaços métricos distintos. Além disso, decorre de (2) que

$$d_{\max}(x,y) \leq d(x,y) \leq d_{\mathrm{s}}(x,y) \leq n d_{\max}(x,y) \ \ \forall x, y \in \mathbb{R}^n.$$

Dados um espaço métrico (M,d) e $M_0 \subset M$, é evidente que $d_0 = d|_{M_0 \times M_0}$ define uma métrica em M_0, isto é, restringindo-se a função distância de um espaço métrico a um qualquer de seus subconjuntos, este último torna-se um espaço métrico.

Analogamente ao que ocorre em $\mathbb{R}^n$, toda norma num espaço vetorial determina neste uma função distância e o torna um espaço métrico. No entanto, nem toda métrica num espaço vetorial é necessariamente proveniente de uma norma (vide Exercício 22). Assim, no universo dos espaços vetoriais, a classe dos espaços métricos inclui propriamente a classe dos espaços normados.

No tocante a sequências, vale salientar que toda a discussão feita na seção anterior pode ser levada a cabo no contexto mais geral dos espaços métricos. Basta substituir a norma pela métrica[ix]. Em particular, pode-se introduzir aí o conceito

[ix] O conceito de convergência de sequências num espaço métrico (M,d), por exemplo, é: Dado $a \in M$, diz-se que uma sequência (x_k) em (M,d) *converge* para a se, para todo $\epsilon > 0$, existe $k_0 \in \mathbb{N}$, tal que $k \geq k_0 \Rightarrow d(x_k, a) < \epsilon$.

de sequência de Cauchy. Neste sentido, merece menção o fato de que uma sequência de Cauchy num espaço métrico arbitrário não é, necessariamente, convergente (vide Exercício 23).

Espaços métricos em que todas as sequências de Cauchy são convergentes são ditos *completos*. Segue-se, portanto, do Teorema 2, que o espaço $\mathbb{R}^n$ munido da métrica euclidiana é um espaço métrico completo.

5.2. Isometrias de $\mathbb{R}^n$. A relação de equivalência mais natural na classe dos espaços métricos é, certamente, aquela estabelecida pelas aplicações (bijetivas) que preservam distância, chamadas de isometrias. Mais precisamente, dados dois espaços métricos (M_1, d_1) e (M_2, d_2), uma aplicação $\varphi : (M_1, d_1) \to (M_2, d_2)$ é dita uma *isometria* se é sobrejetiva e satisfaz

$$d_2(\varphi(x), \varphi(y)) = d_1(x, y) \ \ \forall x, y \in M_1 .$$

No caso afirmativo, diz-se que (M_1, d_1) e (M_2, d_2) são *isométricos*

Note que toda isometria é injetiva (portanto, bijetiva), pois

$$\varphi(x) = \varphi(y) \ \Rightarrow \ d_2(\varphi(x), \varphi(y)) = 0 \ \Rightarrow \ d_1(x, y) = 0 \ \Rightarrow x = y.$$

Deve-se observar também que "ser isométrico a" é, de fato, uma relação de equivalência na classe dos espaços métricos, pois, como se verifica facilmente, dados espaços métricos M_1, M_2, M_3, tem-se:

- A aplicação identidade de M_1 é uma isometria (reflexividade);
- $\varphi : M_1 \to M_2$ é uma isometria se, e somente se, $\varphi^{-1} : M_2 \to M_1$ é uma isometria (simetria);
- se $\varphi : M_1 \to M_2$ e $\psi : M_2 \to M_3$ são isometrias, então $\psi \circ \varphi : M_1 \to M_3$ é uma isometria (transitividade).

As isometrias de $\mathbb{R}^n$ que preservam a métrica euclidiana admitem a caracterização seguinte.

Teorema 3 (isometrias de $\mathbb{R}^n$). *Seja d a métrica euclidiana de $\mathbb{R}^n$. Então, uma aplicação $\varphi : (\mathbb{R}^n, d) \to (\mathbb{R}^n, d)$ é uma isometria se, e somente se,*

$$\varphi(x) = Tx + x_0 \quad \forall x \in \mathbb{R}^n,$$

em que $T \in \mathrm{L}(\mathbb{R}^n)$ é um operador ortogonal e $x_0 \in \mathbb{R}^n$.

Demonstração. Seja $T : \mathbb{R}^n \to \mathbb{R}^n$ um operador ortogonal e $\varphi = T + x_0$, $x_0 \in \mathbb{R}^n$. Neste caso, φ é bijetiva (pois T o é) e, para quaisquer $x, y \in \mathbb{R}^n$, $d(\varphi(x), \varphi(y)) = \|\varphi(x) - \varphi(y)\| = \|Tx - Ty\| = \|T(x - y)\| = \|x - y\| = d(x, y)$. Logo, φ é uma isometria.

Reciprocamente, suponhamos que $\varphi : \mathbb{R}^n \to \mathbb{R}^n$ seja uma isometria. Façamos $x_0 = \varphi(0)$ e consideremos a aplicação $T : \mathbb{R}^n \to \mathbb{R}^n$, definida por $T(x) = \varphi(x) - x_0$. Temos que $\|T(x) - T(y)\| = \|\varphi(x) - \varphi(y)\| = \|x - y\|$, isto é, T preserva distância. Em particular, T preserva norma, pois $T(0) = 0$. Segue-se, portanto, da identidade de polarização (8), que T preserva também produto interno, mesmo não sendo, necessariamente, linear. De fato, dados $x, y \in \mathbb{R}^n$, tem-se

$$\begin{aligned} \langle T(x), T(y) \rangle &= \frac{1}{2}(\|T(x)\|^2 + \|T(y)\|^2 - \|T(x) - T(y)\|^2) \\ &= \frac{1}{2}(\|x\|^2 + \|y\|^2 - \|x - y\|^2) \ = \ \langle x, y \rangle. \end{aligned}$$

Resta-nos, pois, provar que T é linear.

Uma vez que φ é bijetiva, T também o é. Além disso, T^{-1} preserva produto interno, pois, dados $a, b \in \mathbb{R}^n$, existem únicos $x, y \in \mathbb{R}^n$, tais que $T(x) = a$ e $T(y) = b$. Daí, temos que

$$\langle a, b\rangle = \langle T(x), T(y)\rangle = \langle x, y\rangle = \langle T^{-1}(a), T^{-1}(b)\rangle.$$

Tomando-se, desta forma, $x, y \in \mathbb{R}^n$ e $\lambda \in \mathbb{R}$, tem-se, para todo $z \in \mathbb{R}^n$,

$$\begin{aligned}
\langle T(\lambda x + y), z\rangle &= \langle \lambda x + y, T^{-1}(z)\rangle \\
&= \lambda\langle x, T^{-1}(z)\rangle + \langle y, T^{-1}(z)\rangle \\
&= \lambda\langle T(x), z\rangle + \langle T(y), z\rangle \\
&= \langle \lambda T(x) + T(y), z\rangle\,,
\end{aligned}$$

donde $T(\lambda x+y) = \lambda T(x)+T(y)$. Segue-se que T é linear e, portanto, um operador ortogonal de $\mathbb{R}^n$. $\square$

6. O Teorema Ergódico da Média (∗)

Nesta seção final, faremos uma breve discussão sobre um resultado fundamental da Teoria Ergódica, o qual foi estabelecido em 1931 pelo matemático húngaro (naturalizado americano) John von Neumann (1903–1957) e é conhecido como Teorema Ergódico da Média.

A Teoria Ergódica é uma das diversas subáreas da Teoria dos Sistemas Dinâmicos, a qual, por sua vez, estuda o comportamento das iterações (via composição) de aplicações do tipo $f : X \to X$ quando o número de iterações tende a infinito, isto é, o limite

$$\lim_{k\to\infty} f^k(x), \ x \in X.$$

Neste contexto, a especificidade de cada subárea da Teoria dos Sistemas Dinâmicos é determinada pela natureza do conjunto X e pelas propriedades das aplicações f. No que concerne à Teoria Ergódica, o conjunto X pertence a uma classe de espaços ditos *de medida* — na qual se inclui o espaço $\mathbb{R}^n$ quando se considera neste a noção de *volume*, que é uma medida — e a aplicação f tem a propriedade de *preservar medida*.

Dado um espaço de medida, X, a cada aplicação $f : X \to X$ que preserva medida, associa-se naturalmente um operador ortogonal T definido num *espaço de Hilbert*, $\mathbb{H}$, isto é, um espaço vetorial (não necessariamente de dimensão finita) com produto interno e completo com respeito à métrica induzida pelo mesmo. A partir deste fato e através de seu teorema da média, von Neumann estabeleceu, então, uma forte conexão entre a teoria analítica dos operadores ortogonais T e o comportamento geométrico das aplicações f (vide [**12**]).

Apresentaremos a versão particular do Teorema Ergódico da Média no caso em que $\mathbb{H}$ é o espaço $\mathbb{R}^n$. Em sua demonstração, faremos uso de uma ideia do matemático húngaro Frigyes Riesz (1880–1956), a qual se adapta, feitas ligeiras modificações, a diversos casos em que $\mathbb{H}$ é um espaço de Hilbert de dimensão infinita.

Teorema Ergódico da Média em $\mathbb{R}^n$ (von Neumann). *Sejam* $T \in \mathrm{L}(\mathbb{R}^n)$ *um operador ortogonal e* $\mathbb{V} = \{v \in \mathbb{R}^n; Tv = v\}$. *Então, denotando-se por* I *o* *operador identidade de* $\mathbb{R}^n$, *a sequência* $(T_k)_{k\in\mathbb{N}}$ *de* $\mathrm{L}(\mathbb{R}^n)$,

$$T_k = \frac{1}{k}\left(I + T + T^2 + \cdots + T^{k-1}\right),$$

é convergente e seu limite é a projeção ortogonal de $\mathbb{R}^n$ *sobre* $\mathbb{V}$, $P_{\mathbb{V}} : \mathbb{R}^n \to \mathbb{V}$.

Demonstração. Denotemos por $\mathbb{W} \subset \mathbb{R}^n$ o conjunto-imagem do operador $I - T$ e provemos, inicialmente, que $\mathbb{W} = \mathbb{V}^{\perp}$. Com efeito, uma vez que T é ortogonal, dado $v \in \mathbb{V}$, tem-se $v = T^*Tv = T^*v$. Logo, para quaisquer vetores $v \in \mathbb{V}$ e $w = y - Ty \in \mathbb{W}$, $y \in \mathbb{R}^n$, tem-se

$$\langle w, v\rangle = \langle y - Ty, v\rangle = \langle y, v\rangle - \langle Ty, v\rangle = \langle y, v\rangle - \langle y, T^*v\rangle = \langle y, v\rangle - \langle y, v\rangle = 0,$$

donde $\mathbb{W} \subset \mathbb{V}^{\perp}$. Para obtermos a inclusão contrária, basta observarmos que $\mathbb{V}$ é o núcleo de $I - T$. Daí, pelo Teorema do Núcleo e da Imagem, tem-se $\dim \mathbb{W} = n - \dim \mathbb{V} = \dim \mathbb{V}^{\perp}$, donde $\mathbb{W} = \mathbb{V}^{\perp}$. Em particular, $\mathbb{R}^n = \mathbb{V} \oplus \mathbb{W}$.

Assim, dado $x \in \mathbb{R}^n$, existem únicos vetores $v \in \mathbb{V}$ e $w = y - Ty \in \mathbb{W}$, tais que $x = v + w$. Entretanto,

$$T_k v = \frac{1}{k}\left(v + Tv + \cdots + T^{k-1}v\right) = \frac{kv}{k} = v,$$

donde $T_k v \to v$, e

$$T_k w = \frac{1}{k}\left(y - Ty + T(y - Ty) + \cdots + T^{k-1}(y - Ty)\right) = \frac{1}{k}(y - T^k y).$$

Porém, uma vez que T é um operador ortogonal, tem-se

$$\|y - T^k y\| \le \|y\| + \|T^k y\| = 2\|y\|.$$

Assim, $\|T_k w\| \le \frac{2}{k}\|y\|$, o que implica $T_k w \to 0$. Segue-se que

$$\lim T_k x = \lim T_k(v + w) = \lim T_k v + \lim T_k w = v = P_{\mathbb{V}}\, x.$$

Logo, pela Proposição 10, $T_k \to P_{\mathbb{V}}$, como desejávamos provar. □

7. Exercícios

Seções 1 e 2

1. Seja $\|\ \|$ uma norma em $\mathbb{R}$. Mostre que existe um real $\mu > 0$, tal que, para todo $x \in \mathbb{R}$, $\|x\| = \mu|x|$.

2. Prove que, para toda norma $\|\ \|$ num espaço vetorial $\mathbb{V}$, tem-se
$$|\,\|x\| - \|y\|\,| \le \|x - y\| \ \ \forall x, y \in \mathbb{V}.$$

3. Dados $x, y \in \mathbb{R}^n$, mostre que
$$\langle x, y\rangle(\|x\| + \|y\|) \le \|x\|\,\|y\|\,\|x + y\|$$
e verifique que não vale o mesmo se substituirmos $\langle x, y\rangle$ por $|\langle x, y\rangle|$.

4. Mostre que se $x, y \in \mathbb{R}^n$ são não-nulos e tais que $\|x + y\| = \|x\| + \|y\|$, então os vetores x, y são múltiplos um do outro. Mostre ainda que isto é falso nas normas do máximo e da soma.

5. Sejam $x, y, z \in \mathbb{R}^n$, tais que $\|x - z\| = \|x - y\| + \|y - z\|$. Use o exercício anterior para mostrar que existe $t \in [0,1]$ satisfazendo $y = (1-t)x + tz$. Interprete geometricamente.

6. Seja $\mathbb{V}$ um espaço vetorial. Prove que toda norma $\| \ \|$ em $\mathbb{V}$ que provém de um produto interno satisfaz a *identidade do paralelogramo*(x),
$$\|x+y\|^2 + \|x-y\|^2 = 2(\|x\|^2 + \|y\|^2) \ \forall x, y \in \mathbb{V}.$$
Conclua que nenhuma das normas em $\mathbb{R}^n$, da soma ou do máximo, provém de um produto interno.

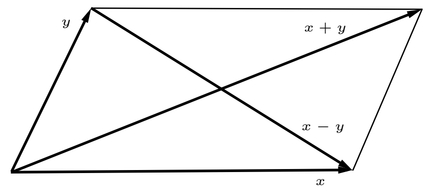

FIGURA 5

7. Prove o *Teorema de Representação de Riesz*: Para toda transformação linear $T \in \mathrm{L}(\mathbb{R}^n, \mathbb{R})$, existe um único vetor $a \in \mathbb{R}^n$, tal que
$$Tx = \langle a, x \rangle \ \forall x \in \mathbb{R}^n.$$

8. Seja $f : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^p$ uma aplicação bilinear. Mostre que existe um real $\mu > 0$, tal que, para todo par $(x, y) \in \mathbb{R}^n \times \mathbb{R}^m$, tem-se
$$\|f(x,y)\| \leq \mu \|x\| \, \|y\|.$$
Generalize para aplicações n-lineares.

9. Dado um isomorfismo linear $T : \mathbb{R}^n \to \mathbb{R}^n$, prove que são equivalentes as seguintes afirmações:
   i) $\measuredangle(Tx, Ty) = \measuredangle(x, y) \ \forall x, y \in \mathbb{R}^n - \{0\}$;
   ii) existe $\mu > 0$, tal que $\|Tx\| = \mu \|x\| \ \forall x \in \mathbb{R}^n$;
   iii) existe $\lambda > 0$, tal que $\langle Tx, Ty \rangle = \lambda \langle x, y \rangle \ \forall x, y \in \mathbb{R}^n$.

   Diz-se que T é um isomorfismo *conforme* se cumpre uma das (e, portanto, todas as) condições acima.

Seção 3

10. Use a igualdade (10) para calcular a norma espectral da transformação linear $T : \mathbb{R}^2 \to \mathbb{R}^2$, dada por $T(x_1, x_2) = (\sqrt{5}x_1 - 2x_2, 4x_1 + 2\sqrt{5}x_2)$.

11. Seja $Z \in \mathrm{L}(\mathbb{R}^n)$ um operador ortogonal. Prove que, para qualquer operador $T \in \mathrm{L}(\mathbb{R}^n)$, valem as igualdades $\|TZ\| = \|ZT\| = \|T\|$.

(x)Em geometria euclidiana, a igualdade dada corresponde ao fato de que a soma dos quadrados dos comprimentos das diagonais de um paralelogramo é igual à soma dos quadrados dos comprimentos dos seus lados (Fig. 5).

12. Dada $T \in \mathrm{L}(\mathbb{R}^n, \mathbb{R}^m)$, mostre que $\|T\| = \inf\{\mu \in \mathbb{R};\ \|Tx\| \leq \mu\|x\|\ \ \forall x \in \mathbb{R}^n\}$.

13. Seja $T \in \mathrm{L}(\mathbb{R}^n)$ um operador linear em $\mathbb{R}^n$, tal que $\|T - I\| < 1$. Prove que T é invertível.

14. Sejam $T \in \mathrm{L}(\mathbb{R}^n)$ e $\mu > 0$, tais que $\|Tx\| \geq \mu\|x\|$ para todo $x \in \mathbb{R}^n$. Mostre que T é invertível e $\|T^{-1}\| \leq 1/\mu$.

15. Dadas transformações lineares $T \in \mathrm{L}(\mathbb{R}^n, \mathbb{R}^m)$ e $Z \in \mathrm{L}(\mathbb{R}^m, \mathbb{R}^n)$, mostre que $\lambda \neq 0$ é autovalor de ZT se, e somente se, é autovalor de TZ. Conclua daí que $\|T\| = \|T^*\|\ \forall T \in \mathrm{L}(\mathbb{R}^n, \mathbb{R}^m)$.

Seção 4

16. Diz-se que um conjunto $X \subset \mathbb{R}^n$ é *denso* em $\mathbb{R}^n$ quando todo ponto de $\mathbb{R}^n$ é limite de uma sequência de pontos de X. Prove que:
    i) $\mathbb{Q}^n = \{(x_1, \ldots, x_n) \in \mathbb{R}^n\,;\ x_i \in \mathbb{Q}\ \forall i = 1, \ldots, n\}$ é denso em $\mathbb{R}^n$;
    ii) se $X \subset \mathbb{R}^n$ é denso em $\mathbb{R}^n$ e (x_k) é uma sequência em $\mathbb{R}^n$ que satisfaz, para um certo $a \in \mathbb{R}^n$ e todo $x \in X$, $\|x_k - x\| \to \|a - x\|$, então $x_k \to a$.

17. Sejam $T \in \mathrm{L}(\mathbb{R}^n, \mathbb{R}^m)$ uma transformação linear injetiva e (x_k) uma sequência em $\mathbb{R}^n$, tais que (Tx_k) é uma sequência convergente em $\mathbb{R}^m$. Prove que (x_k) é convergente.

18. Seja (T_k) uma sequência convergente em $\mathrm{L}(\mathbb{R}^n, \mathbb{R}^m)$, tal que $T_k \to T$. Use o resultado do Exercício 15 para provar que $T_k^* \to T^*$ em $\mathrm{L}(\mathbb{R}^m, \mathbb{R}^n)$. Conclua, então, que toda sequência de operadores ortogonais em $\mathrm{L}(\mathbb{R}^n)$ possui uma subsequência que converge para um operador ortogonal.

19. Seja (A_k) uma sequência em $\mathrm{M}(n)$, tal que A_k é invertível para todo $k \in \mathbb{N}$. Mostre que se $A_k \to A \in \mathrm{M}(n)$ e a sequência (A_k^{-1}) é limitada, então A é invertível.

20. Prove que toda matriz $A \in \mathrm{M}(n)$ é limite de uma sequência (A_k), em $\mathrm{M}(n)$, tal que, para todo $k \in \mathbb{N}$, A_k é invertível e distinta de A.

21. Sejam $\mathbb{V}, \mathbb{W} \subset \mathbb{R}^n$ subespaços vetoriais de $\mathbb{R}^n$ e $P_{\mathbb{V}}$, $P_{\mathbb{W}}$ as respectivas projeções ortogonais de $\mathbb{R}^n$ sobre $\mathbb{V}$ e $\mathbb{W}$. Prove que:
    i) $\mathbb{V} \cap \mathbb{W} = \{v \in \mathbb{R}^n;\ (P_{\mathbb{V}} P_{\mathbb{W}} P_{\mathbb{V}})v = v\}$;
    ii) $\lim\limits_{k \to \infty} (P_{\mathbb{V}} P_{\mathbb{W}} P_{\mathbb{V}})^k = P_{\mathbb{V} \cap \mathbb{W}}$ (Fig. 6).

Seção 5

22. A fim de que uma métrica d num espaço vetorial $\mathbb{V}$ seja proveniente de uma norma, é necessário e suficiente que, para $x, a \in \mathbb{V}$ e $\lambda \in \mathbb{R}$ arbitrários, se tenha
$$d(x + a, y + a) = d(x, y) \qquad \text{e} \qquad d(\lambda x, \lambda y) = |\lambda| d(x, y).$$
Prove esta afirmação e conclua que a função $d' : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$, definida por
$$d'(x, y) = \sqrt{\|x - y\|}\,,$$

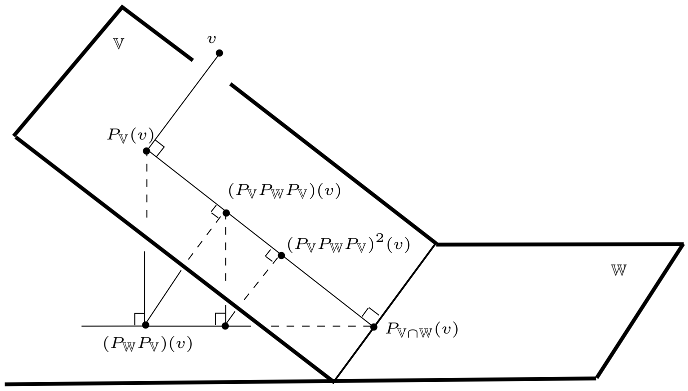

FIGURA 6

é uma métrica em $\mathbb{R}^n$ que não provém de produto interno algum.

23. Sejam $M = (0,1] \subset \mathbb{R}$ e d a métrica euclidiana de $\mathbb{R}$. Mostre que (M,d) não é um espaço métrico completo e que (M,d') é um espaço métrico completo, em que
$$d'(x,y) = \left|\frac{1}{x} - \frac{1}{y}\right|, \quad x,y \in M.$$

24. Sejam $d_{\max}$ a métrica de $\mathbb{R}^2$ proveniente da norma do máximo e d a métrica euclidiana de $\mathbb{R}$. Considere o subespaço $\mathbb{V} = \{(t,at) \in \mathbb{R}^2;\, t \in \mathbb{R}\} \subset \mathbb{R}^2$ e prove que a projeção $P : (\mathbb{V}, d_{\max}) \to (\mathbb{R}, d)$, $P(x_1,x_2) = x_2$, é uma isometria se, e somente se, $|a| \geq 1$.

25. Considere a *esfera* de $\mathbb{R}^n$ com centro na origem e raio igual a 1,
$$S = \{x \in \mathbb{R}^n;\, \|x\| = 1\}.$$
Prove que, para quaisquer pontos distintos $x,y \in S$, existe uma isometria $\varphi : (S,d) \to (S,d)$ (d a métrica euclidiana de $\mathbb{R}^n$), tal que $\varphi(x) = y$.

CAPíTULO 2

O Espaço Topológico $\mathbb{R}^n$

Desde sua origem até meados meados do século XIX, a Geometria consistia, essencialmente, do estudo das propriedades de seus objetos — curvas e superfícies, em geral — que permanecem invariantes quando estes são submetidos a deformações isométricas, isto é, que preservam distância.

A partir desta época, geômetras como August Moebius (1790–1868), Johann Listing (1808–1882) e Bernhard Riemann (1826–1866) perceberam a necessidade de se estabelecer uma teoria geométrica em que se permitissem quaisquer deformações dos objetos, não somente as isométricas, e se observassem, então, as propriedades que permaneciam invariantes com relação às mesmas. A este novo tipo de geometria chamou-se *Topologia*[(i)] (do grego *tópos* = lugar + *lógos* = tratado) que, desde então, desenvolveu-se e consolidou-se como uma das mais fecundas e importantes teorias da Matemática.

Neste surgimento da Topologia, deve-se a Riemann o primeiro grande trabalho, o qual constituiu a sua tese de doutorado, apresentada em 1851. Nela, Riemann incorpora a Topologia à teoria das funções analíticas de variável complexa, introduzindo uma importante classe de superfícies que, presentemente, levam o seu nome.

Os fundamentos da Topologia, tal qual a estudamos hoje, foram estabelecidos, notadamente, por Henry Poincaré (1854–1912), que em 1895, através de conceitos algébricos, introduziu as noções topológicas de *homotopia* e *homologia*, inaugurando assim a *Topologia Algébrica*; Luitzen Brouwer (1881–1966), que por volta de 1910, através de uma nova abordagem à teoria — na qual os argumentos das demonstrações não apelavam à intuição geométrica —, obteve resultados de extrema significância, dentre eles, os célebres *Teorema do Ponto Fixo* e *Teorema da Invariância do Domínio*; e Felix Hausdorff (1868–1942), que em 1914 introduziu o conceito de *espaço topológico*[(ii)].

Os espaços topológicos são, pois, os objetos de estudo da Topologia. Estes são conjuntos munidos de uma certa estrutura denominada, igualmente, *topologia*. Tais espaços, dentre os quais se incluem os espaços métricos, constituem os ambientes mais gerais onde se pode considerar a noção de convergência, que, conforme assinalamos anteriormente, é fundamental em Análise.

Desta forma, introduziremos neste capítulo os conceitos elementares da Topologia que nos permitirão tratar $\mathbb{R}^n$, $\mathrm{L}(\mathbb{R}^n, \mathbb{R}^m)$ e $\mathrm{M}(m,n)$ como espaços topológicos. Constataremos, então, nos capítulos posteriores, a importância e indispensabilidade desta conceituação, uma vez que, de modo geral, o comportamento das funções

[(i)]O termo "topologia" foi introduzido por Listing em seu livro, *Vorstudien zur Topologie*, o qual foi publicado em 1847 e consta como o primeiro sobre o assunto.

[(ii)]O atual conceito de espaço topológico é mais geral que o introduzido por Hausdorff. Os espaços por ele concebidos são ditos, hoje, *espaços de Hausdorff*, sobre os quais discutiremos na Seção 6.

reflete-se nas propriedades topológicas dos conjuntos onde estas são definidas ou tomam valores.

1. Conjuntos Abertos – Topologia

Os intervalos da reta real generalizam-se em $\mathbb{R}^n$ através do conceito de *bola.* Relembremos que, dados $a \in \mathbb{R}^n$ e um real $r > 0$, a *bola aberta* de centro a e raio r, em $\mathbb{R}^n$, é o conjunto

$$B(a,r) = \{x \in \mathbb{R}^n;\ \|x-a\| < r\}.$$

Definem-se ainda a *bola fechada* e a *esfera*, ambos de centro a e raio r, respectivamente por

$$B[a,r] = \{x \in \mathbb{R}^n;\ \|x-a\| \leq r\} \quad \text{e} \quad S[a,r] = \{x \in \mathbb{R}^n;\ \|x-a\| = r\}.$$

Qualquer bola ou esfera de $\mathbb{R}^n$ de raio 1 é dita *unitária* e a esfera unitária de $\mathbb{R}^n$ cujo centro é a origem é denotada por S^{n-1}.

Note que valem as igualdades

$$B(a,r) = \{a+h \in \mathbb{R}^n;\ \|h\| < r\} \quad \text{e} \quad B[a,r] = \{a+h \in \mathbb{R}^n;\ \|h\| \leq r\},$$

e que, em $\mathbb{R}$,

$$B(a,r) = (a-r, a+r), \quad B[a,r] = [a-r, a+r] \quad \text{e} \quad S[a,r] = \{a-r, a+r\}.$$

O conceito de bola (bem como o de esfera) em $\mathbb{R}^n$, conforme introduzido, estende-se facilmente aos espaços métricos, pois podemos também escrever

$$B(a,r) = \{x \in \mathbb{R}^n;\ d(x,a) < r\} \quad \text{e} \quad B[a,r] = \{x \in \mathbb{R}^n;\ d(x,a) \leq r\}.$$

Em particular, podemos considerar normas em $\mathbb{R}^n$ cujas bolas têm configurações geométricas distintas das suas análogas euclidianas. Por exemplo, na norma do máximo em $\mathbb{R}^2$, a bola fechada de raio 1 centrada na origem é o quadrado cujos vértices são os pontos $(-1,-1)$, $(-1,1)$, $(1,-1)$ e $(1,1)$, enquanto na norma da soma, a bola fechada com mesmo centro e raio é o quadrado cujos vértices são os pontos $(-1,0)$, $(0,-1)$, $(1,0)$ e $(0,1)$ (Fig. 1).

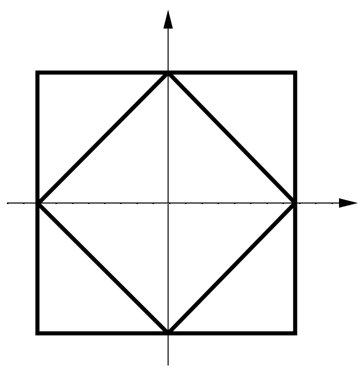

FIGURA 1

Observação 10. Consideremos duas normas $\| \ \|_1$ e $\| \ \|_2$ em $\mathbb{R}^n$. Como sabemos (Proposição 7 - Capítulo 1), elas são equivalentes, isto é, existem reais positivos λ, μ, tais que, para todo $x \in \mathbb{R}^n$,

$$\|x\|_1 \leq \lambda \|x\|_2 \qquad \text{e} \qquad \|x\|_2 \leq \mu \|x\|_1 .$$

Desta forma, designando-se respectivamente por $B_1(a,r)$ e $B_2(a,r)$ as bolas de centro $a \in \mathbb{R}^n$ e raio $r > 0$ com respeito às normas $\| \ \|_1$ e $\| \ \|_2$, tem-se

$$B_1(a, r/\mu) \subset B_2(a,r) \quad \text{e} \quad B_2(a, r/\lambda) \subset B_1(a,r),$$

isto é, para toda bola aberta B_2 na norma $\| \ \|_2$, existe uma bola aberta B_1 (com mesmo centro) na norma $\| \ \|_2$, tal que $B_1 \subset B_2$ e reciprocamente. Na Figura 2 ilustramos este fato no caso em que as normas são a euclidiana e a do máximo.

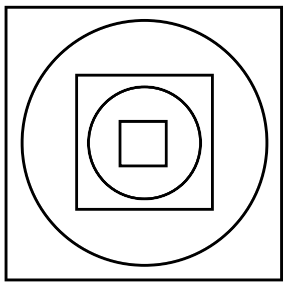

Figura 2

Um conjunto $X \subset \mathbb{R}^n$ é dito *limitado* se existe uma bola que o contém, isto é, se existem $a \in \mathbb{R}^n$ e $r > 0$, tais que $X \subset B(a,r)$. Dado um conjunto A, uma aplicação $f : A \to \mathbb{R}^n$ é dita *limitada* quando seu conjunto-imagem é limitado.

Definição 6 (interior – vizinhança – conjunto aberto). Diz-se que um ponto $a \in X \subset \mathbb{R}^n$ é *interior* ao conjunto X se existe uma bola aberta com centro em a e contida em X. Neste caso, diz-se também que X é uma *vizinhança* de a. O *interior* de X, $\operatorname{int} X$, é o conjunto formado por todos os pontos que são interiores a X. Um conjunto $A \subset \mathbb{R}^n$ é chamado de *aberto* se todos os seus pontos são interiores, isto é, se $\operatorname{int} A = A$.

Considerando-se a igualdade $B(a,r) = \{a + h \in \mathbb{R}^n ; \|h\| < r\}$, tem-se que um ponto $a \in \mathbb{R}^n$ é interior a um conjunto $X \subset \mathbb{R}^n$ quando existe $r > 0$, tal que, para todo $h \in \mathbb{R}^n$ satisfazendo $\|h\| < r$, tem-se $a + h \in X$ isto é, quando a pode "mover-se" em qualquer direção sem sair de X.

Exemplo 10. *Toda bola aberta $B(a,r)$ de $\mathbb{R}^n$ é um conjunto aberto.* De fato, dado $b \in B(a,r)$, temos $\|b - a\| < r$. Logo, existe $\delta \in \mathbb{R}$ satisfazendo $0 < \delta < r - \|b - a\|$. Tomando-se, então, a bola $B(b,\delta)$ e $x \in B(b,\delta)$, tem-se $\|x - b\| < \delta$. Assim, $\|x-a\| \leq \|x-b\| + \|b-a\| < \delta + \|b-a\| < r$, isto é, $x \in B(a,r)$, o que nos dá $B(b,\delta) \subset B(a,r)$ e prova que $B(a,r)$ é aberto (Fig. 3).

De maneira semelhante, prova-se que são abertos de $\mathbb{R}^n$ os seguintes conjuntos:

i) $\mathbb{R}^n - B[a,r] = \{x \in \mathbb{R}^n ; \|x - a\| > r\}$;

ii) $E_i = \{x \in \mathbb{R}^n ; x_i = \langle x, e_i \rangle > 0\}$, $i = 1, \ldots, n$;

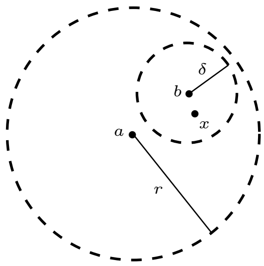

FIGURA 3

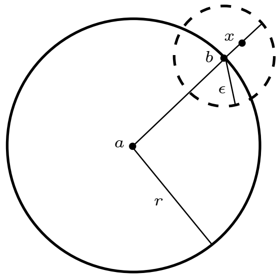

FIGURA 4

iii) $A(a, r_0, r_1) = \{x \in \mathbb{R}^n;\, r_0 < \|x - a\| < r_1,\ a \in \mathbb{R}^n,\ 0 < r_0 < r_1\}$.

Cada conjunto E_i, em (ii), é dito um *semi-espaço aberto*, enquanto $A(a, r_0, r_1)$, em (iii), é chamado de *anel aberto* (com *centro* em a, *raio menor* r_0 e *raio maior* r_1).

EXEMPLO 11. *Nenhuma bola fechada ou esfera de* $\mathbb{R}^n$ *é um conjunto aberto.* De fato, sejam $B[a, r] \subset \mathbb{R}^n$ uma bola fechada e $b \in B[a, r]$, tais que $\|b - a\| = r$, isto é, $b \in S[a, r]$. Dada, então, uma bola $B(b, \epsilon)$, tomemos $x = \lambda(b - a) + a = (\lambda - 1)(b - a) + b \in \mathbb{R}^n$, em que $1 < \lambda < \epsilon/r + 1$ (Fig. 4). Temos que,

$$\|x - a\| = \lambda\|b - a\| = \lambda r > r \quad \text{e} \quad \|x - b\| = \|(\lambda - 1)(b - a)\| = (\lambda - 1)r < \epsilon.$$

Segue-se que $x \in B(b, \epsilon)$ e $x \notin B[a, r]$. Assim, nenhuma bola de $\mathbb{R}^n$ centrada em b está contida em $S[a, r]$ ou $B[a, r]$. Daí e do Exemplo 10, infere-se que

$$\operatorname{int}(S[a, r]) = \emptyset \quad \text{e} \quad \operatorname{int}(B[a, r]) = B(a, r).$$

Em particular, nenhum dos conjuntos, $S[a, r]$ ou $B[a, r]$, é aberto.

De modo análogo, prova-se que o conjunto

$$\mathbb{R}^n - B(a, r) \;=\; \{x \in \mathbb{R}^n; \|x - a\| \geq r\},$$

igualmente, não é aberto.

EXEMPLO 12. Dado $i \in \mathbb{N}$ satisfazendo $1 \leq i \leq n$, denotemos por U_i o subconjunto de $\mathbb{R}^n$ formado pelos pontos $x = (x_1, \dots x_i, \dots, x_n)$, tais que

$$|x_i| > |x_j| \ \forall j \in \{1, \dots, i - 1, i + 1, \dots, n\}.$$

Tomando-se, então, $x \in U_i$ e fazendo-se

$$\delta = \frac{1}{2} \min_j \{|x_i| - |x_j|\,;\, i \neq j\},$$

tem-se, para todo $h = (h_1, \dots, h_n) \in \mathbb{R}^n$ satisfazendo $\|h\| < \delta$, que

$$|x_i| - |x_j| \geq 2\delta > 2\|h\| \geq 2\|h\|_{\max} \geq |h_i| + |h_j| \ \ \forall j \neq i,$$

isto é, para todo $j \neq i$, $|x_i| - |h_i| > |x_j| + |h_j|$. Sendo assim,

$$|x_i + h_i| \geq |x_i| - |h_i| > |x_j| + |h_j| \geq |x_j + h_j| \ \ \forall j \neq i,$$

donde se infere que $x + h \in U_i$. Segue-se que $B(x, \delta) \subset U_i$ e, portanto, que U_i é um subconjunto aberto de $\mathbb{R}^n$.

EXEMPLO 13. Consideremos os seguintes subconjuntos de $\mathbb{R}^n$,

- $\mathbb{Q}^n = \{(x_1, \dots, x_n) \in \mathbb{R}^n\,;\, x_i \in \mathbb{Q}\ \forall i = 1, \dots, n\}$;

- $(\mathbb{R}-\mathbb{Q})^n = \{(x_1,\dots,x_n)\in\mathbb{R}^n\,;\, x_i\in(\mathbb{R}-\mathbb{Q})\ \forall i=1,\dots,n\}$;

e verifiquemos que nenhuma bola aberta de $\mathbb{R}^n$ está contida em $\mathbb{Q}^n$ ou $(\mathbb{R}-\mathbb{Q})^n$, donde se conclui que $\operatorname{int}\mathbb{Q}^n = \operatorname{int}(\mathbb{R}-\mathbb{Q})^n = \emptyset$ e, em particular, que nenhum destes conjuntos é aberto.

Para tanto, tomemos $a=(a_1,\dots,a_n)\in\mathbb{R}^n$ e $r>0$. Uma vez que $\mathbb{R}-\mathbb{Q}$ e $\mathbb{Q}$ são, ambos, densos em $\mathbb{R}$, existem $x_1\in\mathbb{R}-\mathbb{Q}$ e $y_1\in\mathbb{Q}$, tais que $x_1,y_1\in(a_1-r,a_1+r)$. Os pontos $x=(x_1,a_2,\dots,a_n)$ e $y=(y_1,a_2,\dots,a_n)$, desta forma, satisfazem $\|x-a\|=|x_1-a_1|<r$ e $\|y-a\|=|y_1-a_1|<r$, donde $x\in B(a,r)-\mathbb{Q}^n$ e $y\in B(a,r)-(\mathbb{R}-\mathbb{Q})^n$. Logo, a bola aberta $B(a,r)$ não está contida em $\mathbb{Q}^n$ ou $(\mathbb{R}-\mathbb{Q})^n$.

Diz-se que uma aplicação $f:\mathbb{R}^n\to\mathbb{R}^m$ é *aberta* quando, para todo aberto $A\subset\mathbb{R}^n$, $f(A)\subset\mathbb{R}^m$ é aberto.

Exemplo 14. *Todo isomorfismo linear $T\in\mathrm{L}(\mathbb{R}^n)$ é uma aplicação aberta.* Com efeito, seja $A\subset\mathbb{R}^n$ aberto. Então, para todo $a\in A$, existe $r>0$, tal que $B(a,r)\subset A$. Fazendo-se, desta forma, $r_0=r/\|T^{-1}\|$ e tomando-se $y\in B(Ta,r_0)$, tem-se que $x=T^{-1}y$ satisfaz $\|x-a\|=\|T^{-1}y-T^{-1}Ta\|=\|T^{-1}(y-Ta)\|\le\|T^{-1}\|\,\|y-Ta\|<\|T^{-1}\|r_0=r$, isto é, $x\in B(a,r)\subset A$, o que implica $y=Tx\in T(A)$. Segue-se que a bola aberta $B(Ta,r_0)$ está contida em $T(A)$, donde se infere que este conjunto é aberto.

Exemplo 15. *Toda projeção ortogonal sobre um subespaço n-dimensional é uma aplicação aberta.* Mais especificamente, seja $\mathbb{V}$ um subespaço n-dimensional de $\mathbb{R}^{n+m}$. Identificando-se $\mathbb{V}$ e $\mathbb{V}^\perp$ com $\mathbb{R}^n$ e $\mathbb{R}^m$, respectivamente, bem como $\mathbb{R}^{n+m}$ com o produto $\mathbb{V}\times\mathbb{V}^\perp$, temos que a projeção ortogonal de $\mathbb{R}^{n+m}$ sobre $\mathbb{V}$ passa a ser a aplicação

$$\begin{array}{cccc} P: & \mathbb{R}^n\times\mathbb{R}^m & \to & \mathbb{R}^n \\ & (x,y) & \mapsto & x, \end{array}$$

a qual provaremos ser aberta.

Para tanto, verifiquemos inicialmente que, para todo ponto $a=(x_0,y_0)$ de $\mathbb{R}^n\times\mathbb{R}^m$, tem-se $P(B(a,r))=B(P(a),r)$. De fato, dado $x\in P(B(a,r))$, existe $y\in\mathbb{R}^m$, tal que $b=(x,y)\in B(a,r)$, isto é, $P(b)=x$. Desta forma, $r^2>\|b-a\|^2=\|x-x_0\|^2+\|y-y_0\|^2$. Logo, $\|x-x_0\|^2<r^2$, o que implica $x\in B(x_0,r)$, ou seja, $P(B(a,r))\subset B(x_0,r)=B(P(a),r)$. Agora, dado $x'\in B(x_0,r)$, temos $\|x'-x_0\|^2<r^2$, donde $c=(x',y_0)$ satisfaz $\|c-a\|^2<r^2$, isto é, $c\in B(a,r)$. Assim, $x'=P(c)\in P(B(a,r))$, ou seja, $B(x_0,r)\subset P(B(a,r))$ e, portanto, $P(B(a,r))=B(P(a),r)$.

Suponhamos agora que $A\subset\mathbb{R}^n\times\mathbb{R}^m$ seja aberto e tomemos $P(a)\in P(A)$. Neste caso, existe uma bola aberta $B(a,r)$ contida em A, donde se infere que $B(P(a),r))=P(B(a,r))\subset P(A)$. Logo, $P(A)$ é aberto.

O conceito de conjunto aberto pode ser introduzido em qualquer espaço métrico de forma análoga àquela feita em $\mathbb{R}^n$. No que se segue, vamos considerá-lo nos espaços $\mathrm{L}(\mathbb{R}^n,\mathbb{R}^m)$ e $\mathrm{M}(m,n)$ munidos de suas respectivas métricas determinadas pela norma espectral[iii].

[iii]O mesmo se aplica a outros conceitos topológicos que consideraremos neste capítulo, os quais serão introduzidos em $\mathbb{R}^n$ e, eventualmente, considerados em $\mathrm{L}(\mathbb{R}^n,\mathbb{R}^m)$ e $\mathrm{M}(m,n)$ sem mais comentários.

Exemplo 16. *O conjunto* $\mathrm{I}(\mathbb{R}^n) \subset \mathrm{L}(\mathbb{R}^n)$, *dos operadores invertíveis de* $\mathbb{R}^n$, *é um aberto de* $\mathrm{L}(\mathbb{R}^n)$. Provaremos isso mostrando que, dado um operador linear invertível, $T \in \mathrm{I}(\mathbb{R}^n)$, existe $\mu > 0$, tal que, para todo operador linear $H \in \mathrm{L}(\mathbb{R}^n)$ satisfazendo $\|H\| < \mu$, tem-se

$$\|(T+H)x\| \geq \mu\|x\| \, \forall x \in \mathbb{R}^n. \tag{12}$$

Neste caso, para cada tal H, $T+H$ é invertível, por ter núcleo trivial, donde se infere que $B(T,\mu) \subset I(\mathbb{R}^n)$ e, portanto, que $\mathrm{I}(\mathbb{R}^n)$ é aberto.

Tomemos, então, $x \in \mathbb{R}^n$ e observemos que $\|x\| = \|T^{-1}Tx\| \leq \|T^{-1}\| \, \|Tx\|$, o que nos dá $\|Tx\| \geq \|x\|/\|T^{-1}\|$. Logo, uma vez que $\|Hx\| \leq \|H\| \, \|x\|$, tem-se

$$\|(T+H)x\| = \|Tx + Hx\| \geq \|Tx\| - \|Hx\| \geq \left(\frac{1}{\|T^{-1}\|} - \|H\|\right)\|x\|.$$

Assim, fazendo-se $\mu = 1/2\|T^{-1}\|$ obtém-se o desejado, pois, para todo operador $H \in \mathrm{L}(\mathbb{R}^n)$ que cumpre $\|H\| < \mu$, tem-se

$$\frac{1}{\|T^{-1}\|} - \|H\| > \frac{1}{\|T^{-1}\|} - \mu = \frac{1}{\|T^{-1}\|} - \frac{1}{2\|T^{-1}\|} = \frac{1}{\|T^{-1}\|} = \mu.$$

Equivalentemente, *o conjunto* $\mathrm{I}(n) \subset \mathrm{M}(n)$, *das matrizes invertíveis de ordem* n, *é um aberto de* $\mathrm{M}(n)$.

1.1. Topologia em $\mathbb{R}^n$. Na proposição a seguir, verificaremos as propriedades fundamentais dos conjuntos abertos de $\mathbb{R}^n$, as quais nos permitirão caracterizá-lo como espaço topológico.

Teorema 4 (propriedades fundamentais dos abertos). *Os conjuntos abertos de* $\mathbb{R}^n$ *têm as seguintes propriedades:*

i) *O conjunto vazio e o espaço* $\mathbb{R}^n$ *são abertos;*

ii) *a interseção de uma família finita de abertos é um conjunto aberto;*

iii) *a união de uma família qualquer de abertos é um conjunto aberto.*

Demonstração. (i) Se o conjunto vazio não fosse aberto, ele possuiria um elemento x, tal que nenhuma bola centrada em x estaria nele contida, o que é absurdo. Logo, o vazio é aberto. Segue-se diretamente da definição de conjunto aberto que $\mathbb{R}^n$ é aberto.

(ii) Sejam $A_1, A_2, \dots, A_k \subset \mathbb{R}^n$ conjuntos abertos e $A = A_1 \cap \cdots \cap A_k$. Tem-se, pelo item (i), que se $A = \emptyset$, então A é aberto. Caso contrário, seja $a \in A$. Como cada A_i é aberto, para cada $i = 1, \dots, k$, existe $r_i > 0$, tal que $B(a, r_i) \subset A_i$. Tomando-se, então, $r = \min\{r_1, \dots r_k\}$, tem-se $B(a,r) \subset B(a,r_i) \subset A_i$ para todo $i = 1, \dots, k$. Logo, $B(a,r) \subset A$, donde A é aberto.

(iii) Seja $\{A_\lambda\}_{\lambda \in \Lambda}$ uma família de abertos (não todos vazios) de $\mathbb{R}^n$. Dado $a \in A = \bigcup A_\lambda$, existe $\lambda \in \Lambda$, tal que $a \in A_\lambda$. Logo, existe $r > 0$, tal que $B(a,r) \subset A_\lambda \subset A$, donde se conclui que A é aberto. □

Convém observar que a interseção de uma família infinita de abertos pode não ser um aberto. O exemplo clássico é $\{A_k\}_{k \in \mathbb{N}}$, em que $A_k = B(0, 1/k)$. Claramente, $\bigcap A_k = \{0\}$, que não é aberto.

Um *espaço topológico* é um par (X, τ), em que X é um conjunto e τ é uma família de subconjuntos de X, chamados de *abertos*, que satisfazem as condições (i), (ii) e (iii) do Teorema 4 (substituindo-se $\mathbb{R}^n$ por X, evidentemente). Diz-se, então, que a família τ define uma *topologia* em X.

Desta forma, os conjuntos abertos de $\mathbb{R}^n$, conforme introduzidos, definem no mesmo uma topologia e o tornam um espaço topológico. Deve-se observar que esta topologia foi introduzida por meio de bolas abertas da métrica euclidiana, o que a caracteriza como a topologia de $\mathbb{R}^n$ *proveniente* desta métrica.

Duas normas distintas em $\mathbb{R}^n$, embora definam métricas distintas, determinam neste a mesma topologia. De fato, se τ_1 e τ_2 são as topologias de $\mathbb{R}^n$ determinadas por duas tais métricas, pelas considerações da Observação 10, para todo subconjunto A de $\mathbb{R}^n$, tem-se $A \in \tau_1$ se, e somente se, $A \in \tau_2$, donde $\tau_1 = \tau_2$.

Cabe-nos mencionar, ainda, que, conforme discutiremos mais detalhadamente na Seção 6, podem-se definir topologias em $\mathbb{R}^n$ que não são provenientes de métrica alguma.

A proposição seguinte caracteriza topologicamente as sequências convergentes.

PROPOSIÇÃO 11 (TOPOLOGIA E CONVERGÊNCIA). *Uma sequência* (x_k) *em* $\mathbb{R}^n$ *converge para* $a \in \mathbb{R}^n$ *se, e somente se, para toda vizinhança* V *de* a, *existe* $k_0 \in \mathbb{N}$, *tal que* $k \geq k_0 \Rightarrow x_k \in V$.

DEMONSTRAÇÃO. Suponhamos que se tenha $x_k \to a \in \mathbb{R}^n$. Dada, então, uma vizinhança V de a, existe $\epsilon > 0$, tal que $B(a, \epsilon) \subset V$. Porém, para este ϵ, existe $k_0 \in \mathbb{N}$, tal que $k \geq k_0 \Rightarrow x_k \in B(a, \epsilon) \subset V$.

Reciprocamente, suponhamos que, para toda vizinhança V de a, exista $k_0 \in \mathbb{N}$, tal que $k \geq k_0 \Rightarrow x_k \in V$. Neste caso, dado $\epsilon > 0$, podemos fazer $V = B(a, \epsilon)$ e concluir que $x_k \to a$. □

2. Conjuntos Fechados

Consideraremos agora a família de subconjuntos de $\mathbb{R}^n$ formada pelos complementares de seus abertos, ditos fechados, e suas relações com o conceito de convergência de sequências.

DEFINIÇÃO 7 (CONJUNTO FECHADO). Um conjunto $F \subset \mathbb{R}^n$ é dito *fechado* se seu complementar, $\mathbb{R}^n - F$, é aberto.

O conjunto vazio e o espaço $\mathbb{R}^n$ são exemplos imediatos de conjuntos fechados. Uma vez que estes conjuntos são também abertos, vê-se que, do ponto de vista da Topologia, "aberto" e "fechado" não são antônimos, como na linguagem cotidiana.

Outros exemplos de conjuntos fechados são as bolas fechadas, as esferas, os anéis fechados e os semi-espaços fechados (estes dois últimos são definidos como seus análogos abertos, trocando-se os sinais $<$ e $>$ por $\leq$ e $\geq$, respectivamente). É fácil ver também que qualquer subconjunto finito de $\mathbb{R}^n$ é fechado.

Dados $a \in \mathbb{R}^n$ e $r > 0$, o complementar de $F = \{x \in \mathbb{R}^n; \|x - a\| \geq r\}$ é a bola aberta $B(a, r)$. Logo, F é fechado. Por outro lado, sabemos que F não é aberto, donde se conclui que nenhuma bola aberta de $\mathbb{R}^n$ é um conjunto fechado.

Note ainda que existem subconjuntos de $\mathbb{R}^n$ que não são abertos ou fechados, como, por exemplo, $\mathbb{Q}^n$ ou $\{x \in \mathbb{R}^n; r_0 < \|x\| \leq r_1, 0 < r_0 < r_1\}$.

Relembremos que, dada uma família arbitrária $\{F_\lambda\}_{\lambda \in \Lambda}$ de subconjuntos de $\mathbb{R}^n$, valem as igualdades

$$\mathbb{R}^n - \bigcup F_\lambda = \bigcap (\mathbb{R}^n - F_\lambda) \qquad \text{e} \qquad \mathbb{R}^n - \bigcap F_\lambda = \bigcup (\mathbb{R}^n - F_\lambda).$$

Daí e do Teorema 4 segue-se facilmente o resultado seguinte.

TEOREMA 5 (PROPRIEDADES FUNDAMENTAIS DOS FECHADOS). *Os conjuntos fechados de* $\mathbb{R}^n$ *têm as seguintes propriedades:*

i) *O conjunto vazio e o espaço* $\mathbb{R}^n$ *são fechados;*

ii) *a união de uma família finita de fechados é um conjunto fechado;*

iii) *a interseção de uma família qualquer de fechados é um conjunto fechado.*

Vale salientar que a união arbitrária de fechados não é, necessariamente, um fechado. Para constatarmos isto, basta observarmos que um conjunto que não é fechado é dado pela união dos seus pontos, que são conjuntos fechados.

Diz-se que uma aplicação $f : \mathbb{R}^n \to \mathbb{R}^m$ é *fechada* quando leva fechados de $\mathbb{R}^n$ em fechados de $\mathbb{R}^m$.

EXEMPLO 17. *Todo isomorfismo linear* $T \in \mathrm{L}(\mathbb{R}^n)$ *é uma aplicação fechada.* De fato, dado $F \subset \mathbb{R}^n$ fechado, temos que $A = \mathbb{R}^n - F$ é aberto e, pelo Exemplo 14, que $T(A)$ é aberto. Porém, sendo T bijetiva, tem-se $T(A) = T(\mathbb{R}^n - F) = T(\mathbb{R}^n) - T(F) = \mathbb{R}^n - T(F)$. Assim, $\mathbb{R}^n - T(F)$ é aberto e, portanto, $T(F)$ é fechado.

2.1. Aderência – Fronteira – Pontos de Acumulação. Introduziremos agora alguns conceitos topológicos que, dentre outros benefícios, nos permitirão caracterizar os conjuntos fechados de forma independente do conceito de conjunto aberto.

DEFINIÇÃO 8 (ADERÊNCIA – FECHO). Um ponto $a \in \mathbb{R}^n$ é dito *aderente* a um conjunto $X \subset \mathbb{R}^n$ se existe uma sequência de pontos de X que converge para a. O *fecho* de X, denotado por $\overline{X}$, é o conjunto formado por todos os pontos de $\mathbb{R}^n$ que são aderentes a X.

Sejam $X \subset \mathbb{R}^n$ e $a \in X$. Tomando-se a sequência constante $x_k = a$, vemos que $a \in \overline{X}$. Logo, $X \subset \overline{X}$ para todo $X \subset \mathbb{R}^n$. Além disso, segue-se diretamente da definição de fecho que

$$X \subset Y \quad \Rightarrow \quad \overline{X} \subset \overline{Y}.$$

EXEMPLO 18 (FECHO DE BOLAS E ESFERAS). O fecho de qualquer bola aberta de $\mathbb{R}^n$ é a bola fechada que tem mesmo centro e raio. Para verificarmos isto, consideremos uma bola aberta $B(a,r)$ de $\mathbb{R}^n$ e tomemos $x \in S[a,r]$. A sequência $x_k = (1 - 1/k)(x - a) + a$, $k \in \mathbb{N}$, claramente, converge para x. Além disso, para todo $k \in \mathbb{N}$, tem-se

$$\|x_k - a\| = \left(1 - \frac{1}{k}\right)\|x - a\| = \left(1 - \frac{1}{k}\right) r < r,$$

isto é, $x_k \in B(a,r)$. Logo, $x \in \overline{B(a,r)}$, donde $S[a,r] \subset \overline{B(a,r)}$. Uma vez que $B(a,r) \subset \overline{B(a,r)}$, tem-se, então,

$$B[a,r] = B(a,r) \cup S[a,r] \subset \overline{B(a,r)}.$$

Por outro lado, dado $y \in \mathbb{R}^n - B[a,r]$, é fácil ver que a bola com centro em y e raio $\epsilon = (\|y - a\| - r)/2$ é disjunta de $B(a,r)$. Sendo assim, nenhuma sequência de pontos de $B(a,r)$ pode convergir para y, donde se infere que $y \notin \overline{B(a,r)}$ e, portanto, que $\overline{B(a,r)} = B[a,r]$.

De forma análoga, verifica-se que toda bola fechada ou esfera coincide com seu fecho. Assim, dados $a \in \mathbb{R}^n$ e $r > 0$, valem as igualdades

- $\overline{B(a,r)} = \overline{B[a,r]} = B[a,r]$;
- $\overline{S(a,r)} = S(a,r)$.

Proposição 12 (propriedades do fecho). *Para todo conjunto $X \subset \mathbb{R}^n$, são verdadeiras as seguintes afirmações:*

i) $a \in \overline{X}$ *se, e somente se,* $B(a,r) \cap X \neq \emptyset \ \forall r > 0$;

ii) $\overline{X} = \mathbb{R}^n - \operatorname{int}(\mathbb{R}^n - X)$;

iii) $\overline{\overline{X}} = \overline{X}$.

Demonstração. (i) Dados $a \in \overline{X}$ e $r > 0$, tem-se que toda bola aberta $B(a,r)$ contém pelo menos um elemento de X, uma vez que existe uma sequência de elementos deste conjunto que converge para a. Reciprocamente, se $a \in \mathbb{R}^n$ é tal que toda bola aberta de $\mathbb{R}^n$ com centro em a intersecta $X \subset \mathbb{R}^n$, então, para cada $k \in \mathbb{N}$, existe $x_k \in X$, tal que $x_k \in B(a, 1/k)$. Logo, $\|x_k - a\| < 1/k \ \forall k \in \mathbb{N}$, donde $x_k \to a$ e, portanto, $a \in \overline{X}$.

(ii) Pelo item (i), temos que $a \in \overline{X}$ se, e somente se, para todo $r > 0$, $B(a,r)$ intersecta X. Esta última condição, entretanto, equivale a de que a não pertence ao interior de $\mathbb{R}^n - X$, isto é, a de que $a \in \mathbb{R}^n - \operatorname{int}(\mathbb{R}^n - X)$.

(iii) Dados $a \in \overline{\overline{X}}$ e $r > 0$, por (i), temos que existe $x \in B(a,r) \cap \overline{X}$. Como $B(a,r)$ é aberto, existe também $r_0 > 0$, tal que $B(x, r_0) \subset B(a,r)$. Porém, novamente por (i), $B(x,r_0) \cap X \neq \emptyset$, pois $x \in \overline{X}$. Logo, $B(a,r) \cap X \neq \emptyset$ e, portanto, $a \in \overline{X}$, isto é, $\overline{\overline{X}} = \overline{X}$. □

Exemplo 19 (fecho do complementar de bolas e esferas). Segue-se da Proposição 12-(ii) e das considerações dos exemplos 10 e 11 que:

- $\overline{\mathbb{R}^n - B[a,r]} = \mathbb{R}^n - \operatorname{int}(B[a,r]) = \mathbb{R}^n - B(a,r)$;
- $\overline{\mathbb{R}^n - B(a,r)} = \mathbb{R}^n - \operatorname{int}(B(a,r)) = \mathbb{R}^n - B(a,r)$;
- $\overline{\mathbb{R}^n - S[a,r]} = \mathbb{R}^n - \operatorname{int}(S[a,r]) = \mathbb{R}^n$.

Exemplo 20 (fecho do complementar de $\mathbb{Q}^n$ e de $(\mathbb{R} - \mathbb{Q})^n$). Infere-se dos resultados do Exemplo 13 e da Proposição 12-(ii) que:

- $\overline{\mathbb{R}^n - \mathbb{Q}^n} = \mathbb{R}^n - \operatorname{int}\mathbb{Q}^n = \mathbb{R}^n$;
- $\overline{\mathbb{R}^n - (\mathbb{R} - \mathbb{Q})^n} = \mathbb{R}^n - \operatorname{int}(\mathbb{R} - \mathbb{Q})^n = \mathbb{R}^n$.

Diz-se que um conjunto $X \subset \mathbb{R}^n$ é *denso* em $\mathbb{R}^n$ quando $\overline{X} = \mathbb{R}^n$.

Como se sabe, $\mathbb{Q}$ e $\mathbb{R} - \mathbb{Q}$ são, ambos, densos em $\mathbb{R}$. Deste fato e do resultado da Proposição 5 do Capítulo 1, conclui-se facilmente que

$$\overline{\mathbb{Q}^n} = \overline{(\mathbb{R} - \mathbb{Q})^n} = \mathbb{R}^n,$$

isto é, $\mathbb{Q}^n$ e $(\mathbb{R} - \mathbb{Q})^n$ são, ambos, densos em $\mathbb{R}^n$ (compare com o Exercício 16 do Capítulo 1). Além disso, pelo Exemplo 20, os complementares desses conjuntos em $\mathbb{R}^n$ são, igualmente, densos em $\mathbb{R}^n$.

Na proposição seguinte, obtemos a preterida caracterização dos conjuntos fechados de $\mathbb{R}^n$.

TEOREMA 6. *Um subconjunto de* $\mathbb{R}^n$ *é fechado se, e somente se, contém todos os seus pontos aderentes, isto é,*

$$F \subset \mathbb{R}^n \text{ é fechado} \quad \Leftrightarrow \quad \overline{F} = F.$$

DEMONSTRAÇÃO. Dado $F \subset \mathbb{R}^n$ façamos $A = \mathbb{R}^n - F$. Pela Proposição 12-(ii), tem-se $\overline{F} = \mathbb{R}^n - \operatorname{int} A$, donde se infere que $A = \operatorname{int} A$, isto é, A é aberto, se, e somente se, $F = \overline{F}$. Porém, por definição, a condição de A ser aberto equivale à de F ser fechado. □

Segue-se da Proposição 12-(iii) e do Teorema 6 que o fecho de todo subconjunto de $\mathbb{R}^n$ é fechado.

DEFINIÇÃO 9 (FRONTEIRA). A *fronteira* (ou *bordo*) de um conjunto $X \subset \mathbb{R}^n$ é o conjunto $\partial X = \overline{X} \cap \overline{\mathbb{R}^n - X}$.

É imediato que, para todo conjunto $X \subset \mathbb{R}^n$, $\partial X = \partial(\mathbb{R}^n - X)$ e que ∂X é fechado. Também, segue-se da Proposição 12-(i) que $a \in \mathbb{R}^n$ é um ponto da fronteira de X se, e somente se, toda bola aberta de $\mathbb{R}^n$ centrada em a contém pontos de X e de $\mathbb{R}^n - X$ (Fig. 5), isto é,

$$a \in \partial X \quad \Leftrightarrow \quad B(a, \epsilon) \cap X \neq \emptyset \quad \text{e} \quad B(a, \epsilon) \cap (\mathbb{R}^n - X) \neq \emptyset \ \ \forall \epsilon > 0.$$

Daí, conclui-se facilmente que $A \subset \mathbb{R}^n$ é aberto se, e somente se, $A \cap \partial A = \emptyset$ e também que, para todo $X \subset \mathbb{R}^n$, tem-se

$$\mathbb{R}^n = \operatorname{int} X \cup \operatorname{int}(\mathbb{R}^n - X) \cup \partial X \quad \text{(união disjunta)}.$$

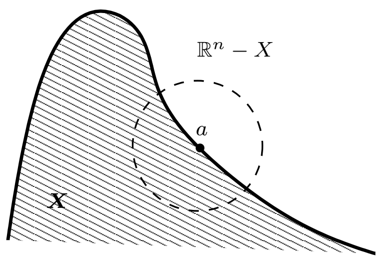

FIGURA 5

Considerando-se a densidade de $\mathbb{Q}^n$ e $(\mathbb{R} - \mathbb{Q})^n$ em $\mathbb{R}^n$, bem como a dos seus respectivos complementares em $\mathbb{R}^n$, obtém-se:

- $\partial \mathbb{Q}^n = \overline{\mathbb{Q}^n} \cap \overline{\mathbb{R}^n - \mathbb{Q}^n} = \mathbb{R}^n$;
- $\partial(\mathbb{R} - \mathbb{Q})^n = \overline{(\mathbb{R} - \mathbb{Q})^n} \cap \overline{\mathbb{R}^n - (\mathbb{R} - \mathbb{Q})^n} = \mathbb{R}^n$.

EXEMPLO 21 (FRONTEIRAS DE BOLAS E ESFERAS). Combinando-se os resultados dos exemplos 18 e 19 , tem-se, para quaisquer $a \in \mathbb{R}^n$ e $r > 0$, que:

- $\partial B(a, r) = \overline{B(a, r)} \cap \overline{\mathbb{R}^n - B(a, r)} = B[a, r] \cap (\mathbb{R}^n - B(a, r)) = S[a, r]$;
- $\partial B[a, r] = \overline{B[a, r]} \cap \overline{\mathbb{R}^n - B[a, r]} = B[a, r] \cap (\mathbb{R}^n - B(a, r)) = S[a, r]$;
- $\partial S[a, r] = \overline{S[a, r]} \cap \overline{\mathbb{R}^n - S[a, r]} = S[a, r] \cap \mathbb{R}^n = S[a, r]$.

Abordaremos agora um outro conceito relacionado com o de aderência, o de ponto de acumulação, que, dentre outras características, é essencial para que se introduza apropriadamente o conceito de limite de aplicações.

Definição 10 (ponto de acumulação). Dado $X \subset \mathbb{R}^n$, um ponto $a \in \mathbb{R}^n$ é dito um *ponto de acumulação* de X se $a \in \overline{X - \{a\}}$, isto é, se existe uma sequência (x_k) em $\mathbb{R}^n$, tal que:

- $x_k \in X - \{a\} \, \forall k \in \mathbb{N}$;
- $x_k \to a$.

Denota-se o conjunto dos pontos de acumulação de um conjunto X por X'. Note que um conjunto X não está necessariamente contido em X', que pode, inclusive, ser vazio. Este é o caso de qualquer subconjunto finito de $\mathbb{R}^n$, por exemplo. Observe também que $X' \subset \overline{X}$ para todo $X \subset \mathbb{R}^n$.

Mais geralmente, vale, para todo $X \subset \mathbb{R}^n$, a igualdade $\overline{X} = X \cup X'$, a qual é de verificação imediata.

É fácil ver também que, em $\mathbb{R}^n$, o conjunto dos pontos de acumulação de qualquer bola ou esfera coincide com o seu respectivo fecho. O mesmo vale para o complementar, em $\mathbb{R}^n$, de qualquer um destes conjuntos.

No Exercício 20 do Capítulo 1, verificamos que toda matriz A de $\mathrm{M}(n)$ é limite de uma sequência de matrizes invertíveis de $\mathrm{M}(n)$ e distintas de A, isto é, $\mathrm{I}(n)' = \mathrm{M}(n)$.

Em $\mathbb{R}$, consideram-se também as noções de ponto de acumulação à direita, bem como à esquerda, isto é, dado $a \in X \subset \mathbb{R}$, diz-se que a é *ponto de acumulação à direita* de X se existe uma sequência (x_k) em X, tal que $x_k > a$ e $x_k \to a$. De forma análoga, define-se *ponto de acumulação à esquerda*.

Quando $a \in X - X'$, dizemos que a é um *ponto isolado* de X. Se todos os pontos de X são isolados, dizemos que X é um conjunto *discreto*. Exemplos triviais de conjuntos discretos são $\mathbb{Z}^n$ e os subconjuntos finitos de $\mathbb{R}^n$.

Proposição 13. *Um ponto de $\mathbb{R}^n$ é ponto de acumulação de um conjunto $X \subset \mathbb{R}^n$ se, e somente se, toda vizinhança deste ponto contém uma infinidade de elementos de X.*

Demonstração. Sejam $X \subset \mathbb{R}^n$ e $a \in X'$. Então, existe uma sequência (x_k) em $X - \{a\}$, tal que $x_k \to a$. Dada, então, uma vizinhança V de a, a partir de um certo k_0, todos os termos de (x_k) estão em V. Além disso, o conjunto $\{x \in X;\, x = x_k \text{ para algum } k \geq k_0\}$ é infinito. Caso contrário, a partir de um certo índice $k_1 \geq k_0$, todos os x_k seriam iguais a a, o que constitui, claramente, uma contradição.

Reciprocamente, se toda vizinhança de a contém uma infinidade de pontos de X, temos que $a \in X'$, pois, assim sendo, para cada $k \in \mathbb{N}$, podemos escolher um $x_k \in X$ satisfazendo $x_k \in B(a, 1/k)$ e $x_k \neq a$, donde se terá $a \in \overline{X - \{a\}}$. □

Teorema 7 (Weierstrass). *Todo subconjunto infinito e limitado de $\mathbb{R}^n$ possui um ponto de acumulação.*

Demonstração. Seja $X \subset \mathbb{R}^n$ infinito e limitado. Desta forma, existe uma sequência (x_k) em X, tal que $x_k \neq x_l \; \forall k, l \in \mathbb{N}$ (caso contrário X seria finito). Além disso, como X é limitado, a sequência (x_k) é limitada. Segue-se do Teorema de Bolzano-Weierstrass que (x_k) possui uma subsequência $(x_k)_{k\in\mathbb{N}_0}$ que converge para um ponto $a \in \mathbb{R}^n$. Note que a pode ser igual a, no máximo, um termo da subsequência $(x_k)_{k\in\mathbb{N}_0}$. Neste caso, nós podemos simplesmente exclui-lo e concluir que $a \in X'$. $\square$

3. Topologia Relativa

A estrutura vetorial do espaço $\mathbb{R}^n$, como se sabe, é herdada por apenas alguns de seus subconjuntos, ditos subespaços vetoriais. Veremos agora que, nesse sentido, a estrutura topológica de $\mathbb{R}^n$ é menos rígida, isto é, ela permite que *qualquer* um de seus subconjuntos seja um *subespaço topológico.* Este fato é relevante por, pelo menos, dois motivos. Primeiro, o importante conceito de conexidade de conjuntos, que veremos na Seção 5, pressupõe que estes tenham uma estrutura de espaço topológico. Segundo, até o ponto em que estejam envolvidas apenas questões topológicas, podemos considerar aplicações definidas, e tomando valores, em quaisquer subconjuntos de $\mathbb{R}^n$. Isto será especialmente útil no estudo das aplicações contínuas.

Definição 11 (aberto relativo – vizinhança relativa). Sejam X um subconjunto de $\mathbb{R}^n$ e $A \subset X$. Diz-se que A é *aberto em* X, ou *aberto relativamente a* X, quando existe um aberto U de $\mathbb{R}^n$, tal que $A = U \cap X$. Um conjunto $V \subset X$ é dito uma *vizinhança* de $a \in V$ *em* X se existe $A \subset V$, tal que $a \in A$ e A é aberto em X (Fig. 6).

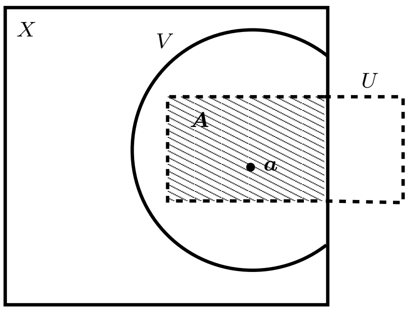

Figura 6

Verifica-se facilmente que a família dos abertos relativos de um subconjunto X de $\mathbb{R}^n$ define neste uma topologia, dita *relativa* à de $\mathbb{R}^n$ (vide Exercício 13). Grosso modo, podemos dizer que a topologia relativa de $X \subset \mathbb{R}^n$ é a restrição da topologia de $\mathbb{R}^n$ ao conjunto X.

É imediato que um subconjunto de $\mathbb{R}^n$ é aberto se, e somente se, é aberto relativamente a $\mathbb{R}^n$. Também, se A é um aberto de $\mathbb{R}^n$ e $A \subset X$, então $A = A\cap X$, donde A é aberto em X. Mais geralmente, tem-se

$$A, X \subset Y \subset \mathbb{R}^n \quad \text{e} \quad A \text{ aberto em } Y \quad \Rightarrow \quad A \cap X \text{ aberto em } X.$$

Com efeito, sendo A aberto em Y, existe $U \subset \mathbb{R}^n$, aberto, tal que $A = U \cap Y$. Logo, $A \cap X = (U \cap Y) \cap X = U \cap (Y \cap X) = U \cap X$, donde $A \cap X$ é aberto em X.

Exemplo 22. Sejam $X = [0, 2)$ e $A = [0, 1)$. Uma vez que $\operatorname{int} A = (0, 1) \neq A$, tem-se que A não é um aberto de $\mathbb{R}$. No entanto, tomando-se o aberto $U = (-1, 1)$ de $\mathbb{R}$, tem-se $A = U \cap X$, isto é, A é aberto em X.

Exemplo 23. Os conjuntos

$$A_1 = \{x \in S^n \,;\, \langle x, e_{n+1}\rangle > 0\} \quad \text{e} \quad A_2 = \{x \in S^n \,;\, \langle x, e_{n+1}\rangle < 0\}$$

cumprem as igualdades $A_1 = U_1 \cap S^2$ e $A_2 = U_2 \cap S^2$, em que U_1 e U_2 denotam, respectivamente, os semi-espaços abertos $\{x \in \mathbb{R}^{n+1} \ \langle x, e_{n+1}\rangle > 0\}$ e $\{x \in \mathbb{R}^{n+1} \,;\, \langle x, e_{n+1}\rangle < 0\}$. Logo, A_1 e A_2 são, ambos, abertos em S^2. Note que A_1 e A_2, como subconjuntos de $\mathbb{R}^3$, têm interior vazio. Em particular, nenhum deles é um aberto de $\mathbb{R}^3$.

Exemplo 24. Considere o plano $X = \{(x_1, x_2, x_3) \in \mathbb{R}^3; x_3 = 0\} \subset \mathbb{R}^3$ e $A = \{(x_1, x_2, 0) \in X; x_1^2 + x_2^2 < 1\} \subset X$. Temos que $A = B(0, 1) \cap X \subset \mathbb{R}^3$. Logo, A é aberto em X. Como no exemplo anterior, o interior de A é vazio e, portanto, A não é aberto em $\mathbb{R}^3$ (Fig. 7).

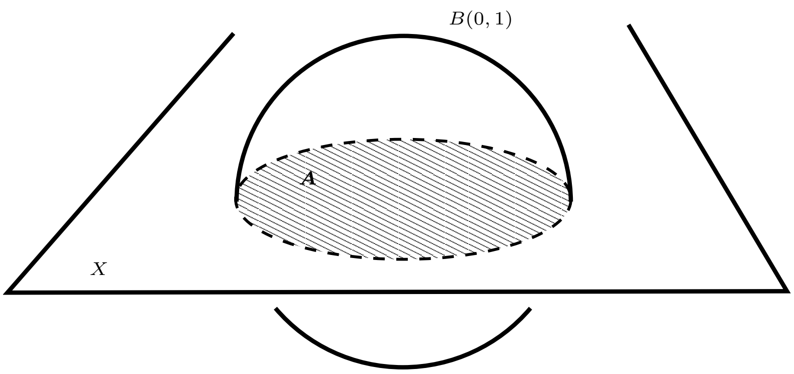

Figura 7

A proposição seguinte caracteriza os abertos relativos por meio de bolas.

Proposição 14. *Um subconjunto A de $X \subset \mathbb{R}^n$ é aberto em X se, e somente se, para todo $a \in A$, existe $r > 0$, tal que*

$$B(a, r) \cap X \subset A.$$

Demonstração. Suponhamos que A seja aberto em X. Então, existe um aberto U de $\mathbb{R}^n$, tal que $A = U \cap X$. Agora, dado $a \in A$, tem-se $a \in U$. Assim, existe $r > 0$, tal que $B(a, r) \subset U$, pois U é aberto. Logo,

$$B(a, r) \cap X \subset U \cap X = A.$$

Reciprocamente, suponhamos que, para cada $a \in A$, exista $r_a > 0$, tal que $B(a, r_a) \cap X \subset A$. Uma vez que a união arbitrária de abertos é aberta, temos que $U = \bigcup_{a \in A} B(a, r_a)$ é um aberto de $\mathbb{R}^n$. Além disso,

$$U \cap X = \left(\bigcup_{a \in A} B(a, r_a)\right) \cap X = \bigcup_{a \in A} (B(a, r_a) \cap X) = A,$$

donde se infere que A é aberto em X. $\square$

Corolário 1. *Seja* $X \subset \mathbb{R}^n$. *Então,* $V \subset X$ *é uma vizinhança de* $a \in V$ *em* X *se, e somente se, existe* $r > 0$, *tal que* $B(a,r) \cap X \subset V$.

Introduziremos agora os conjuntos fechados de subconjuntos de $\mathbb{R}^n$ com respeito à topologia relativa e, em seguida, veremos como caracterizá-los através de seus respectivos fechos.

Definição 12 (fechado relativo). Um conjunto $F \subset X \subset \mathbb{R}^n$ é dito *fechado em* X, ou *fechado relativamente a* X, se seu complementar em X é aberto relativamente a X.

Segue-se das considerações do Exemplo 23 que os hemisférios

$$F_1 = \{x \in S^2; \langle x, e_3 \rangle \geq 0\} \quad \text{e} \quad F_2 = \{x \in S^2; \langle x, e_3 \rangle \leq 0\}$$

são fechados em S^2.

Proposição 15. *Um subconjunto* F *de* $X \subset \mathbb{R}^n$ *é fechado em* X *se, e somente se, contém todos os seus pontos aderentes que estão em* X, *isto é,*

$$F \text{ é fechado em } X \;\Leftrightarrow\; F = \overline{F} \cap X.$$

Demonstração. Suponhamos que F seja fechado em X. Então, $X - F$ é aberto em X. Logo, existe $U \subset \mathbb{R}^n$, aberto, tal que $X - F = U \cap X$. Daí, temos que $F \subset \mathbb{R}^n - U$. Como $\mathbb{R}^n - U$ é fechado em $\mathbb{R}^n$, temos $\overline{F} \subset \overline{\mathbb{R}^n - U} = \mathbb{R}^n - U$, donde $\overline{F} \cap X \subset (\mathbb{R}^n - U) \cap X = X - U = F$. Além disso, as inclusões $F \subset \overline{F}$ e $F \subset X$ nos dão $F \subset \overline{F} \cap X$. Logo, $F = \overline{F} \cap X$.

Suponhamos agora que $F = \overline{F} \cap X$ se cumpra. Neste caso, $X - F = X - \overline{F} = (\mathbb{R}^n - \overline{F}) \cap X$, donde $X - F$ é aberto em X, pois $\overline{F}$ é fechado e, consequentemente, $\mathbb{R}^n - \overline{F}$ é aberto. Logo, F é fechado em X. □

Dados $F \subset X \subset \mathbb{R}^n$, segue-se da Proposição 15 que se F é fechado, então F é fechado em X, pois, nestas condições, tem-se $F = F \cap X = \overline{F} \cap X$.

Exemplo 25. O intervalo $(1,2]$ é fechado em $(1,3]$, pois

$$\overline{(1,2]} \cap (1,3] = [1,2] \cap (1,3] = (1,2].$$

Note que $(1,2]$ não é fechado em $\mathbb{R}$.

4. Compacidade

Introduziremos agora o importante conceito topológico de compacidade de conjuntos. Veremos, então, que um conjunto compacto de $\mathbb{R}^n$ pode ser caracterizado de diversas maneiras, e constataremos, nos capítulos posteriores, que as aplicações (contínuas ou diferenciáveis) cujos domínios são conjuntos compactos de $\mathbb{R}^n$ têm propriedades especiais.

Dado um conjunto $X \subset \mathbb{R}^n$, uma família $\mathscr{A} = \{A_\lambda\}_{\lambda \in \Lambda}$ de subconjuntos de $\mathbb{R}^n$ é dita uma *cobertura* de X se $X \subset \bigcup A_\lambda$. Uma *subcobertura* de $\mathscr{A}$ é uma subfamília $\{A_\lambda\}_{\lambda \in \Lambda_0}$, $\Lambda_0 \subset \Lambda$, tal que $X \subset \bigcup_{\lambda \in \Lambda_0} A_\lambda$.

Uma cobertura $\mathscr{A} = \{A_\lambda\}_{\lambda \in \Lambda}$ é dita *aberta* se cada A_λ é aberto e *finita* se Λ é um conjunto finito.

Definição 13 (conjunto compacto). Diz-se que um conjunto $X \subset \mathbb{R}^n$ é *compacto* quando toda cobertura aberta $\mathscr{A} = \{A_\lambda\}_{\lambda\in\Lambda}$ de X admite uma subcobertura finita, isto é, quando existem $\lambda_1, \dots, \lambda_i \in \Lambda$, tais que

$$X \subset A_{\lambda_1} \cup \cdots \cup A_{\lambda_i}.$$

É imediato que qualquer subconjunto finito de $\mathbb{R}^n$ é compacto.

Dados $a \in \mathbb{R}^n$ e $r > 0$, façamos, para cada $k \in \mathbb{N}$, $A_k = B(a, r - \frac{r}{k+1})$. Claramente, $\mathscr{A} = \{A_k\}_{k\in\mathbb{N}}$ é uma cobertura aberta de $B(a,r)$, tal que, para cada $k \in \mathbb{N}$, $A_k \subset A_{k+1} \subset B(a,r)$. Segue-se daí que esta cobertura não admite subcobertura finita. Com efeito, dada qualquer quantidade finita de abertos de $\mathscr{A}$, dentre eles, existirá um A_k que conterá todos os outros. Porém, para nenhum $k \in \mathbb{N}$, $B(a,r)$ está contida em A_k. Logo, nenhuma bola aberta de $\mathbb{R}^n$ é um conjunto compacto.

Seja $X \subset \mathbb{R}^n$ um subconjunto ilimitado de $\mathbb{R}^n$. A família $\mathscr{B} = \{B(0,k)\}_{k\in\mathbb{N}}$ constitui uma cobertura aberta de $\mathbb{R}^n$, em particular, de X. Uma vez que, para todo $k \in \mathbb{N}$, $B(0,k) \subset B(0,k+1)$, conclui-se daí, como no exemplo do parágrafo anterior, que $\mathscr{B}$ não admite subcobertura finita, pois nenhuma bola de $\mathbb{R}^n$ contém X. Deste argumento, segue-se o resultado seguinte.

Proposição 16. *Todo conjunto compacto de $\mathbb{R}^n$ é, necessariamente, limitado.*

Exemplo 26. Seja (x_k) uma sequência convergente em $\mathbb{R}^n$ cujo limite é $a \in \mathbb{R}^n$. Então, $K = \{x_k\,;\, k \in \mathbb{N}\} \cup \{a\} \subset \mathbb{R}^n$ é um subconjunto compacto de $\mathbb{R}^n$. De fato, dada uma cobertura aberta de K, $\mathscr{A} = \{A_\lambda\}_{\lambda\in\Lambda}$, existe $\lambda_0 \in \Lambda$, tal que $a \in A_{\lambda_0}$. Então, sendo a o limite de (x_k), existe $k_0 \in \mathbb{N}$, tal que $x_k \in A_{\lambda_0}$ para todo $k \geq k_0$. Além disso, para cada $i = 1, \dots, k_0 - 1$, existe um aberto A_{λ_i} de $\mathscr{A}$ que contém x_i. Desta forma, $K \subset A_{\lambda_0} \cup A_{\lambda_1} \cup \cdots \cup A_{\lambda_{k_0-1}}$.

Exemplo 27. *Todo isomorfismo linear $T \in \mathrm{L}(\mathbb{R}^n)$ leva compactos de $\mathbb{R}^n$ em compactos de $\mathbb{R}^n$.* Com efeito, dados um compacto $K \subset \mathbb{R}^n$ e uma cobertura aberta $\mathscr{O} = \{O_\lambda\}_{\lambda\in\Lambda}$ de $T(K) \subset \mathbb{R}^n$, temos, para cada $\lambda \in \Lambda$, que $A_\lambda = T^{-1}(O_\lambda)$ é um aberto de $\mathbb{R}^n$, pois T^{-1} é uma aplicação aberta (vide Exemplo 14). Uma vez que $T(K) \subset \bigcup O_\lambda$, tem-se, então,

$$K = T^{-1}(T(K)) \subset T^{-1}\left(\bigcup O_\lambda\right) = \bigcup T^{-1}(O_\lambda) = \bigcup A_\lambda,$$

isto é, $\mathscr{A} = \{A_\lambda\}_{\lambda\in\Lambda}$ é uma cobertura aberta de K. Sendo este compacto, tem-se que $\mathscr{A}$ admite uma subcobertura finita, donde existem $\lambda_1, \dots, \lambda_i \in \Lambda$, tais que $K \subset A_{\lambda_1} \cup \cdots \cup A_{\lambda_i}$. Desta forma,

$$T(K) \subset T(A_{\lambda_1} \cup \cdots \cup A_{\lambda_i}) = T(A_{\lambda_1}) \cup \cdots \cup T(A_{\lambda_i}) = O_{\lambda_1} \cup \cdots \cup O_{\lambda_i}.$$

Segue-se que $\mathscr{O}$ admite uma subcobertura finita e, portanto, que $T(K)$ é compacto.

Um subconjunto de $\mathbb{R}^n$ que é dado pelo produto cartesiano de n intervalos fechados de $\mathbb{R}$ é chamado de *paralelepípedo*.

Dado um paralelepípedo $K = [a_1, b_1] \times \cdots \times [a_n, b_n] \subset \mathbb{R}^n$, temos que

$$K = \{(x_1, \dots, x_n) \in \mathbb{R}^n;\, a_i \leq x_i \leq b_i \ \forall i = 1, \dots, n\}.$$

Logo, fazendo-se

$$\rho = \sqrt{\sum_{i=1}^{n} (a_i - b_i)^2},$$

tem-se, para quaisquer $x = (x_1, \ldots, x_n)$, $y = (y_1, \ldots, y_n) \in K$, que

$$\|x - y\| = \sqrt{\sum_{i=1}^{n} (x_i - y_i)^2} \leq \sqrt{\sum_{i=1}^{n} (a_i - b_i)^2} = \rho.$$

Chamaremos ρ a *diagonal* de K.

PROPOSIÇÃO 17. *Todo paralelepípedo de $\mathbb{R}^n$ é compacto.*

DEMONSTRAÇÃO. Sejam $K = [a_1, b_1] \times \cdots \times [a_n, b_n]$ um paralelepípedo em $\mathbb{R}^n$ e ρ a sua diagonal. Suponhamos, por absurdo, que K não seja compacto. Então, existe uma cobertura aberta de K, $\mathscr{A} = \{A_\lambda\}_{\lambda \in \Lambda}$, que não admite subcobertura finita.

Consideremos, para cada $i = 1, \ldots, n$, o ponto $c_i = (a_i + b_i)/2$ e observemos que este decompõe o intervalo $[a_i, b_i]$ em dois intervalos fechados de igual comprimento. Desta forma, os 2^n paralelepípedos

$$[\alpha_1, \beta_1] \times \cdots \times [\alpha_n, \beta_n], \quad [\alpha_i, \beta_i] = [a_i, c_i] \text{ ou } [\alpha_i, \beta_i] = [c_i, b_i], \quad i = 1, \ldots n,$$

decompõem K. Note que cada um deles tem diagonal igual a $\rho/2$, pois

$$\sqrt{\sum_{i=1}^{n} (\alpha_i - \beta_i)^2} = \sqrt{\sum_{i=1}^{n} (a_i - b_i)^2/4} = \frac{\rho}{2} \cdot$$

Além disso, pelo menos um desses paralelepípedos, que denotaremos por K_1, não pode ser coberto por um número finito de abertos de $\mathscr{A}$, caso contrário, poderíamos extrair de $\mathscr{A}$ uma subcobertura finita de K.

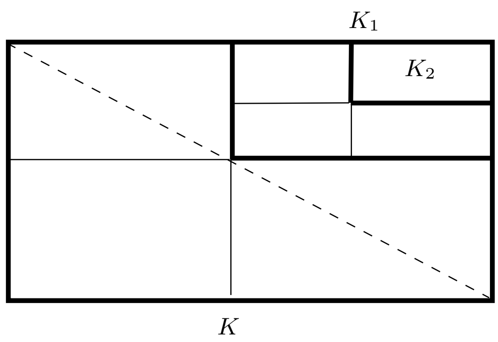

FIGURA 8

Repetindo-se este processo de decomposição indefinidamente, obtém-se uma sequência decrescente de paralelepípedos (Fig. 8),

$$K = K_0 \supset K_1 \supset K_2 \supset \cdots \supset K_k \supset \cdots,$$

tais que, para todo $k \in \mathbb{N}$,

i) o paralelepípedo K_k não pode ser coberto por um número finito de abertos de $\mathscr{A}$;

ii) a diagonal de K_k é igual a $\rho/2^k$ ($\Rightarrow \|x - y\| \leq \rho/2^k \ \forall x, y \in K_k$).

Tomemos agora uma sequência (x_k) em K, tal que $x_k \in K_k\, \forall k \in \mathbb{N}$. Evidentemente, K é limitado. Logo, (x_k) é limitada. Então, aplicando-se o Teorema de Bolzano-Weierstrass e passando-se a uma subsequência, se necessário, podemos supor que $x_k \to a \in \mathbb{R}^n$. Porém, K é dado por um produto cartesiano de conjuntos fechados, o que implica que K é fechado (vide Exercício 8) e, consequentemente, que $a \in K$.

Observemos que, para todo $k_0 \in \mathbb{N}$, a subsequência $(x_k)_{k \geq k_0}$ é uma sequência em K_{k_0} que converge para a. Assim, $a \in \overline{K_{k_0}} = K_{k_0}\ \forall k_0 \in \mathbb{N}$, isto é, $a \in \bigcap K_k$.

Agora, basta vermos que $a \in A_\lambda$ para algum $\lambda \in \Lambda$. Logo, existe $r > 0$, tal que $B(a,r) \subset A_\lambda$. Tomando-se, então, $k \in \mathbb{N}$ satisfazendo $\rho/2^k < r$, temos, por (ii), que $\|x - a\| < r\ \forall x \in K_k$ (pois $a \in K_k\ \forall k \in \mathbb{N}$). Segue-se que $K_k \subset B(a,r) \subset A_\lambda$, o que contradiz (i) e prova que K é compacto. □

Proposição 18. *Todo subconjunto fechado de um conjunto compacto é compacto.*

Demonstração. Sejam $K \subset \mathbb{R}^n$ um conjunto compacto e $F \subset K$ fechado em $\mathbb{R}^n$. Então, $\mathbb{R}^n - F$ é aberto. Assim, dada uma cobertura aberta de F, $\mathscr{A} = \{A_\lambda\}_{\lambda \in \Lambda}$, temos que os abertos de $\mathscr{A}$, juntamente com $\mathbb{R}^n - F$, constituem uma cobertura aberta de K. Como K é compacto, desta cobertura podemos extrair uma subcobertura finita, $A_{\lambda_1}, \dots, A_{\lambda_i}, \mathbb{R}^n - F$, que, consequentemente, cobrirá F. Porém, $\mathbb{R}^n - F$ é disjunto de F. Logo, $\{A_{\lambda_1}, \dots, A_{\lambda_i}\}$ é uma subcobertura finita de $\mathscr{A}$, donde se infere que F é compacto. □

Decorre do fato de $\mathbb{R}^n$ ser um espaço vetorial de dimensão finita e de estarmos considerando no mesmo uma topologia proveniente de uma métrica, que os conjuntos compactos deste espaço admitem uma simples caracterização. Mais precisamente, provaremos que a condição de ser compacto é equivalente à de ser fechado e limitado. Note que "ser limitado" é uma condição métrica.

Os exemplos dados acima nos sugerem, de certa maneira, essa caracterização dos compactos. As bolas abertas, que não são compactas, são limitadas, porém não são fechadas. O espaço $\mathbb{R}^n$, que tampouco é compacto, é fechado mas não é limitado. Por outro lado, os paralelepípedos de $\mathbb{R}^n$, que provamos serem compactos, são todos fechados e limitados (note que este fato foi fundamental na prova da compacidade desses conjuntos).

Teorema de Heine-Borel. *Um subconjunto de $\mathbb{R}^n$ é compacto se, e somente se, é fechado e limitado.*

Demonstração. Seja $K \subset \mathbb{R}^n$ compacto. Pela Proposição 16, K é limitado. Provemos, então, que $\mathbb{R}^n - K$ é aberto e, consequentemente, que K é fechado. Para isto, tomemos $a \in \mathbb{R}^n - K$ e, para cada $x \in K$, consideremos a bola aberta $B(x, r_x)$, com centro em x e raio $r_x = \|x - a\|/2$. A família $\mathscr{B} = \{B(x, r_x)\,;\, x \in K\}$, claramente, constitui uma cobertura aberta de K, donde existem $B_{x_1}, \dots, B_{x_i} \in \mathscr{B}$, tais que $K \subset B_{x_1} \cup \dots \cup B_{x_i}$. Observando-se que, para todo $x \in K$, $B(x, r_x) \cap B(a, r_x) = \emptyset$, e tomando-se $\epsilon > 0$, tal que $\epsilon < \min\{r_{x_1}, \dots, r_{x_i}\}$, conclui-se facilmente que $B(a, \epsilon)$ é disjunta de K, isto é, $B(a, \epsilon) \subset \mathbb{R}^n - K$. Logo, $\mathbb{R}^n - K$ é aberto.

Reciprocamente, suponhamos que $K \subset \mathbb{R}^n$ seja fechado e limitado. Desta última propriedade, segue-se que existe um paralelepípedo que contém K. Como K é fechado, pelas proposições 17 e 18, K é compacto. □

Segue-se do Teorema de Heine-Borel que todas as bolas fechadas, anéis fechados e esferas de $\mathbb{R}^n$ são conjuntos compactos. Já a interseção de qualquer bola de $\mathbb{R}^n$ com $\mathbb{Q}^n$, embora seja um conjunto limitado, não é compacto por não ser fechado. Pelo mesmo motivo, uma bola fechada menos um número finito de pontos não é compacta.

A demonstração da Proposição 17 nos sugere o teorema abaixo, que generaliza aquele da reta real conhecido como o Teorema dos Intervalos Encaixados.

Teorema dos Compactos Encaixados (Cantor). *Seja*

$$K_1 \supset K_2 \supset \cdots \supset K_k \supset K_{k+1} \supset \cdots$$

uma família enumerável e decrescente de compactos não-vazios de $\mathbb{R}^n$. *Então, existe* $a \in \mathbb{R}^n$, *tal que* $a \in K_k$ *para todo* $k \in \mathbb{N}$, *isto é, a interseção* $\bigcap K_k$ *é não-vazia.*

Demonstração. Seja (x_k) uma sequência em $\mathbb{R}^n$, tal que $x_k \in K_k$ para todo $k \in \mathbb{N}$. Em particular, (x_k) é uma sequência no compacto K_1. Pelo Teorema de Heine-Borel, K_1 é fechado e limitado. Usando-se, então, o Teorema de Bolzano-Weierstrass e passando-se a uma subsequência, se necessário, podemos supor que $x_k \to a \in K_1$. Afirmamos que $a \in K_k$ para todo $k \in \mathbb{N}$. Com efeito, dado $k_0 \in \mathbb{N}$, a subsequência $(x_k)_{k \geq k_0}$ é uma sequência em K_{k_0} que converge para a. Logo, $a \in \overline{K_{k_0}} = K_{k_0}$. Como k_0 é arbitrário, tem-se $a \in \bigcap K_k$. □

No Teorema do Compactos Encaixados, a hipótese de compacidade da família $(K_k)_{k\in\mathbb{N}}$ não pode ser substituída pela de ser limitada ou fechada. Tome-se, por exemplo, a família $\{B_k\}_{k\in\mathbb{N}}$, em que $B_k = B[0, 1/k] - \{0\}$. Vê-se facilmente que $B_1 \supset B_2 \supset \cdots \supset B_k \supset B_{k+1} \supset \cdots$ e que cada B_k é limitado (e não-compacto). No entanto, $\bigcap B_k = \emptyset$. Igualmente, a família de fechados e ilimitados $\{F_k\}_{k\in\mathbb{N}}$, em que $F_k = [k, +\infty) \subset \mathbb{R}$, satisfaz $F_1 \supset F_2 \supset \cdots \supset F_k \supset \cdots$ e $\bigcap F_k = \emptyset$.

Vejamos agora que, como consequência do Teorema de Bolzano-Weierstrass e do Teorema de Heine-Borel, os compactos de $\mathbb{R}^n$ podem ser caracterizados por meio de sequências, conforme a definição e o teorema que se seguem.

Definição 14 (compacidade sequencial). Diz-se que um conjunto $K \subset \mathbb{R}^n$ é *sequencialmente compacto* quando toda sequência em K possui uma subsequência que converge para um elemento de K.

Teorema 8. *Um subconjunto de* $\mathbb{R}^n$ *é compacto se, e somente se, é sequencialmente compacto.*

Demonstração. Suponha que $K \subset \mathbb{R}^n$ seja compacto. Então, K é limitado e, portanto, toda sequência de pontos de K é limitada. Logo, pelo Teorema de Bolzano-Weierstrass, toda sequência em K possui uma subsequência convergente. Além disso, o limite de uma tal subsequência tem que pertencer a K, uma vez que, pelo Teorema de Heine-Borel, este é fechado. Logo, K é sequencialmente compacto.

Reciprocamente, se $K \subset \mathbb{R}^n$ é sequencialmente compacto, então K é limitado. Caso contrário, para cada $k \in \mathbb{N}$, poderíamos escolher $x_k \in K$, tal que $\|x_k\| > k$. A sequência (x_k) não teria, desta forma, subsequências limitadas e, portanto, nenhuma delas seria convergente, o que vai de encontro à hipótese. Além disso, $\overline{K} = K$, pois se existisse $a \in \overline{K} - K$, existiria uma sequência de elementos de

K convergindo para um elemento não pertencente à K, o que também contraria a hipótese. Desta forma, K é fechado e limitado. Segue-se, então, do Teorema de Heine-Borel, que K é compacto. □

Uma vez que o conceito de compacidade em $\mathbb{R}^n$ envolve apenas as noções de cobertura e conjunto aberto, ele pode ser introduzido, de forma análoga, em qualquer espaço topológico, em particular, em espaços vetoriais normados munidos das topologias determinadas por suas respectivas normas. No caso de estes terem dimensão finita, generalizam-se, neste contexto, o Teorema de Heine-Borel e o Teorema 8. Vejamos, sucintamente, como isto se dá.

Sejam $(\mathbb{V}, \| \ \|)$ um espaço vetorial normado de dimensão n e $T : V \to \mathbb{R}^n$ um isomorfismo linear. Aplicando-se a T os argumentos dos exemplos 14, 17 e 27, conclui-se facilmente que as aplicações T e T^{-1} são, ambas, abertas e fechadas, bem como levam compactos em compactos. Além disso, segue-se diretamente das propriedades da norma espectral que estas aplicações levam conjuntos limitados em conjuntos limitados. Assim, dado $K \subset \mathbb{V}$, tem-se que K é compacto $\Leftrightarrow$ $T(K)$ é compacto $\Leftrightarrow$ $T(K)$ é fechado e limitado $\Leftrightarrow$ $K = T^{-1}(T(K))$ é fechado e limitado, donde vale, em $\mathbb{V}$, o Teorema de Heine-Borel. Como consequência, podemos substituir, no enunciado e na demonstração do Teorema 8, o espaço $\mathbb{R}^n$ por $\mathbb{V}$. Vale, portanto, o resultado seguinte.

Teorema 9. *Seja $\mathbb{V}$ um espaço vetorial normado de dimensão finita munido da topologia proveniente de sua norma. Então, as seguintes afirmações a respeito de um subconjunto K de $\mathbb{V}$ são equivalentes:*

- *K é compacto;*
- *K é fechado e limitado;*
- *K é sequencialmente compacto.*

Pelo Exercício 18 do Capítulo 1, o conjunto $\mathrm{O}(\mathbb{R}^n)$, formado pelos operadores ortogonais de $\mathrm{L}(\mathbb{R}^n)$, é sequencialmente compacto. Logo, $\mathrm{O}(\mathbb{R}^n)$ *é compacto.* Por outro lado, o conjunto dos operadores auto-adjuntos de $\mathrm{L}(\mathbb{R}^n)$, por ser um espaço vetorial, é fechado (vide Exercício 5), porém, é ilimitado e, portanto, não é compacto.

Convém observar que se $\mathbb{V}$ é um espaço vetorial normado arbitrário, não necessariamente de dimensão finita, tomando-se neste a topologia proveniente da norma, valem as implicações (com demonstrações análogas àquelas em que $\mathbb{V} = \mathbb{R}^n$):

i) $K \subset \mathbb{V}$ compacto $\Rightarrow$ K fechado e limitado.

ii) $K \subset \mathbb{V}$ compacto $\Rightarrow$ K sequencialmente compacto.

Vale, em qualquer espaço métrico, em particular em $\mathbb{V}$, a recíproca de (ii). No entanto, em geral, a recíproca de (i) não se cumpre. Há, por exemplo, espaços vetoriais normados (de dimensão infinita) em que as esferas, que são conjuntos fechados e limitados, não são compactas, conforme o exemplo que se segue.

Exemplo 28 (esferas não-compactas). Considere o espaço vetorial $\mathbb{V}$ formado pelas sequências de números reais em que todos os termos, exceto um número finito deles, são iguais a zero. Dado um elemento $v = (x_i)_{i \in \mathbb{N}}$ de $\mathbb{V}$, a função $\|v\| = \max\{|x_i|; i \in \mathbb{N}\}$ define uma norma neste espaço (verifique!). Seja S a esfera de $\mathbb{V}$ com centro na origem e raio 1. Para cada $k \in \mathbb{N}$, denote por v_k a sequência de números reais em que o k-ésimo termo é igual a 1 e todos os demais são iguais a zero. Então, (v_k) é uma sequência em S que satisfaz $\|v_k - v_l\| = 1\ \forall k \neq l \in \mathbb{N}$. Logo, subsequência alguma de (v_k) é de Cauchy, o que implica que (v_k) não tem subsequências convergentes. Desta forma, S não é compacto, pois não é sequencialmente compacto.

Dado um conjunto $X \subset \mathbb{R}^n$, consideremo-lo como um espaço topológico munido da topologia relativa à de $\mathbb{R}^n$. Conforme discutimos na Seção 3, os abertos e fechados de X são relativos, no sentido de que um subconjunto de X pode ser aberto (ou fechado) relativamente a X, mesmo que não seja um aberto (ou fechado) de $\mathbb{R}^n$. Vejamos agora que o mesmo não ocorre com subconjuntos compactos de X — os quais, por definição, são aqueles que admitem uma subcobertura finita de toda cobertura dos mesmos por abertos relativos de X —, isto é,

$$K \subset X \text{ é compacto em } X \quad \Leftrightarrow \quad K \text{ é compacto em } \mathbb{R}^n.$$

Com efeito, seja $\mathscr{A} = \{A_\lambda\}_{\lambda \in \Lambda}$ uma cobertura aberta de K. Uma vez que, para cada $\lambda \in \Lambda$, $A_\lambda \cap X$ é aberto em X e $K \subset X$, temos que a família $\{A_\lambda \cap X\}_{\lambda \in \Lambda}$ constitui uma cobertura aberta de K com respeito à topologia relativa de X. Assim, se K for compacto em X, teremos $\lambda_1, \dots, \lambda_i \in \Lambda$, tais que

$$K \subset (A_{\lambda_1} \cap X) \cup \cdots \cup (A_{\lambda_i} \cap X) \subset A_{\lambda_1} \cup \cdots \cup A_{\lambda_i}\,,$$

donde K será compacto em $\mathbb{R}^n$.

Reciprocamente, suponhamos que K seja compacto em $\mathbb{R}^n$ e tomemos uma cobertura aberta de K em X, $\mathscr{O} = \{O_\lambda\}_{\lambda \in \Lambda}$. Neste caso, para cada $\lambda \in \Lambda$, tem-se $O_\lambda = A_\lambda \cap X$, em que A_λ é um conjunto aberto de $\mathbb{R}^n$. Desta forma,

$$K \subset \bigcup O_\lambda = \bigcup (A_\lambda \cap X) \subset \bigcup A_\lambda\,,$$

isto é, a família $\{A_\lambda\}_{\lambda \in \Lambda}$ é uma cobertura aberta de K. Sendo este compacto, existem $\lambda_1, \dots, \lambda_i \in \Lambda$, tais que $K \subset A_{\lambda_1} \cup \cdots \cup A_{\lambda_i}$. Logo,

$$K \subset (A_{\lambda_1} \cup \cdots \cup A_{\lambda_i}) \cap X = (A_{\lambda_1} \cap X) \cup \cdots \cup (A_{\lambda_i} \cap X) = O_{\lambda_1} \cup \cdots \cup O_{\lambda_i}\,,$$

donde se infere que K é compacto em X.

Segue-se destas considerações que, em $\mathbb{R}^n$, a noção de compacidade é absoluta, o que nos permite chamar seus subconjuntos compactos de *espaços (topológicos) compactos.*

5. Conexidade

No estudo da diferenciabilidade das funções reais de uma variável, verifica-se que toda função diferenciável definida num intervalo aberto cuja derivada se anula em todos os pontos é constante. Esta afirmação não permanece válida se dela suprimirmos a palavra "intervalo". Se considerarmos, por exemplo, a função $f : \mathbb{R} - \{0\} \to \mathbb{R}$, tal que $f(x) = 1$ para todo $x > 0$ e $f(x) = -1$ para todo $x < 0$, vemos que f tem, certamente, derivada nula em todos os pontos do seu domínio, que é um subconjunto aberto de $\mathbb{R}$. No entanto, f não é constante. A patologia deste exemplo reside no fato de $\mathbb{R} - \{0\}$ não possuir uma propriedade topológica que, em $\mathbb{R}$, só os intervalos têm, chamada conexidade, sobre a qual passaremos a discutir.

Definição 15 (cisão). Uma *cisão* de um conjunto $X \subset \mathbb{R}^n$ é uma decomposição do mesmo em dois conjuntos disjuntos que são, ambos, abertos em X, isto é, tem-se $A, B \subset X \subset \mathbb{R}^n$, tais que:

- $X = A \cup B$;
- $A \cap B = \emptyset$;
- A e B são abertos em X.

Todo conjunto $X \subset \mathbb{R}^n$ admite uma cisão. Basta fazer $X = X \cup \emptyset$. Esta é chamada de *cisão trivial*.

A decomposição $\mathbb{R} - \{0\} = (-\infty, 0) \cup (0, +\infty)$ é uma cisão de $\mathbb{R} - \{0\}$, pois $(-\infty, 0)$ e $(0, +\infty)$ são, ambos, abertos em $\mathbb{R} - \{0\}$ (por serem abertos em $\mathbb{R}$).

Seja $X = \{a, b\} \subset \mathbb{R}^n$, em que $a \neq b$. Se considerarmos uma bola aberta $B(a, r)$, disjunta de $\{b\}$, teremos $\{a\} = B(a, r) \cap X$. Logo $\{a\}$ é aberto em X. Analogamente, $\{b\}$ é aberto em X. Segue-se que $X = \{a\} \cup \{b\}$ é uma cisão de X.

Com um argumento análogo ao do parágrafo anterior, mostra-se que se $X \subset \mathbb{R}^n$ é discreto e possui mais de um elemento, então toda decomposição de X determinada por dois conjuntos disjuntos e não-vazios é uma cisão.

Note que se A e B são disjuntos e $X = A \cup B$, então $A = X - B$ e $B = X - A$. Logo, A e B são abertos em X se, e somente se, são fechados em X, isto é, na definição de cisão, pode-se trocar "abertos" por "fechados".

Proposição 19. *Uma decomposição $X = A \cup B$ de um conjunto $X \subset \mathbb{R}^n$ é uma cisão se, e somente se, não há pontos de aderência de A em B ou pontos de aderência de B em A, isto é,*

$$X = A \cup B \text{ é uma cisão } \Leftrightarrow A \cap \overline{B} = B \cap \overline{A} = \emptyset.$$

Demonstração. Se $X = A \cup B$ é uma cisão, então A e B são fechados em X. Daí e da Proposição 15, segue-se que $A = \overline{A} \cap X$ e $B = \overline{B} \cap X$. Além disso, A e B são subconjuntos disjuntos de X, donde

$$A \cap \overline{B} = (A \cap X) \cap \overline{B} = A \cap (X \cap \overline{B}) = A \cap B = \emptyset.$$

De modo análogo, verifica-se que $B \cap \overline{A} = \emptyset$.

Reciprocamente, suponhamos que se tenha $A \cap \overline{B} = B \cap \overline{A} = \emptyset$. Então,

$$\overline{A} \cap X = \overline{A} \cap (A \cup B) = (\overline{A} \cap A) \cup (\overline{A} \cap B) = A$$

e, igualmente, $\overline{B} \cap X = B$, donde se infere, novamente pela Proposição 15, que A e B são, ambos, fechados em X. Logo, $X = A \cup B$ é uma cisão. □

Definição 16 (conexidade). Um conjunto $X \subset \mathbb{R}^n$ é dito *conexo* se a única cisão que ele admite é a trivial. Caso contrário, ele é dito *desconexo*.

Segue-se das considerações acima que $\mathbb{R} - \{0\}$ é desconexo, assim como o é qualquer conjunto discreto de $\mathbb{R}^n$ com mais de um elemento. Em particular, toda esfera de $\mathbb{R}$ é desconexa. Para todo $a \in \mathbb{R}^n$, o conjunto $X = \{a\}$ é conexo, o que decorre do fato de este ter apenas dois subconjuntos, o próprio X e o conjunto vazio.

Convém observar que se $A \subset X$ é um subconjunto próprio e não-vazio que é aberto e fechado em X, então a decomposição $X = A \cup (X - A)$ é uma cisão não-trivial de X. Logo, vale o resultado seguinte.

Proposição 20. *Um conjunto $X \subset \mathbb{R}^n$ é conexo se, e somente se, os únicos subconjuntos de X que são abertos e fechados em X são o próprio X e o conjunto vazio.*

Dados $a, b \in \mathbb{R}^n$, o *segmento de reta* (*fechado*) determinado por a e b é, por definição(iv), o conjunto

$$[a,b] = \{x \in \mathbb{R}^n;\, x = (1-t)a + tb,\ t \in [0,1]\}.$$

Diz-se que um conjunto $X \subset \mathbb{R}^n$ é *convexo* quando, para quaisquer $a, b \in X$, tem-se $[a,b] \subset X$ (Fig. 9).

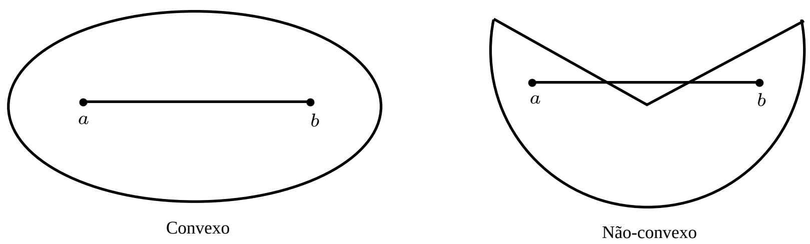

FIGURA 9

O espaço $\mathbb{R}^n$ é, obviamente, convexo. Toda bola de $\mathbb{R}^n$ é também convexa. De fato, seja B uma bola (aberta ou fechada) de $\mathbb{R}^n$ com centro em c e raio $r > 0$. Tomando-se $a, b \in B$, $x_t = (1-t)a + tb \in [a,b]$ e escrevendo-se $c = (1-t)c + tc$, tem-se $x_t - c = (1-t)(a-c) + t(b-c)$. Desta forma,

$$\|x_t - c\| \le (1-t)\|a-c\| + t\|b-c\| \le (1-t)r + tr = r,$$

sendo que, no caso de B ser uma bola aberta, a última desigualdade é estrita. Logo, $x_t \in B$, isto é, $[a,b] \subset B$, donde B é convexa(v).

Dada uma bola B de $\mathbb{R}^2$, aberta ou fechada, verifica-se, igualmente, que o *cilindro sólido* $C = B \times \mathbb{R} \subset \mathbb{R}^3$ é convexo. É fácil ver também que qualquer paralelepípedo, subespaço vetorial ou semi-espaço de $\mathbb{R}^n$ é um conjunto convexo.

PROPOSIÇÃO 21. *Todo conjunto convexo de $\mathbb{R}^n$ é conexo.*

DEMONSTRAÇÃO. Tomemos um conjunto convexo $X \subset \mathbb{R}^n$ e suponhamos, por absurdo, que $X = A \cup B$ seja uma cisão não-trivial do mesmo. Dados, então, $a \in A$ e $b \in B$, escrevamos, para cada $t \in [0,1]$, $x_t = (1-t)a + tb \in [a,b]$.

Consideremos o conjunto $\Omega = \{t \in [0,1];\, x_t \in A\}$ e observemos que $\Omega \neq \emptyset$, pois $a = x_0 \in A$, isto é, $0 \in \Omega$. Como Ω é limitado, existe $s = \sup \Omega$ e, em particular, uma sequência (t_k) em Ω, tal que $t_k \le s$ e $t_k \to s$. Uma vez que Ω é subconjunto de $[0,1]$ e este é fechado, tem-se $s \in [0,1]$. Assim, $(x_{t_k})_{k \in \mathbb{N}}$ é uma sequência em A que, pelas propriedades operatórias das sequências (vide Proposição 6 – Capítulo 1), converge para $x_s \in [a,b]$ (Fig. 10). Logo, $x_s \in \overline{A}$ e, então, pela Proposição 19, $x_s \notin B$. Em particular, $s < 1$.

Agora, pela definição de supremo, para todo $t \in (s,1]$, tem-se $x_t \notin A$, isto é, $x_t \in B$. Tomando-se, desta forma, uma sequência (t_i) em $(s,1]$ convergindo para s ($t_i = s + \frac{1-s}{i}$, por exemplo), teremos $x_{t_i} \to x_s$, isto é, $x_s \in \overline{B}$. Novamente pela

(iv)No caso $n = 1$, isto é, quando $a, b \in \mathbb{R}$, tem-se que o *segmento* $[a,b]$ coincide com o *intervalo* fechado $[a,b]$ se $a < b$, e com o *intervalo* fechado $[b,a]$ se $b < a$.

(v)Um argumento inteiramente análogo mostra que qualquer bola de um espaço vetorial normado é convexa. Em particular, as bolas de $\mathbb{R}^n$ relativas a quaisquer normas são convexas.

Proposição 19, $x_s \notin A$. Assim, $x_s \notin A \cup B = X$, isto é, $[a, b] \not\subset X$, o que contradiz o fato de X ser convexo. □

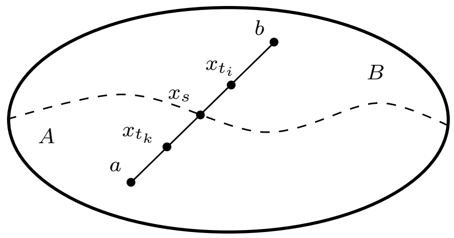

FIGURA 10

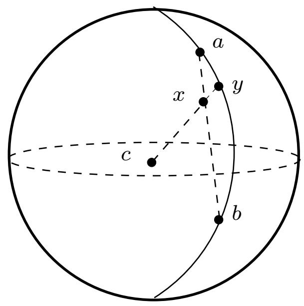

FIGURA 11

Como corolário da demonstração da Proposição 21, obtemos o resultado seguinte.

PROPOSIÇÃO 22. *Toda esfera de* $\mathbb{R}^{n+1}$ *é um conjunto conexo.*

DEMONSTRAÇÃO. Tomemos uma esfera $S = S[c, r] \subset \mathbb{R}^{n+1}$ e suponhamos que $S = A \cup B$ seja uma cisão não-trivial de S. Como S não é uma esfera de $\mathbb{R}$, pelo menos um dos conjuntos, A ou B, contém mais de um ponto. Assim, podemos escolher $a \in A$ e $b \in B$, de tal modo que os pontos a, b e c não sejam colineares, bastando, para tanto, escolhê-los de tal modo que $c \neq (a + b)/2$ (Fig. 11).

Desta forma, podemos definir o conjunto

$$[a, b]_S = \left\{ y \in S\,;\ y = r\frac{x - c}{\|x - c\|} + c,\ \ x \in [a, b] \right\}.$$

Fazendo-se, para cada $t \in [0, 1]$,

$$y_t = r\frac{x_t - c}{\|x_t - c\|} + c, \quad x_t = (1 - t)a + tb,$$

tem-se $y_t \in [a, b]_S \, \forall t \in [0, 1]$. Sendo assim, considerando-se o conjunto

$$\Omega = \{t \in [0, 1]\,;\ y_t \in A\},$$

podemos proceder como na demonstração da Proposição 21 (substituindo-se, no argumento, o conjunto $[a, b]$ por $[a, b]_S$) e concluir, desta forma, que a esfera S é conexa. □

Provemos agora uma das afirmações feitas no início desta seção.

PROPOSIÇÃO 23 (CONEXOS DE $\mathbb{R}$). *Um subconjunto não-vazio de* $\mathbb{R}$ *é conexo se, e somente se, é um intervalo.*

DEMONSTRAÇÃO. Todo intervalo de $\mathbb{R}$ é, obviamente, um conjunto convexo, portanto, conexo.

Reciprocamente, suponhamos que $X \subset \mathbb{R}$ seja conexo e não-vazio. Se X contém apenas um ponto, então X é um intervalo (degenerado). Caso contrário, sejam $a, b \in X$, tais que $a < b$. Então, dado c entre a e b, devemos ter $c \in X$. De fato, se não fosse assim, fazendo-se $A = X \cap (-\infty, c)$ e $B = X \cap (c, +\infty)$, a decomposição $X = A \cup B$ seria uma cisão não-trivial de X, pois A e B são, ambos, abertos não-vazios de X (note que $a \in A$ e $b \in B$). Logo, X é um intervalo. $\square$

Exemplo 29. *Todo isomorfismo linear* $T \in \mathrm{L}(\mathbb{R}^n)$ *leva conexos* de $\mathbb{R}^n$ *em conexos* de $\mathbb{R}^n$. De fato, seja $X \subset \mathbb{R}^n$ um conjunto conexo e $T(X) = A \cup B$ uma cisão de $T(X)$. Neste caso, temos que $A = U \cap T(X)$, em que $U \subset \mathbb{R}^n$ é aberto, e, então, $T^{-1}(A) = T^{-1}(U \cap T(X)) = T^{-1}(U) \cap T^{-1}(T(X)) = T^{-1}(U) \cap X$, donde $T^{-1}(A)$ é aberto em X, pois (vide Exemplo 14) $T^{-1}(U)$ é aberto em $\mathbb{R}^n$. Analogamente, $T^{-1}(B) \subset X$ é aberto em X. Além disso,

$$X = T^{-1}(T(X)) = T^{-1}(A \cup B) = T^{-1}(A) \cup T^{-1}(B)$$

e $T^{-1}(A) \cap T^{-1}(B) = T^{-1}(A \cap B) = T^{-1}(\emptyset) = \emptyset$, isto é, a decomposição $X = T^{-1}(A) \cup T^{-1}(B)$ é uma cisão de X. Sendo este conexo, devemos ter $T^{-1}(A) = \emptyset$ ou $T^{-1}(B) = \emptyset$, o que nos dá $A = \emptyset$ ou $B = \emptyset$. Logo, $T(X)$ é conexo.

Proposição 24. *A união de uma família arbitrária de conjuntos conexos de* $\mathbb{R}^n$ *que têm um ponto em comum é um conjunto conexo.*

Demonstração. Seja $\{X_\lambda\}_{\lambda \in \Lambda}$ uma família de conexos de $\mathbb{R}^n$, tal que $a \in X_\lambda \ \forall \lambda \in \Lambda$. Ponha $X = \bigcup X_\lambda$ e considere uma cisão $X = A \cup B$. Como $a \in X$, tem-se $a \in A$ ou $a \in B$. Suponhamos então, sem perda de generalidade, que $a \in A$. Uma vez que A e B são abertos em X e, para todo $\lambda \in \Lambda$, $X_\lambda \subset X$, temos que $A \cap X_\lambda$ e $B \cap X_\lambda$ são abertos em X_λ, donde, para todo $\lambda \in \Lambda$, $X_\lambda = (A \cap X_\lambda) \cup (B \cap X_\lambda)$ é uma cisão de X_λ. Sendo cada X_λ conexo, devemos ter $A \cap X_\lambda = \emptyset$ ou $B \cap X_\lambda = \emptyset$. Porém, $a \in X_\lambda$ para todo λ, o que nos dá $B \cap X_\lambda = \emptyset$. Logo, $B = B \cap X = B \cap \bigcup X_\lambda = \bigcup (X_\lambda \cap B) = \emptyset$, donde X é conexo. $\square$

Note que a interseção de conexos não é, necessariamente, conexa. Tome, por exemplo, uma esfera e uma reta que passa pelo seu centro.

Exemplo 30 (conexidade da bola menos um ponto). Seja B uma bola de $\mathbb{R}^n$, aberta ou fechada, centrada na origem. Como aplicação da Proposição 24, provemos que o conjunto $X = B - \{0\}$ é um conexo de $\mathbb{R}^n$ quando $n > 1$. Por simplicidade de notação, consideremos o caso $n = 2$. Claramente, os conjuntos

- $Q_1 = \{(x_1, x_2) \in \mathbb{R}^2; x_1, x_2 \geq 0\} - \{(0,0)\}$;
- $Q_2 = \{(x_1, x_2) \in \mathbb{R}^2; x_1 \leq 0, x_2 \geq 0\} - \{(0,0)\}$;

são, ambos, convexos. É fácil ver também que a interseção de convexos é convexa. Desta forma, os conjuntos $C_1 = Q_1 \cap B$ e $C_2 = Q_2 \cap B$ são convexos e, portanto, conexos (Fig. 12). Como $C_1 \cap C_2$ é, obviamente, não-vazio, segue-se da Proposição 24 que $X_1 = C_1 \cup C_2 = \{(x_1, x_2) \in X;\ x_2 \geq 0\}$ é conexo. Analogamente, o conjunto $X_2 = \{(x_1, x_2) \in X;\ x_2 \leq 0\}$ é conexo. Uma vez que

$$X = X_1 \cup X_2 \quad \text{e} \quad X_1 \cap X_2 = \{(x_1, x_2) \in X;\ x_2 = 0\} \neq \emptyset,$$

infere-se da Proposição 24 que X é conexo.

Exemplo 31 (união disjunta e conexa). Sejam $p = (1, 0) \in \mathbb{R}^2$ e (x_k) uma sequência em $S^1 - \{p\} \subset \mathbb{R}^2$, tais que $x_k \to p$. Uma tal sequência existe, pois nenhum ponto de S^1 é isolado (vide Exercício 12). Para cada $k \in \mathbb{N}$, considere

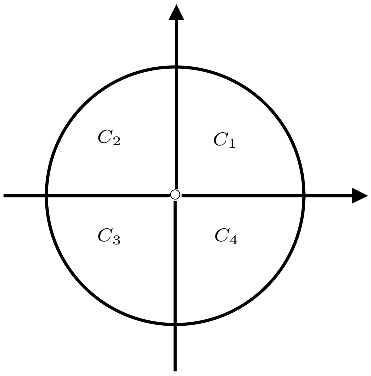

FIGURA 12

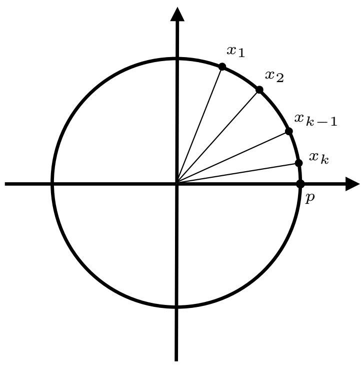

FIGURA 13

o segmento $X_k = [0, x_k] \subset \mathbb{R}^2$ (Fig. 13). Cada X_k é conexo e $\bigcap X_k = \{0\}$. Logo, pela Proposição 24, $X = \bigcup X_k$ é conexo. Consideremos, agora, o conjunto $C = X \cup \{p\} \subset \mathbb{R}^2$ e provemos que, contrariamente à nossa intuição, C é conexo. Para isto, observemos inicialmente que a união que define C, apesar de ser disjunta, não é uma cisão, pois $x_k \to p$, isto é, $p \in \overline{X}$. Desta forma, uma cisão de C é, necessariamente, da forma $C = (A \cup \{p\}) \cup B$, em que $A, B \subset X$, $A \neq \emptyset$ e $A \cup \{p\}$ e B são disjuntos e abertos em C. Em particular, tem-se $X = A \cup B$ e $A \cap B = \emptyset$. Agora, uma vez que $X \subset C$, $A = (A \cup \{p\}) \cap X$ e $B = B \cap X$, temos que A e B são, ambos, abertos em X. Sendo este último conexo, isto implica que $B = \emptyset$ e $A = X$, donde C é conexo (compare com o resultado do Exercício 22).

DEFINIÇÃO 17 (COMPONENTE CONEXA). Dado um conjunto $X \subset \mathbb{R}^n$, a *componente conexa* de um elemento $x \in X$, a qual denotamos por C_x, é a união de todos os subconjuntos de X que são conexos e contêm x.

Infere-se da Proposição 24 que, para quaisquer $X \subset \mathbb{R}^n$ e $x \in X$, o conjunto C_x é conexo. Além disso, é imediato da definição que C_x é o maior subconjunto conexo de X que contém x, isto é, se C é um subconjunto conexo de X e $x \in C$, então $C \subset C_x$. Daí, segue-se que X se exprime como uma união disjunta de suas componentes conexas, isto é,

$$X = \bigcup_{x \in X} C_x \quad \text{e} \quad C_x \cap C_y \neq \emptyset \;\Rightarrow\; C_x = C_y\,.$$

De fato, se $y \in C_x$, então $C_x \subset C_y$, donde $x \in C_y$ e, portanto, $C_y \subset C_x$, ou seja, $C_x = C_y$. Assim, duas componentes conexas de X são disjuntas ou iguais.

Se X é conexo, então $C_x = X \; \forall x \in X$. É fácil ver também que $\mathbb{R} - \{0\}$ tem duas componentes conexas, $C_{-1} = (-\infty, 0)$ que é a componente conexa de qualquer real negativo, e $C_1 = (0, +\infty)$, que é a componente conexa de qualquer real positivo. Observe ainda que se X é um conjunto discreto, $C_x = \{x\} \; \forall x \in X$.

Aplicaremos agora a noção de componente conexa para caracterizar os conjuntos abertos de $\mathbb{R}$, conforme a proposição seguinte.

PROPOSIÇÃO 25 (CARACTERIZAÇÃO DOS ABERTOS DE $\mathbb{R}$). *Todo conjunto aberto e não-vazio de $\mathbb{R}$ se exprime como uma união disjunta, finita ou enumerável, de intervalos abertos.*

DEMONSTRAÇÃO. Seja $A \subset \mathbb{R}$ um conjunto aberto e não-vazio. Uma vez que os únicos subconjuntos conexos e não-vazios de $\mathbb{R}$ são os intervalos, cada componente conexa de A é um intervalo I_x, $x \in A$. Ademais, cada intervalo I_x é aberto, pois se tivéssemos um extremo a de I_x no conjunto A, existiria um intervalo aberto J contendo a e contido em A. Neste caso, $I_x \cup J$ seria um intervalo contendo I_x propriamente e contido em A, o que vai de encontro ao fato de I_x ser uma componente conexa de A.

Assim, o conjunto das componentes conexas de A, $\Lambda = \{I_x\,;\, x \in A\}$, constitui-se de intervalos abertos e disjuntos que decompõem A. Resta-nos, então, mostrar que Λ é finito ou enumerável. Para tanto, basta observarmos que, devido à densidade de $\mathbb{Q}$ em $\mathbb{R}$, para cada $x \in A$, existe um racional q_x no intervalo I_x. A associação $I_x \mapsto q_x$ define, claramente, uma aplicação bijetiva de Λ num subconjunto de $\mathbb{Q}$, donde se infere que Λ é, de fato, finito ou enumerável. $\square$

Em conclusão, gostaríamos de salientar que em $\mathbb{R}^n$, assim como a compacidade, a conexidade é um conceito absoluto, não relativo. Isto segue-se do fato de, na definição de conjunto conexo, fazer-se referência apenas à topologia do mesmo e não à de $\mathbb{R}^n$. Por esta razão, os subconjuntos conexos de $\mathbb{R}^n$ podem ser chamados de *espaços (topológicos) conexos.*

6. Espaços Topológicos

Relembremos que um *espaço topológico* é um par (X, τ), em que X é um conjunto e τ é uma família de subconjuntos de X, chamados de *abertos*, que satisfazem as condições seguintes:

i) O conjunto X e o conjunto vazio são elementos de τ;

ii) a interseção de uma subfamília finita de τ é um elemento de τ;

iii) a união de uma subfamília qualquer de τ é um elemento de τ.

A família τ é dita, então, uma *topologia* em X.

Relembremos ainda que, a exemplo de $\mathbb{R}^n$ (munido de uma métrica), em todo espaço métrico (M, d) existe uma topologia natural, τ_d, determinada pela métrica d. Neste caso, (M, τ_d) é dito um espaço topológico *metrizável* e a topologia τ_d é dita *proveniente* de d. Há, no entanto, inúmeros espaços topológicos que não são metrizáveis, isto é, espaços topológicos (X, τ) em que a topologia τ não é proveniente de métrica alguma de X (vide Exercício 28). Este fato está ligado à noção de *espaço de Hausdorff*.

Um espaço topológico (X, τ) é dito *de Hausdorff* se, para quaisquer dois pontos distintos de X, a e b, existem abertos disjuntos, $A_1, A_2 \in \tau$, tais que $a \in A_1$ e $b \in A_2$.

Todo espaço topológico metrizável (M, τ_d) é de Hausdorff, pois, dados $a, b \in M$, tomando-se $r = d(a,b)/2$, tem-se que os abertos $A_1 = B(a,r) \ni a$ e $A_2 = B(b,r) \ni b$ são, claramente, disjuntos. Assim, uma condição necessária (mas não suficiente) para um espaço topológico ser metrizável é a de que ele seja de Hausdorff.

Chama-se, então, *Topologia* a teoria que estuda os espaços topológicos. Um dos aspectos de grande relevância da Topologia é o de que os espaços topológicos constituem os ambientes gerais em que se pode tratar a questão de convergência de sequências, isto é, esse conceito se estende (conforme sugerido pelo conteúdo da

Proposição 11) dos espaços métricos para os espaços topológicos. Mais especificamente, diz-se que uma sequência $(x_k)_{k\in\mathbb{N}}$ num espaço topológico (X,τ) *converge* para $a \in X$ se cumpre a seguinte condição:

$$\boxed{\text{Para toda vizinhança } V \text{ de } a, \text{ existe } k_0 \in \mathbb{N}, \text{ tal que } k \geq k_0 \Rightarrow x_k \in V.}$$

Por *vizinhança* de $a \in X$, entenda-se um conjunto $V \subset X$ que contém um aberto $A \ni a$ de τ. Neste caso, diz-se que a é um ponto *interior* a V e denota-se o conjunto dos pontos interiores de $V \subset X$ por $\operatorname{int} V$.

Uma vez que, intuitivamente, convergir significa "estar cada vez mais próximo", uma topologia num conjunto X pode ser vista como uma abstração do conceito de proximidade. Deve-se observar, entretanto, que essa "proximidade abstrata" é pontual, isto é, ela diz respeito a certos pontos de X estarem "próximos" de um outro, o qual foi predeterminado. Já a noção de proximidade, entre si, dos pontos de uma sequência, não se caracteriza topologicamente, isto é, ela não se estende dos espaços métricos aos espaços topológicos (vide Exercício 29). Em suma, o conceito de convergência — o mais fundamental da Análise — é métrico e topológico, enquanto o de sequência de Cauchy é métrico e não-topológico.

Do ponto de vista da Análise em $\mathbb{R}^n$, deve-se notar que a complexidade das funções de mais de uma variável, comparadas àquelas de uma única, reside no fato de a topologia usual de $\mathbb{R}^n$, isto é, aquela proveniente da métrica euclidiana, ser mais rica que a de $\mathbb{R}$ quando $n > 1$. Sejamos mais claros quanto a isto.

Conforme constatamos na Proposição 25, os únicos abertos (da topologia usual) de $\mathbb{R}$ são aqueles dados por uma união finita ou enumerável de bolas abertas — intervalos abertos, no caso —, duas a duas disjuntas. Os abertos de $\mathbb{R}^n$, $n > 1$, todavia, não admitem essa caracterização. Para vermos isto, basta considerarmos uma bola aberta qualquer de $\mathbb{R}^2$, por exemplo. Deformando-a convenientemente, obtemos facilmente um aberto deste espaço que não é uma bola aberta ou uma união disjunta, finita ou enumerável, destas (Fig. 14). Note que uma deformação qualquer de uma bola de $\mathbb{R}$ produz nada mais que uma outra bola.

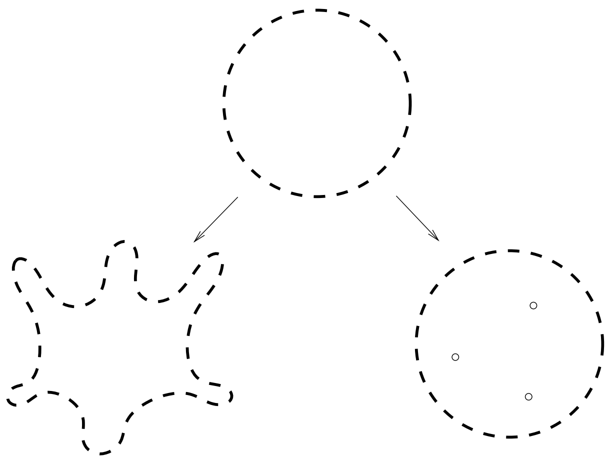

FIGURA 14

Outra maneira de se obter abertos de $\mathbb{R}^n$ a partir de bolas abertas, consiste em remover destas uma quantidade finita de pontos. Neste caso, quando $n > 1$, obtém-se um aberto que, claramente, não é uma bola aberta ou uma união disjunta, finita ou enumerável, de bolas abertas. No entanto, quando $n = 1$, obtém-se daí uma união disjunta de uma quantidade finita de bolas abertas.

Desta forma, podemos dizer que quando $n > 1$, há na topologia usual de $\mathbb{R}^n$ uma diversidade de abertos que não há na de $\mathbb{R}$.

No que concerne à classificação dos espaços topológicos, esta é feita, como esperado, via bijeções entre estes que preservam as suas devidas estruturas. Para tanto, estas bijeções devem ser aplicações abertas cujas inversas são também abertas. Mais precisamente, diz-se que dois espaços topológicos (X_1, τ_1) e (X_2, τ_2) são *homeomorfos* quando existe um *homeomorfismo* entre os mesmos, isto é, uma bijeção $\varphi : X_1 \to X_2$, tal que, para quaisquer abertos $A_1 \in \tau_1$, $A_2 \in \tau_2$, tem-se $\varphi(A_1) \in \tau_2$ e $\varphi^{-1}(A_2) \in \tau_1$.

É imediato que são homeomorfismos:

- A aplicação identidade de um espaço topológico;
- a inversa de um homeomorfismo;
- a composta de homeomorfismos;

donde se conclui que "ser homeomorfo a" é uma relação de equivalência na classe dos espaços topológicos.

As noções de conjunto fechado, topologia relativa, conjunto compacto e conjunto conexo introduzem-se em espaços topológicos de forma análoga à que fizemos em $\mathbb{R}^n$. Um subconjunto de um espaço topológico (X, τ) munido da topologia relativa a τ é dito um *subespaço topológico* de X. Verifica-se, então, que, como em $\mathbb{R}^n$, compacidade e conexidade são conceitos absolutos, isto é, são intrínsecos aos subespaços e não relativos aos espaços que os contêm. Assim, dizemos que um espaço topológico (X, τ) é *compacto* se toda cobertura (que, neste caso, é uma decomposição) de X por abertos de τ admite uma subcobertura finita, e *conexo* se a única cisão (por meio de abertos de τ) que admite é a trivial.

Pelas considerações do Exemplo 14, tem-se que todo isomorfismo linear de $\mathbb{R}^n$ em si mesmo é um homeomorfismo. Além disso, com os mesmos argumentos dos exemplos 27 e 29, em que se prova que tais isomorfismos levam compactos em compactos e conexos em conexos, obtém-se o resultado seguinte.

Teorema 10. *Sejam X e Y espaços topológicos homeomorfos. Então,*

- *X é compacto se, e somente se, Y é compacto.*
- *X é conexo se, e somente se, Y é conexo.*

O Teorema 10 estabelece, então, a invariância das propriedades de compacidade e conexidade por homeomorfismos, constituindo-se, desta forma, num instrumento para distinguir espaços topológicos. Dele se infere, por exemplo, que o espaço $\mathbb{R}^n$ não é homeomorfo a uma esfera, uma vez que esta é compacta e $\mathbb{R}^n$ não é compacto. Um argumento semelhante mostra também que uma bola aberta de $\mathbb{R}^2$ não é homeomorfa a um intervalo aberto de $\mathbb{R}$. De fato, se houvesse um homeomorfismo $\varphi : B(x_0, r) \subset \mathbb{R}^2 \to (a, b) \subset \mathbb{R}$, então $\varphi : B(x_0, r) - \{x_0\} \to (a, b) - \{\varphi(x_0)\}$ seria também um homeomorfismo (vide Exercício 30). Isto, porém, contradiz o Teorema 10, pois $B(x_0, r) - \{x_0\}$ é conexo (vide Exemplo 30) e $(a, b) - \{\varphi(x_0)\}$ é desconexo.

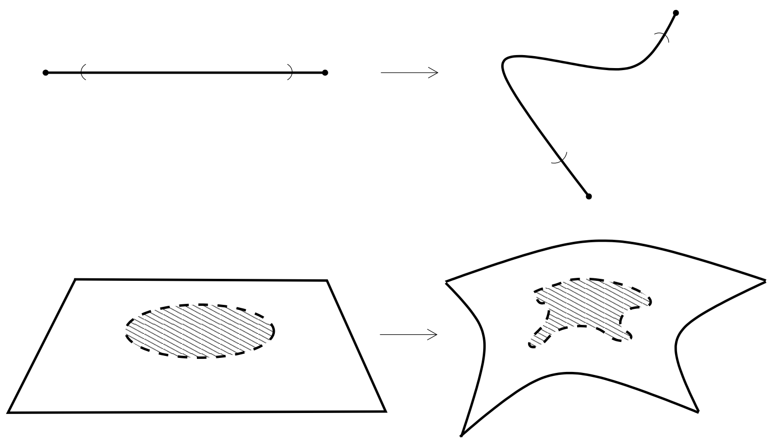

FIGURA 15

A intuição geométrica nos diz que, em $\mathbb{R}^n$, abertos relativos de curvas ou superfícies são invariantes por deformações (Fig. 15). Sendo assim, enquanto espaços topológicos, duas curvas ou superfícies de $\mathbb{R}^n$ que são obtidas uma da outra através de uma deformação são homeomorfas, como, por exemplo, uma reta e uma parábola de $\mathbb{R}^2$ ou uma esfera e um elipsoide de $\mathbb{R}^3$ (vide Exercício 30). Conforme mencionamos à introdução deste capítulo, este fato remete às origens da Topologia e justifica a sua designação de *geometria da folha de borracha.*

7. Topologia ⇔ Álgebra (∗)

A fim de ilustrar a riqueza e generalidade da Topologia, bem como da Álgebra, faremos um interessante intercâmbio entre estas teorias apresentando uma demonstração topológica de um teorema algébrico e uma demonstração algébrica de um teorema topológico.

7.1. Topologia e a Infinitude dos Números Primos. Iniciemos, pois, exibindo uma bela demonstração da infinitude dos números primos usando argumentos topológicos. Ela é devida ao matemático israelense Harry Furstenberg (1935–), que a publicou em 1955 (vide [**10**]).

Designemos por $\mathbb{P} \subset \mathbb{N}$ o conjunto dos números primos, isto é, aqueles que possuem dois, e somente dois, divisores em $\mathbb{N}$.

TEOREMA 11. $\mathbb{P}$ *é um conjunto infinito.*

DEMONSTRAÇÃO. Dados $a \in \mathbb{Z}$ e $r \in \mathbb{N}$, defina

$$B(a,r) = \{x \in \mathbb{Z};\, x = a + kr,\, k \in \mathbb{Z}\}$$

e note que, para quaisquer $a \in \mathbb{Z}$ e $r, s \in \mathbb{N}$, $B(a, sr) \subset B(a, r)$.

Diremos, então, que um conjunto $A \subset \mathbb{Z}$ é *aberto* se, para todo $a \in A$, existe $r \in \mathbb{N}$, tal que $B(a,r) \subset A$. Vejamos que, com esta definição, o conjunto τ, formado pelos abertos de $\mathbb{Z}$, definem neste uma topologia.

Com efeito, é imediato que $\mathbb{Z}$ e o conjunto vazio (por vacuidade) são abertos. Segue-se também diretamente da definição que a união arbitrária de conjuntos abertos é aberta. Agora, se A_1, A_2 são abertos não-disjuntos e $a \in A_1 \cap A_2$, então existem $r_1, r_2 \in \mathbb{N}$, tais que $B(a, r_1) \subset A_1$ e $B(a, r_2) \subset A_2$. Logo, $B(a, r_1 r_2) \subset A_1 \cap A_2$, implicando que $A_1 \cap A_2$ é aberto. Por indução, conclui-se facilmente que a interseção finita de abertos é aberta, donde τ é uma topologia em $\mathbb{Z}$.

Destaquemos duas propriedades especiais desta topologia. Primeiro, cada conjunto aberto e não-vazio é infinito, pois contém algum conjunto $B(a, r)$, que, por sua vez, é infinito. Segundo, para quaisquer $a \in \mathbb{Z}$ e $r \in \mathbb{N}$, $B(a, r)$ é, evidentemente, aberto. Porém, é também fechado. Para vermos isto, basta notarmos que

$$\mathbb{Z} - B(a, r) = \bigcup_{i=1}^{r-1} B(a + i, r).$$

Assim, o complementar de $B(a, r)$ em $\mathbb{Z}$ é aberto (pois é dado por uma união de abertos). Logo, $B(a, r)$ é fechado.

Agora, exceto 1 e -1, cada inteiro a admite um divisor primo p, isto é, $a \in B(0, p)$ para algum primo $p \in \mathbb{P}$. Desta forma, podemos escrever

$$\mathbb{Z} - \{-1, 1\} = \bigcup_{p \in \mathbb{P}} B(0, p),$$

donde se infere que $\mathbb{P}$ deve ser infinito. Caso contrário, o conjunto $\mathbb{Z} - \{-1, 1\}$ seria fechado (por ser uma união finita de fechados) e $\{-1, 1\}$ seria, então, aberto. Isto, porém, contradiria o fato de os abertos de τ serem todos infinitos. □

7.2. O Teorema do Fecho-Complemento de Kuratowski. Dado um espaço topológico (X, τ), designemos por M o conjunto formado por todos os operadores (funções) $f : P(X) \to P(X)$, em que $P(X)$ é o conjunto das partes de X, isto é, aquele formado por todos os subconjuntos de X.

O curioso resultado que apresentaremos a seguir, estabelecido em 1922 pelo matemático polonês Kazimierz Kuratowski (1896–1980) (vide [**16**]), envolve especialmente dois desses operadores, o *complemento* e o *fecho*, dados respectivamente por

$$A \mapsto X - A \quad \text{e} \quad A \mapsto \overline{A}, \quad A \in P(X).$$

Teorema do Fecho-Complemento (Kuratowski). *Sejam (X, τ) um espaço topológico e $A \subset X$ um subconjunto arbitrário de X. Então, aplicando-se a A sucessões aleatórias de fechos e complementos, é possível obter no máximo 14 subconjuntos de X que sejam distintos entre si. Além disso, existem um espaço topológico X e um subconjunto A deste, em que o limite máximo de 14 subconjuntos distintos é atingido.*

Para uma devida demonstração do Teorema do Fecho-Complemento e consequente clarificação do misterioso número 14 que aparece em seu enunciado, faz-se natural uma abordagem algébrica. Isto se dá a partir da observação de que o par $(M, \circ)$, em que $\circ$ denota a composição em M, compõe um monoide[(vi)] cujo elemento neutro é o operador identidade de $P(X)$, o qual denotaremos por e.

Mais especificamente, dados $f, g \in M$, definindo-se o *produto* $fg = f \circ g \in M$, tem-se, para quaisquer $f, g, h \in M$,

[(vi)]Um *monoide* é uma estrutura algébrica que consiste de um conjunto não-vazio e uma operação binária associativa, neste definida, que possui um elemento neutro. Deve-se observar que a estrutura de monoide é mais fraca que a de grupo, pois prescinde da propriedade de existência do elemento inverso.

- $(fg)h = f(gh)$ (associatividade);
- $fe = ef = f$ (existência de elemento neutro).

Neste contexto, adotam-se as notações f^k, para a composta $f \circ f \circ \cdots \circ f$ (k vezes), e $f^0 = e$.

Denotemos por a e b, respectivamente, os operadores complemento e fecho, isto é, dado $A \subset X$, tem-se

$$a(A) = X - A \qquad \text{e} \qquad b(A) = \overline{A},$$

e observemos que estes têm as seguintes propriedades:

i) $a^2 = e$;

ii) $b^2 = b$;

iii) $i = aba$;

em que i é o operador *interior*: $A \mapsto \operatorname{int} A$, $A \subset X$.

As propriedades (i) e (ii) são evidentes. Quanto a (iii), basta observarmos que (vide Proposição 12-(ii)) $i(A) = X - \overline{X - A} = a(\overline{X - A}) = a(b(a(A)))$, donde $i = aba$.

É imediato que o conjunto $K \subset M$, de todos os produtos cujos fatores são a ou b, é também um monoide (dito, *de Kuratowski*), pois $a^2 = e \in K$. A demonstração do Teorema do Fecho-Complemento reduz-se, desta forma, a mostrar que K possui, precisamente, 14 elementos.

Com isto em mente, definimos uma relação de ordem parcial em K da seguinte forma: Dadas $f, g \in K$, dizemos que $f \leq g$ se, para todo $A \subset X$, $f(A) \subset g(A)$. Tem-se, por exemplo, que $i \leq e \leq b$, pois, para todo $A \subset X$,

$$i(A) = \operatorname{int} A \subset A = e(A) \subset \overline{A} = b(A).$$

Dados $f, g \in K$, verificam-se, então, as seguintes propriedades:

iv) $f \leq g \quad \text{e} \quad g \leq f \quad \Rightarrow \quad f = g$;

v) $f \leq g \quad \Rightarrow \quad bf \leq bg$;

vi) $f \leq g \quad \Rightarrow \quad ag \leq af$.

Demonstração do Teorema do Fecho-Complemento. Adotemos a notação introduzida acima e observemos que, pelas propriedades (i) e (ii), todo elemento de K se exprime de uma das seguintes formas:

$$e, \quad a, \quad b, \quad (ab)^k, \quad (ba)^k, \quad (ab)^k a, \quad (ba)^k b, \tag{13}$$

em que $k \in \mathbb{N}$.

Suponhamos, então, que valha a igualdade

$$bab = (ba)^3 b. \tag{14}$$

Multiplicando-se ambos os seus membros[vii] por a à esquerda ou à direita, obtém-se, em cada caso, as respectivas igualdades

$$(ab)^2 = (ab)^4 \qquad \text{e} \qquad (ba)^2 = (ba)^4.$$

[vii] Aqui, estamos usando a seguinte propriedade: $f, g \in M$ e $f = g \Rightarrow hf = hg$ e $fh = gh \, \forall h \in M$.

Daí, segue-se que existem 3, e somente 3, elementos distintos de K que se escrevem como $(ab)^k$, quais sejam, ab, $(ab)^2$ e $(ab)^3$. Analogamente, existem apenas três elementos distintos de K que se escrevem como $(ba)^k$ e o mesmo ocorre com os produtos da forma $(ab)^k a$. Finalmente, devido à igualdade (14), os produtos da forma $(ba)^k b$ produzem apenas dois elementos distintos de K, bab e $(ba)^2 b$.

Considerando-se, desta forma, a lista (13), vê-se que, uma vez verificada a igualdade (14), pode-se concluir que K tem 14 elementos. Especificamente,

$$K = \{e, a, b, ab, (ab)^2, (ab)^3, ba, (ba)^2, (ba)^3, aba, (ab)^2 a, (ab)^3 a, bab, (ba)^2 b\}.$$

Provemos, então, a validez de (14). Para tanto, observemos que, pela propriedade (iii), acima, tem-se

$$(ab)^3 = (aba)(bab) = i(bab),$$

donde $(ab)^3 \leq bab$, já que $i \leq e$. Agora, partindo-se da desigualdade $i \leq b$, obtém-se, de (ii) e (v), $ib \leq b^2 = b$. Daí e de (vi), tem-se $ba \leq iba$. Desta desigualdade e de (v), obtém-se, finalmente, $bab \leq i(bab) = (ab)^3$. Portanto, pela propriedade (iv), vale a igualdade $(ab)^3 = bab$, da qual se obtém (14) por multiplicação à esquerda por b.

Para concluirmos a demonstração, tomemos como espaço topológico (X, τ) o conjunto dos números reais $\mathbb{R}$ munido de sua topologia usual e consideremos o conjunto

$$A = (0,1) \cup (1,2) \cup \{3\} \cup (\mathbb{Q} \cap [4,5]) \subset \mathbb{R}.$$

Aplicando-se a A cada um dos 14 operadores do monoide de Kuratowski K, obtém-se 14 conjuntos distintos entre si, conforme ilustrado na Tabela 1.

OPERADOR f	CONJUNTO $f(A)$
e	$(0,1) \cup (1,2) \cup \{3\} \cup (\mathbb{Q} \cap [4,5])$
a	$(-\infty,0] \cup \{1\} \cup [2,3) \cup (3,4) \cup ((\mathbb{R} - \mathbb{Q}) \cap [4,5]) \cup (5,+\infty)$
b	$[0,1] \cup [1,2] \cup \{3\} \cup [4,5]$
ab	$(-\infty,0) \cup (2,3) \cup (3,4) \cup (5,+\infty)$
$(ab)^2$	$(0,2) \cup (4,5)$
$(ab)^3$	$(-\infty,0) \cup (2,4) \cup (5,+\infty)$
ba	$(-\infty,0] \cup \{1\} \cup [2,+\infty)$
$(ba)^2$	$[0,2]$
$(ba)^3$	$(-\infty,0] \cup [2,+\infty)$
aba	$(0,1) \cup (1,2)$
$(ab)^2 a$	$(-\infty,0) \cup (2,+\infty)$
$(ab)^3 a$	$(0,2)$
bab	$(-\infty,0] \cup [2,4] \cup [5,+\infty)$
$(ba)^2 b$	$[0,2] \cup [4,5]$

TABELA 1

□

8. Exercícios

Seções 1 e 2

1. Mostre que, para todo $X \subset \mathbb{R}^n$, $\operatorname{int} X$ é aberto e contém qualquer aberto que esteja contido em X.

2. Dados subconjuntos X, Y de $\mathbb{R}^n$, verifique as seguintes relações:
   i) $\operatorname{int}(X \cap Y) = \operatorname{int} X \cap \operatorname{int} Y$;
   ii) $\operatorname{int}(X \cup Y) \supset \operatorname{int} X \cup \operatorname{int} Y$;
   iii) $\overline{X \cup Y} = \overline{X} \cup \overline{Y}$;
   iv) $\overline{X \cap Y} \subset \overline{X} \cap \overline{Y}$.

   Exiba exemplos em que as inclusões dos itens (ii) e (iv) não se reduzam a igualdades.

3. Mostre que $A \subset \mathbb{R}^n$ é aberto se, e somente se, $A \cap \overline{X} \subset \overline{A \cap X}$ para todo $X \subset \mathbb{R}^n$.

4. Sejam $F \subset \mathbb{R}^n$ um conjunto fechado e $r > 0$. Prove que
$$X = \{x \in \mathbb{R}^n;\ \|x - a\| = r \text{ para algum } a \in F\}$$
   é um subconjunto fechado de $\mathbb{R}^n$.

5. Prove que todo subespaço vetorial próprio de $\mathbb{R}^n$ é fechado e tem interior vazio.

6. Dada $T \in \mathrm{L}(\mathbb{R}^n, \mathbb{R}^m)$, mostre que:
   i) Para todo aberto A de $\mathbb{R}^m$, $T^{-1}(A)$ é um aberto de $\mathbb{R}^n$;
   ii) para todo fechado F de $\mathbb{R}^m$, $T^{-1}(F)$ é um fechado de $\mathbb{R}^n$.

7. Prove que:
   i) Toda aplicação linear injetiva $T : \mathbb{R}^n \to \mathbb{R}^{n+m}$ é fechada;
   ii) toda aplicação linear sobrejetiva $T : \mathbb{R}^{n+m} \to \mathbb{R}^n$ é aberta.

8. Sejam X e Y subconjuntos de $\mathbb{R}^n$ e $\mathbb{R}^m$, respectivamente. Considere o resultado do Exercício 6 e mostre que:
   i) $X \times Y$ é um subconjunto aberto de $\mathbb{R}^n \times \mathbb{R}^m$ se, e somente se, X e Y são abertos;
   ii) $X \times Y$ é um subconjunto fechado de $\mathbb{R}^n \times \mathbb{R}^m$ se, e somente se, X e Y são fechados.

9. Prove que cada um dos seguintes conjuntos:
   i) $A_1 = \{T \in \mathrm{L}(\mathbb{R}^n, \mathbb{R}^m)\,;\, T \text{ é injetiva}\}\,;$
   ii) $A_2 = \{T \in \mathrm{L}(\mathbb{R}^n, \mathbb{R}^m)\,;\, T \text{ é sobrejetiva}\}\,;$

   é um aberto de $\mathrm{L}(\mathbb{R}^n, \mathbb{R}^m)$.

10. Mostre que a fronteira de todo aberto de $\mathbb{R}^n$ tem interior vazio.

11. Seja a um ponto isolado de um conjunto fechado $F \subset \mathbb{R}^n$. Suponha que $F - \{a\}$ é não-vazio e prove que existe $r > 0$, tal que
    i) $B(a, r) \cap F = \{a\}$;
    ii) $S[a, r] \cap F \neq \emptyset$.

    Mostre, através de um contra-exemplo, que a hipótese de F ser fechado é necessária.

12. Prove que nenhum ponto de $S^n \subset \mathbb{R}^{n+1}$ é isolado.

Seção 3

13. Mostre que a família dos abertos relativos de um conjunto $X \subset \mathbb{R}^n$ define uma topologia em X.

14. Prove que todo conjunto $X \subset \mathbb{Z}^n$ é aberto e fechado relativamente a $\mathbb{Z}^n$.

15. Dados um subconjunto X de $\mathbb{R}^n$ e $F \subset X$, mostre que F é fechado em X se, e somente se, existe um fechado $G \subset \mathbb{R}^n$, tal que $F = G \cap X$.

Seção 4

16. Sejam $X \subset \mathbb{R}^n$ e $Y \subset \mathbb{R}^m$. Considere o Exercício 8 e prove que $X \times Y$ é compacto se, e somente se, X e Y são compactos.

17. Mostre que a fronteira de todo conjunto limitado de $\mathbb{R}^n$ é compacta.

18. Sejam $T \in \mathrm{L}(\mathbb{R}^n, \mathbb{R}^m)$ e $K \subset \mathbb{R}^n$ um conjunto compacto. Mostre que $T(K)$ é compacto.

19. Seja X um subconjunto de $\mathbb{R}^n$, tal que, para todo compacto $K \subset \mathbb{R}^n$, $X \cap K$ é compacto. Prove que X é fechado.

20. Seja $K \subset \mathbb{R}^n$ compacto e $\mathscr{A} = \{A_\lambda\}_{\lambda \in \Lambda}$ uma cobertura aberta de K. Mostre que existe $\epsilon > 0$ (dito um *número de Lebesgue* de $\mathscr{A}$), tal que, para todo $x \in K$, $B(x, \epsilon) \subset A_\lambda$ para algum $\lambda \in \Lambda$.

21. Use o Teorema dos Compactos Encaixados para provar o *Teorema de Baire*: A interseção de uma família enumerável de abertos densos de $\mathbb{R}^n$ é densa em $\mathbb{R}^n$. Conclua que a união de uma família enumerável de subconjuntos fechados e de interior vazio de $\mathbb{R}^n$ tem interior vazio.

Seção 5

22. Suponha que $X \subset \mathbb{R}^n$ seja conexo e que $X \subset Y \subset \overline{X}$. Prove que Y é conexo. Conclua, então, que:
    i) O fecho de um conjunto conexo é conexo;
    ii) as componentes conexas de um subconjunto X de $\mathbb{R}^n$ são fechadas em X.

23. Prove que o conjunto formado pelas matrizes invertíveis de ordem n, $\mathrm{I}(n)$, é desconexo.

24. Seja $X \subset \mathbb{R}^n$, tal que $X \neq \emptyset$. Mostre que $\partial X = \emptyset$ se, e somente se, $X = \mathbb{R}^n$.

25. Sejam $X, Y \subset \mathbb{R}^n$ conjuntos conexos. Prove que se $\partial X \subset Y$, então $X \cup Y$ é conexo.

26. Seja $\{X_k\}_{k \in \mathbb{N}}$ uma família enumerável de conjuntos conexos de $\mathbb{R}^n$, tal que, para todo $k \in \mathbb{N}$, $X_k \cap X_{k+1} \neq \emptyset$. Prove que $X = \bigcup X_k$ é conexo.

27. Prove o *Teorema da Alfândega*: Seja $C \subset \mathbb{R}^n$ um conjunto conexo e $X \subset \mathbb{R}^n$ um conjunto arbitrário. Se C contém pontos de X e de $\mathbb{R}^n - X$, então C contém algum ponto da fronteira de X.

Seção 6

28. Seja τ a família de subconjuntos de $\mathbb{R}^n$ formada pelo conjunto vazio e pelos complementares dos subconjuntos finitos de $\mathbb{R}^n$. Mostre que τ define em $\mathbb{R}^n$ uma topologia, dita *do complementar finito*, e que $(\mathbb{R}^n, \tau)$ é um espaço topológico não-metrizável por não ser de Hausdorff. Mostre também que, em $(\mathbb{R}^n, \tau)$, todo subespaço (topológico) é compacto e todo subespaço infinito é conexo.

29. Sejam $\mathbb{R}_+$ o conjunto dos números reais positivos, d a métrica euclidiana de $\mathbb{R}$ e d' a métrica de $\mathbb{R}_+$ definida por $d'(x,y) = d(1/x, 1/y)$ (vide Exercício 23 – Capítulo 1). Prove que a topologia de $\mathbb{R}_+$ proveniente de d coincide com aquela proveniente de d'. Em seguida, prove que a sequência $(1/k)_{k \in \mathbb{N}}$ é de Cauchy em $(\mathbb{R}_+, d)$, porém não é de Cauchy em $(\mathbb{R}_+, d')$.

30. Sejam X um subconjunto de $\mathbb{R}^n$ e $\varphi : \mathbb{R}^n \to \mathbb{R}^n$ um homeomorfismo. Prove que $\varphi|_X : X \to \varphi(X)$ é um homeomorfismo. Conclua, então, que são homeomorfos os seguintes subconjuntos de $\mathbb{R}^3$ (Fig. 16):
    i) A esfera $S = \{(x,y,z) \in \mathbb{R}^3;\ x^2 + y^2 + z^2 = 1\}$;
    ii) o elipsoide $E = \{(x,y,z) \in \mathbb{R}^3;\ ax^2 + by^2 + cz^2 = 1,\ a, b, c > 0\}$.

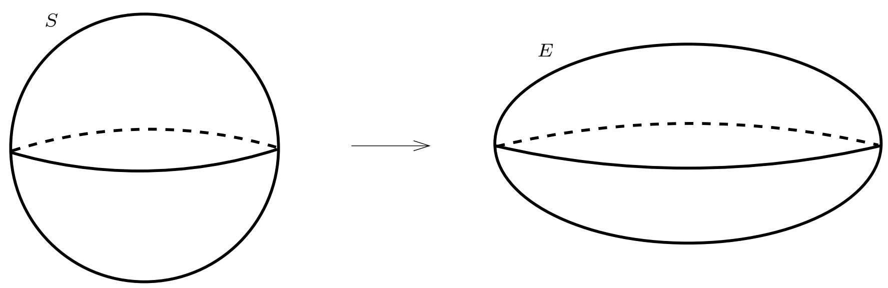

FIGURA 16

CAPíTULO 3

Aplicações Contínuas

Do ponto de vista geométrico, o conjunto dos números reais, $\mathbb{R}$, é o *continuum* formado pelos pontos de uma reta. Desta forma, ao considerarmos uma função f, de variável real x e que toma valores em $\mathbb{R}$, é natural nos perguntarmos se, quando x varia ao longo do *continuum*, isto é, passa de um ponto a outro sem "pular" nenhum, o mesmo acontece com $f(x)$. No caso afirmativo, a função f é dita contínua.

Mais precisamente, a continuidade de uma função $f : \mathbb{R} \to \mathbb{R}$ exprime-se através da propriedade — para todo $a \in \mathbb{R}$, pode-se ter $f(x)$ arbitrariamente próximo de $f(a)$, desde que se tome x suficientemente próximo de a.

Em seu *Cours d'analyse*, publicado em 1821, o matemático francês Augustin Louis Cauchy (1789–1857) estabeleceu parâmetros de organização e conceituação em Análise, os quais culminaram numa abordagem — proposta por Karl Weierstrass (1815–1897) — denominada *aritmetização da Análise*, segundo a qual, a fundamentação desta teoria reduz-se à fundamentação dos números reais. Como consequência, introduziram-se os célebres ϵ e δ, que, dito de forma simples, traduzem matematicamente as expressões "arbitrariamente próximo" e "suficientemente próximo" do parágrafo anterior.

Sob esta ótica, introduziremos neste capítulo os conceitos de continuidade e limite de aplicações entre espaços euclidianos. Constataremos, então, que estes conceitos são topológicos, fato que concede às aplicações contínuas, em Topologia, o mesmo *status* das aplicações lineares em Álgebra Linear, isto é, pode-se dizer que a Topologia é o estudo dos espaços topológicos e das aplicações contínuas entre os mesmos. Neste contexto, veremos que as propriedades topológicas de compacidade e conexidade são preservadas por aplicações contínuas e que homeomorfismos são bijeções contínuas cujas inversas são também contínuas. Concluiremos, então, fazendo uma breve apresentação do Teorema de Borsuk-Ulam, o qual constitui um belo resultado da Topologia envolvendo o conceito de continuidade.

1. Continuidade em $\mathbb{R}^n$

Dados conjuntos $X \subset \mathbb{R}^n$ e $Y \subset \mathbb{R}^n$, diz-se que uma aplicação $f : X \to Y$ é *contínua no ponto* $a \in X$ se, para todo $\epsilon > 0$, existe $\delta > 0$, tal que

$$x \in X \quad \text{e} \quad \|x - a\| < \delta \quad \Rightarrow \quad \|f(x) - f(a)\| < \epsilon\,. \tag{15}$$

Diz-se que f é *contínua* quando é contínua em cada um dos pontos de X.

Note que a implicação (15) é equivalente a cada uma das seguintes afirmações (Fig. 1):

- $x \in B(a, \delta) \cap X \quad \Rightarrow \quad f(x) \in B(f(a), \epsilon) \cap Y;$
- $f(B(a, \delta) \cap X) \ \subset \ B(f(a), \epsilon) \cap Y.$

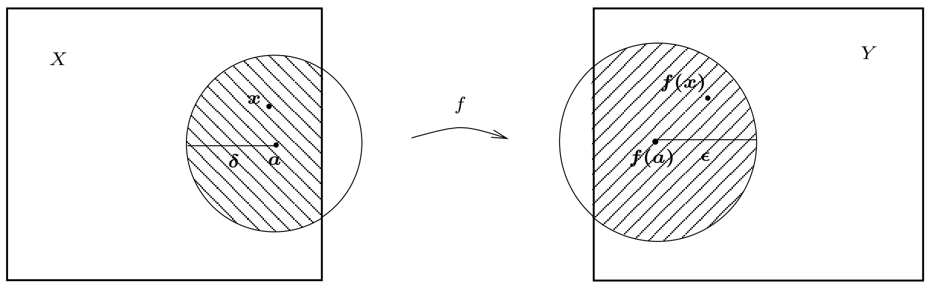

Figura 1

Dado $X \subset \mathbb{R}^n$, verifica-se facilmente que são contínuas:

- A aplicação identidade de X;
- qualquer aplicação constante $f : X \to Y \subset \mathbb{R}^m$;
- a restrição, a um subconjunto de X, de qualquer aplicação contínua $f : X \to Y \subset \mathbb{R}^m$.

Exemplo 32. Se $a \in X \subset \mathbb{R}^n$ é um ponto isolado de X, então qualquer aplicação $f : X \to Y \subset \mathbb{R}^m$ é contínua em a. Com efeito, sendo a isolado, existe $\delta > 0$, tal que $B(a, \delta) \cap X = \{a\}$. Logo, para todo $\epsilon > 0$,

$$f(B(a, \delta) \cap X) = f(\{a\}) = \{f(a)\} \subset B(f(a), \epsilon) \cap Y,$$

donde f é contínua em a. Em particular, se X for discreto, toda aplicação $f : X \to \mathbb{R}^m$ é contínua.

Observemos que o conceito de continuidade em $\mathbb{R}^n$, conforme introduzido, estende-se facilmente aos espaços métricos, em particular, aos espaços vetoriais normados. Também é fácil ver que a continuidade de uma aplicação entre espaços vetoriais normados não é afetada quando se substituem suas normas por outras que lhes sejam respectivamente equivalentes.

Exemplo 33 (normas). Toda norma $\| \ \|$ definida num espaço vetorial $\mathbb{V}$ é uma função contínua. De fato, fazendo-se $f(x) = \|x\|$, $x \in \mathbb{V}$, tem-se, para todo $a \in \mathbb{V}$, $|f(x) - f(a)| = |\,\|x\| - \|a\|\,| \leq \|x - a\|$. Logo, dado $\epsilon > 0$, fazendo-se $\delta = \epsilon$, obtém-se $\|x - a\| < \delta \Rightarrow |f(x) - f(a)| < \epsilon$, isto é, f é contínua.

Exemplo 34 (aplicações lineares). Dados $a \in \mathbb{R}^n$ e $T \in \mathrm{L}(\mathbb{R}^n, \mathbb{R}^m)$, $T \neq 0$, tem-se, para todo $x \in \mathbb{R}^n$, $\|Tx - Ta\| = \|T(x - a)\| \leq \|T\|\,\|x - a\|$. Assim, dado $\epsilon > 0$, tomando-se $\delta = \epsilon/\|T\|$, para todo $x \in \mathbb{R}^n$ satisfazendo $\|x - a\| < \delta$, tem-se $\|Tx - Ta\| < \epsilon$. Logo, T é contínua.

Uma aplicação $f : X \subset \mathbb{R}^n \to Y \subset \mathbb{R}^m$ que não é contínua em $a \in X$ é dita *descontínua* neste ponto.

Assim, uma função $f : X \subset \mathbb{R}^n \to Y \subset \mathbb{R}^m$ é descontínua em a se

$$\forall \delta > 0, \exists \epsilon > 0, x \in X, \text{ tais que } \|x - a\| < \delta \text{ e } \|f(x) - f(a)\| \geq \epsilon.$$

Exemplo 35. A função $f : \mathbb{R}^n \to \mathbb{R}$, definida por

$$f(x) = \begin{cases} \dfrac{1}{\|x\|} & \text{se } \ x \neq 0 \\ \\ 1 & \text{se } \ x = 0, \end{cases}$$

não é contínua em $x = 0$. Com efeito, dado $x \in \mathbb{R}^n$ satisfazendo $0 < \|x\| < 1/2$, temos que $\frac{1}{\|x\|} > 2$. Logo,

$$|f(x) - f(0)| = \left| \frac{1}{\|x\|} - 1 \right| > \frac{1}{\|x\|} - 1 > 1.$$

Desta forma, tomando-se $\epsilon = 1$, para todo $\delta > 0$, existe um $x \in \mathbb{R}^n$, tal que $\|x\| < \delta$ e $|f(x) - f(0)| > \epsilon$. Basta tomar $x \in \mathbb{R}^n - \{0\}$ satisfazendo $\|x\| < \min\{\delta, 1/2\}$.

Teorema 12 (caracterização topológica da continuidade). *Uma aplicação $f : X \subset \mathbb{R}^n \to Y \subset \mathbb{R}^m$ é contínua em $a \in X$ se, e somente se, para toda vizinhança V de $f(a)$ em Y, $f^{-1}(V)$ é uma vizinhança de a em X. Em particular, f é contínua se, e somente se, para todo aberto U de Y, $f^{-1}(U)$ é aberto em X.*

Demonstração. Suponhamos que f seja contínua em $a \in X$. Dada uma vizinhança V de $f(a)$ em Y, temos que existe $\epsilon > 0$, tal que $B(f(a), \epsilon) \cap Y \subset V$ (Corolário 1 – Capítulo 2) e, pela continuidade de f em a, para este ϵ, existe $\delta > 0$, tal que $f(B(a, \delta) \cap X) \subset B(f(a), \epsilon) \cap Y \subset V$. Logo, $B(a, \delta) \cap X \subset f^{-1}(V)$, donde $f^{-1}(V)$ é uma vizinhança de a em X.

Reciprocamente, suponhamos que, para toda vizinhança V de $f(a)$ em Y, $f^{-1}(V)$ seja uma vizinhança de a em X. Neste caso, dado $\epsilon > 0$, fazendo-se $V = B(f(a), \epsilon) \cap Y$, temos que $f^{-1}(B(f(a), \epsilon) \cap Y)$ é uma vizinhança de a em X. Desta forma, existe $\delta > 0$, tal que $B(a, \delta) \cap X \subset f^{-1}(B(f(a), \epsilon) \cap Y)$, isto é, $f(B(a, \delta) \cap X) \subset B(f(a), \epsilon) \cap Y$. Logo, f é contínua em a.

Observemos agora que todo conjunto $U \subset Y$, aberto em Y, é uma vizinhança de cada um de seus pontos. Logo, pelo estabelecido, $f : X \to Y$ será contínua se, e somente se, para cada tal U, o conjunto $f^{-1}(U)$, quando não-vazio, for uma vizinhança, em X, de todos os seus pontos, isto é, se, e somente se, $f^{-1}(U)$ for aberto em X. Isto prova a última asserção do enunciado do teorema e conclui a sua demonstração. $\square$

O Teorema 12 sugere a definição de aplicação contínua entre espaços topológicos, isto é, dados espaços topológicos X, Y, diz-se que uma aplicação $f : X \to Y$ é *contínua* num ponto $a \in X$ se, para toda vizinhança V de $f(a)$ em Y, $f^{-1}(V)$ é uma vizinhança de a em X.

Observação 11. No enunciado do Teorema 12, pode-se substituir "aberto" por "fechado", isto é, pode-se afirmar que: $f : X \subset \mathbb{R}^n \to Y \subset \mathbb{R}^m$ *é contínua se, e somente se, para todo fechado F de Y, $f^{-1}(F)$ é fechado em X.* Para provarmos isto, basta notarmos que, pelo Teorema 12, f é contínua se, e somente se, para todo subconjunto fechado F de Y, $f^{-1}(Y - F) \subset X$ é aberto em X. Porém, $f^{-1}(Y - F) = f^{-1}(Y) - f^{-1}(F) = X - f^{-1}(F)$, isto é, $f^{-1}(Y - F)$ é aberto em X se, e somente se, $f^{-1}(F)$ é fechado em X.

Segue-se das considerações do Exemplo 33 e da Observação 11 que toda esfera S de um espaço vetorial normado $(\mathbb{V}, \|\ \|)$, de raio $r > 0$ e com centro na origem, é fechada. De fato, a função $f(x) = \|x\|$, $x \in \mathbb{V}$, é contínua e $\{r\} \subset \mathbb{R}$ é fechado. Logo, $S = f^{-1}(\{r\})$ é fechado em $\mathbb{V}$.

Dados $a \in \mathbb{R}$ e uma função contínua, $f : X \subset \mathbb{R}^n \to \mathbb{R}$, façamos

$$U_1 = \{x \in X;\, f(x) > a\}, \quad U_2 = \{x \in X;\, f(x) < a\} \quad \text{e} \quad F = \{x \in X;\, f(x) = a\}.$$

Uma vez que $U_1 = f^{-1}(a, +\infty)$, $U_2 = f^{-1}(-\infty, a)$ e $F = f^{-1}(\{a\})$, segue-se do Teorema 12 que, relativamente a X, U_1 e U_2 são abertos, enquanto F é fechado.

Conforme vimos assinalando neste capítulo e no anterior, convergência e continuidade são conceitos essenciais da Topologia. É legítimo, portanto, indagarmos se há uma relação entre os mesmos e, no caso afirmativo, como ela se dá. A proposição seguinte elucida esta questão.

Proposição 26 (continuidade e convergência). *Dados conjuntos $X \subset \mathbb{R}^n$ e $Y \subset \mathbb{R}^m$, uma aplicação $f : X \to Y$ é contínua em $a \in X$ se, e somente se, para toda sequência $(x_k)_{k \in \mathbb{N}}$ em X satisfazendo $x_k \to a$, tem-se $f(x_k) \to f(a)$.*

Demonstração. Suponhamos que f seja contínua em a. Então, pelo Teorema 12, dada uma vizinhança V de $f(a)$ em Y, $f^{-1}(V)$ é uma vizinhança de a em X. Assim, pela Proposição 11 do Capítulo 2, se (x_k) é uma sequência em X que converge para a, existe $k_0 \in \mathbb{N}$, tal que $x_k \in f^{-1}(V)$ para todo $k \geq k_0$, o que implica $f(x_k) \in V$ para todo $k \geq k_0$. Logo, ainda pela Proposição 11 do Capítulo 2, $f(x_k) \to f(a)$.

Suponhamos agora, por contraposição, que f não seja contínua em a. Neste caso, existe uma vizinhança V de $f(a)$ em Y, tal que $f^{-1}(V)$ não é uma vizinhança de a em X. Então, para todo $k \in \mathbb{N}$, existe $x_k \in B(a, 1/k) \cap X$, tal que $x_k \notin f^{-1}(V)$. Logo, $x_k \to a$ e, para todo $k \in \mathbb{N}$, $f(x_k) \notin V$, donde $(f(x_k))$ não converge para $f(a)$. □

Observação 12. Pelas considerações do segundo parágrafo da demonstração acima, segue-se que $f : X \to Y$ será contínua em $a \in X$, desde que, para toda sequência (x_k) em X convergindo para a, exista uma subsequência de $(f(x_k))$ que convirja para $f(a)$. Com efeito, a sequência $(f(x_k))$, obtida a partir da hipótese de f não ser contínua em a, não possui subsequências que convirjam para $f(a)$.

Exemplo 36 (aplicações n-lineares). Segue-se da Proposição 9 do Capítulo 1 e da Proposição 26, acima, que toda aplicação n-linear entre espaços vetoriais de dimensão finita é contínua. Em particular são contínuas as seguintes aplicações:

- o produto escalar de $\mathbb{R}^n$, $\langle\, , \,\rangle : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$;
- a multiplicação de matrizes, $f(X, Y) = XY$, $X \in \mathrm{M}(m, n)$, $Y \in \mathrm{M}(n, p)$;
- o determinante $\det : \mathbb{R}^n \times \cdots \times \mathbb{R}^n \to \mathbb{R}$.

O resultado a seguir nos diz que as operações elementares de funções preservam continuidade, isto é, somas e produtos de aplicações contínuas resultam em aplicações contínuas. A demonstração é imediata a partir dos resultados da Proposição 26 e da Proposição 6 do Capítulo 1.

Proposição 27 (propriedades operatórias). *Sejam $f, g : X \subset \mathbb{R}^n \to \mathbb{R}^m$ aplicações contínuas e $\lambda : X \to \mathbb{R}$ uma função contínua. Então, são contínuas as aplicações:*

i) $(f + g) : X \to \mathbb{R}^m$, *em que* $(f + g)(x) = f(x) + g(x)$;

ii) $(\lambda f) : X \to \mathbb{R}^m$, *em que* $(\lambda f)(x) = \lambda(x) f(x)$;

iii) $(\frac{1}{\lambda}) : X \to \mathbb{R}$, *em que* $(\frac{1}{\lambda})(x) = \frac{1}{\lambda(x)}$ *e* $\lambda(x) \neq 0\ \forall x \in X$.

Observação 13. O resultado do item (iii) da proposição acima é uma consequência do fato, de fácil constatação, de que se (x_k) é uma sequência de termos não-nulos em $\mathbb{R}$ que converge para um real $a \neq 0$, então $1/x_k \to 1/a$.

Dado $X \subset \mathbb{R}^n$, denotemos por $\mathrm{F}(X, \mathbb{R}^m)$ o conjunto formado por todas as aplicações $f : X \to \mathbb{R}^m$. Analogamente a $\mathrm{L}(\mathbb{R}^n, \mathbb{R}^m)$, $\mathrm{F}(X, \mathbb{R}^m)$ é um espaço vetorial quando se consideram aí as operações usuais de soma de funções e produto de escalar por função. Segue-se, então, dos itens (i) e (ii) da Proposição 27, que o conjunto

$$\mathrm{C}(X, \mathbb{R}^n) = \{f : X \to \mathbb{R}^m;\ f \text{ é contínua}\}$$

é um subespaço vetorial de $\mathrm{F}(X, \mathbb{R}^m)$.

Consideremos uma aplicação $f : X \subset \mathbb{R}^n \to Y \subset \mathbb{R}^m \in \mathrm{F}(X, Y)$. Temos, então, que a cada $x \in X$ está associado um único $y = (y_1, \dots, y_m) \in Y$. As m funções $f_i : X \to \mathbb{R}$, definidas por $f_i(x) = y_i$, são chamadas de *coordenadas* de f. Desta forma, tem-se

$$f(x) \,=\, (\, f_1(x), \dots, f_m(x)\,),\ \ x \in X.$$

O resultado abaixo segue-se diretamente da Proposição 26 e da Proposição 5 do Capítulo 1.

Proposição 28 (continuidade das coordenadas de uma aplicação). *Uma aplicação $f : X \subset \mathbb{R}^n \to Y \subset \mathbb{R}^m$ é contínua em $a \in X$ se, e somente se, cada uma de suas coordenadas é contínua em a.*

Analogamente, dados espaços vetoriais normados $\mathbb{V}, \mathbb{W}_1, \dots, \mathbb{W}_m$, cada aplicação $f : X \subset \mathbb{V} \to \mathbb{W}_1 \times \cdots \times \mathbb{W}_m$ determina m aplicações $f_i : X \to \mathbb{W}_i$, $i = 1, \dots, m$, em que $f(x) = (y_1, \dots, y_m)$ se, e somente se, $f_i(x) = y_i$. Como se pode verificar facilmente, vale, neste contexto, o mesmo resultado da Proposição 28, isto é, f é contínua em $a \in X$ se, e somente se, cada uma de suas coordenadas, f_i, é contínua em a.

Exemplo 37 (aplicações com coordenadas polinomiais). Decorre das proposições 27 e 28 que é contínua toda aplicação $f : \mathbb{R}^n \to \mathbb{R}^m$ cujas coordenadas são polinômios. Para vermos isto, basta observarmos que, para cada $k = 1, \dots, n$, a projeção ortogonal $P_k(x_1, \dots, x_n) = x_k$, por ser linear, é contínua. Ora, sendo cada coordenada f_i de f um polinômio, temos, para cada $i = 1, \dots, m$, que f_i é dada por uma soma de produtos de funções do tipo $\mu_k (P_k)^q$, em que $\mu_k \in \mathbb{R}$ e $q \in \mathbb{N} \cup \{0\}$. Uma vez que somas e produtos de funções contínuas são contínuas, temos que cada coordenada de f é contínua, o que implica, pela Proposição 28, que f é contínua.

Por exemplo, a aplicação $f : \mathbb{R}^2 \to \mathbb{R}^3$, definida por

$$f(x_1, x_2) = (x_1 x_2, 2x_1^2 - 1, x_2^3 - 5x_1),$$

é tal que

$$f_1 = P_1 P_2\,, \quad f_2 = 2(P_1)^2 - (P_2)^0 \quad \text{e} \quad f_3 = (P_2)^3 - 5P_1\,.$$

Consideremos agora as funções $f, g : \mathbb{R}^2 \to \mathbb{R}$, definidas por

$$f(x,y) = \begin{cases} \frac{x^2 y}{x^2 + y^2} & \text{se} \quad (x,y) \neq 0 \\ 0 & \text{se} \quad (x,y) = 0 \end{cases} \qquad \text{e} \qquad g(x,y) = \begin{cases} \frac{xy}{x^2 + y^2} & \text{se} \quad (x,y) \neq 0 \\ 0 & \text{se} \quad (x,y) = 0. \end{cases}$$

Em $\mathbb{R}^2 - \{(0,0)\}$, f e g são dadas por quocientes de polinômios. Logo, ambas estas funções são contínuas em $\mathbb{R}^2 - \{(0,0)\}$. Verifiquemos agora que f é contínua em $(0,0)$, enquanto g é descontínua neste ponto. Com efeito, dado $\epsilon > 0$, temos, para todo $(x,y) \in \mathbb{R}^2 - (0,0)$ satisfazendo $\|(x,y)\| < \epsilon$, que

$$|f(x,y) - f(0,0)| = \left|\frac{x^2y}{x^2+y^2}\right| = \frac{x^2}{x^2+y^2}|y| \leq |y| \leq \|(x,y)\| < \epsilon,$$

donde f é contínua em $(0,0)$.

Por outro lado, tomando-se a sequência $z_k = (1/k, 1/k)$ em $\mathbb{R}^2$, temos que $z_k \to (0,0)$. Porém,

$$g(z_k) = \frac{x_k y_k}{x_k^2 + y_k^2} = \frac{1/k^2}{2/k^2} = \frac{1}{2}\,.$$

Assim, $g(z_k) \to 1/2 \neq g(0,0)$. Segue-se, então, da Proposição 26, que g não é contínua em $(0,0)$.

Vejamos agora que a composição de aplicações preserva continuidade, conforme a proposição a seguir.

Proposição 29 (continuidade e composição). *Suponha que as aplicações $f : X \subset \mathbb{R}^n \to Y \subset \mathbb{R}^m$ e $g : Y \to Z \subset \mathbb{R}^p$ sejam tais que f é contínua em $a \in X$ e g é contínua em $b = f(a) \in Y$. Então, a aplicação composta $g \circ f : X \to \mathbb{R}^p$ é contínua em a. Em particular, a composta de aplicações contínuas é contínua.*

Demonstração. Sejam (x_k) uma sequência em X que converge para a e (y_k) a sequência em Y, tal que $y_k = f(x_k)$. Como f é contínua em a, temos que $y_k = f(x_k) \to f(a) = b$. Agora, pela continuidade de g em b, $g(y_k) \to g(b)$, isto é, $(g \circ f)(x_k) \to (g \circ f)(a)$. Logo, $g \circ f$ é contínua em a. □

Sejam $f, g : X \subset \mathbb{R}^n \to Y \subset \mathbb{R}^m$ aplicações contínuas em $a \in X$. Então, a aplicação $x \mapsto (f(x), g(x))\,,\ x \in X$, é contínua em a, pois suas coordenadas o são. Compondo-a com o produto interno de $\mathbb{R}^m$, que é contínuo, obtemos a aplicação $x \mapsto \langle f(x), g(x)\rangle$ que, então, pela Proposição 29, é contínua em a.

Exemplo 38. Consideremos a aplicação

$$\begin{array}{rccc} f: & \mathrm{M}(n) & \to & \mathrm{M}(n) \\ & X & \mapsto & XX^*. \end{array}$$

Então, $f = \psi \circ \varphi$, em que

$$\begin{array}{rccc} \varphi: & \mathrm{M}(n) & \to & \mathrm{M}(n) \times \mathrm{M}(n) \\ & X & \mapsto & (X, X^*) \end{array} \quad \text{e} \quad \begin{array}{rccc} \psi: & \mathrm{M}(n) \times \mathrm{M}(n) & \to & \mathrm{M}(n) \\ & (X,Y) & \mapsto & XY. \end{array}$$

As coordenadas de φ são a aplicação identidade de $\mathrm{M}(n)$, $X \mapsto X$, e a aplicação $X \mapsto X^*$. Ambas são lineares e, portanto, contínuas. Logo, φ é contínua. O mesmo vale para a aplicação ψ, pois esta é bilinear. Segue-se, então, da Proposição 29, que f é contínua.

Deste exemplo, conclui-se imediatamente que o conjunto $\mathrm{O}(n)$, das matrizes ortogonais de $\mathrm{M}(n)$, é fechado em $\mathrm{M}(n)$ (fato já provado no Capítulo 2). Com efeito, $\mathrm{O}(n) = f^{-1}(\{I\})$, em que I denota a matriz identidade de $\mathrm{M}(n)$.

2. Continuidade Uniforme

O conceito de continuidade uniforme, que introduziremos a seguir, surge naturalmente a partir da observação de que, na definição de continuidade de uma aplicação f num ponto a de seu domínio, o δ que existe para cada ϵ dado, em geral, depende não só de ϵ, mas também do ponto a, conforme ilustrado no exemplo seguinte.

Consideremos a função $f : \mathbb{R}^n \to \mathbb{R}$, definida por $f(x) = \|x\|^2$, e provemos, usando a definição, que f é contínua em todo ponto $a \in \mathbb{R}^n$. Para isto, observemos inicialmente que, dado $x \in \mathbb{R}^n$, tem-se

$$|f(x) - f(a)| \,=\, |\,\|x\|^2 - \|a\|^2| \,=\, |\,\|x\| - \|a\|\,|\,(\|x\| + \|a\|) \,\le\, \|x - a\|\,(\|x\| + \|a\|).$$

Fazendo-se $\|x\| = \|x - a + a\|$ e usando-se a desigualdade triangular, obtém-se

$$|f(x) - f(a)| \,\le\, \|x - a\|\,(\|x - a\| + 2\|a\|). \tag{16}$$

Agora, dado $\epsilon > 0$, procuremos encontrar um $\delta > 0$, tal que $\|x - a\| < \delta$ implique $|f(x) - f(a)| < \epsilon$. Para tanto, pela desigualdade (16), é suficiente que δ seja positivo e satisfaça $\delta(\delta + 2\|a\|) \le \epsilon$, isto é,

$$\delta^2 + 2\|a\|\delta - \epsilon \le 0. \tag{17}$$

Basta, então, tomarmos $\delta = \sqrt{\epsilon + \|a\|^2} - \|a\|$ (veja que a desigualdade (17) é uma inequação quadrática de variável δ).

Notemos que, neste exemplo, o δ encontrado depende, de fato, do ponto a e de ϵ. Isto contrasta com o comportamento da função norma, considerada no Exemplo 33. Lá, dados a e ϵ, o δ encontrado depende apenas de ϵ, não de a. Neste caso, que é especial, diz-se que esta função é uniformemente contínua.

Definição 18 (continuidade uniforme). Dados conjuntos $X \subset \mathbb{R}^n$ e $Y \subset \mathbb{R}^m$, diz-se que uma aplicação $f : X \to Y$ é *uniformemente contínua* quando, para todo $\epsilon > 0$, existe $\delta > 0$, tal que

$$x, y \in X \quad \text{e} \quad \|x - y\| < \delta \quad \Rightarrow \quad \|f(x) - f(y)\| < \epsilon.$$

Na definição seguinte, introduziremos uma ampla classe de aplicações uniformemente contínuas, a qual inclui as aplicações lineares.

Definição 19 (aplicações lipschitzianas). Diz-se que uma aplicação $f : X \subset \mathbb{R}^n \to Y \subset \mathbb{R}^m$ é *lipschitziana* quando existe $\lambda > 0$ (dito uma *constante de Lipschitz* de f), tal que, para quaisquer $x, y \in X$, tem-se

$$\|f(x) - f(y)\| \,\le\, \lambda\|x - y\|.$$

Toda aplicação lipschitziana $f : X \subset \mathbb{R}^n \to Y \subset \mathbb{R}^m$ é uniformemente contínua, pois, dado $\epsilon > 0$, tomando-se $\delta = \epsilon/\lambda$, em que λ é uma constante de Lipschitz de f, tem-se

$$x, y \in X \text{ e } \|x - y\| < \delta \Rightarrow \|f(x) - f(y)\| \,\le\, \lambda\|x - y\| \,<\, \lambda\delta \,=\, \epsilon.$$

Exemplos simples de aplicações lipschitzianas são as translações em $\mathbb{R}^n$, isto é, as aplicações $f : \mathbb{R}^n \to \mathbb{R}^n$, tais que $f(x) = x + a$, $a \in \mathbb{R}^n$.

Uma aplicação lipschitziana com uma constante de Lipschitz $\lambda < 1$ (respectivamente, $\lambda > 1$) é dita uma *contração* (respectivamente, *dilatação*).

Exemplo 39. Para todo $r > 0$, a função

$$f(x) = \sqrt{\|x\|}, \quad x \in \mathbb{R}^n - B(0,r),$$

é lipschitziana e, portanto, uniformemente contínua. Com efeito, para quaisquer $x, y \in \mathbb{R}^n - B(0,r)$, tem-se

$$|f(x) - f(y)| = |\sqrt{\|x\|} - \sqrt{\|y\|}\,| = \frac{|\,\|x\| - \|y\|\,|}{\sqrt{\|x\|} + \sqrt{\|y\|}} \le \frac{1}{2\sqrt{r}}\|x - y\|. \tag{18}$$

Exemplo 40 (função distância a um conjunto). Define-se a *distância* de um ponto $x \in \mathbb{R}^n$ a um conjunto $A \subset \mathbb{R}^n$, $d(x, A)$, por

$$d(x, A) = \inf\{\|x - a\|;\, a \in A\}. \tag{19}$$

Tomando-se uma sequência (a_k), em A, satisfazendo

$$\|x - a_k\| \to d(x, A),$$

e levando-se em conta que $\|a_k\| \le \|a_k - x\| + \|x\|$, concluímos que (a_k) é limitada e possui, desta forma, uma subsequência convergente, (a_{k_i}). Logo, pela continuidade da norma,

$$d(x, A) = \lim_{k\to\infty} \|x - a_k\| = \lim_{i\to\infty} \|x - a_{k_i}\| = \|x - a\|,$$

em que $a = \lim_{i\to\infty} a_{k_i} \in \overline{A}$.

Destas considerações, segue-se que se $F \subset \mathbb{R}^n$ é um conjunto fechado, então:

- Para todo $x \in \mathbb{R}^n$, existe $a \in F$, tal que $d(x, F) = \|x - a\|$;
- $d(x, F) = 0$ se, e somente se, $x \in F$.

Dados, então, $x, y \in \mathbb{R}^n$, para quaisquer pontos $a, b \in F$ satisfazendo $d(x, F) = \|x - a\|$ e $d(y, F) = \|y - b\|$, tem-se $\|x - a\| \le \|x - b\|$ e $\|y - b\| \le \|y - a\|$. Logo,

$$-\|x - y\| \le \|x - a\| - \|y - a\| \le \|x - a\| - \|y - b\| \le \|x - b\| - \|y - b\| \le \|x - y\|,$$

isto é,

$$|d(x, F) - d(y, F)| = |\,\|x - a\| - \|y - b\|\,| \le \|x - y\|.$$

Desta forma, a *função distância* ao conjunto fechado $F \subset \mathbb{R}^n$,

$$\begin{array}{rccc} d_F : & \mathbb{R}^n & \to & \mathbb{R} \\ & x & \mapsto & d(x, F), \end{array}$$

é lipschtziana. Em particular, d_F é uniformemente contínua.

Note que, se $\mathbb{V} \subset \mathbb{R}^n$ é um subespaço vetorial de $\mathbb{R}^n$, pela desigualdade (5) do Capítulo 1, tem-se

$$d_{\mathbb{V}}(x) = \|x - P_{\mathbb{V}}(x)\| \ \forall x \in \mathbb{R}^n.$$

É imediato que toda aplicação uniformemente contínua é contínua. Não vale, entretanto, a recíproca, isto é, existem aplicações contínuas que não são uniformemente contínuas. Um exemplo é a função que consideramos no início desta seção, $f(x) = \|x\|^2$, $x \in \mathbb{R}^n$.

Quando discutimos a continuidade desta função num ponto $a \in \mathbb{R}^n$, o δ obtido para cada ϵ dado dependia de a e ϵ. No entanto, isto não prova que f não é uniformemente contínua pois, naquele argumento, nada garante que não haja um outro $\delta > 0$ que satisfaça as condições da definição de continuidade e que dependa apenas de ϵ.

Para provarmos que $f(x) = \|x\|^2$ não é uniformemente contínua, devemos obter um $\epsilon > 0$, tal que, para todo $\delta > 0$, existam $x, y \in \mathbb{R}^n$ satisfazendo $\|x - y\| < \delta$ e $\|f(x) - f(y)\| \geq \epsilon$.

Tomemos $\epsilon = 2$. Dado, então, $\delta > 0$, consideremos $\lambda > 0$, tal que $1/\lambda < \delta$. Fazendo-se $x = \lambda u$, em que $u \in \mathbb{R}^n$ é um vetor unitário, e $y = x + \frac{1}{\lambda}u = (\lambda + 1/\lambda)u$, tem-se $\|x - y\| = \frac{1}{\lambda} < \delta$. Além disso,

$$|f(x) - f(y)| = |\,\|x\|^2 - \|y\|^2| = \left(\lambda + \frac{1}{\lambda}\right)^2 - \lambda^2 = 2 + \frac{1}{\lambda^2} > 2 = \epsilon,$$

donde se infere que f não é uniformemente contínua.

Analogamente à continuidade, a continuidade uniforme pode ser caracterizada por meio de convergência de sequências, conforme a proposição seguinte.

Proposição 30 (continuidade uniforme e convergência). *Dados conjuntos $X \subset \mathbb{R}^n$ e $Y \subset \mathbb{R}^m$, uma aplicação $f : X \to Y$ é uniformemente contínua se, e somente se, dadas sequências (x_k) e (y_k) em X satisfazendo $\|x_k - y_k\| \to 0$, tem-se $\|f(x_k) - f(y_k)\| \to 0$.*

Demonstração. Suponhamos que f seja uniformemente contínua e tomemos sequências (x_k) e (y_k) em X, tais que $\|x_k - y_k\| \to 0$. Da continuidade uniforme de f, segue-se que, dado $\epsilon > 0$, existe $\delta > 0$, tal que $\|x_k - y_k\| < \delta \Rightarrow \|f(x_k) - f(y_k)\| < \epsilon$. Porém, para este δ, existe $k_0 \in \mathbb{N}$, tal que $k \geq k_0 \Rightarrow \|x_k - y_k\| < \delta$, pois $\|x_k - y_k\| \to 0$. Logo, $k \geq k_0 \Rightarrow \|f(x_k) - f(y_k)\| < \epsilon$, donde $\|f(x_k) - f(y_k)\| \to 0$.

Suponhamos agora, por contraposição, que f não seja uniformemente contínua. Então, existe $\epsilon > 0$, tal que, para todo $k \in \mathbb{N}$, existem $x_k, y_k \in X$ satisfazendo $\|x_k - y_k\| < \frac{1}{k}$ e $\|f(x_k) - f(y_k)\| \geq \epsilon$. Neste caso, $\|x_k - y_k\| \to 0$ e $\|f(x_k) - f(y_k)\|$ não converge para 0. □

Corolário 2. *Toda aplicação uniformemente contínua leva sequências de Cauchy em sequências de Cauchy.*

Demonstração. Sejam $f : X \subset \mathbb{R}^n \to Y \subset \mathbb{R}^m$ uma aplicação uniformemente contínua e (x_k) uma sequência de Cauchy em X. Então, para todo $p \in \mathbb{N}$, $\|x_{k+p} - x_k\| \to 0$, donde, pela continuidade uniforme de f, se obtém $\|f(x_{k+p}) - f(x_k)\| \to 0$, isto é, $(f(x_k))$ é de Cauchy. □

Exemplo 41 (projeção radial). Consideremos a aplicação *projeção radial* de $\mathbb{R}^n - \{0\}$ sobre a esfera unitária S^{n-1},

$$f(x) = \frac{x}{\|x\|}, \quad x \in \mathbb{R}^n - \{0\}.$$

É fácil ver que f é contínua. Agora, tomando-se vetores distintos $u, v \in S^{n-1}$, temos que as sequências $x_k = \frac{1}{k}u$ e $y_k = \frac{1}{k}v$, em $\mathbb{R}^n$, são tais que

$$x_k - y_k = \frac{1}{k}(u - v) \to 0 \quad \text{e} \quad f(x_k) - f(y_k) = u - v \neq 0.$$

Logo, pela Proposição 30, f não é uniformemente contínua. Além disso, fazendo-se $z_k = ((-1)^k/k)u$, tem-se $f(z_k) = (-1)^k u$. Claramente, (z_k) é convergente e $(f(z_k))$ é divergente. Em particular, (z_k) é de Cauchy e $(f(z_k))$ não o é. Isto mostra que, no Corolário 2, a hipótese de continuidade uniforme não pode ser substituída pela de continuidade.

Por outro lado, para todo $r > 0$ e $x, y \in \mathbb{R}^n - B(0,r)$, tem-se

$$\begin{aligned}
\|f(x) - f(y)\| &= \left| \frac{x}{\|x\|} - \frac{y}{\|y\|} \right| = \frac{1}{\|x\|\,\|y\|} \|\, \|y\|x - \|x\|y \,\| \\
&= \frac{1}{\|x\|\,\|y\|} \|\, (\|y\| - \|x\|)x + \|x\|(x-y) \,\| \\
&\leq \frac{1}{\|x\|\,\|y\|} (\, \|x-y\|\,\|x\| + \|x\|\,\|x-y\|) \\
&= \frac{2}{\|y\|} \|x-y\| \,\leq\, 2r\|x-y\|,
\end{aligned}$$

donde $f|_{\mathbb{R}^n - B(0,r)}$ é lipschtziana e, portanto, uniformemente contínua.

Todas as aplicações uniformemente contínuas que exibimos são lipschtzianas, o que pode causar a falsa impressão de que continuidade uniforme é uma condição equivalente à de ser lipschtziana. A fim de evitar este mal-entendido, na Seção 4, daremos um exemplo de uma função uniformemente contínua e não-lipschtziana.

Seguem-se diretamente da Proposição 30 e das propriedades das sequências convergentes os dois resultados abaixo, cujas demonstrações deixamos a cargo do leitor.

Proposição 31. *A composta de aplicações uniformemente contínuas é uniformemente contínua.*

Proposição 32. *Uma aplicação é uniformemente contínua se, e somente se, suas coordenadas o são.*

Convém observar que o conceito de continuidade é local, isto é, ele é considerado num ponto, enquanto o conceito de continuidade uniforme é global, pois diz respeito ao comportamento da função como um todo. Não faz sentido, portanto, falar-se da continuidade uniforme de uma função num ponto de seu domínio.

Vale também mencionar que a noção de continuidade uniforme, conforme introduzida, estende-se facilmente aos espaços métricos. Para isto, basta substituir, na conceituação, a norma pela métrica. No entanto, ela não se estende dos espaços métricos aos espaços topológicos, pois existem espaços métricos distintos, (M,d) e (M,d'), cujas métricas determinam a mesma topologia em M, bem como aplicações $f : (M,d) \to (M,d)$, $g : (M,d') \to (M,d)$, tais que $f(x) = g(x)\, \forall x \in M$, f não é uniformemente contínua e g é uniformemente contínua. Com efeito, conforme constatamos no Exercício 29 do Capítulo 2, a topologia de $\mathbb{R}_+$ proveniente da métrica $d'(x,y) = d(1/x, 1/y)$ coincide com a topologia proveniente da métrica euclidiana d. Porém, como se verifica facilmente, a função $f : (\mathbb{R}_+\,, d) \to (\mathbb{R}_+\,, d)$, $f(x) = 1/x$, não é uniformemente contínua, enquanto a função $g : (\mathbb{R}_+\,, d') \to (\mathbb{R}_+\,, d)$, $g(x) = 1/x$, é uma isometria e, portanto, é uniformemente contínua.

3. Homeomorfismos

Relembremos que, conforme discutido na Seção 6 do capítulo anterior, as aplicações que preservam a estrutura de espaço topológico são chamadas de homeomorfismos. Mais precisamente, uma aplicação bijetiva entre espaços topológicos X e Y, $f: X \to Y$, é dita um *homeomorfismo* quando ambas, f e f^{-1}, são abertas.

Segue-se, então, do Teorema 12, o resultado seguinte.

TEOREMA 13. *Dados $X \subset \mathbb{R}^n$ e $Y \subset \mathbb{R}^m$, uma bijeção $f: X \to Y$ é um homeomorfismo se, e somente se, f e f^{-1} são contínuas.*

Este teorema, dentre outras características, nos permite provar que são homeomorfos muitos conjuntos que, por intuição geométrica, já os sabemos serem. Isto será ilustrado nos exemplos a seguir.

EXEMPLO 42 (BOLAS). Sejam B_0 a bola unitária de $\mathbb{R}^n$ com centro na origem e $B = B(a, r)$ uma bola aberta qualquer de $\mathbb{R}^n$. A aplicação

$$\begin{array}{rccc} f: & B_0 & \to & B \\ & x & \mapsto & rx + a \end{array}$$

é, claramente, bijetiva e contínua. Sua inversa, $f^{-1}: B \to B_0$, é dada por $f^{-1}(y) = \frac{1}{r}(y - a)$, donde se vê que f^{-1} é contínua. Logo, f é um um homeomorfismo.

Note que, geometricamente, a função f corresponde a uma translação que leva a origem ao ponto a, seguida de uma dilatação (no caso de ser $r > 1$) ou contração (no caso de ser $r < 1$).

Observe também que, pela transitividade da relação de homeomorfismo, podemos concluir daí que duas bolas abertas quaisquer de $\mathbb{R}^n$ são homeomorfas. Um argumento análogo prova que vale o mesmo para duas bolas, ambas, fechadas.

EXEMPLO 43 (BOLAS ABERTAS E $\mathbb{R}^n$). Consideremos novamente a bola aberta unitária de $\mathbb{R}^n$ com centro na origem, $B_0 = B(0, 1)$, e verifiquemos que a aplicação

$$\begin{array}{rccc} f: & B_0 & \to & \mathbb{R}^n \\ & x & \mapsto & \frac{x}{1-\|x\|} \end{array}$$

é um homeomorfismo entre B_0 e $\mathbb{R}^n$ (geometricamente, f consiste numa deformação da bola aberta B_0 em $\mathbb{R}^n$ ao longo dos seus raios). Para isto, tomemos $y \in \mathbb{R}^n$ e provemos que a equação (vetorial) $f(x) = y$ admite uma única solução $x \in B_0$, o que implicará a bijetividade de f. Devemos, então, determinar $x \in B_0$, tal que $x/(1 - \|x\|) = y$. Daí, deve-se ter $\|x\| = \|y\|(1 - \|x\|)$, donde $\|x\| = \|y\|/(1 + \|y\|)$, resultando em $1 - \|x\| = 1/(1 + \|y\|)$. Logo, $x = y/(1 + \|y\|)$ é a solução procurada. Desta forma, f é invertível e

$$f^{-1}(y) = \frac{y}{1 + \|y\|}, \quad y \in \mathbb{R}^n.$$

Uma vez que f e f^{-1} são, claramente, contínuas, temos que f é um homeomorfismo. Segue-se daí e do exemplo anterior que qualquer bola aberta de $\mathbb{R}^n$ é homeomorfa a $\mathbb{R}^n$.

EXEMPLO 44 (BOLAS E CUBOS). Seja $B_{\max}[0, 1] \subset \mathbb{R}^n$ a bola fechada de $\mathbb{R}^n$ com respeito à norma do máximo, de raio 1 e com centro na origem, isto é,

$$B_{\max}[0, 1] = \{x \in \mathbb{R}^n;\ \|x\|_{\max} \leq 1\}.$$

Note que $B_{\max}[0,1]$ nada mais é que o *cubo n-dimensional*, obtido pelo produto cartesiano de n cópias do intervalo $[-1,1]$.

Consideremos, então, a aplicação

$$\begin{array}{rccc} f: & B[0,1] & \to & B_{\max}[0,1] \\ & x & \mapsto & \lambda(x)\,x, \end{array}$$

em que $\lambda : B[0,1] \to \mathbb{R}$ é tal que

$$\lambda(x) = \begin{cases} \|x\|/\|x\|_{\max} & x \neq 0 \\ \\ 0 & x = 0. \end{cases}$$

Assim, temos que

$$\|f(x)\|_{\max} = \lambda(x)\|x\|_{\max} = \|x\|, \tag{20}$$

donde $\|f(x)\|_{\max} \leq 1$ se, e somente se, $\|x\| \leq 1$, isto é, f está bem definida.

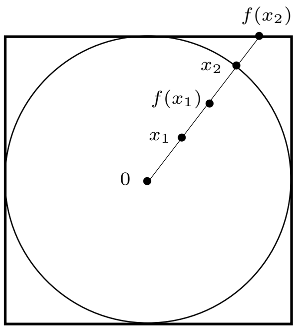

Figura 2

É evidente que f é contínua em $B[0,1]-\{0\}$. Além disso, para toda sequência (x_k) em $B[0,1]$ satisfazendo $x_k \to 0$, por (20), tem-se $f(x_k) \to 0$, donde se conclui que f é contínua, também, em $x = 0$.

Procedendo-se como no exemplo anterior, conclui-se facilmente que f é bijetiva e

$$f^{-1}(y) = \begin{cases} (\frac{\|y\|_{\max}}{\|y\|})y & y \neq 0 \\ \\ 0 & y = 0\,, \end{cases}$$

donde f^{-1} é contínua. Logo, f é um homeomorfismo.

Observe que, denotando-se a bola aberta unitária e a esfera unitária de $\mathbb{R}^n$ com respeito à norma do máximo, ambas com centro na origem, respectivamente por $B_{\max}(0,1)$ e $S_{\max}[0,1]$, tem-se

- $f(B(0,1)) = B_{\max}(0,1)$;
- $f(S[0,1]) = S_{\max}[0,1]$.

Logo, as restrições

$$f|_{S[0,1]} : S[0,1] \to S_{\max}[0,1] \quad \text{e} \quad f|_{B(0,1)} : B(0,1) \to B_{\max}(0,1)$$

estão bem definidas e são homeomorfismos (Fig. 2).

Exemplo 45 (gráficos de aplicações contínuas e seus domínios). Dada uma aplicação $f : X \subset \mathbb{R}^n \to Y \subset \mathbb{R}^m$, relembremos que seu *gráfico* é o conjunto

$$\operatorname{graf}(f) = \{(x, f(x));\, x \in X\} \subset X \times Y \subset \mathbb{R}^n \times \mathbb{R}^m.$$

Vejamos que, quando f é contínua, o gráfico de f é homeomorfo a X. De fato, a aplicação $\varphi : X \to \operatorname{graf}(f) \subset \mathbb{R}^n \times \mathbb{R}^m$, dada por $\varphi(x) = (x, f(x))$, é contínua, pois suas coordenadas o são, e bijetiva. Uma vez que $\varphi^{-1} = P|_{\operatorname{graf}(f)}$, em que $P : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^n$ é a projeção ortogonal de $\mathbb{R}^n \times \mathbb{R}^m$ sobre $\mathbb{R}^n$, e P é contínua, por ser linear, temos que φ^{-1} é contínua. Logo, φ é um homeomorfismo.

Em particular, o *paraboloide de revolução*, gráfico da função

$$f(x, y) = x^2 + y^2, \ (x, y) \in \mathbb{R}^2,$$

é homeomorfo a $\mathbb{R}^2$ (Fig. 3).

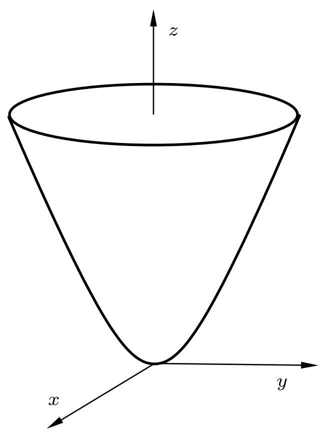

Figura 3

Exemplo 46 (projeção estereográfica). Consideremos a decomposição $\mathbb{R}^{n+1} = \mathbb{R}^n \times \mathbb{R}$ e denotemos um ponto $p \in \mathbb{R}^{n+1}$ por $p = (x, s)$, em que $x \in \mathbb{R}^n$ e $s \in \mathbb{R}$. Tomemos a esfera unitária S^n de $\mathbb{R}^{n+1}$ juntamente com seu "polo-norte", $p_0 = (0, 1)$, e identifiquemos o subespaço $\mathbb{V} = \mathbb{R}^n \times \{0\} = \{(x, s) \in \mathbb{R}^{n+1};\, s = 0\}$ com $\mathbb{R}^n$. A *projeção estereográfica* é a aplicação $f : S^n - \{p_0\} \to \mathbb{R}^n$, definida geometricamente da seguinte forma: Dado $p \in S^n - \{p_0\}$, $f(p)$ é o ponto de interseção entre a reta ℓ, determinada por p e p_0, e $\mathbb{V}$ (Fig. 4). Provemos, então, que f é um homeomorfismo.

Com este fim, façamos $p = (x, s)$ e observemos que um ponto $q \in \mathbb{R}^{n+1}$ pertence à reta ℓ se, e somente se, $q = (p - p_0)t + p_0$ para algum $t \in \mathbb{R}$, isto é, se, e somente se, $q = (tx, t(s-1) + 1)$. Para que q pertença também à $\mathbb{V}$, devemos ter $t(s-1) + 1 = 0$, isto é, $t = 1/(1-s)$. Logo,

$$f(x, s) = \left(\frac{x}{1-s}, 0\right), \ (x, s) \in S^n.$$

Dado $(a, 0) \in \mathbb{V}$, verifiquemos que a equação $f(x, s) = (a, 0)$ admite uma única solução, provando, desta forma, que f é bijetiva (o que, geometricamente, é evidente). Ora, um ponto (x, s) é solução se, e só se, $\frac{x}{1-s} = a$ e $\|x\|^2 + s^2 = 1$.

Tomando-se a norma ao quadrado em ambos os membros da primeira equação e combinando-se com a segunda, obtém-se facilmente $s = (\|a\|^2 - 1)/(\|a\|^2 + 1)$ e, consequentemente, $x = \frac{2a}{\|a\|^2+1}$. Logo, f é bijetiva e

$$f^{-1}(a, 0) = \left(\frac{2a}{\|a\|^2 + 1}, \frac{\|a\|^2 - 1}{\|a\|^2 + 1}\right), \ a \in \mathbb{R}^n.$$

Como se vê, f e f^{-1} são contínuas, pois suas coordenadas são contínuas, donde f é um homeomorfismo. Em particular, $S^n - \{p_0\}$ é conexo e não-compacto.

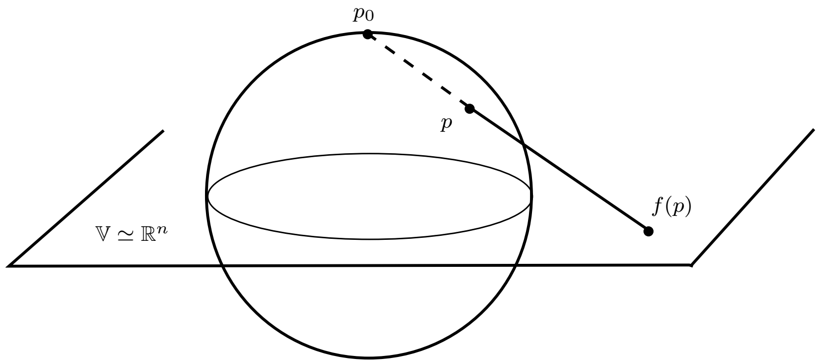

FIGURA 4

EXEMPLO 47 (INVERSÕES). Dados $a \in \mathbb{R}^n$ e $r > 0$, a *inversão* de $\mathbb{R}^n - \{a\}$ com respeito à esfera $S[a, r] \subset \mathbb{R}^n$ é a aplicação

$$\begin{array}{rccc} f: & \mathbb{R}^n - \{a\} & \to & \mathbb{R}^n - \{a\} \\ & x & \mapsto & r^2 \frac{x-a}{\|x-a\|^2} + a\,. \end{array}$$

É imediato que f é contínua. Além disso, um cálculo direto mostra que f é uma *involução*, isto é, $f \circ f$ é a aplicação identidade de $\mathbb{R}^n - \{a\}$. Daí, segue-se que f é bijetiva e $f^{-1} = f$, donde f é um homeomorfismo. Observe ainda que a aplicação f tem as seguintes propriedades:

- $f|_{S[a,r]} : S[a, r] \to S[a, r]$ é a aplicação identidade de $S[a, r]$;
- $f(B(a, r) - \{a\}) = \mathbb{R}^n - B[a, r]$ e $f(\mathbb{R}^n - B[a, r]) = B(a, r) - \{a\}$.

Em particular, $\mathbb{R}^n - B[a, r]$ e $B(a, r) - \{a\}$ são homeomorfos, donde $\mathbb{R}^n - B[a, r]$ é conexo (vide Exemplo 30 – Capítulo 2).

3.1. Homeomorfismos e Pontos Fixos. Um dos célebres resultados relacionados ao conceito de homeomorfismo é o Teorema da Invariância do Domínio, devido ao matemático holandês Luitzen Egbertus Jan Brouwer (1881-1966). Ele atesta que, se $U \subset \mathbb{R}^n$ é aberto e $f : U \to \mathbb{R}^n$ é injetiva e contínua, então $f(U)$ é um aberto de $\mathbb{R}^n$ e $f : U \to f(U)$ é um homeomorfismo. O corolário clássico do Teorema da Invariância do Domínio estabelece que $\mathbb{R}^n$ é homeomorfo a $\mathbb{R}^m$ se, e somente se, $m = n$.

Na demonstração do Teorema da Invariância do Domínio, Brouwer faz uso de um outro dos seus importantes resultados, o Teorema do Ponto Fixo, segundo o qual, toda aplicação contínua $f : B[0,1] \subset \mathbb{R}^n \to B[0,1]$ possui, pelo menos, um *ponto fixo*, isto é, um ponto $x \in B[0,1]$, tal que $f(x) = x$.

As demonstrações destes teoremas requerem conceitos e resultados da Topologia Algébrica que estão muito além dos nossos propósitos aqui. No entanto, provaremos um caso particular do Teorema da Invariância do Domínio, dito Teorema da Perturbação da Identidade[(i)], em cuja demonstração usaremos um caso particular do Teorema do Ponto Fixo (Lema 1, abaixo).

Deve-se observar que, no Teorema do Ponto Fixo, é necessário que a função em questão esteja definida e tome valores num (mesmo) conjunto fechado. Na verdade, para todo $n \in \mathbb{R}^n$, existem não apenas aplicações contínuas, mas também homeomorfismos $f : B(0,1) \subset \mathbb{R}^n \to B(0,1)$ sem pontos fixos (vide Exercício 16).

Lema 1 (teorema do ponto fixo para contrações). *Sejam $K \subset \mathbb{R}^n$ compacto e $\varphi : K \to K$ uma contração. Então, φ possui um único ponto fixo.*

Demonstração. Sejam $\lambda < 1$ uma constante de Lipschitz de φ e f a função

$$\begin{array}{rccc} f: & K & \to & \mathbb{R} \\ & x & \mapsto & \|\varphi(x) - x\|. \end{array}$$

Claramente, f é contínua. Além disso, para todo $x \in K$, tem-se

$$f(\varphi(x)) = \|\varphi(\varphi(x)) - \varphi(x)\| \leq \lambda\|\varphi(x) - x\| = \lambda f(x). \tag{21}$$

Daí, escrevendo-se $\mu = \inf\{f(x);\, x \in K\}$, obtém-se, para todo $x \in K$,

$$0 \leq \mu \leq f(\varphi(x)) \leq \lambda f(x),$$

isto é, $\mu/\lambda \leq f(x)$. Logo, devemos ter $\mu/\lambda \leq \mu$, o que nos dá $\mu = 0$, já que $0 < \lambda < 1$.

Existe, então, uma sequência em K, (x_k), tal que $f(x_k) \to 0$. Da compacidade de K, segue-se que existe uma subsequência (x_{k_i}), de (x_k), tal que $x_{k_i} \to a \in K$. Sendo f contínua, tem-se

$$f(a) = f(\lim x_{k_i}) = \lim f(x_{k_i}) = 0,$$

isto é, $a = \varphi(a)$ é um ponto fixo de φ.

Finalmente, suponhamos que b seja um ponto fixo de φ. Neste caso,

$$\|a - b\| \;=\; \|\varphi(a) - \varphi(b)\| \;\leq\; \lambda\|a - b\|,$$

donde $(1 - \lambda)\|a - b\| \leq 0$ e, portanto, $a = b$. □

Teorema da Perturbação da Identidade. *Sejam $U \subset \mathbb{R}^n$ um aberto e $\varphi : U \to \mathbb{R}^n$ uma contração. Então, a aplicação $f : U \to f(U) \subset \mathbb{R}^n$, dada por $f(x) = x + \varphi(x)$, é um homeomorfismo e $f(U)$ é aberto em $\mathbb{R}^n$.*

Demonstração. Seja $\lambda < 1$ uma constante de Lipschitz de φ. Dados, então, $x, y \in U$, tem-se $\|f(x) - f(y)\| = \|(x - y) - (\varphi(y) - \varphi(x))\|$. Logo,

$$\|f(x) - f(y)\| \geq \|x - y\| - \|\varphi(x) - \varphi(y)\| \geq \|x - y\| - \lambda\|x - y\| = (1 - \lambda)\|x - y\|.$$

[(i)] Este teorema será crucial na demonstração de um dos resultados mais importantes que estabeleceremos, a saber, o Teorema da Função Inversa.

Desta forma, $f(x) = f(y)$ implica $\|x - y\| = 0$, donde f é injetiva. Além disso, fazendo-se $x = f^{-1}(a)$ e $y = f^{-1}(b)$, $a, b \in f(U)$, tem-se

$$\|f^{-1}(a) - f^{-1}(b)\| \le \frac{1}{1-\lambda}\|a - b\|.$$

Logo, f^{-1} é lipschitziana e, portanto, (uniformemente) contínua, donde f é um homeomorfismo de U sobre $f(U)$.

Resta-nos, pois, mostrar que $f(U)$ é aberto em $\mathbb{R}^n$.

Para tanto, consideremos $x_0 \in U$ e $\delta > 0$, tais que a bola fechada $B[x_0, \delta]$ esteja contida em U. Provaremos que a bola aberta $B(f(x_0), \epsilon)$, em que $\epsilon = (1-\lambda)\delta$, está contida em $f(U)$, provando, desta forma, que $f(U)$ é aberto.

É suficiente, pois, mostrar que, dado $a \in B(f(x_0), \epsilon)$, existe $x \in B[x_0, \delta]$ satisfazendo $f(x) = a$. Para isto, consideremos a aplicação

$$\begin{array}{rccc} \psi : & B[x_0, \delta] & \to & \mathbb{R}^n \\ & x & \mapsto & x - f(x) + a \end{array}$$

e observemos que $x \in B[x_0, \delta]$ é ponto fixo de ψ se, e somente se, $f(x) = a$, o que reduz a nossa tarefa a encontrar um ponto fixo de ψ.

Notando-se que, para todo $x \in B[x_0, \delta]$, $\psi(x) = a - \varphi(x)$, conclui-se que ψ é uma contração, pois φ o é. Além disso, dado $x \in B[x_0, \delta]$, tem-se

$$\|\psi(x) - x_0| = \|a - \varphi(x) - x_0\| = \|a - f(x_0) + \varphi(x_0) - \varphi(x)\|.$$

Desta forma,

$$\|\psi(x) - x_0| \le \|a - f(x_0)\| + \|\varphi(x_0) - \varphi(x)\| < \epsilon + \lambda\|x_0 - x\| \le (1-\lambda)\delta + \lambda\delta = \delta,$$

isto é, $\psi(B[x_0, \delta]) \subset B[x_0, \delta]$. Uma vez que a bola $B[x_0, \delta]$ é compacta, segue-se do Lema 1 que ψ possui um (único) ponto fixo, como desejávamos mostrar. □

4. Continuidade e Compacidade

Estudaremos, nesta seção, o comportamento das aplicações contínuas definidas em conjuntos compactos e constataremos que estas têm propriedades especiais.

TEOREMA 14. *A imagem de um conjunto compacto por uma aplicação contínua é um conjunto compacto.*

DEMONSTRAÇÃO. Consideremos um conjunto compacto $K \subset \mathbb{R}^n$ e uma aplicação contínua $f : K \to \mathbb{R}^m$. Dada uma cobertura aberta $\mathscr{A} = \{A_\lambda\}_{\lambda \in \Lambda}$, de $f(K)$, tem-se, pelo Teorema 12, que $f^{-1}(A_\lambda)$ é aberto em K qualquer que seja $\lambda \in \Lambda$. Desta forma, a família $\{f^{-1}(A_\lambda)\}_{\lambda \in \Lambda}$ constitui uma cobertura aberta de K (ii). Sendo este compacto, temos que existem índices $\lambda_1, \dots, \lambda_i \in \Lambda$, tais que $K \subset f^{-1}(A_{\lambda_1}) \cup \dots \cup f^{-1}(A_{\lambda_i})$. Daí, tem-se

$$f(K) \subset f(f^{-1}(A_{\lambda_1}) \cup \dots \cup f^{-1}(A_{\lambda_i})) \subset A_{\lambda_1} \cup \dots \cup A_{\lambda_i},$$

donde $f(K)$ é compacto. □

Como consequência do Teorema 14, temos o seguinte resultado clássico da Topologia.

(ii) Note que os abertos desta cobertura são relativos. No entanto, conforme vimos no capítulo anterior, a compacidade de K independe do fato de usarmos a topologia de $\mathbb{R}^n$ ou a topologia relativa de K.

Teorema de Weierstrass. *Toda função contínua definida num compacto é limitada e atinge seus valores mínimo e máximo, isto é, se* $K \subset \mathbb{R}^n$ *é compacto e* $f : K \to \mathbb{R}$ *é contínua, então existem* $a, b \in K$, *tais que*

$$f(a) \leq f(x) \leq f(b) \; \forall x \in K.$$

Demonstração. Temos, pelo teorema anterior, que $f(K)$ é compacto e, portanto, fechado e limitado. Daí, segue-se que f é limitada. Sejam λ e μ o ínfimo e o supremo de $f(K)$, respectivamente. Então, existem sequências (x_k), (y_k), em K, tais que $f(x_k) \to \lambda$ e $f(y_k) \to \mu$. Uma vez que K é compacto, passando-se a subsequências, se necessário, podemos supor que $x_k \to a \in K$ e $y_k \to b \in K$, donde $f(x_k) \to f(a)$ e $f(y_k) \to f(b)$, pois f é contínua. Logo, $\lambda = f(a)$ e $\mu = f(b)$. □

Proposição 33. *Toda bijeção contínua definida num compacto é um homeomorfismo.*

Demonstração. Sejam $K \subset \mathbb{R}^n$ compacto e $f : K \to f(K) \subset \mathbb{R}^m$ uma bijeção contínua. Pelo Teorema 13, a aplicação f, sendo bijetiva e contínua, é um homeomorfismo se, e somente se, f^{-1} é contínua. Uma vez que a inversa de f^{-1} é a própria f, basta provarmos que f é uma aplicação aberta.

Tomemos, então, $A \subset K$ aberto em K. Assim, $K - A$ é fechado em K. Porém, K é fechado em $\mathbb{R}^n$. Logo, $K - A$ é fechado, também, em $\mathbb{R}^n$. Segue-se que $K - A$, por ser um subconjunto fechado de um compacto, é compacto. Agora, pelo Teorema 14, $f(K - A)$ é compacto, portanto, fechado. No entanto, a bijetividade de f nos dá $f(K - A) = f(K) - f(A)$. Logo, $f(K) - f(A)$ é fechado em $\mathbb{R}^n$ e, portanto, em $f(K)$. Desta forma, $f(A)$ é aberto em $f(K)$, donde se conclui que f é aberta. □

Nesta proposição, a hipótese de compacidade é essencial. De fato, a aplicação

$$\begin{array}{rccc} f: & [0, 2\pi) & \to & S^1 \subset \mathbb{R}^2 \\ & t & \mapsto & (\cos t, \operatorname{sen} t) \end{array}$$

é bijetiva e contínua. Entretanto, f^{-1} não é contínua, pois S^1 é compacto e $f^{-1}(S^1) = [0, 2\pi)$ não é compacto. Logo, f não é um homeomorfismo.

Proposição 34. *Toda aplicação contínua definida num conjunto compacto é uniformemente contínua.*

Demonstração. Consideremos uma aplicação contínua $f : K \to \mathbb{R}^m$, em que $K \subset \mathbb{R}^n$ é compacto, e fixemos $\epsilon > 0$. Então, para todo $p \in K$, existe $\delta(p) > 0$, tal que

$$x \in K \text{ e } \|x - p\| < \delta(p) \Rightarrow \|f(x) - f(p)\| < \epsilon. \tag{22}$$

Claramente, a família $\{B(p, \delta(p)/2)\}_{p \in K}$ constitui uma cobertura aberta de K. Logo, existem $p_1, \dots, p_i \in K$, tais que

$$K \subset B(p_1, \delta(p_1)/2) \cup \dots \cup B(p_i, \delta(p_i)/2). \tag{23}$$

Seja $\delta = \min\{\delta(p_1)/2, \dots, \delta(p_i)/2\}$. Tomando-se $x, y \in K$ satisfazendo $\|x - y\| < \delta$, tem-se, por (23), que $x \in B(p_j, \delta(p_j)/2)$ para algum $j \in \{1, \dots, i\}$. Em particular, $\|x - p_j\| < \delta(p_j)/2$. Além disso,

$$\|y - p_j\| \leq \|y - x\| + \|x - p_j\| < \delta + \delta(p_j)/2 \leq \delta(p_j).$$

Daí e de (22) segue-se que

$$\|f(x) - f(y)\| \leq \|f(x) - f(p_j)\| + \|f(p_j) - f(y)\| < 2\epsilon.$$

Logo, f é uniformemente contínua. □

Exemplo 48 (função não-lipschtziana e uniformemente contínua). Consideremos a função $f : \mathbb{R}^n \to \mathbb{R}$, definida por

$$f(x) = \sqrt{\|x\|}\,,$$

e provemos que f não é lipschtziana, porém, é uniformemente contínua.

Para tanto, consideremos as sequências $x_k = \frac{1}{k}u$ e $y_k = \frac{1}{2k}u$, em que $u \in S^{n-1}$. Temos, então,

$$\frac{|f(x_k) - f(y_k)|}{\|x_k - y_k\|} = \frac{\sqrt{\|x_k\|} - \sqrt{\|y_k\|}}{\|x_k - y_k\|} = 2\left(1 - \frac{\sqrt{2}}{2}\right)\sqrt{k},$$

donde se infere que a função $(x, y) \mapsto \frac{|f(x)-f(y)|}{\|x-y\|}$, $x \neq y$, é ilimitada e, portanto, que f não é lipschtziana.

Por outro lado, escrevendo-se $K = B[0, 1]$ e $F = \mathbb{R}^n - B(0, 1)$, temos, pela Proposição 34, que $f|_K$ é uniformemente contínua, pois K é compacto. Além disso, conforme discutimos no Exemplo 39, a restrição $f|_F$, igualmente, é uniformemente contínua. Logo, dado $\epsilon > 0$, existem $\delta_0, \delta_1 > 0$, tais que

$$x, y \in K \ \text{ e } \ \|x - y\| < \delta_0 \quad \text{ou} \quad x, y \in F \ \text{ e } \ \|x - y\| < \delta_1 \quad \Rightarrow \quad |f(x) - f(y)| < \epsilon.$$

Façamos, então, $\delta = \min\{\delta_1, \delta_2\}$ e tomemos $x, y \in \mathbb{R}^n$, tais que $\|x - y\| < \delta$. Se $x, y \in K$ ou $x, y \in F$, é imediato que $|f(x) - f(y)| < \epsilon$. Caso contrário, o segmento $[x, y]$ terá um extremo em K e o outro em F e intersectará a esfera $S^{n-1} = K \cap F$ num único ponto a. Uma vez que $\|x - a\|$ e $\|y - a\|$ são, ambos, menores que $\|x - y\| < \delta$, tem-se

$$|f(x) - f(y)| \leq |f(x) - f(a)| + |f(a) - f(y)| < 2\epsilon,$$

donde f é uniformemente contínua.

5. Continuidade e Conexidade

Em analogia à seção anterior, veremos agora como se comportam as funções contínuas definidas em conjuntos conexos.

Teorema 15. *A imagem de um conjunto conexo por uma aplicação contínua é um conjunto conexo.*

Demonstração. Consideremos um conexo $X \subset \mathbb{R}^n$ e $f : X \to f(X) \subset \mathbb{R}^m$ uma aplicação contínua. Dada uma cisão $f(X) = A \cup B$, temos que A e B são abertos disjuntos de $f(X)$. Isto implica que $f^{-1}(A)$ e $f^{-1}(B)$ são disjuntos e, sendo f contínua, abertos em X. Além disso,

$$X = f^{-1}(f(X)) = f^{-1}(A \cup B) = f^{-1}(A) \cup f^{-1}(B).$$

Uma vez que X é conexo, devemos ter, então, $f^{-1}(A) = \emptyset$ ou $f^{-1}(B) = \emptyset$, implicando que A ou B é vazio e, portanto, que $f(X)$ é conexo. □

A invariância da conexidade por aplicações contínuas implica no resultado seguinte, que, como o Teorema de Weierstrass, constitui um clássico da Topologia.

TEOREMA DO VALOR INTERMEDIÁRIO. *Sejam* $X \subset \mathbb{R}^n$ *conexo e* $f : X \to \mathbb{R}$ *uma função contínua. Se* $x_1, x_2 \in X$ *são tais que* $f(x_1) < f(x_2)$, *então, para todo* $c \in \mathbb{R}$ *satisfazendo* $f(x_1) < c < f(x_2)$, *existe* $x \in X$, *tal que* $f(x) = c$.

DEMONSTRAÇÃO. Pelo Teorema 15, $f(X) \subset \mathbb{R}$ é conexo. Logo, é um intervalo (Proposição 23 - Capítulo 2). Desta forma, dados $x_1, x_2 \in X$, a função f assume todos os valores entre $f(x_1)$ e $f(x_2)$. $\square$

EXEMPLO 49 (TEOREMA DE BORSUK-ULAM EM DIMENSÃO 1). Dados $n \in \mathbb{N}$ e uma função contínua $f : S^n \to \mathbb{R}^n$, existe pelo menos um $x_0 \in S^n$, tal que $f(x_0) = f(-x_0)$. Em particular, S^n *não é homeomorfo a subconjunto algum de* $\mathbb{R}^n$. Este resultado é conhecido como Teorema de Borsuk-Ulam, sobre o qual discutiremos na Seção 7. No caso em que $n = 1$, no entanto, ele é uma mera consequência do Teorema do Valor Intermediário. Com efeito, considerando-se a função $\varphi(x) = f(x) - f(-x)$, $x \in S^1$, vemos que $\varphi(x) = -\varphi(-x)$. Desta forma, φ é identicamente nula (implicando $f(x) = f(-x)\, \forall x \in S^1$) ou existe $x \in S^1$, tal que $\varphi(x)$ e $\varphi(-x)$ têm sinais contrários. Neste último caso, uma vez que S^1 é conexo, podemos aplicar o Teorema do Valor Intermediário e concluir que existe $x_0 \in S^1$ satisfazendo $\varphi(x_0) = 0$, o que nos dá $f(x_0) = f(-x_0)$.

EXEMPLO 50 (CONEXIDADE DE S^n). Vimos, no capítulo anterior, que, para todo $n \in \mathbb{N}$, a esfera unitária $S^n \subset \mathbb{R}^{n+1}$ é conexa. Podemos provar isso, igualmente, considerando-se a projeção radial $f : B(0,1) - \{0\} \subset \mathbb{R}^{n+1} \to S^n$, dada por $f(x) = x/\|x\|$. Com efeito, f é, claramente, contínua e sobrejetiva. Além disso, o conjunto $B(0,1) - \{0\}$ é conexo (Exemplo 30 – Capítulo 2). Logo, pelo Teorema 15, a esfera S^n é conexa.

PROPOSIÇÃO 35 (PRODUTO CARTESIANO DE CONEXOS). *Sejam* X *e* Y *subconjuntos de* $\mathbb{R}^n$ *e* $\mathbb{R}^m$, *respectivamente. Então,* $X \times Y \subset \mathbb{R}^n \times \mathbb{R}^m$ *é conexo se, e somente se,* X *e* Y *são conexos.*

DEMONSTRAÇÃO. As projeções

$$\begin{array}{ccc} X \times Y & \to & X \\ (x,y) & \mapsto & x \end{array} \quad \text{e} \quad \begin{array}{ccc} X \times Y & \to & Y \\ (x,y) & \mapsto & y \end{array}$$

são contínuas, por serem (restrições de) aplicações lineares. Assim, se $X \times Y$ for conexo, pelo Teorema 15, X e Y são conexos.

Reciprocamente, suponhamos que X e Y sejam conexos e consideremos uma cisão $X \times Y = A \cup B$. Suponhamos ainda que $A \neq \emptyset$ e tomemos $(x_0, y_0) \in A$. Claramente, os subconjuntos $X \times \{y_0\}$ e $\{x_0\} \times Y$ de $\mathbb{R}^n \times \mathbb{R}^m$ são homeomorfos a X e Y, respectivamente (basta considerar as aplicações $(x, y_0) \mapsto x$ e $(x_0, y) \mapsto y$). Logo, são conexos. Além disso, os conjuntos $(X \times \{y_0\}) \cap A$ e $(X \times \{y_0\}) \cap B$ são disjuntos e abertos em $X \times \{y_0\}$ pois A e B são disjuntos e abertos em $X \times Y$ e $X \times \{y_0\} \subset X \times Y$. Segue-se que a igualdade

$$X \times \{y_0\} = ((X \times \{y_0\}) \cap A) \cup ((X \times \{y_0\}) \cap B)$$

define uma cisão de $X \times \{y_0\}$. Uma vez que $(X \times \{y_0\}) \cap A$ é não-vazio e $X \times \{y_0\}$ é conexo, devemos ter $(X \times \{y_0\}) \cap B = \emptyset$, isto é, $X \times \{y_0\} \subset A$.

Suponhamos agora, por absurdo, que $B \neq \emptyset$ e tomemos $(x_1, y_1) \in B$. Um procedimento análogo ao do parágrafo anterior nos levaria a concluir que $\{x_1\} \times Y \subset B$. Sendo assim, teríamos $(x_1, y_0) \in A \cap B$, contradizendo o fato de A e B serem disjuntos. Desta forma, $B = \emptyset$, donde $X \times Y$ é conexo. $\square$

Exemplo 51. Consideremos o *cilindro circular reto*

$$C = \{(x,y,z) \in \mathbb{R}^3;\, x^2+y^2=1\} \subset \mathbb{R}^3.$$

Tomando-se a decomposição $\mathbb{R}^3 = \mathbb{R}^2 \times \mathbb{R}$, vê-se facilmente que $C = S^1 \times \mathbb{R}$. Logo, pela Proposição 35, C é conexo.

5.1. Conexidade por Caminhos. Dado $X \subset \mathbb{R}^n$, chama-se *curva* (ou *caminho*) em X a toda aplicação contínua $\alpha : I \subset \mathbb{R} \to X$, em que $I \subset \mathbb{R}$ é um intervalo de $\mathbb{R}$. O *traço* de uma curva α é o seu conjunto-imagem, $\alpha(I) \subset X$, e, no caso em que $I = [0,1]$, $\alpha(0), \alpha(1) \in X$ são chamados de *extremos* de α e ditos, também, *o ponto inicial* e *o ponto final* de α, respectivamente.

As aplicações $\alpha, \beta : [0,1] \to \mathbb{R}^2$, por exemplo, dadas por

$$\alpha(t) = (1-t)x + ty, \;\; x,y \in \mathbb{R}^2, \quad \text{e} \quad \beta(t) = (\cos 2\pi t, \operatorname{sen} 2\pi t),$$

são curvas em $\mathbb{R}^2$ que têm como traços, respectivamente, o segmento fechado $[x,y]$ e a esfera S^1.

Cabe-nos advertir que, a despeito destes (e tantos outros) exemplos e contrariamente à intuição, o traço de uma curva, não necessariamente, é um objeto unidimensional. Mais precisamente, existem curvas $\alpha : [0,1] \to \mathbb{R}^2$ cujo traço é o quadrado $[0,1] \times [0,1] \subset \mathbb{R}^2$. Elas são chamadas de *curvas de Peano*, em consideração ao matemático italiano Giuseppe Peano (1858-1932) que, em 1890, surpreendeu a comunidade matemática ao exibir o primeiro exemplo de uma tal curva.

Definição 20 (conexidade por caminhos). Diz-se que um conjunto $X \subset \mathbb{R}^n$ é *conexo por caminhos* se, para quaisquer $x, y \in X$, existe uma curva $\alpha : [0,1] \to X$, tal que $\alpha(0) = x$ e $\alpha(1) = y$. Neste caso, diz-se que a curva α (ou o seu traço) *liga* o ponto x ao ponto y.

Todo conjunto convexo $X \subset \mathbb{R}^n$ é, claramente, conexo por caminhos. Em particular, o espaço $\mathbb{R}^n$, suas bolas, semi-espaços e paralelepípedos são conexos por caminhos.

Sejam $X \subset \mathbb{R}^n$ e $\alpha, \beta : [0,1] \to X$ curvas em X, tais que $\alpha(1) = \beta(0)$. Fazendo-se $\alpha(0) = x$ e $\beta(1) = y$, tem-se que a aplicação $\gamma : [0,1] \to X$, dada por

$$\gamma(t) = \begin{cases} \alpha(2t) & \text{se} \quad t \in [0,1/2] \\ \\ \beta(2t-1) & \text{se} \quad t \in [1/2,1], \end{cases}$$

é, claramente, contínua e satisfaz $\gamma(0) = x$, $\gamma(1) = y$. Logo, γ é uma curva que liga x a y, a qual é denotada por $\alpha \vee \beta$ e dita a *justaposição* de α e β.

Destas considerações, segue-se facilmente que, para todo $n \geq 2$, $\mathbb{R}^n - \{0\}$ é conexo por caminhos. Com efeito, dados $x, y \in \mathbb{R}^n - \{0\}$, se o segmento $[x,y]$ não intersecta a origem, então ele liga x a y. Caso contrário, tomamos $z \in \mathbb{R}^n$, tal que z e $x-y$ sejam linearmente independentes e justapomos as curvas correspondentes aos segmentos $[x,z]$ e $[z,y]$. O mesmo argumento se aplica aos conjuntos $B[0,1] - \{0\} \subset \mathbb{R}^n$ e $B(0,1) - \{0\} \subset \mathbb{R}^n$ quando $n \geq 2$.

Exemplo 52. *Para todo* $n \in \mathbb{N}$, *a esfera* $S^n \subset \mathbb{R}^{n+1}$ *é conexa por caminhos.* De fato, dados $x, y \in S^n$, se $x \neq -y$, então, para todo $t \in [0,1]$, $(1-t)x + ty \neq 0$, donde

$$\begin{array}{rccc} \alpha: & [0,1] & \to & S^n \\ & t & \mapsto & \frac{(1-t)x+ty}{\|(1-t)x+ty\|} \end{array}$$

é uma curva em S^n (dada pela composta da projeção radial $x \mapsto x/\|x\|$ com a curva $t \mapsto (1-t)x + ty$) que liga x a y. Agora, se $x = -y$, tomamos $z \in S^n$, tal que $z \neq \pm x$. Neste caso, existem curvas $\alpha, \beta : [0,1] \to S^n$, tais que α liga x a z e β liga z a y, donde a justaposição $\alpha \vee \beta$ liga x a y.

Proposição 36. *Todo conjunto conexo por caminhos é conexo.*

Demonstração. Sejam $X \subset \mathbb{R}^n$ conexo por caminhos e $a \in X$. Então, para todo $x \in X$, existe um caminho $\alpha_x : [0,1] \to X$ ligando a a x. Sendo $[0,1]$ conexo e α_x contínua, temos que $\alpha_x([0,1])$ é conexo. No entanto, para todo $x \in X$, $a \in \alpha_x([0,1])$ e, obviamente,

$$X = \bigcup_{x \in X} \alpha_x([0,1]).$$

Logo, X é conexo, por ser a união de conexos com um ponto em comum. □

Não vale, entretanto, a recíproca desta proposição, conforme ilustrado no exemplo seguinte.

Exemplo 53 (um conexo que não é conexo por caminhos). Consideremos o ponto $p = (1,0) \in \mathbb{R}^2$ e uma sequência (x_k) em $S^1 - \{p\} \subset \mathbb{R}^2$, tal que $x_k \to p$. Conforme constatamos no Exemplo 31 do Capítulo 2, fazendo-se $X = \bigcup X_k$, em que $X_k = [0, x_k] \subset \mathbb{R}^2$, tem-se que $C = X \cup \{p\} \subset \mathbb{R}^2$ é conexo. Verifiquemos, então, que C não é conexo por caminhos. Com este intuito, consideremos a função (contínua)

$$\begin{array}{cccc} f: & C - \{0\} & \to & \mathbb{R} \\ & x & \mapsto & \langle \frac{x}{\|x\|}, p \rangle \end{array}$$

e observemos que $f(x) = 1$ se, e somente se, $x = p$. Supondo-se, sem perda de generalidade, que os termos da sequência (x_k) são distintos entre si, tem-se que cada vetor de $X - \{0\}$ está em um, e somente um, dos segmentos $[0, x_k]$. Logo, $f(C - \{0\}) = \{1, f(x_1), \dots, f(x_k), \dots\}$, donde $f(C - \{0\})$ é enumerável.

Tomemos uma curva $\alpha : [0,1] \to C$ com ponto inicial p e consideremos o conjunto

$$\Omega = \{t \in [0,1];\ \alpha(t) = p\}.$$

Uma vez que $0 \in \Omega$ e α é contínua, temos que Ω é não-vazio e fechado em $[0,1]$.

Agora, dados $t_0 \in \Omega$ e uma bola aberta $B = B(\alpha(t_0), r) \subset \mathbb{R}^2 - \{0\}$, pela continuidade de α, o conjunto $A = \alpha^{-1}(B \cap C)$ é um aberto de $[0,1]$ que contém t_0. Em particular, existe um intervalo $I \subset [0,1]$, aberto em $[0,1]$, tal que $t_0 \in I \subset A$. Sendo I um conjunto conexo e $f \circ \alpha$ uma aplicação contínua, temos que $f(\alpha(I))$ é um subconjunto conexo de $\mathbb{R}$ e, portanto, um intervalo. No entanto, $f(\alpha(I))$ é enumerável, donde se conclui que este intervalo é degenerado, isto é, $f \circ \alpha|_I$ é constante e igual a $f(\alpha(t_0)) = 1$. Daí, infere-se que $\alpha(t) = p\, \forall t \in I$, pois $f(x) = 1$ se, e somente se, $x = p$. Logo, $I \subset \Omega$, isto é, Ω é aberto em $[0,1]$.

Segue-se destas considerações e da conexidade de $[0,1]$ que $\Omega = [0,1]$. Sendo assim, toda curva em C com ponto inicial p é constante, donde se conclui que C não é conexo por caminhos.

Sejam $X \subset \mathbb{R}^n$ um conjunto conexo por caminhos e $f : X \to f(X) \subset \mathbb{R}^m$ uma aplicação contínua. Dados $x, y \in X$ e uma curva $\alpha : [0,1] \to X$ ligando x

a y, temos que $\beta = f \circ \alpha : [0,1] \to f(X)$ é contínua, pois é uma composta de aplicações contínuas. Além disso, $\beta(0) = f(x)$ e $\beta(1) = f(y)$, donde β é uma curva em $f(X)$ que liga $f(x)$ a $f(y)$. Logo, vale o resultado seguinte.

Proposição 37. *A imagem de um conjunto conexo por caminhos por uma aplicação contínua é um conjunto conexo por caminhos.*

Segue-se da Proposição 37 que $S^n - \{-p, p\}$, em que p é o polo-norte de S^n, é conexo por caminhos. De fato, vimos no Exemplo 46 que a projeção estereográfica de $S^n - \{p\}$ sobre $\mathbb{R}^n$, $f : S^n - \{p\} \to \mathbb{R}^n$, é um homeomorfismo. Em particular, f^{-1} é contínua. Porém, $S^n - \{-p, p\} = f^{-1}(\mathbb{R}^n - \{0\})$, donde $S^n - \{-p, p\}$ é conexo por caminhos, pois $\mathbb{R}^n - \{0\}$ o é.

No exemplo a seguir, usaremos a noção de conexidade por caminhos e o Teorema do Valor Intermediário para provar que $\mathbb{R}^n - \{0\}$ e $\mathbb{R}^n$ não são homeomorfos. Para tanto, adaptaremos uma engenhosa ideia de Tiago Ribeiro (vide [**27**]).

Exemplo 54 ($\mathbb{R}^n - \{0\}$ não é homeomorfo a $\mathbb{R}^n$). É imediato que $\mathbb{R} - \{0\}$ e $\mathbb{R}$ não são homeomorfos, pois $\mathbb{R}$ é conexo e $\mathbb{R} - \{0\}$ é desconexo. Tomemos, então, $n \geq 2$ e suponhamos, por absurdo, que exista um homeomorfismo $f : \mathbb{R}^n - \{0\} \to \mathbb{R}^n$. Consideremos os conjuntos $A = \{x \in \mathbb{R}^n; 0 < \|x\| < 1\}$ e $B = \{x \in \mathbb{R}^n; \|x\| > 1\}$ e verifiquemos que suas imagens por f são conjuntos ilimitados de $\mathbb{R}^n$. Com efeito, dada uma sequência (x_k) em A, tal que $x_k \to 0$, temos que a sequência $(f(x_k))$ é ilimitada. Caso contrário, para alguma subsequência $(f(x_{k_i}))$ de $(f(x_k))$, teríamos $f(x_{k_i}) \to b \in \mathbb{R}^n$. Pela continuidade de f^{-1}, teríamos então, $0 = \lim x_{k_i} = f^{-1}(b)$, o que é absurdo, já que $f^{-1}(\mathbb{R}^n) = \mathbb{R}^n - \{0\}$. Logo, $f(A)$ é ilimitado.

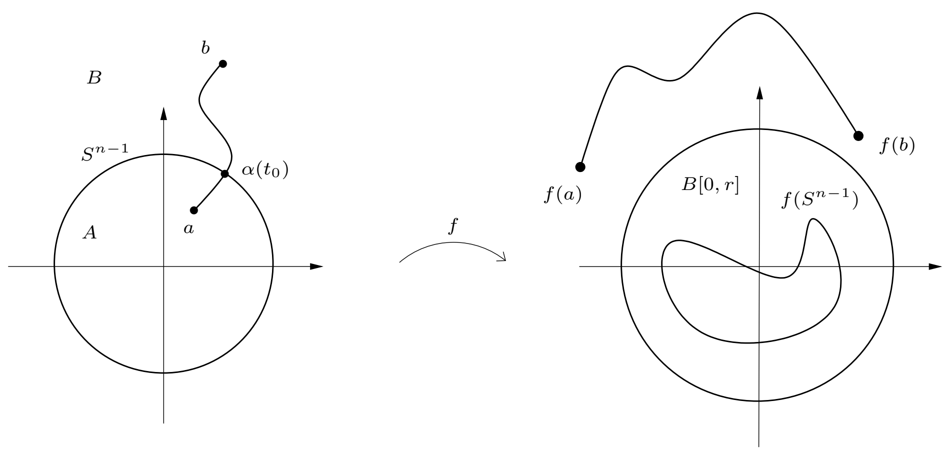

Figura 5

Tomando-se agora uma sequência (y_k) em B, tal que, para todo $k \in \mathbb{N}$, tenha-se $\|y_k\| > k$, concluímos, de forma análoga, que $(f(y_k))$ é ilimitada, donde $f(B)$ é ilimitado.

Por outro lado, uma vez que S^{n-1} é compacto e f é contínua, temos que $f(S^{n-1})$ é compacto e, em particular, limitado. Logo, existe $r > 0$, tal que $f(S^{n-1}) \subset B[0, r]$ e, como $f(A)$ e $f(B)$ são ilimitados, existem $a \in A$ e $b \in B$, tais que $f(a), f(b) \in \mathbb{R}^n - B[0, r]$.

Observemos agora que $\mathbb{R}^n - B[0,r]$ é conexo por caminhos, pois é homeomorfo ao conjunto $B(0,r) - \{0\}$ (vide Exemplo 47), que, conforme constatamos, é conexo por caminhos. Logo, existe uma curva $\beta : [0,1] \to \mathbb{R}^n - B[0,r]$, tal que $\beta(0) = f(a)$ e $\beta(1) = f(b)$. Sendo f um homeomorfismo, a aplicação $\alpha : [0,1] \to \mathbb{R}^n - \{0\}$, $\alpha(t) = (f^{-1} \circ \beta)(t)$, é uma curva em $\mathbb{R}^n - \{0\}$, tal que $\alpha(0) = a$ e $\alpha(1) = b$. Logo, a função $\lambda : [0,1] \to \mathbb{R}$, dada por $\lambda(t) = \|\alpha(t)\|$, é contínua. Além disso, $\lambda(0) = \|a\| < 1$ e $\lambda(1) = \|b\| > 1$. Desta forma, pelo Teorema do Valor Intermediário, existe $t_0 \in (0,1)$, tal que $\lambda(t_0) = 1$, donde $\alpha(t_0) \in S^{n-1}$. Logo, $\beta(t_0) = f(\alpha(t_0)) \in f(S^{n-1}) \subset B[0,r]$, o que contradiz o fato de β ser uma curva em $\mathbb{R}^n - B[0,r]$. Segue-se, portanto, desta contradição, que um tal homeomorfismo f não existe.

6. Limites

Introduziremos agora o conceito mais essencial da Análise, o de limite de aplicações, que generaliza o de limite de sequências e relaciona-se intimamente com a noção de continuidade. Veremos, então, nos capítulos posteriores, que muitas propriedades e conceitos importantes acerca de aplicações, como a derivada, são estabelecidos a partir desta noção geral de limite.

Definição 21 (limite). Sejam $f : X \subset \mathbb{R}^n \to Y \subset \mathbb{R}^m$ uma aplicação e $a \in \mathbb{R}^n$ um ponto de acumulação de X. Diz-se que $b \in \mathbb{R}^m$ é *limite* de f quando x tende para a, e escreve-se

$$\lim_{x \to a} f(x) = b,$$

se, para todo $\epsilon > 0$, existe $\delta > 0$, tal que

$$x \in X \quad \text{e} \quad 0 < \|x - a\| < \delta \ \Rightarrow\ \|f(x) - b\| < \epsilon.$$

Verifica-se facilmente, como no caso das sequências, que um limite de uma aplicação, quando existe, é único.

As definições de continuidade de f num ponto a e de limite de f quando x tende para a, apesar de semelhantes, guardam diferenças cruciais. De fato, na definição acima, o ponto a deve ser um ponto de acumulação de X, porém, não necessariamente um ponto de X, enquanto na definição de continuidade, o ponto a deve pertencer a X mas não necessariamente ser um ponto de acumulação de X. Assim, o limite de f num ponto a, quando existe, nos permite avaliar o comportamento de f numa vizinhança de a, mesmo que f não esteja definida em a. Neste fato residem a importância e efetividade do conceito de limite.

Como reflexo da relação entre os conceitos de limite e continuidade, tem-se o resultado seguinte, cuja verificação é imediata.

Proposição 38 (limite e continuidade). *Sejam $X \subset \mathbb{R}^n$, $Y \subset \mathbb{R}^m$ e $a \in X \cap X'$. Então, $f : X \to Y$ é contínua em a se, e somente se,*

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

Observemos que $a \in \mathbb{R}^n$ é ponto de acumulação de $X \subset \mathbb{R}^n$ se, e somente se, 0 é ponto de acumulação do conjunto $\{h \in \mathbb{R}^n ; a + h \in X\}$. Fazendo-se, então, $x = a + h$, as equivalências de cada uma das colunas do diagrama

$$\begin{array}{ccc} 0 < \|x-a\| < \delta & \Rightarrow & \|f(x)-b\| < \epsilon \\ \Updownarrow & & \Updownarrow \\ 0 < \|h\| < \delta & \Rightarrow & \|f(a+h)-b\| < \epsilon \end{array}$$

são, evidentemente, verdadeiras. Logo, a implicação da primeira linha é equivalente à da segunda, isto é,

$$\lim_{x\to a} f(x) = b \ \Leftrightarrow\ \lim_{h\to 0} f(a+h) = b.$$

É pertinente mencionarmos também que a desigualdade $|\,\|f(x)\| - \|b\|\,| \le \|f(x)-b\|$ nos dá

$$\lim_{x\to a} f(x) = b \ \Rightarrow\ \lim_{x\to a} \|f(x)\| = \|b\|.$$

Além disso, como no caso das sequências, verifica-se facilmente que vale a recíproca quando $b = 0$, isto é,

$$\lim_{x\to a} f(x) = 0 \ \Leftrightarrow\ \lim_{x\to a} \|f(x)\| = 0.$$

Definição 22 (limites laterais). Sejam X um subconjunto de $\mathbb{R}$ e a um ponto de acumulação à direita de X. Dada uma função $f : X \to \mathbb{R}^m$, diz-se que $b \in \mathbb{R}^n$ é *limite à direita* de f quando x tende a a, e escreve-se

$$\lim_{x\to a+} f(x) = b,$$

quando, para todo $\epsilon > 0$, existe $\delta > 0$, tal que

$$x \in X \quad \text{e} \quad a < x < a + \delta \ \Rightarrow\ \|f(x) - b\| < \epsilon.$$

Quando a é ponto de acumulação à esquerda de X, introduz-se de forma análoga a noção de *limite à esquerda* de f quando x tende a a.

As seguintes propriedades dos limites laterais são facilmente verificadas:

i) Um limite lateral de uma aplicação, à direita ou à esquerda, quando existe, é único;

ii) um ponto $b \in \mathbb{R}^m$ é limite à direita (respectivamente, à esquerda) de $f : X \subset \mathbb{R} \to \mathbb{R}^m$ quando x tende a $a \in X'_+$ (respectivamente, $a \in X'_-$) se, e somente se, para toda sequência x_k em X, tal que $x_k > a$ (respectivamente, $x_k < a$) e $x_k \to a$, se tenha $f(x_k) \to b$;

iii) se a é ponto de acumulação à direita e à esquerda de $X \subset \mathbb{R}$, então o limite de $f : X \to \mathbb{R}^m$ quando x tende a a existe e é igual a $b \in \mathbb{R}^m$ se, e somente se, ambos os limites laterais de f quando x tende a a existem e são iguais a b.

Vejamos agora alguns resultados sobre limites de funções que envolvem as relações de ordem em $\mathbb{R}$.

Proposição 39 (teorema do confronto). *Sejam $a \in \mathbb{R}^n$ um ponto de acumulação de $X \subset \mathbb{R}^n$ e $f, g, h : X \to \mathbb{R}$ funções, tais que:*

i) $f(x) \le h(x) \le g(x)\, \forall x \in X$;

ii) $\lim_{x\to a} f(x) = \lim_{x\to a} g(x) = \lambda \in \mathbb{R}$.

Então, $\lim\limits_{x\to a} h(x)$ *existe e é igual a* λ.

Demonstração. Dado $\epsilon > 0$, segue-se da hipótese (ii) que existem $\delta_1 > 0$ e $\delta_2 > 0$, tais que:

- $x \in X$ e $0 < \|x - a\| < \delta_1 \Rightarrow -\epsilon < f(x) - \lambda < \epsilon$;
- $x \in X$ e $0 < \|x - a\| < \delta_2 \Rightarrow -\epsilon < g(x) - \lambda < \epsilon$.

Desta forma, para todo $x \in X$ satisfazendo $0 < \|x-a\| < \delta = \min\{\delta_1, \delta_2\}$, vale $-\epsilon < f(x)-\lambda \le h(x)-\lambda \le g(x)-\lambda < \epsilon$, donde se conclui que $\lim\limits_{x\to a} h(x) = \lambda$. □

Proposição 40 (permanência da desigualdade). *Sejam* $X \subset \mathbb{R}^n$, $a \in \mathbb{R}^n$ *um ponto de acumulação de* X *e* $f, g : X \to \mathbb{R}$ *funções, tais que*

$$\lim_{x\to a} f(x) = \lambda_1 \qquad e \qquad \lim_{x\to a} g(x) = \lambda_2 .$$

Então, se $f(x) \le g(x)$ *para todo* $x \in X$, *tem-se* $\lambda_1 \le \lambda_2$.

Demonstração. Suponhamos, por contraposição, que $\lambda_1 > \lambda_2$ e tomemos $\epsilon > 0$, tal que $0 < \epsilon < (\lambda_1 - \lambda_2)/2$ (isto é, $\lambda_1 - \epsilon > \lambda_2 + \epsilon$). Pela hipótese, existem $\delta_1 > 0$ e $\delta_2 > 0$, tais que:

- $x \in X$ e $0 < \|x - a\| < \delta_1 \Rightarrow \lambda_1 - \epsilon < f(x) < \lambda_1 + \epsilon$;
- $x \in X$ e $0 < \|x - a\| < \delta_2 \Rightarrow \lambda_2 - \epsilon < g(x) < \lambda_2 + \epsilon$.

Desta forma, tomando-se $\delta = \min\{\delta_1, \delta_2\}$ e $x \in X \cap B(a, \delta)$, tem-se

$$f(x) > \lambda_1 - \epsilon > \lambda_2 + \epsilon > g(x),$$

donde se conclui o desejado. □

A proposição a seguir relaciona os conceitos de limite de aplicações e sequências. Sua demonstração é análoga à da Proposição 26 e a deixamos a cargo do leitor.

Proposição 41 (limite e convergência de sequências). *Sejam* X *um subconjunto de* $\mathbb{R}^n$ *e* $f : X \to \mathbb{R}^m$ *uma aplicação. Se* $a \in \mathbb{R}^n$ *é um ponto de acumulação de* X, *então* $\lim_{x\to a} f(x) = b$ *se, e somente se, para toda sequência* (x_k) *em* $X - \{a\}$ *satisfazendo* $x_k \to a$, *tem-se* $f(x_k) \to b$.

Observação 14. No tocante à proposição acima, cabe aqui um comentário análogo àquele feito na Observação 12, isto é, para que se tenha $\lim_{x\to a} f(x) = b$ é suficiente que, para toda sequência (x_k) em $X - \{a\}$ satisfazendo $x_k \to a$, exista uma subsequência convergente de $(f(x_k))$ cujo limite seja b.

Combinando-se os resultados da Proposição 41 e da Proposição 6 do Capítulo 1, obtemos facilmente o resultado seguinte.

Proposição 42 (propriedades operatórias). *Consideremos aplicações* $f, g : X \subset \mathbb{R}^n \to \mathbb{R}^m$ *e uma função* $\lambda : X \to \mathbb{R}$. *Suponhamos que* $a \in X'$ *e que existam os limites*

$$\lim_{x\to a} f(x), \quad \lim_{x\to a} g(x) \quad e \quad \lim_{x\to a} \lambda(x).$$

Então, existem os limites

$$\lim_{x\to a}(f+g)(x), \quad \lim_{x\to a}(\lambda f)(x) \quad e \quad \lim_{x\to a}(1/\lambda)(x) \ \ (\textit{supondo-se}\ \lambda(x) \ne 0\, \forall x \in X).$$

Além disso, tem-se

- $\lim\limits_{x\to a}(f+g)(x) = \lim\limits_{x\to a} f(x) + \lim\limits_{x\to a} g(x)$;
- $\lim\limits_{x\to a}(\lambda f)(x) = \lim\limits_{x\to a} \lambda(x) \lim\limits_{x\to a} f(x)$;
- $\lim\limits_{x\to a}(1/\lambda)(x) = 1/\lim\limits_{x\to a} \lambda(x)$ (*supondo-se* $\lim\limits_{x\to a} \lambda(x) \neq 0$).

Segue-se da Proposição 41 e da Proposição 5 do Capítulo 1 que o limite de uma aplicação existe se, e somente se, existe o limite correspondente de cada uma de suas funções-coordenada, conforme a proposição abaixo.

PROPOSIÇÃO 43 (LIMITE DAS COORDENADAS DE UMA APLICAÇÃO). *Considere uma aplicação* $f : X \subset \mathbb{R}^n \to Y \subset \mathbb{R}^m$ *e* $a \in X'$. *Então,*

$$\lim_{x\to a} f(x) = b = (b_1, \dots, b_m) \quad \Leftrightarrow \quad \lim_{x\to a} f_i(x) = b_i \ \ \forall\, i \in \{1, \dots, m\},$$

em que $f_1, \dots, f_m$ *são as coordenadas de* f.

Para futura referência, verificaremos na proposição seguinte uma propriedade que envolve limites de aplicações que tomam valores em $\mathrm{L}(\mathbb{R}^n, \mathbb{R}^m)$.

PROPOSIÇÃO 44. *Considere uma aplicação* $\Phi : X \subset \mathbb{R}^n \to \mathrm{L}(\mathbb{R}^n, \mathbb{R}^m)$. *Dado* $a \in X'$, *suponha que existe* $T \in \mathrm{L}(\mathbb{R}^n, \mathbb{R}^m)$, *tal que*

$$\lim_{x\to a} \Phi(x)v \to Tv \ \ \forall v \in \mathbb{R}^n.$$

Então, $\lim\limits_{x\to a} \Phi(x) = T$.

DEMONSTRAÇÃO. Seja (x_k) uma sequência em $X - \{a\}$, tal que $x_k \to a$. Segue-se da hipótese e da Proposição 41 que a sequência $(\Phi(x_k))$ converge pontualmente para T. Logo, pela Proposição 5 do Capítulo 1, $\Phi(x_k) \to T$. Daí e, uma vez mais, da Proposição 41, conclui-se que $\lim\limits_{x\to a} \Phi(x) = T$, como desejado. $\square$

Tomemos o conjunto das transformações lineares invertíveis de $\mathbb{R}^n$, $\mathrm{I}(\mathbb{R}^n)$, e provemos, usando limites, que a aplicação

$$\begin{array}{cccc} \varphi : & \mathrm{I}(\mathbb{R}^n) & \to & \mathrm{I}(\mathbb{R}^n) \\ & T & \mapsto & T^{-1} \end{array}$$

é contínua.

Antes, porém, observemos que se $f : X \subset \mathbb{R}^n \to \mathbb{R}^m$ é limitada por uma constante $\mu > 0$ e $\lambda : X \to \mathbb{R}$ é tal que, para $a \in X'$, $\lim_{x\to a} \lambda(x) = 0$, então

$$\lim_{x\to a}(\lambda f)(x) = \lim_{x\to a} \lambda(x) f(x) = 0, \tag{24}$$

mesmo que o limite de f quando x tende para a não exista. De fato, pela hipótese, dado $\epsilon > 0$, existe $\delta > 0$, tal que $|\lambda(x)| < \epsilon$ para todo $x \in X$ satisfazendo $\|x - a\| < \delta$. Logo, para estes valores de x, tem-se

$$\|\lambda(x) f(x)\| = |\lambda(x)| \|f(x)\| < \mu\epsilon,$$

o que implica (24).

Retomemos agora a aplicação φ. No Exemplo 16 do Capítulo 2, vimos que, dada uma transformação $T \in \mathrm{I}(\mathbb{R}^n)$, existe uma constante $\mu > 0$, tal que, para toda $H \in \mathrm{L}(\mathbb{R}^n)$ satisfazendo $\|H\| < \mu$, tem-se $\|(T+H)x\| \geq \mu\|x\| \, \forall x \in \mathbb{R}^n$. Daí segue-se que $T + H$ é invertível e (vide Exercício 14 – Capítulo 1)

$$\|(T+H)^{-1}\| \leq \frac{1}{\mu} \cdot \tag{25}$$

Além disso, para toda tal H, tem-se

(26) $$(T+H)^{-1} - T^{-1} = (T+H)^{-1}\left(I - (T+H)T^{-1}\right) = -(T+H)^{-1}HT^{-1},$$

em que I é a aplicação identidade de $\mathbb{R}^n$. Logo,

(27) $$\|\varphi(T+H) - \varphi(T)\| = \|(T+H)^{-1} - T^{-1}\| \leq \|(T+H)^{-1}\| \, \|H\| \, \|T^{-1}\|.$$

Porém, por (25), $(T+H)^{-1}$ é limitada. Desta forma,

$$\lim_{H\to 0} \|(T+H)^{-1}\| \, \|H\| \, \|T^{-1}\| = 0.$$

Segue-se, então, de (27) e da Proposição 39, que $\lim_{H\to 0} \|\varphi(T+H) - \varphi(T)\| = 0$, o que nos dá

$$\lim_{H\to 0} \varphi(T+H) = \varphi(T),$$

provando que φ é contínua.

Note que, igualmente, a inversão de matrizes é uma aplicação contínua.

7. O Teorema de Borsuk-Ulam $(*)$

Em conclusão ao capítulo, dedicamos esta seção a um consagrado teorema da Topologia, o qual enunciamos.

Teorema de Borsuk-Ulam. *Dada uma aplicação contínua* $f : S^n \to \mathbb{R}^n$, *existe* $x \in S^n$, *tal que* $f(x) = f(-x)$.

Esta bela propriedade da esfera foi conjecturada por Stanislaw Ulam (1909–1984) no célebre Scottish Café, na cidade de Lvov, Ucrânia (antes, Polônia), que, entre 1930 e 1940, foi lugar de inúmeras reuniões dos mais renomados matemáticos poloneses. Entre eles, encontrava-se Karol Borsuk (1905–1982), que, em 1933 (vide [**2**]), provou a conjectura de Ulam e, desta forma, compartilhou com este o nome do teorema.

Não dispomos, lamentavelmente, dos instrumentos necessários para fornecer uma demonstração do Teorema de Borsuk-Ulam, pois, para isto, salvo no caso unidimensional (Exemplo 49), são necessários invariantes topológicos mais sofisticados que compacidade e conexidade. No entanto, podemos abordar um dos vários aspectos importantes deste teorema, qual seja, a sua vasta aplicabilidade.

Com este intuito, demonstraremos inicialmente o Teorema de Lyusternik (1899–1981) e Shnirelman (1905–1938), que estabelece, a partir do Teorema de Borsuk-Ulam, uma interessante propriedade topológica da esfera. Em seguida, mostraremos que o Teorema de Borsuk-Ulam implica o Teorema do Ponto Fixo, de Brouwer.

Teorema de Lyusternik-Shnirelman. *Se* $n+1$ *conjuntos fechados cobrem* S^n, *então pelo menos um deles contém um par de pontos antípodas*[iii].

Demonstração. Sejam $F_1, \dots, F_{n+1} \subset S^n$ $n+1$ conjuntos fechados, tais que $F_1 \cup \dots \cup F_{n+1} = S^n$. Definamos, então, a função

$$\begin{array}{rccc} f: & S^n & \to & \mathbb{R}^n \\ & x & \mapsto & (d(x,F_1), \dots, d(x,F_n)). \end{array}$$

Uma vez que, para cada $i \in \{1, \dots, n\}$, a função $x \mapsto d(x, F_i)$ é contínua (vide Exemplo 40), temos que f é contínua. Podemos, desta forma, aplicar o Teorema

[iii] Dois pontos $x, y \in S^n$ são ditos *antípodas* se $y = -x$.

de Borsuk-Ulam e concluir que existe $x \in S^n$, tal que $f(x) = f(-x)$, donde, para todo $i \in \{1, \ldots, n\}$, $d(x, F_i) = d(-x, F_i)$. Se $x \in F_i$ para algum $i \neq n+1$, então, $0 = d(x, F_i) = d(-x, F_i)$, isto é, $-x \in F_i$, pois cada F_i é fechado. Logo, se para todo $i \in \{1, \ldots, n\}$ x não pertence a F_i, tem-se $-x, x \in F_{n+1}$. Em todo caso, x e $-x$ estão, ambos, num mesmo fechado F_i. $\square$

Em consideração ao Teorema de Lyusternik-Shnirelman, convém mencionar que toda esfera S^n pode ser coberta por $n+2$ fechados de uma maneira tal que nenhum deles contenha pontos antípodas. De fato, no caso $n = 1$, basta inscrever um triângulo equilárero em S^1 e considerar os três fechados obtidos pelas projeções radiais, sobre S^1, de cada um dos lados deste triângulo, conforme indicado na Figura 6.

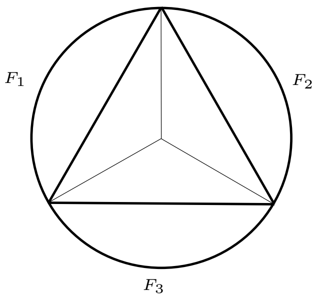

Figura 6

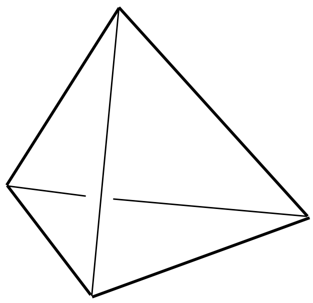

Figura 7

Este procedimento se generaliza facilmente inscrevendo-se em S^n um análogo $(n+1)$-dimensional do triângulo equilátero, o dito *simplexo regular* de dimensão $n+1$,

$$\Delta^{n+1} = \left\{ x \in \mathbb{R}^{n+1};\, x = \sum_{i=1}^{n+2} t_i p_i\,, \quad \sum_{i=1}^{n+2} t_i \leq 1,\;\; t_i \geq 0 \right\},$$

em que $p_1, \ldots, p_{n+2}$ são pontos de S^n convenientemente escolhidos, os quais são chamados de *vértices* de Δ^{n+1}. Note que o simplexo tridimensional, Δ^3, é o tetraedro regular (Fig. 7).

Segue-se da existência de uma tal cobertura de S^n, não somente que o número $n+1$ no Teorema de Lyusternik-Shnirelman é justo, mas também que este teorema é equivalente ao Teorema de Borsuk-Ulam! Com efeito, supondo-se, por contraposição, que existe $f : S^n \to \mathbb{R}^n$ contínua, tal que, para todo $x \in S^n$, tenha-se $f(x) \neq f(-x)$, podemos definir uma aplicação contínua $g : S^n \to S^{n-1}$, tal que $g(x) = -g(-x)\, \forall x \in S^n$, isto é, g leva antípodas em antípodas. Basta fazer

$$g(x) = \frac{f(x) - f(-x)}{\|f(x) - f(-x)\|}\,.$$

Desta forma, tomando-se $n+1$ fechados $F_1, \ldots, F_{n+1} \subset S^{n-1}$ que cobrem S^{n-1} e tais que nenhum deles contenha um par de antípodas, tem-se, então, $S^n = g^{-1}(F_1) \cup \cdots \cup g^{-1}(F_{n+1})$ e nenhum dos fechados $g^{-1}(F_i)$ contém um par de antípodas, donde se conclui que o Teorema de Lyusternik-Shnirelman implica o Teorema de Borsuk-Ulam.

Passemos ao Teorema do Ponto Fixo. A maneira usual de demonstrá-lo consiste em reduzi-lo ao fato de que S^{n-1} não é um *retrato* da bola $B^n = B[0,1] \subset \mathbb{R}^n$.

Isto significa que não existe uma *retração* de B^n em S^{n-1}, isto é, uma aplicação contínua $\rho : B^n \to S^{n-1}$, tal que $\rho(x) = x\,\forall x \in S^{n-1}$.

Observe que, geometricamente, é fácil ver que não se pode transformar continuamente a bola B^n na esfera S^{n-1} mantendo-se esta última fixa. Usaremos, então, o Teorema de Borsuk-Ulam para estabelecer este fato algebricamente, conforme o lema seguinte.

LEMA 2. *S^{n-1} não é um retrato de B^n.*

DEMONSTRAÇÃO. Suponhamos, por absurdo, que exista uma retração $\rho : B^n \to S^{n-1}$ e consideremos a decomposição $\mathbb{R}^{n+1} = \mathbb{R}^n \times \mathbb{R}$. Definamos, então, a função $\varphi : S^n \to \mathbb{R}^n$, em que

$$\varphi(x,t) = \begin{cases} \rho(x) & \text{se} \quad t \geq 0 \\ x & \text{se} \quad t < 0. \end{cases}$$

Uma vez que ρ, bem como a projeção $(x,t) \to x$, são aplicações contínuas, e a restrição $\rho|_{S^{n-1}}$ é a aplicação identidade de S^{n-1}, temos que φ é contínua.

Considerando-se, então, a igualdade

$$\varphi(x,t) = \varphi(-x,-t),\ (x,t) \in S^n, \tag{28}$$

temos que, para $t = 0$, ela equivale a $\rho(x) = \rho(-x)$, $x \in S^{n-1}$, a qual nunca se cumpre, pois, qualquer que seja $x \in S^{n-1}$, tem-se $\rho(x) = x \neq -x = \rho(-x)$. Agora, para $t \neq 0$, (28) nos dá $\rho(x) = -x$ (se $t > 0$) ou $\rho(-x) = x$ (se $t < 0$). Entretanto, nenhuma destas duas igualdades pode ocorrer, pois, qualquer que seja o ponto $(x,t) \in S^n$ cuja coordenada t é não-nula, tem-se $\|x\| < 1 = \|\rho(x)\|$ (Fig. 8).

Logo, não existe $(x,t) \in S^n$, tal que $\varphi(x,t) = \varphi(-x,-t)$, o que contradiz o Teorema de Borsuk-Ulam. Segue-se desta contradição que a retração ρ não existe e, portanto, que S^{n-1} não é um retrato de B^n, como desejávamos provar. □

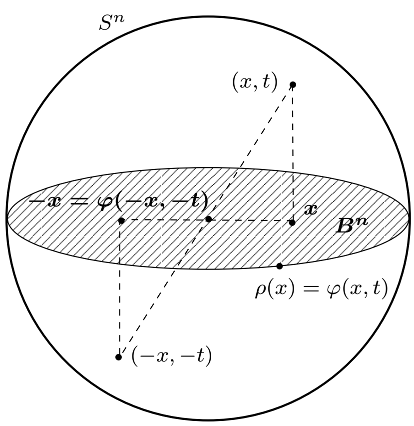

FIGURA 8

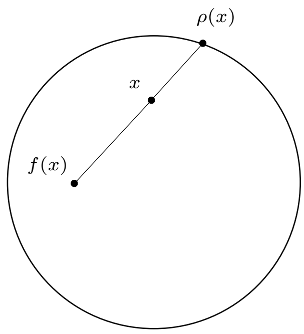

FIGURA 9

TEOREMA DO PONTO FIXO (BROUWER). *Toda aplicação contínua $f : B^n \to B^n$ possui um ponto fixo.*

Demonstração. Suponhamos, por absurdo, que f não possua pontos fixos. Neste caso, podemos definir a aplicação $\rho : B^n \to S^{n-1}$, em que $\rho(x) \in S^{n-1}$ é tal que $x \in [f(x), \rho(x)]$ (Fig. 9). Mostraremos que ρ, assim definida, é uma retração, o que irá de encontro ao Lema 2 e provará, desta forma, o teorema.

A fim de descrever a aplicação ρ algebricamente e, desta forma, concluir o desejado, façamos, para cada $x \in B^n$, $u = u(x) = (x - f(x))/\|x - f(x)\|$. Devemos, então, determinar um número real $\lambda = \lambda(x) \geq 0$, tal que

$$\rho(x) = x + \lambda u$$

seja um vetor de S^{n-1}. Assim, devemos ter $\|x + \lambda u\|^2 = 1$, o que nos dá a equação quadrática

$$\lambda^2 + 2\langle x, u\rangle\lambda - (1 - \|x\|^2) = 0,$$

cuja única raiz não-negativa é

$$\lambda = \sqrt{\langle x, u\rangle^2 + 1 - \|x\|^2} - \langle x, u\rangle.$$

Claramente, λ é uma função contínua de x, donde ρ é uma retração. $\square$

8. Exercícios

Seção 1

1. Sejam $f, g : X \subset \mathbb{R}^n \to \mathbb{R}^m$ aplicações contínuas no ponto $a \in X$. Prove que se $f(a) \neq g(a)$, então existe uma bola aberta $B = B(a, r) \subset \mathbb{R}^n$, tal que

$$x \in X \cap B \Rightarrow f(x) \neq g(x).$$

2. Mostre que uma aplicação $f : \mathbb{R}^n \to \mathbb{R}^m$ é contínua se, e somente se, para todo conjunto $Y \subset \mathbb{R}^m$, tem-se $\partial f^{-1}(Y) \subset f^{-1}(\partial Y)$.

3. Dada uma aplicação contínua $f : \mathbb{R}^n \to \mathbb{R}^m$, prove que, para todo $X \subset \mathbb{R}^n$, $f(\overline{X}) \subset \overline{f(X)}$, e que esta inclusão reduz-se a uma igualdade no caso em que f é fechada.

4. Sejam $F \subset \mathbb{R}^n$ fechado e $f : F \to \mathbb{R}^m$ uma aplicação contínua. Mostre que f leva subconjuntos limitados de F em subconjuntos limitados de $\mathbb{R}^m$. Prove, exibindo um contra-exemplo, que não se conclui o mesmo removendo-se a hipótese de F ser fechado.

5. Seja $f : \mathbb{R}^n \to \mathbb{R}^m$ uma aplicação contínua e sobrejetiva. Mostre que se $X \subset \mathbb{R}^n$ é denso em $\mathbb{R}^n$, então $f(X)$ é denso em $\mathbb{R}^m$.

6. Prove que uma aplicação $T : \mathbb{R}^n \to \mathbb{R}^m$ é linear se é contínua e satisfaz

$$T(x + y) = T(x) + T(y)\ \forall x, y \in \mathbb{R}^n.$$

7. Seja $B = B(0, 1)$ a bola aberta unitária de $\mathbb{R}^n$ com centro na origem. Suponha que $f : B \to B$ seja uma aplicação contínua que satisfaz $\|f(x)\| < \|x\|$ qualquer que seja $x \in B - \{0\}$. Dado $x_1 \in B$, considere a sequência (x_k), em B, tal que $x_{k+1} = f(x_k)$. Prove que $\lim x_k = 0$.

8. Suponha que $f, g : X \subset \mathbb{R}^n \to \mathbb{R}$ sejam contínuas. Prove que são contínuas as funções $\mu(x) = \min\{f(x), g(x)\}$ e $\lambda(x) = \max\{f(x), g(x)\}$, $x \in X$.

Seção 2

9. Dadas funções limitadas e uniformemente contínuas, $f, g : X \subset \mathbb{R}^n \to \mathbb{R}$, prove que fg é uniformemente contínua. Prove, em seguida, que, para todo $r > 0$, a função

$$f(x_1, \dots, x_n) = \frac{x_1 \dots x_n}{(x_1^2 + \cdots + x_n^2)^{\frac{n}{2}}}\,, \ (x_1, \dots, x_n) \in \mathbb{R}^n - B(0, r),$$

é uniformemente contínua.

10. Use a função distância (a um conjunto) para mostrar o *Teorema de Urysohn* em $\mathbb{R}^n$: Dados dois subconjuntos de $\mathbb{R}^n$ fechados e disjuntos, F_1 e F_2, existe uma função contínua $f : \mathbb{R}^n \to [0, 1]$, tal que $f(x) = 0\, \forall x \in F_1$ e $f(x) = 1\, \forall x \in F_2$.

11. Seja $U \subset \mathbb{R}^n$ um subconjunto aberto, próprio e não-vazio de $\mathbb{R}^n$. Mostre que existe uma função $f : \mathbb{R}^n \to \mathbb{R}$, tal que f é descontínua em todos os pontos de U e uniformemente contínua e não-constante em $\mathbb{R}^n - U$.

Seção 3

12. Mostre que são homeomorfos:
    i) O cone $C = \{(x, y, z) \in \mathbb{R}^3;\, z = \sqrt{x^2 + y^2}\}$ e $\mathbb{R}^2$;
    ii) $\mathbb{R}^{n+1} - \{0\}$ e o cilindro $S^n \times \mathbb{R} \subset \mathbb{R}^{n+2}$.

13. Sejam $U \subset \mathbb{R}^n$ aberto e $f : U \to \mathbb{R}^n$ um homeomorfismo uniformemente contínuo. Prove que $U = \mathbb{R}^n$.

14. Sejam $\|\ \|_0$ e $\|\ \|_1$ duas normas em $\mathbb{R}^n$ e B_0, B_1 bolas com respeito a $\|\ \|_0$ e $\|\ \|_1$, respectivamente. Prove que se B_0 e B_1 são ambas abertas ou ambas fechadas, então ela são homeomorfas.

15. Diz-se que um conjunto $X \subset \mathbb{R}^n$ é *topologicamente homogêneo* se, para quaisquer $x, y \in X$, existe um homeomorfismo $f : X \to X$, tal que $f(x) = y$. Prove que são topologicamente homogêneos:
    i) O intervalo $(-1, 1) \subset \mathbb{R}$;
    ii) a bola aberta unitária $B_{\max}(0, 1) = \{x \in \mathbb{R}^n; \|x\|_{\max} < 1\}$.

    Conclua daí e do exercício anterior que, para toda norma $\|\ \|_0$ em $\mathbb{R}^n$, qualquer bola aberta com respeito a $\|\ \|_0$ é topologicamente homogênea.

16. Prove que:
    i) Existe um homeomorfismo $\lambda : (-1, 1) \subset \mathbb{R} \to (-1, 1)$ sem pontos fixos;
    ii) existe um homeomorfismo $f : B_{\max}(0, 1) \subset \mathbb{R}^n \to B_{\max}(0, 1)$ sem pontos fixos.

    Conclua daí e do Exercício 14 que, para toda norma $\|\ \|_0$ em $\mathbb{R}^n$ e toda bola aberta B_0 com respeito a $\|\ \|_0$, existe um homeomorfismo $f : B_0 \to B_0$ sem pontos fixos.

Seção 4

17. Prove a recíproca do Teorema de Weierstrass: Se $K \subset \mathbb{R}^n$ é tal que, toda função contínua $f : K \to \mathbb{R}$ é limitada, então K é compacto.

18. Sejam $K \subset \mathbb{R}^n$ compacto e $f : K \to \mathbb{R}^m$ contínua. Mostre que, para todo $\epsilon > 0$, existe $\mu > 0$, tal que
$$\|f(x) - f(y)\| \le \mu\|x - y\| + \epsilon \ \forall x, y \in K.$$

19. Considere um conjunto compacto $K \subset \mathbb{R}^n$. Prove que uma aplicação $f : K \to \mathbb{R}^m$ é contínua se, e somente se, o gráfico de f é compacto.

20. Sejam $X \subset \mathbb{R}^n$ e $f : X \to X$ uma isometria de X, isto é,
$$\|f(x) - f(y)\| = \|x - y\| \ \forall x, y \in X.$$
Dado $x_0 \in X$, prove que a sequência (x_k), em que $x_k = f^k(x_0)$ [(iv)], satisfaz
$$\|x_{k+l} - x_k\| \ge d(x_0, f(X)) \,\forall k, l \in \mathbb{N}.$$
Conclua que f é um homeomorfismo se X for compacto.

21. Uma aplicação $f : \mathbb{R}^n \to \mathbb{R}^m$ é dita *própria* se é contínua e, para todo compacto $K \subset \mathbb{R}^m$, $f^{-1}(K)$ é compacto. Prove que:
    i) Toda aplicação própria $f : \mathbb{R}^n \to \mathbb{R}^m$ é fechada;
    ii) toda aplicação $f : \mathbb{R}^n \to \mathbb{R}^m$, contínua, injetiva e fechada, é própria.

Seção 5

22. Prove que o conjunto $\mathrm{O}(n)$, das matrizes ortogonais de ordem n, não é conexo.

23. Prove que um conjunto $X \subset \mathbb{R}^n$ é conexo se, e somente se, toda função contínua $f : X \to \{0, 1\}$ é constante. Demonstre, a partir deste fato, que:
    i) Qualquer união de uma família de conexos de $\mathbb{R}^n$ com um ponto em comum é conexa;
    ii) o fecho de um subconjunto conexo de $\mathbb{R}^n$ é conexo.

24. Sejam $f : \mathbb{R}^n \to \mathbb{R}$ uma função definida em $\mathbb{R}^n$ e $\mathrm{graf}(f)$ o seu gráfico. Mostre que uma, e somente uma, das implicações abaixo é verdadeira:
    i) f contínua $\Rightarrow$ $\mathrm{graf}(f)$ conexo;
    ii) $\mathrm{graf}(f)$ conexo $\Rightarrow$ f contínua.

25. Mostre que S^1 não é homeomorfo a S^2.

26. Prove as seguintes afirmações:
    i) $X \subset \mathbb{R}^n$ e $Y \subset \mathbb{R}^m$ são conexos por caminhos se, e somente se, $X \times Y$ é conexo por caminhos;
    ii) se $\{X_\lambda\}_{\lambda \in \Lambda}$ é uma família de subconjuntos de $\mathbb{R}^n$, todos conexos por caminhos, tal que $\bigcap X_\lambda \ne \emptyset$, então $\bigcup X_\lambda$ é conexo por caminhos;
    iii) todo subconjunto de $\mathbb{R}^n$ que é aberto e conexo é conexo por caminhos.

[(iv)] Aqui, f^k denota a aplicação obtida por k composições sucessivas da aplicação f consigo mesma.

Seção 6

27. Dados $X \subset \mathbb{R}^n$ e $a \in X'$, suponha que as aplicações $f : X \to \mathbb{R}^n$ e $\Phi : X \to \mathrm{L}(\mathbb{R}^n, \mathbb{R}^m)$ cumpram
$$\lim_{x\to a} f(x) = x_0 \quad \text{e} \quad \lim_{x\to a} \Phi(x) = T.$$
Prove que $\lim\limits_{x\to a} \Phi(x)f(x) = Tx_0$.

28. Sejam $f, g : X \subset \mathbb{R}^n \to \mathbb{R}$ funções, tais que
$$\lim_{x\to a} f(x) = \lambda_1 \quad \text{e} \quad \lim_{x\to a} g(x) = \lambda_2\,.$$
Prove que se $\lambda_1 < \lambda_2$, então existe uma vizinhança de a em X, V, tal que $f(x) < g(x)\ \forall\, x \in V - \{a\}$.

29. Considere uma aplicação $f : X \subset \mathbb{R}^n \to Y \subset \mathbb{R}^m$ e suponha que, para todo $a \in X' \neq \emptyset$, existe o limite $\lim_{x\to a} f(x)$. Mostre que se X é limitado, então $f(X)$ é limitado.

30. Seja $f : X \subset \mathbb{R}^n \to \mathbb{R}^m$ uma aplicação uniformemente contínua. Prove que:
    i) Para todo $a \in X'$, existe o limite $\lim\limits_{x\to a} f(x)$;
    ii) se X é denso em $\mathbb{R}^n$, então existe uma única aplicação contínua $\varphi : \mathbb{R}^n \to \mathbb{R}^m$, tal que $f = \varphi|_X$ e, além disso, φ é uniformemente contínua.

CAPíTULO 4

Aplicações Diferenciáveis

Iniciaremos, neste capítulo, o estudo de um dos conceitos mais importantes da Análise, o de derivada. No caso das funções reais de uma variável real, como é sabido, a derivada está intimamente ligada à noção geométrica de reta tangente e também à de taxa de variação entre grandezas.

Em virtude destas características, a derivada incide em inúmeras teorias matemáticas, desempenhando, inclusive, um papel fundamental na modelagem de diversos fenômenos naturais, como o movimento dos planetas ou a desintegração de substâncias radioativas.

Ainda no caso das funções de uma variável, a derivada admite uma outra interpretação, que parte da ideia de se considerar a reta tangente ao gráfico de uma função (diferenciável) num determinado ponto, como o gráfico de uma função linear que a aproxima numa vizinhança deste ponto (vide Seção 1, abaixo). Valendo-se disto, estende-se, de forma natural, o conceito de diferenciabilidade às aplicações entre espaços euclidianos arbitrários, isto é, uma tal aplicação é diferenciável num ponto de seu domínio, quando existe uma determinada transformação linear que a aproxima numa vizinhança deste ponto, a qual chama-se derivada da aplicação neste ponto.

Essencialmente, o conceito de derivada caracteriza-se pelo fato de permitir, em muitos aspectos, deduzir o comportamento da aplicação numa vizinhança de um ponto a partir das propriedades de uma transformação linear — a derivada da aplicação neste ponto —, as quais são de fácil verificação. A riqueza e força deste conceito generalizado de derivada, dentre outros fatos, reflete-se na profusa atividade científica de teorias como Geometria Diferencial, Equações Diferenciais e suas correlatas.

Ao longo deste capítulo, após sua definição, deduziremos algumas propriedades elementares da derivada e introduziremos as noções de derivada parcial e derivada de ordem superior. Estabeleceremos, em seguida, o resultado mais fundamental ligado à derivada, a Regra da Cadeia, e introduziremos o conceito de difeomorfismo, o qual estabelece uma relação de equivalência entre os conjuntos abertos de $\mathbb{R}^n$. A título de aplicação, encerraremos apresentando um dos teoremas fundamentais da Análise Convexa, conhecido como Teorema de Motzkin. Reservamos, então, ao próximo capítulo, a apresentação dos principais teoremas do cálculo diferencial das aplicações entre espaços euclidianos.

1. Diferenciabilidade em $\mathbb{R}^n$

Relembremos que uma função $f : I \subset \mathbb{R} \to \mathbb{R}$, definida num intervalo aberto I, é dita *diferenciável* num ponto $x \in I$ se existe o limite

$$\lim_{h\to 0} \frac{f(x+h)-f(x)}{h}, \tag{29}$$

o qual, no caso afirmativo, chamamos *derivada* de f em x e denotamos por $f'(x)$.

Nesta definição, a estrutura algébrica de $\mathbb{R}$, qual seja, a de corpo, desempenha um papel fundamental, visto que a expressão (29) envolve uma divisão. Uma vez que, em geral, a divisão de vetores não está definida, a fim de generalizar-se este conceito de diferenciabilidade para aplicações definidas em $\mathbb{R}^n$, convém reinterpretar a derivada em $\mathbb{R}$ a partir de sua estrutura vetorial.

Para isto, consideremos uma vizinhança $V_\delta = (x-\delta, x+\delta) \subset I$, de x em I, e a função

$$\omega(h) = f(x+h) - f(x), \ \ h \in (-\delta, \delta),$$

cujo comportamento, vale salientar, reflete aquele da função f em V_δ [(i)].

Dado $a \in \mathbb{R}$, definamos a *função resto*, $r : (-\delta,\delta) \to \mathbb{R}$, por

$$r(h) = \omega(h) - ah, \tag{30}$$

a qual pode ser vista como uma medida do "erro" cometido quando se substitui a função ω pela função linear $h \mapsto ah$.

Supondo-se f contínua em x, tem-se $\lim_{h\to 0} r(h) = 0$. Neste caso, diz-se que o resto r é um *infinitésimo* e que a função $h \mapsto ah$ é uma *aproximação* de ω em $(-\delta,\delta)$ (ou, equivalentemente, de f em V_δ).

Assim, podemos dizer que, quando f é contínua em x, toda função linear restrita a $(-\delta,\delta)$ constitui uma aproximação de f em V_δ. No entanto, quando f é diferenciável em x, a aproximação linear definida pela derivada de f em x, isto é, a função $h \mapsto f'(x)h$, define uma função resto r, tal que $r(h)$ tende a zero mais rapidamente que h, sendo, inclusive, a única aproximação linear de f a ter esta propriedade. Mais precisamente, tem-se a equivalência seguinte, a qual se obtém facilmente dividindo-se ambos os membros de (30) por h e tomando-se o limite com h tendendo a zero.

$$\lim_{h\to 0} \frac{r(h)}{h} = 0 \quad \Leftrightarrow \quad f \text{ é diferenciável em } x \text{ e } a = f'(x).$$

Quando uma função r cumpre o limite acima, diz-se dela que é um *infinitésimo de ordem superior a* 1 (com respeito à variável h). Segue-se, portanto, desta equivalência, que, para que f seja diferenciável em x, é necessário e suficiente que exista uma aproximação linear de f numa vizinhança de x, a qual define uma função resto que é um infinitésimo de ordem superior a 1 (Fig. 1).

Observando-se ainda que

$$\lim_{h\to 0} \frac{r(h)}{h} = 0 \ \ \Leftrightarrow \ \ \lim_{h\to 0} \frac{r(h)}{|h|} = 0,$$

obtém-se destas considerações o resultado seguinte.

[(i)] Por exemplo, se ω tem limite nulo quando h tende a zero, então f é contínua em x, se ω é não-negativa, então $f|_{V_\delta}$ assume um valor mínimo em x, etc.

Teorema 16 (derivada como aproximação linear). *Dado um intervalo aberto $I \subset \mathbb{R}$, uma função $f : I \to \mathbb{R}$ é diferenciável em $x \in I$ se, e somente se, existe uma função linear $h \mapsto ah$, $a \in \mathbb{R}$, tal que a função $r = r(h)$, definida pela igualdade $f(x+h) - f(x) = ah + r(h)$, cumpre*

$$\lim_{h \to 0} \frac{r(h)}{|h|} = 0.$$

No caso afirmativo, tem-se $a = f'(x)$.

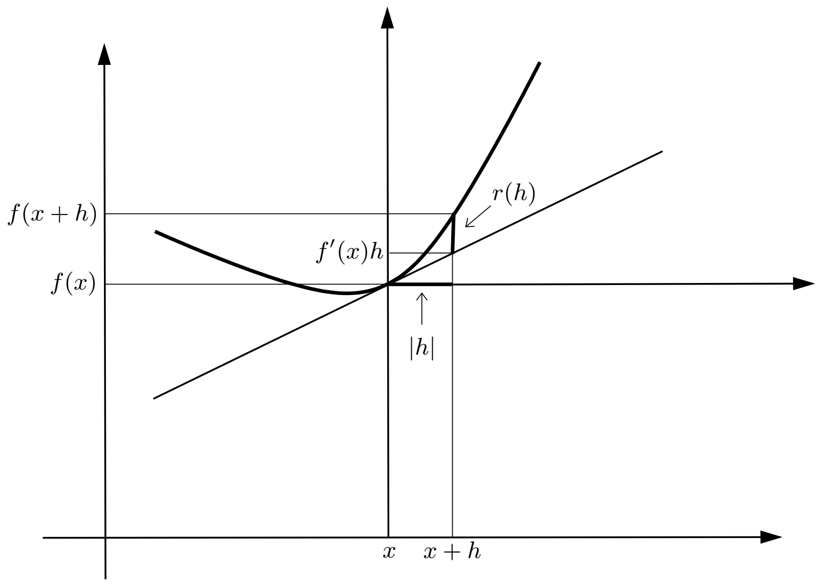

Figura 1

No teorema acima, a única estrutura considerada em $\mathbb{R}$ é a de espaço vetorial normado. Assim, faz-se natural a definição seguinte.

Definição 23 (diferenciabilidade). Seja $U \subset \mathbb{R}^n$ um conjunto aberto. Dizemos que uma aplicação $f : U \to \mathbb{R}^m$ é *diferenciável* em $x \in U$ se existe uma transformação linear $T : \mathbb{R}^n \to \mathbb{R}^m$, tal que a aplicação *resto*, $r = r(h)$, definida pela igualdade

$$f(x+h) - f(x) = Th + r(h),$$

satisfaz

$$\lim_{h \to 0} \frac{r(h)}{\|h\|} = 0.$$

A aplicação f é dita *diferenciável* em $X \subset U$ quando é diferenciável em cada ponto de X. Diz-se, simplesmente, que f é *diferenciável* quando é diferenciável em U.

É essencial, nesta definição, que o conjunto U seja aberto. De fato, neste caso, para todo $x \in U$, existe $\delta > 0$, tal que $B(x, \delta) \subset U$. Assim, para todo $h \in \mathbb{R}^n$ satisfazendo $\|h\| < \delta$, tem-se $x + h \in U$ e, portanto, $f(x+h)$ está bem definido.

A aplicação T da Definição 23, quando existe, é única. Isto decorre do seguinte fato. Dado $t \in \mathbb{R}$, $t \neq 0$, tem-se,

$$Th = \frac{T(th)}{t} = \frac{f(x+th)-f(x)}{t} \pm \frac{r(th)}{\|th\|}\|h\|.$$

Logo,

$$Th = \lim_{t\to 0} \frac{f(x+th)-f(x)}{t}, \tag{31}$$

o que prova a unicidade de T, uma vez que, nesta igualdade, o segundo membro depende apenas de f, x e h.

A aplicação linear T é, então, chamada de *derivada* de f em x e é denotada por $f'(x)$.

Segue-se imediatamente da definição de diferenciabilidade que qualquer aplicação constante $f : U \subset \mathbb{R}^n \to \mathbb{R}^m$ é diferenciável, sendo a derivada, em cada $x \in U$, a transformação linear nula $f'(x) = 0$.

Exemplo 55. Verifiquemos que a aplicação

$$\begin{array}{cccc} f: & \mathbb{R}^2 & \to & \mathbb{R}^2 \\ & (x,y) & \mapsto & (y,x^2) \end{array}$$

é diferenciável em $(1,2)$ e determinemos a sua derivada neste ponto. Para isto, devemos inicialmente encontrar um candidato a derivada, isto é, uma transformação linear $T : \mathbb{R}^2 \to \mathbb{R}^2$ que cumpra as condições da definição. Fazendo-se, então, $(x,y) = (1,2)$ e $h = (h_1,h_2)$, se f for diferenciável em $(1,2)$, teremos, por (31),

$$Th = T(h_1,h_2) \;=\; \lim_{t\to 0} \frac{f\left((1,2)+t(h_1,h_2)\right)-f(1,2)}{t} \;=\; (h_2, 2h_1).$$

Neste caso, a aplicação resto é dada por

$$r(h) = r(h_1,h_2) = f((1,2)+(h_1,h_2)) - f(1,2) - T(h_1,h_2) = (0,h_1^2).$$

Uma vez que

$$\frac{h_1^2}{\|(h_1,h_2)\|} = |h_1|\frac{|h_1|}{\|(h_1,h_2)\|} \leq |h_1|,$$

tem-se

$$\lim_{h\to 0}\frac{r(h)}{\|h\|} = 0.$$

Desta forma, f é diferenciável em $(1,2)$ e $f'(1,2) : \mathbb{R}^2 \to \mathbb{R}^2$ é a transformação linear dada por $f'(1,2)(h_1,h_2) = (h_2, 2h_1)$.

Observação 15 (diferenciabilidade implica continuidade). Consideremos uma aplicação $f : U \subset \mathbb{R}^n \to \mathbb{R}^m$ definida num aberto $U \subset \mathbb{R}^n$ e difereniciável em $x \in U$. Neste caso, temos que a aplicação resto, $r(h) = f(x+h) - f(x) - f'(x)h$, satisfaz $\lim_{h\to 0} r(h)/\|h\| = 0$. Logo,

$$\lim_{h\to 0} r(h) = \lim_{h\to 0}\frac{r(h)}{\|h\|}\|h\| = 0.$$

Sendo assim,

$$\lim_{h\to 0} f(x+h) = \lim_{h\to 0}\left(f(x) + f'(x)h + r(h) \right) = f(x),$$

donde f é contínua em x.

Observação 16 (derivada direcional e diferenciabilidade). O limite indicado na igualdade (31), quando existe, é chamado de *derivada direcional* de f em x na direção h e é denotado por $\frac{\partial f}{\partial h}(x)$. Desta forma, se f é diferenciável em x, a derivada direcional $\frac{\partial f}{\partial h}(x)$ está bem definida e satisfaz

$$\frac{\partial f}{\partial h}(x) = f'(x)h.$$

No entanto, a existência das derivadas direcionais em x não assegura a diferenciabilidade de f neste ponto. Com efeito, considere a função $f : \mathbb{R}^2 \to \mathbb{R}$,

$$f(x,y) = \begin{cases} \frac{x^2y}{x^2+y^2} & \text{se} \quad (x,y) \neq (0,0), \\ 0 & \text{se} \quad (x,y) = (0,0). \end{cases}$$

Fazendo-se $h = (h_1, h_2) \in \mathbb{R}^2$, tem-se

$$\frac{\partial f}{\partial h}(0,0) = \frac{h_1^2 h_2}{h_1^2 + h_2^2}. \tag{32}$$

Segue-se que f não é diferenciável em $(0,0)$, pois, se o fosse, valeria a igualdade $\frac{\partial f}{\partial h}(0,0) = f'(0,0)h$. Entretanto, conforme indicado em (32), $\frac{\partial f}{\partial h}(0,0)$ não é uma função linear de h.

Vale salientar que, para uma dada aplicação $f : U \subset \mathbb{R}^n \to \mathbb{R}^m$, mesmo que a aplicação $h \mapsto T(h) = \frac{\partial f}{\partial h}(x)$ esteja bem definida e seja linear, não se pode concluir daí que f seja diferenciável em x (vide Exercício 2).

Feitas as devidas alterações, introduz-se, de forma análoga à que fizemos acima, o conceito de diferenciabilidade de aplicações entre espaços vetoriais normados arbitrários. Neste sentido, podemos investigar, por exemplo, a diferenciabilidade da aplicação

$$\begin{array}{rccc} f: & \mathrm{M}(n) & \to & \mathrm{M}(n) \\ & X & \mapsto & X^2. \end{array}$$

Com este fim, tomemos $X \in \mathrm{M}(n)$ e tentemos obter uma "parte linear" da expressão $f(X+H) - f(X)$, que será, desta forma, um candidato a derivada de f em X. Temos que

$$f(X+H) - f(X) = (X+H)^2 - X^2 = XH + HX + H^2.$$

Fazendo-se, então, $TH = XH + HX$, $H \in \mathrm{M}(n)$, tem-se que T é linear e a aplicação resto definida por T é $R(H) = H^2$. Logo, para $H \neq 0$,

$$\frac{\|R(H)\|}{\|H\|} = \frac{\|H^2\|}{\|H\|} \leq \frac{\|H\|^2}{\|H\|} = \|H\|,$$

donde

$$\lim_{H \to 0} \frac{\|R(H)\|}{\|H\|} = 0.$$

Segue-se que f é diferenciável em $\mathrm{M}(n)$ e, para todo $X \in \mathrm{M}(n)$,

$$f'(X)H = XH + HX, \ H \in \mathrm{M}(n).$$

O teorema seguinte, atribuído ao matemático francês Jacques Hadamard (1865–1963), estabelece a equivalência entre a diferenciabilidade de uma aplicação (num ponto) e a existência e continuidade (no mesmo ponto) de uma outra aplicação, a ela vinculada, que toma valores num espaço de transformações lineares. Este resultado, devidamente aplicado, simplifica consideravelmente diversas demonstrações que envolvem a questão de uma aplicação ser ou não diferenciável num determinado ponto de seu domínio. Por esta razão, ele se prestará à demonstração de teoremas importantes que apresentaremos, como a Regra da Cadeia, neste capítulo, e o Teorema da Função Inversa, no capítulo seguinte.

Teorema de Hadamard. *Uma aplicação* $f : U \subset \mathbb{R}^n \to \mathbb{R}^m$, *definida num aberto* U *de* $\mathbb{R}^n$, *é diferenciável em* $x \in U$ *se, e somente se, existe uma aplicação* $\Phi : U \to \mathrm{L}(\mathbb{R}^n, \mathbb{R}^m)$, *contínua em* x, *e que satisfaz*

$$f(x+h) - f(x) = \Phi(x+h)h, \;\; x+h \in U.$$

No caso afirmativo, tem-se $\Phi(x) = f'(x)$.

Demonstração. Suponhamos que f seja diferenciável em $x \in U$ e, para cada $y = x + h$ de U, definamos a transformação linear $\Phi(x+h) \in \mathrm{L}(\mathbb{R}^n, \mathbb{R}^m)$ por

$$\Phi(x+h)v = \begin{cases} f'(x)v + \left\langle v, \frac{h}{\|h\|} \right\rangle \frac{r(h)}{\|h\|} & \text{se} \quad h \neq 0 \\ f'(x)v & \text{se} \quad h = 0, \end{cases}$$

em que $r = r(h)$ é a aplicação resto, $r(h) = f(x+h) - f(x) - f'(x)h$.

A aplicação $\Phi : U \to \mathrm{L}(\mathbb{R}^n, \mathbb{R}^m)$, assim definida, cumpre

$$\Phi(x+h)h = f'(x)h + r(h) = f(x+h) - f(x).$$

Além disso, uma vez que $\lim_{h\to 0} r(h)/\|h\| = 0$, tem-se, para todo $v \in \mathbb{R}^n$,

$$\lim_{h\to 0} \Phi(x+h)v = f'(x)v.$$

Daí e da Proposição 44 do Capítulo 3, obtém-se

$$\lim_{h\to 0} \Phi(x+h) = f'(x) = \Phi(x),$$

donde Φ é contínua em x.

Reciprocamente, suponhamos que exista uma aplicação $\Phi : U \to \mathrm{L}(\mathbb{R}^n, \mathbb{R}^m)$ que é contínua em $x \in U$ e satisfaz $f(x+h) - f(x) = \Phi(x+h)h$ para todo $x + h \in U$. Fazendo-se, então,

$$r(h) = f(x+h) - f(x) - \Phi(x)h,$$

tem-se $r(h) = (\Phi(x+h) - \Phi(x))h$ e, desta forma,

$$\frac{\|r(h)\|}{\|h\|} = \left\| (\Phi(x+h) - \Phi(x)) \frac{h}{\|h\|} \right\| \leq \|\Phi(x+h) - \Phi(x)\| \;\; \forall h \neq 0.$$

Daí e da continuidade de Φ em x, segue-se que $\lim_{h\to 0} r(h)/\|h\| = 0$, donde se conclui que f é diferenciável em x e $f'(x) = \Phi(x)$. □

De toda aplicação $\Phi : U \to \mathrm{L}(\mathbb{R}^n, \mathbb{R}^m)$ que satisfaz as condições do Teorema de Hadamard, diremos que é uma *representação de Hadamard* de f em x.

Como primeira aplicação do Teorema de Hadamard, verifiquemos que diferenciabilidade é uma condição preservada pelas operações de soma de funções, multiplicação de escalar por função e quociente de funções, conforme a proposição seguinte.

Proposição 45 (propriedades operatórias da derivada). *Sejam* $U \subset \mathbb{R}^n$ *aberto,* $f, g : U \to \mathbb{R}^m$ *aplicações diferenciáveis em* $x \in U$ *e* $\xi : U \to \mathbb{R}$ *uma função diferenciável em* x*. Então, as aplicações* $(f+g), (\xi f) : U \to \mathbb{R}^m$ *e a função* $(1/\xi) : U \to \mathbb{R}$ *(quando definida) são diferenciáveis em* x*. Além disso,*

i) $(f+g)'(x) = f'(x) + g'(x)$;

ii) $(\xi f)'(x) = \xi'(x) f(x) + \xi(x) f'(x)$ [(ii)];

iii) $(\frac{1}{\xi})'(x) = -\frac{\xi'(x)}{(\xi(x))^2}$.

Demonstração. A propriedade relativa à soma $f+g$ é de demonstração imediata e a deixamos a cargo do leitor.

Sejam $\Gamma : U \to \mathrm{L}(\mathbb{R}^n, \mathbb{R})$ e $\Phi : U \to \mathrm{L}(\mathbb{R}^n, \mathbb{R}^m)$ representações de Hadamard de ξ e f, respectivamente, no ponto x. Temos, então, que Γ e Φ são contínuas em x, satisfazem $\Gamma(x) = \xi'(x)$, $\Phi(x) = f'(x)$, e cumprem as igualdades

$$\xi(x+h) - \xi(x) = \Gamma(x+h)h \quad \text{e} \quad f(x+h) - f(x) = \Phi(x+h)h.$$

Assim,

$$\begin{aligned}(\xi f)(x+h) - (\xi f)(x) &= \xi(x+h) f(x+h) - \xi(x) f(x) \\ &= (\xi(x+h) - \xi(x)) f(x+h) + \xi(x)(f(x+h) - f(x)) \\ &= (\Gamma(x+h)h) f(x+h) + \xi(x)\Phi(x+h)h.\end{aligned}$$

Definindo-se, então, $\Psi(x+h) \in \mathrm{L}(\mathbb{R}^m, \mathbb{R}^n)$ por

$$\Psi(x+h)v = (\Gamma(x+h)v) f(x+h) + \xi(x)\Phi(x+h)v, \;\; v \in \mathbb{R}^n,$$

segue-se que $(\xi f)(x+h) - (\xi f)(x) = \Psi(x+h)h$ e que Ψ é contínua em x, pois f, Γ e Φ o são. Logo, Ψ é uma representação de Hadamard de (ξf) em x, donde (ξf) é diferenciável em x e $(\xi f)'(x)$ satisfaz, para todo $v \in \mathbb{R}^n$,

$$(\xi f)'(x)v = \Psi(x)v = (\Gamma(x)v) f(x) + \xi(x)\Phi(x)v = (\xi'(x)v) f(x) + \xi(x) f'(x)v,$$

o que prova (ii).

Supondo-se agora que ξ nunca se anule em U, tem-se

$$\left(\frac{1}{\xi}\right)(x+h) - \left(\frac{1}{\xi}\right)(x) = \frac{1}{\xi(x+h)} - \frac{1}{\xi(x)} = -\frac{\xi(x+h) - \xi(x)}{\xi(x+h)\xi(x)} = -\frac{\Gamma(x+h)h}{\xi(x+h)\xi(x)}\,.$$

Daí, infere-se que a aplicação $\Theta : U \to \mathrm{L}(\mathbb{R}^n, \mathbb{R})$, em que

$$\Theta(x+h) = -\frac{\Gamma(x+h)}{\xi(x+h)\xi(x)}\,,$$

é uma representação de Hadamard de $\frac{1}{\xi}$ em x. Logo, $\frac{1}{\xi}$ é diferenciável em x e

$$\left(\frac{1}{\xi}\right)'(x) = \Theta(x) = -\frac{\Gamma(x)}{(\xi(x))^2} = -\frac{\xi'(x)}{(\xi(x))^2}\,.$$

Isto prova (iii) e conclui, desta forma, a demonstração. □

[(ii)] Esta igualdade deve ser interpretada como: $(\xi f)'(x)v = (\xi'(x)v) f(x) + \xi(x)(f'(x)v)$, $v \in \mathbb{R}^n$.

PROPOSIÇÃO 46 (DIFERENCIABILIDADE DAS COORDENADAS DE UMA APLICAÇÃO). *Uma aplicação* $f : U \subset \mathbb{R}^n \to \mathbb{R}^m$ *é diferenciável num ponto* x *do aberto* U *se, e somente se, cada função-coordenada* f_i *é diferenciável em* x*. No caso afirmativo, tem-se* $f_i'(x) = (f'(x))_i$ *, isto é, no ponto* x*, a derivada de cada coordenada de* f *é a coordenada correspondente da derivada de* f*.*

DEMONSTRAÇÃO. Sejam $T : \mathbb{R}^n \to \mathbb{R}^m$ uma transformação linear e r a aplicação resto definida por f, T e x, isto é, $r(h) = f(x+h) - f(x) - Th$. Temos, para todo $i = 1, \dots, m$, que $r_i(h) = f_i(x+h) - f_i(x) - T_i h$. Além disso, pela Proposição 43 do Capítulo 3,

$$\lim_{h\to 0} \frac{r(h)}{\|h\|} = 0 \quad \Leftrightarrow \quad \lim_{h\to 0} \frac{r_i(h)}{\|h\|} = 0 \quad \forall i = 1, \dots, m.$$

Logo, f é diferenciável em x com derivada $f'(x) = T$ se, e somente se, f_i é diferenciável em x com derivada $f_i' = T_i$. □

2. Exemplos Especiais de Aplicações Diferenciáveis

EXEMPLO 56 (NORMA AO QUADRADO). Seja $f : \mathbb{R}^n \to \mathbb{R}$ a função *norma ao quadrado*, $f(x) = \|x\|^2$. Temos que

$$f(x+h) - f(x) = \langle x+h, x+h\rangle - \langle x, x\rangle = 2\langle x, h\rangle + \|h\|^2.$$

Fazendo-se $r(h) = \|h\|^2$, tem-se, claramente, $\lim_{h\to 0} r(h)/\|h\| = 0$, donde f é diferenciável e $f'(x)h = 2\langle x, h\rangle \ \forall x, h \in \mathbb{R}^n$.

EXEMPLO 57 (APLICAÇÕES LINEARES). Seja $T : \mathbb{R}^n \to \mathbb{R}^m$ uma aplicação linear. Segue-se da igualdade $T(x+h) - T(x) = Th$ que, para todo $x \in \mathbb{R}^n$, T é diferenciável em x e $T'(x)h = Th \ \forall h \in \mathbb{R}^n$. Com efeito, neste caso, a aplicação resto, $r = r(h)$, é identicamente nula. Logo, a condição $\lim_{h\to 0} r(h)/\|h\| = 0$ é trivialmente satisfeita.

EXEMPLO 58 (APLICAÇÕES n-LINEARES). Seja $f : \mathbb{R}^{j_1} \times \cdots \times \mathbb{R}^{j_n} \to \mathbb{R}^m$ uma aplicação n-linear. Dados $x = (x_1, \dots, x_n)$, $h = (h_1, \dots, h_n) \in \mathbb{R}^{j_1} \times \cdots \times \mathbb{R}^{j_n}$, tem-se, então,

$$f(x+h) - f(x) = \sum_{i=1}^{n} f(x_1, \dots, x_{i-1}, h_i, x_{i+1}, \dots, x_n) + r(h),$$

em que $r(h)$ é a soma de todos os vetores $v \in \mathbb{R}^m$ que se escrevem como

$$v = v(h) = f(y_1, \dots, y_{i-1}, h_i, y_{i+1}, \dots, y_{j-1}, h_j, y_{j+1}, \dots, y_n),$$

sendo $1 \le i < j \le n$ e, para cada $k \in \{1, i-1, i+1, \dots, j-1, j+1, \dots, n\}$, $y_k = x_k$ ou $y_k = h_k$. Valendo-se, então, do resultado do Exercício 8 do Capítulo 1, tem-se, para um certo $\mu > 0$ e todo $h \ne 0$,

$$\begin{aligned}
\frac{\|v(h)\|}{\|h\|} &= \frac{\|f(y_1, \dots, y_{i-1}, h_i, y_{i+1}, \dots, y_{j-1}, h_j, y_{j+1}, \dots, y_n)\|}{\|h\|} \\
&\le \frac{\mu\|y_1\| \dots \|y_{i-1}\| \, \|h_i\| \, \|y_{i+1}\| \dots \|y_{j-1}\| \, \|h_j\| \, \|y_{j+1}\| \dots \|y_n\|}{\|h\|} \\
&\le \mu\|y_1\| \dots \|y_{i-1}\| \, \|y_{i+1}\| \dots \|y_{j-1}\| \, \|h_j\| \, \|y_{j+1}\| \dots \|y_n\|,
\end{aligned}$$

pois $\|h_i\| \le \|h\| \, \forall i \in \{1, \dots, n\}$.

Segue-se que $\lim_{h\to 0} v(h)/\|h\| = 0$, donde $\lim_{h\to 0} r(h)/\|h\| = 0$. Logo, f é diferenciável em x e

$$f'(x)h = \sum_{i=1}^{n} f(x_1, \dots, x_{i-1}, h_i, x_{i+1}, \dots, x_n).$$

Em particular, são diferenciáveis:

- o produto interno $\langle\, , \rangle : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$;
- o determinante $\det : \mathbb{R}^n \times \cdots \times \mathbb{R}^n \to \mathbb{R}$.

Além disso, tem-se

$$\langle\, x, y\,\rangle'(v, w) = \langle\, x, w\rangle + \langle\, v, y\rangle \;\forall x, y, v, w \in \mathbb{R}^n$$

e, para quaisquer $(x_1, \dots, x_n), (h_1, \dots, h_n) \in \mathbb{R}^n \times \cdots \times \mathbb{R}^n$,

$$\det{}'(x_1, \dots, x_n)(h_1, \dots, h_n) = \sum_{i=1}^{n} \det(x_1, x_2, \dots, x_{i-1}, h_i, x_{i+1}, \dots, x_n).$$

Exemplo 59 (aplicações com coordenadas polinomiais). Combinando-se os resultados do Exemplo 57 e das proposições 45 e 46, conclui-se facilmente que são diferenciáveis as aplicações cujas coordenadas são polinômios (vide Exemplo 37 – Capítulo 3).

Exemplo 60 (inversão com respeito a S^n). Como consequência do Exemplo 56 e da Proposição 45, tem-se que a inversão de $\mathbb{R}^{n+1} - \{0\}$ com respeito à esfera S^n (vide Exemplo 47 – Capítulo 3),

$$\begin{array}{rccc} f: & \mathbb{R}^{n+1} - \{0\} & \to & \mathbb{R}^{n+1} - \{0\} \\ & x & \mapsto & \frac{x}{\|x\|^2}\,, \end{array}$$

é diferenciável. Além disso, dado $h \in \mathbb{R}^{n+1}$, fazendo-se $\xi(x) = \|x\|^2$, $x \neq 0$, e $\mu(x) = \frac{1}{\xi(x)}$, tem-se $f(x) = \mu(x)x$ e

$$\mu'(x)h = \left(\frac{1}{\xi}\right)'(x)h = -\frac{\xi'(x)h}{(\xi(x))^2} = -\frac{2\langle x, h\rangle}{\|x\|^4}\,.$$

Logo, para todo $x \in \mathbb{R}^{n+1} - \{0\}$, tem-se

$$f'(x)h = (\mu'(x)h)x + \mu(x)h = -\frac{2\langle x, h\rangle}{\|x\|^4}x + \frac{1}{\|x\|^2}h = \frac{1}{\|x\|^2}\left(h - 2\left\langle \frac{x}{\|x\|}, h\right\rangle \frac{x}{\|x\|}\right).$$

Em particular, para todo $x \in \mathbb{R}^{n+1} - \{0\}$, $f'(x)$ é injetiva e, desta forma, um isomorfismo. Com efeito, se $h \in \mathbb{R}^n$ é tal que $f'(x)h = 0$, então

$$2\left\langle \frac{x}{\|x\|}, h\right\rangle \frac{x}{\|x\|} = h.$$

Efetuando-se o produto interno por $x/\|x\|$ em ambos os membros dessa igualdade, obtém-se $\langle \frac{x}{\|x\|}, h\rangle = 0$, donde se conclui que $h = 0$ e, portanto, que $f'(x)$ é injetiva.

Exemplo 61 (inversão de matrizes). Seja $\mathrm{I}(n) \subset \mathrm{M}(n)$ o conjunto das matrizes invertíveis de ordem n, o qual, conforme constatamos, é aberto em $\mathrm{M}(n)$. Vimos também, no capítulo anterior, que a aplicação

$$\begin{array}{rccc} \varphi: & \mathrm{I}(n) & \to & \mathrm{I}(n) \\ & X & \mapsto & X^{-1} \end{array}$$

é contínua. Provemos, então, que φ é diferenciável.

Com efeito, dada $X \in \mathrm{I}(n)$, para toda matriz $X + H \in \mathrm{I}(n)$, tem-se

$$(X+H)^{-1} - X^{-1} = (X+H)^{-1}\left(I - (X+H)X^{-1}\right) = -(X+H)^{-1}HX^{-1},$$

em que I é a matriz identidade de ordem n. Assim, definindo-se a aplicação linear $\Phi(X+H) : \mathrm{I}(n) \to \mathrm{I}(n)$ por

$$\Phi(X+H)A = -(X+H)^{-1}AX^{-1},$$

tem-se $\varphi(X+H) - \varphi(X) = \Phi(X+H)H$. Além disso, pela continuidade de φ, para toda matriz $A \in \mathrm{I}(n)$, tem-se

$$\lim_{H\to 0} \Phi(X+H)A = -X^{-1}AX^{-1} = \Phi(X)A,$$

donde $\lim_{H\to 0} \Phi(X+H) = \Phi(X)$ e, portanto, Φ é contínua em X.

Desta forma, Φ é uma representação de Hadamard de φ em X, donde se infere que φ é diferenciável em X e

$$\varphi'(X)H = \Phi(X)H = -X^{-1}HX^{-1} \ \ \forall H \in \mathrm{M}(n).$$

Exemplo 62 (curvas). Seja $\alpha : I \subset \mathbb{R} \to \mathbb{R}^n$ uma curva em $\mathbb{R}^n$ que é diferenciável no ponto t do intervalo aberto I. Escrevendo-se, então,

$$\alpha(t+h) - \alpha(t) = \alpha'(t)h + r(h), \tag{33}$$

tem-se

$$\lim_{h\to 0} \frac{r(h)}{h} = \lim_{h\to 0} \frac{r(h)}{|h|} = 0.$$

Logo, dividindo-se ambos os membros de (33) por $h \neq 0$ e fazendo-se h tender a zero, obtém-se

$$\lim_{h\to 0} \frac{\alpha(t+h) - \alpha(t)}{h} = \alpha'(t)1.$$

Identificando-se os espaços $\mathrm{L}(\mathbb{R}, \mathbb{R}^n)$ e $\mathbb{R}^n$ por meio do isomorfismo natural entre os mesmos, podemos considerar $\alpha'(t)$ como um vetor de $\mathbb{R}^n$, dito *vetor velocidade* de α em t (Fig. 2), e, desta forma, escrever

$$\alpha'(t) = \lim_{h\to 0} \frac{\alpha(t+h) - \alpha(t)}{h}.$$

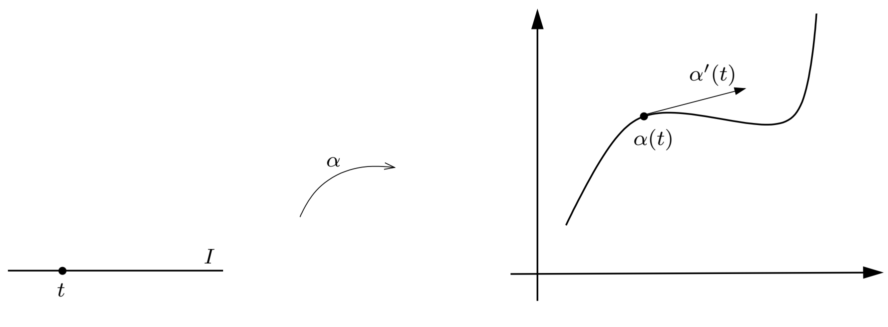

Figura 2

3. Derivadas Parciais – Matriz Jacobiana

Relembremos que a derivada direcional de uma função $f : U \subset \mathbb{R}^n \to \mathbb{R}^m$ num ponto x do aberto U e na direção $h \in \mathbb{R}^n$, quando definida, é o vetor de $\mathbb{R}^m$ dado por

$$\frac{\partial f}{\partial h}(x) = \lim_{t\to 0} \frac{f(x+th) - f(x)}{t}\,.$$

Quando h é o i-ésimo vetor da base canônica de $\mathbb{R}^n$, a derivada direcional $\frac{\partial f}{\partial e_i}(x)$ é chamada de *i-ésima derivada parcial* de f em x e é denotada por $\frac{\partial f}{\partial x_i}(x)$. Assim,

$$\frac{\partial f}{\partial x_i}(x) = \lim_{t\to 0} \frac{f(x+te_i) - f(x)}{t}\,.$$

Supondo-se $f = (f_1, \dots, f_m)$ diferenciável em x, decorre da Proposição 46 que

$$f'(x)e_i = (f_1'(x)e_i, \dots, f_m'(x)e_i) = \left(\frac{\partial f_1}{\partial x_i}(x), \dots, \frac{\partial f_m}{\partial x_i}(x)\right).$$

Segue-se que a matriz de $f'(x)$ relativa às respectivas bases canônicas de $\mathbb{R}^n$ e $\mathbb{R}^m$, a qual chamamos *matriz jacobiana de f* em x e denotamos por $J_f(x)$, é dada por

$$J_f(x) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(x) & \frac{\partial f_1}{\partial x_2}(x) & \cdots & \frac{\partial f_1}{\partial x_n}(x) \\ \frac{\partial f_2}{\partial x_1}(x) & \frac{\partial f_2}{\partial x_2}(x) & \cdots & \frac{\partial f_2}{\partial x_n}(x) \\ \cdots & \cdots & \cdots & \cdots \\ \frac{\partial f_m}{\partial x_1}(x) & \frac{\partial f_m}{\partial x_2}(x) & \cdots & \frac{\partial f_m}{\partial x_n}(x) \end{pmatrix}.$$

Exemplo 63. Consideremos o aberto $U = \{(x_1, x_2, x_3) \in \mathbb{R}^3;\, x_1 > 0\} \subset \mathbb{R}^3$ e a aplicação $f : U \to \mathbb{R}^2$, $f(x_1, x_2, x_3) = (x_3 \log x_1, e^{x_2} \cos x_3)$. Temos que f é diferenciável, pois suas coordenadas são funções diferenciáveis. Além disso, dado $x = (x_1, x_2, x_3) \in U$, tem-se

$$J_f(x) = \begin{pmatrix} x_3/x_1 & 0 & \log x_1 \\ 0 & e^{x_2}\cos x_3 & -e^{x_2}\,\mathrm{sen}\, x_3 \end{pmatrix}.$$

No caso particular de $x = (x_1, x_2, x_3)$ ser um ponto de U, tal que $x_1 \neq 1$ ou $x_3 \neq 0$, tem-se que $J_f(x)$ tem posto 2 e, portanto, $f'(x)$ é sobrejetiva.

Consideremos uma função diferenciável $f : U \subset \mathbb{R}^n \to \mathbb{R}$, definida num aberto U de $\mathbb{R}^n$. Dado $x \in U$, pelo Teorema de Representação de Riesz (vide Exercício 7 – Capítulo 1), existe um único vetor de $\mathbb{R}^n$, o qual designamos por *gradiente* de f em x e denotamos por $\nabla f(x)$, tal que

$$f'(x)h = \langle \nabla f(x), h\rangle \;\; \forall h \in \mathbb{R}^n.$$

Em particular, as coordenadas do gradiente de f em x (com respeito à base canônica de $\mathbb{R}^n$) são as respectivas derivadas parciais de f em x, isto é,

$$\nabla f(x) = \left(\frac{\partial f}{\partial x_1}(x), \dots, \frac{\partial f}{\partial x_n}(x)\right).$$

Exemplo 64 (gradiente da função determinante). Consideremos a função determinante $\det : \mathrm{M}(n) \to \mathbb{R}$ e definamos, para cada $i, j = 1, \dots, n$, a matriz E_{ij} como aquela cujos vetores-coluna são todos nulos, exceto o j-ésimo, que é o i-ésimo vetor da base canônica de $\mathbb{R}^n$, e_i. O conjunto das matrizes E_{ij} corresponde à base canônica de $\mathbb{R}^{n^2}$ quando se identifica este espaço com $\mathrm{M}(n)$. Assim, denotando-se por x_{ij} as coordenadas de uma matriz genérica X de $\mathrm{M}(n)$ com respeito a esta base, tem-se

$$\frac{\partial \det}{\partial x_{ij}}(X) = \det{}'(X)E_{ij}.$$

Indicando-se, então, por $X = [v_1\ v_2\ \cdots\ v_n]$ a matriz X de $M(n)$ cujos vetores-coluna são $v_1, v_2, \dots, v_n \in \mathbb{R}^n$ e fazendo-se

$$X_{ij} = [v_1\ \cdots\ v_{j-1}\ e_i\ v_{j+1}\ \cdots\ v_n],$$

tem-se, pela n-linearidade do determinante (vide Exemplo 58), que

$$\frac{\partial \det}{\partial x_{ij}}(X) = \det{}'(X)E_{ij} = \det{}'[v_1\ \cdots\ v_n][0\ \cdots\ e_i\ \cdots\ 0] = \det X_{ij}.$$

Assim, o gradiente da função determinante em $X \in M(n)$ é a matriz dos cofatores de X (vide Seção 2 – Capítulo 1) e, portanto, satisfaz

$$X(\nabla \det(X))^* = (\det X)I,$$

em que I é a matriz identidade de $\mathrm{M}(n)$.

Exemplo 65 (aplicações do tipo conforme). Seja $f : U \subset \mathbb{R}^n \to \mathbb{R}^n$ uma aplicação diferenciável, tal que, para cada $x \in U$, $f'(x)$ é um isomorfismo. Uma tal f é dita *conforme* se sua derivada, em cada ponto, preserva ângulo, isto é, para todo $x \in U$, existe um real positivo $\mu = \mu(x)$ satisfazendo (vide Exercício 9 – Capítulo 1)

$$\langle f'(x)u, f'(x)v\rangle = \mu^2\langle u, v\rangle \quad \forall u, v \in \mathbb{R}^n. \tag{34}$$

Verifiquemos que, no caso em que $n = 2$ e $f(x_1, x_2) = (f_1(x_1, x_2), f_2(x_1, x_2))$ *preserva orientação* (isto é, $\det f'(x_1, x_2) > 0\,\forall (x_1, x_2) \in U$), a condição de ser conforme é equivalente à validez, para todo $x = (x_1, x_2) \in U$, das igualdades

$$\frac{\partial f_1}{\partial x_1}(x) = \frac{\partial f_2}{\partial x_2}(x) \qquad \text{e} \qquad \frac{\partial f_1}{\partial x_2}(x) = -\frac{\partial f_2}{\partial x_1}(x), \tag{35}$$

denominadas *equações de Cauchy-Riemann*.

Com efeito, tomemos $x \in U$ e escrevamos

$$w_1 = f'(x)e_1 = \left(\frac{\partial f_1}{\partial x_1}(x), \frac{\partial f_2}{\partial x_1}(x)\right) \quad \text{e} \quad w_2 = f'(x)e_2 = \left(\frac{\partial f_1}{\partial x_2}(x), \frac{\partial f_2}{\partial x_2}(x)\right).$$

Então, se f satisfaz (34), tem-se $\langle w_1, w_2\rangle = 0$ e $\|w_1\| = \|w_2\| = \mu$, donde $w_2 = \pm Jw_1$, em que $J : \mathbb{R}^2 \to \mathbb{R}^2$ é o operador ortogonal $J(x_1, x_2) = (-x_2, x_1)$ (rotação de $\pi/2$ no sentido anti-horário). Porém, uma vez que estamos supondo que f preserva orientação, devemos ter $w_2 = Jw_1$, o que implica (35).

Reciprocamente, se f cumpre as equações de Cauchy-Riemann, temos que $\langle w_1, w_2\rangle = 0$ e $\|w_1\| = \|w_2\|$, isto é, $\{w_1, w_2\}$ é uma base ortogonal de $\mathbb{R}^2$ cujos vetores têm mesma norma. Logo, tomando-se coordenadas com respeito a esta base e fazendo-se $\mu = \|w_1\| = \|w_2\|$, segue-se facilmente da bilinearidade do produto interno que (34) se cumpre, donde f é conforme.

Uma aplicação $f : U \subset \mathbb{R}^2 \to \mathbb{R}^2$ que satisfaz as equações de Cauchy-Riemann é dita *holomorfa*. Do ponto de vista da Análise Complexa, as funções holomorfas são diferenciáveis e somente elas o são[(iii)]. Pelas nossas considerações, entretanto, uma tal f é holomorfa se, e somente se, é conforme e preserva orientação, donde se conclui que, em $\mathbb{R}^2$, os respectivos conceitos de diferenciabilidade real e complexa não são equivalentes.

4. Derivadas de Ordem Superior

Consideremos uma aplicação diferenciável f definida num aberto U de $\mathbb{R}^n$ e tomando valores em $\mathbb{R}^m$. Uma vez que, para todo $x \in U$, $f'(x)$ é uma transformação linear de $\mathbb{R}^n$ em $\mathbb{R}^m$, fica definida a aplicação

$$\begin{array}{rccc} f': & U & \to & \mathrm{L}(\mathbb{R}^n, \mathbb{R}^m) \\ & x & \to & f'(x), \end{array}$$

a qual chamamos *derivada* de f.

Podemos, pois, indagar sobre a continuidade ou diferenciabilidade de f'. Dizemos, então, que f é de *classe* C^1 se f' for contínua. Quando f' é diferenciável em $x \in U$, dizemos que f é *duas vezes diferenciável* em x e indicamos a derivada de f' em x por $f''(x)$. Dizemos simplesmente que f é *duas vezes diferenciável* se for duas vezes diferenciável em todos os pontos de U. Neste caso, fica bem definida a aplicação

$$\begin{array}{rccc} f'': & U & \to & \mathrm{L}(\mathbb{R}^n, \mathrm{L}(\mathbb{R}^n, \mathbb{R}^m)) \\ & x & \to & f''(x), \end{array}$$

a qual chamamos *derivada segunda* de f. Diz-se que f é de *classe* C^2 quando é duas vezes diferenciável e f'' é contínua.

A fim de fazer uma descrição mais simples da derivada segunda, observemos que o conjunto

$$\mathrm{L}_2(\mathbb{R}^n, \mathbb{R}^m) = \{g : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^m;\ g \text{ é bilinear}\},$$

munido das operações usuais de soma de funções e multiplicação de escalar por função, é um espaço vetorial. Além disso, dada $T \in \mathrm{L}(\mathbb{R}^n, \mathrm{L}(\mathbb{R}^n, \mathbb{R}^m))$, definindo-se $g_T \in \mathrm{L}_2(\mathbb{R}^n, \mathbb{R}^m)$ por $g_T(x, y) = T(x)y$, verifica-se facilmente que a aplicação

$$\begin{array}{ccc} \mathrm{L}(\mathbb{R}^n, \mathrm{L}(\mathbb{R}^n, \mathbb{R}^m)) & \to & \mathrm{L}_2(\mathbb{R}^n, \mathbb{R}^m) \\ T & \mapsto & g_T \end{array}$$

é um isomorfismo linear.

Identificando-se os espaços $\mathrm{L}(\mathbb{R}^n, \mathrm{L}(\mathbb{R}^n, \mathbb{R}^m))$ e $\mathrm{L}_2(\mathbb{R}^n, \mathbb{R}^m)$ através desse isomorfismo, podemos interpretar a derivada segunda de f como a aplicação que a cada $x \in U$ associa uma forma bilinear $g = f''(x) \in \mathrm{L}_2(\mathbb{R}^n, \mathbb{R}^m)$.

Exemplo 66. Considere a aplicação

$$\begin{array}{rccc} f: & \mathbb{R}^2 & \to & \mathbb{R}^2 \\ & (x, y) & \mapsto & (x^3 + y^3,\ x^3 - y^3). \end{array}$$

Temos que f é diferenciável e, para todo $(x, y) \in \mathbb{R}^2$,

$$J_f(x, y) = \begin{pmatrix} 3x^2 & 3y^2 \\ 3x^2 & -3y^2 \end{pmatrix}.$$

(iii) A diferenciabilidade de uma função complexa de variável complexa $z \in \mathbb{C} \simeq \mathbb{R}^2$ define-se como em (29), substituindo-se x por z.

Identificando-se M(2) com $\mathbb{R}^4$, podemos escrever

$$f'(x,y) = 3(x^2, y^2, x^2, -y^2), \ (x,y) \in \mathbb{R}^2,$$

donde f' é diferenciável, pois suas coordenadas são polinômios. Além disso,

$$J_{f'}(x,y) = 6 \begin{pmatrix} x & 0 \\ 0 & y \\ x & 0 \\ 0 & -y \end{pmatrix}. \tag{36}$$

Logo,

$$f''(x,y)(h_1,h_2) = 6(h_1x, h_2y, h_1x, -h_2y) \sim 6 \begin{pmatrix} h_1x & h_2y \\ h_1x & -h_2y \end{pmatrix},$$

em que $\sim$ denota a identificação de M(2) com $\mathbb{R}^4$. Como aplicação bilinear, $f''(x,y)$ assume a forma

$$f''(x,y)\left((h_1,h_2),(k_1,k_2)\right) = 6 \begin{pmatrix} h_1x & h_2y \\ h_1x & -h_2y \end{pmatrix} \begin{pmatrix} k_1 \\ k_2 \end{pmatrix} = 6 \begin{pmatrix} h_1k_1x + h_2k_2y \\ h_1k_1x - h_2k_2y \end{pmatrix},$$

isto é,

$$f''(x,y)\left((h_1,h_2),(k_1,k_2)\right) = 6\left(h_1k_1x + h_2k_2y, h_1k_1x - h_2k_2y\right).$$

Note que as entradas da matriz (36) são polinômios. Sendo assim, f'' é também diferenciável. Em particular, f é de classe C^2.

Convém observar que, assim como no exemplo acima, se as coordenadas de uma aplicação f são polinômios, as de sua derivada também o são, donde se conclui que toda aplicação cujas coordenadas são polinomiais é de classe C^2.

Exemplo 67 (a derivada segunda de uma aplicação linear). Consideremos uma aplicação linear $T : \mathbb{R}^n \to \mathbb{R}^m$. Conforme constatamos na Seção 2, para todo $x \in \mathbb{R}^n$, $T'(x) = T$. Desta forma, T' é constante e, portanto, $T'' = 0$.

Exemplo 68 (a derivada segunda de uma aplicação bilinear). Consideremos agora uma aplicação bilinear $f : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^p$. Vimos, no Exemplo 58, que f é diferenciável e

$$f'(x,y)(h,k) = f(x,k) + f(h,y) \ \forall (h,k) \in \mathbb{R}^n \times \mathbb{R}^m.$$

Daí, vê-se que f' é linear com respeito à (x,y), isto é, dados $(x_1,y_1),(x_2,y_2)$ em $\mathbb{R}^n \times \mathbb{R}^m$ e $\lambda \in \mathbb{R}$, tem-se

$$f'[(x_1,y_1) + \lambda(x_2,y_2)] = f'(x_1,y_1) + \lambda f'(x_2,y_2).$$

Logo, $f' : \mathbb{R}^n \times \mathbb{R}^m \to \mathrm{L}(\mathbb{R}^n \times \mathbb{R}^m, \mathbb{R}^p)$ é diferenciável e $f''(x,y)(h,k) = f'(h,k)$, donde f'' é constante. Em particular, f é de classe C^2. Como aplicação bilinear, $f''(x,y)$ assume, então, a forma

$$f''(x,y)[(h_1,k_1),(h_2,k_2)] = f'(h_1,k_1)(h_2,k_2) = f(h_1,k_2) + f(h_2,k_1).$$

Dada uma aplicação $f : U \subset \mathbb{R}^n \to \mathbb{R}^m$, definida num aberto U de $\mathbb{R}^n$, suponhamos que, para um dado $k \in \mathbb{R}^n$, existam todas as derivadas direcionais $\frac{\partial f}{\partial k}(x)$. Neste caso, podemos definir a aplicação

$$\begin{array}{rccc} \frac{\partial f}{\partial k}: & U & \to & \mathbb{R}^m \\ & x & \mapsto & \frac{\partial f}{\partial k}(x). \end{array}$$

A derivada direcional de $\frac{\partial f}{\partial k}$ num ponto $x \in U$ e numa direção $h \in \mathbb{R}^n$, quando existe, é dita a *derivada direcional de segunda ordem* de f nas direções h e k, a qual se denota por $\frac{\partial^2 f}{\partial h \partial k}(x)$, isto é

$$\frac{\partial^2 f}{\partial h \partial k}(x) = \frac{\partial \left(\frac{\partial f}{\partial k}\right)}{\partial h}(x).$$

Em particular, quando pertinente, isto é, quando existem as derivadas direcionais envolvidas, escreve-se

$$\frac{\partial^2 f}{\partial x_i \partial x_j}(x) = \frac{\partial \left(\frac{\partial f}{\partial x_j}\right)}{\partial x_i}(x), \ i, j = 1, \dots, n.$$

Vejamos agora que se $f : U \subset \mathbb{R}^n \to \mathbb{R}^m$ é diferenciável no aberto U e duas vezes diferenciável em $x \in U$, então existem todas as derivadas direcionais de segunda ordem de f em x e

$$\frac{\partial^2 f}{\partial h \partial k}(x) \ = \ f''(x)(h, k) \ \ \forall h, k \in \mathbb{R}^n. \tag{37}$$

Para vermos isto, consideremos o caso mais geral de uma aplicação $g : U \to \mathbb{R}^m$, dada por $g(x) = \Phi(x)k$, em que $\Phi : U \subset \mathbb{R}^n \to \mathrm{L}(\mathbb{R}^n, \mathbb{R}^m)$ é diferenciável no ponto $x \in U$ e $k \in \mathbb{R}^n$. Verifiquemos que g é diferenciável em x e

$$g'(x)h = (\Phi'(x)h)k. \tag{38}$$

De fato, sendo Φ diferenciável em x, a aplicação $R = R(h)$, dada por

$$R(h) = \Phi(x + h) - \Phi(x) - \Phi'(x)h,$$

satisfaz $\displaystyle\lim_{h \to 0} \frac{R(h)}{\|h\|} = 0$. Logo, escrevendo-se $r(h) = g(x + h) - g(x) - (\Phi'(x)h)k$, tem-se

$$r(h) = (\Phi(x + h) - \Phi(x) - (\Phi'(x)h))k = R(h)k.$$

Daí,

$$\frac{\|r(h)\|}{\|h\|} \ = \ \frac{\|R(h)k\|}{\|h\|} \ \leq \ \frac{\|R(h)\|}{\|h\|}\|k\| \, ,$$

o que nos dá $\displaystyle\lim_{h \to 0} \frac{r(h)}{\|h\|} = 0$, provando que g é diferenciável em x e satisfaz (38).

Fazendo-se, então, $\Phi = f'$ e $g(x) = \Phi(x)k = f'(x)k = \frac{\partial f}{\partial k}(x)$, tem-se

$$f''(x)(h, k) = (\Phi'(x)h)k = g'(x)h = \frac{\partial g}{\partial h}(x) = \frac{\partial \left(\frac{\partial f}{\partial k}\right)}{\partial h}(x) = \frac{\partial^2 f}{\partial h \partial k}(x).$$

Dada uma função $f : U \subset \mathbb{R}^n \to \mathbb{R}$, duas vezes diferenciável em $x \in U$, chamamos a matriz de $f''(x)$ com respeito às bases canônicas de $\mathbb{R}^n$ e $\mathbb{R}$, isto é, a matriz $A = (a_{ij})$, $a_{ij} = f''(x)(e_i, e_j)$, de *hessiana* de f em x, a qual denotamos por $\operatorname{hess} f(x)$. Segue-se, então, da igualdade (37), que

$$\operatorname{hess} f(x) = \left(\frac{\partial^2 f}{\partial x_i \partial x_j}(x)\right), \ \ i, j = 1, \dots, n.$$

Note que, se $T \in \mathrm{L}(\mathbb{R}^n)$ é o operador linear cuja matriz com respeito à base canônica de $\mathbb{R}^n$ é $\operatorname{hess} f(x)$, então

$$f''(x)(h,k) = \langle Th, k\rangle \ \forall h,k \in \mathbb{R}^n.$$

As derivadas de ordem superior a dois são definidas de forma indutiva, isto é, define-se a derivada terceira a partir da derivada segunda, a quarta a partir da terceira e assim por diante.

Mais precisamente, sejam $k \in \mathbb{N}$, $k \geq 2$, e $f : U \subset \mathbb{R}^n \to \mathbb{R}^m$ uma aplicação definida num aberto U de $\mathbb{R}^n$. Dado $q \in \mathbb{N}$, escrevamos

$$\mathrm{L}_q(\mathbb{R}^n, \mathbb{R}^m) = \{g : \mathbb{R}^n \times \cdots \times \mathbb{R}^n \to \mathbb{R}^m \,;\ g \text{ é } q\text{-linear}\}.$$

Supondo-se, então, definida a derivada de ordem $k-1$ da aplicação f, $f^{(k-1)} : U \to \mathrm{L}_{k-1}(\mathbb{R}^n, \mathbb{R}^m)$, diz-se que f é k *vezes diferenciável* em $x \in U$ se $f^{(k-1)}$ é diferenciável em x.

Assim, escrevendo-se $f^{(k)}(x) = (f^{(k-1)})'(x)$ e identificando-se $\mathrm{L}_k(\mathbb{R}^n, \mathbb{R}^m)$ com $\mathrm{L}(\mathbb{R}^n, \mathrm{L}_{k-1}(\mathbb{R}^n, \mathbb{R}^m))$, quando $f^{(k-1)}$ é diferenciável em U, fica definida a aplicação

$$\begin{array}{rccc} f^{(k)}: & U & \to & \mathrm{L}_k(\mathbb{R}^n, \mathbb{R}^m) \\ & x & \to & f^{(k)}(x), \end{array}$$

a qual chamamos k-*ésima derivada* de f ou *derivada* de f *de ordem* k.

Definição 24 (classes de diferenciabilidade). Dado $k \in \mathbb{N}$, diz-se que uma aplicação f é de *classe* C^k quando é k vezes diferenciável e sua derivada de ordem k, $f^{(k)}$, é contínua. Diz-se que f é de *classe* C^∞ quando tem derivadas de todas as ordens, isto é, quando f é de classe C^k $\forall k \in \mathbb{N}$.

Considerando-se os resultados das proposições 45 e 46 e usando-se indução, verifica-se facilmente que:

i) Uma aplicação é de classe C^k se, e somente se, suas coordenadas o são;

ii) somas de aplicações de classe C^k, bem como produtos de funções de classe C^k por aplicações de classe C^k e quocientes de funções de classe C^k, são de classe C^k.

Segue-se, então, da propriedade (i), que aplicações cujas coordenadas são funções polinomiais são de classe C^∞. Com efeito, verificamos anteriormente que tais aplicações são diferenciáveis. Logo, uma vez que qualquer derivada parcial de um polinômio p de variáveis $x_1, \dots, x_n$, $\frac{\partial p}{\partial x_i}$, é ainda um polinômio, conclui-se, por indução, que f possui derivadas de todas as ordens (vide Exemplo 66).

Também, considerando-se os exemplos 67 e 68 e usando-se indução, verifica-se que aplicações n-lineares são de classe C^∞, em que suas derivadas de ordem $n-1$ são aplicações lineares e, portanto, suas derivadas de ordem n são constantes e as de ordem maior que n são nulas.

Exemplo 69 (a inversão de matrizes é C^∞). Verifiquemos que a inversão de matrizes, $\varphi(X) = X^{-1}$, $X \in \mathrm{I}(n)$, considerada no Exemplo 61, é uma aplicação de classe C^∞. Com efeito, dada $X \in \mathrm{I}(n)$, fazendo-se $X = (x_{ij})$ e $X^{-1} = (y_{kl})$, tem-se, pela fórmula dos cofatores (vide Seção 2 – Capítulo 1), que cada entrada y_{kl} de X^{-1} é dada por

$$y_{kl} = \frac{\det X_{lk}}{\det X}\,,$$

em que X_{lk} é a matriz cujos vetores-coluna são os mesmos de X, exceto pelo k-ésimo, que é o l-ésimo vetor da base canônica de $\mathbb{R}^n$, e_l. Temos que a função determinante, por ser n-linear, é de classe C^∞. Além disso, uma vez que, para quaisquer $k, l \in \{1, \dots, n\}$, $\det X_{lk}$ é um polinômio de variáveis x_{ij}, tem-se que a função $X \mapsto \det X_{kl}$ é, igualmente, de classe C^∞. Logo, por ser o quociente de funções de classe C^∞, cada coordenada de φ é de classe C^∞, donde φ é de classe C^∞.

5. A Regra da Cadeia

A Regra da Cadeia é aquela que estabelece a relação entre a derivada de duas aplicações dadas com a derivada da composta dessas aplicações. Ela nos diz, essencialmente, que a derivada da composta é a composta das derivadas.

Regra da Cadeia. *Sejam $U \subset \mathbb{R}^n$ e $V \subset \mathbb{R}^m$ conjuntos abertos. Dadas aplicações $f : U \to V$ e $g : V \to \mathbb{R}^p$, se f é diferenciável em $x \in U$ e g é diferenciável em $f(x) \in V$, então $g \circ f : U \to \mathbb{R}^p$ é diferenciável em x e*

$$(g \circ f)'(x) = g'(f(x)) \circ f'(x).$$

Demonstração. Sejam $\Phi : U \to \mathrm{L}(\mathbb{R}^n, \mathbb{R})$ e $\Psi : V \to \mathrm{L}(\mathbb{R}^n, \mathbb{R}^m)$ representações de Hadamard de f e g nos pontos x e $f(x)$, respectivamente. Sendo assim, Φ é contínua em x, Ψ é contínua em $f(x)$, $\Phi(x) = f'(x)$ e $\Psi(f(x)) = g'(f(x))$. Além disso, para quaisquer $x + h \in U$ e $f(x) + k \in V$, valem as igualdades

$$f(x+h) - f(x) = \Phi(x+h)h \quad \text{e} \quad g(f(x)+k) - g(f(x)) = \Psi(f(x)+k)k.$$

Desta forma,

$$\begin{aligned} (g \circ f)(x+h) - (g \circ f)(x) &= g(f(x+h)) - g(f(x)) \\ &= g(f(x) + \Phi(x+h)h) - g(f(x)) \\ &= \Psi(f(x) + \Phi(x+h)h)\Phi(x+h)h. \end{aligned}$$

Logo, fazendo-se, para todo $x + h \in U$,

$$\Theta(x+h) = \Psi(f(x) + \Phi(x+h)h)\Phi(x+h),$$

tem-se $(g \circ f)(x+h) - (g \circ f)(x) = \Theta(x+h)h$. Agora, levando-se em conta que Φ é contínua em x, Ψ é contínua em $f(x)$ e (vide Exercício 27 – Capítulo 3)

$$\lim_{h \to 0} \Phi(x+h)h = \lim_{h \to 0} \Phi(x+h) \lim_{h \to 0} h = f'(x).0 = 0,$$

conclui-se facilmente que $\lim_{h \to 0} = \Theta(x+h) = \Theta(x)$, donde Θ é contínua em x.

Segue-se que $\Theta : U \to \mathrm{L}(\mathbb{R}^n, \mathbb{R}^p)$ é uma representação de Hadamard de $g \circ f$ em x. Logo, $g \circ f$ é diferenciável em x e

$$(g \circ f)'(x) = \Theta(x) = \Psi(f(x))\Phi(x) = g'(f(x))f'(x),$$

como desejado. □

Corolário 3. *Seja $f : U \subset \mathbb{R}^n \to \mathbb{R}^m$ diferenciável em $x \in U$ e contínua numa vizinhança de x. Nestas condições, se $\alpha : I \to U$ é uma curva diferenciável em $t = 0$ satisfazendo $\alpha(0) = x$ e $\alpha'(0) = h$, então a curva $\beta = f \circ \alpha$ é diferenciável em $t = 0$ e $\beta'(0) = f'(x)h$.*

Demonstração. Pela Regra da Cadeia, β é diferenciável em $t=0$ e

$$\beta'(0) = (f\circ\alpha)'(0) = f'(\alpha(0))\alpha'(0) = f'(x)h. \qquad \square$$

O Corolário 3 revela uma propriedade interessante da derivada, a saber, o vetor velocidade da curva $f\circ\alpha$ em $t=0$ não depende da curva α (que satisfaz $\alpha(0)=x$ e $\alpha'(0)=h$), mas somente de x e h (Fig. 3).

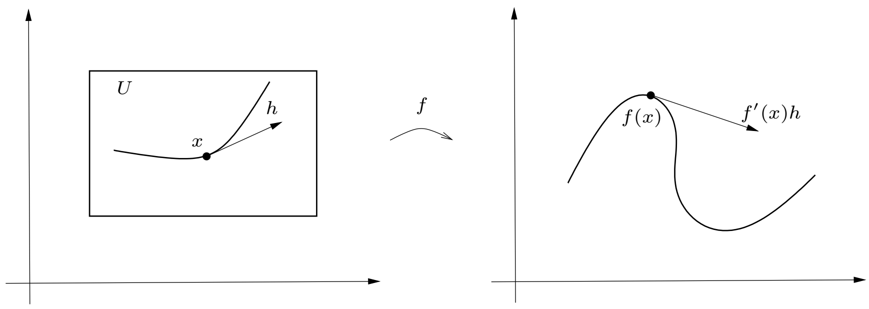

Figura 3

Deve-se notar também que, quando $f: U\subset\mathbb{R}^n\to\mathbb{R}^m$ é diferenciável em $x\in U$, a igualdade

$$f'(x)h = \lim_{t\to 0}\frac{f(x+th)-f(x)}{t},$$

anteriormente estabelecida, é um caso particular do Corolário 3, em que $\alpha(t) = x+th$.

Nos exemplos seguintes, ilustraremos a efetividade da Regra da Cadeia na determinação das derivadas de algumas aplicações especiais.

Exemplo 70 (produtos). Sejam $U\subset\mathbb{R}^n$ um conjunto aberto de $\mathbb{R}^n$ e $f,g: U\subset\mathbb{R}^n\to\mathbb{R}^m$ aplicações diferenciáveis em $a\in U$. Dada uma aplicação bilinear $\varphi:\mathbb{R}^m\times\mathbb{R}^m\to\mathbb{R}^p$, temos que a composta

$$\begin{array}{rccc}\sigma: & U & \to & \mathbb{R}^p\\ & x & \mapsto & \varphi(f(x),g(x))\end{array}$$

é diferenciável em a e vale a igualdade

$$\sigma'(a)h = \varphi(f'(a)h, g(a)) + \varphi(f(a), g'(a)h)$$

(isto é, do ponto de vista da diferenciação, σ é um *produto* das funções $f(x)$ e $g(x)$).

Com efeito, temos que $\sigma = \varphi\circ\psi$, em que $\psi: U\to\mathbb{R}^m\times\mathbb{R}^m$ é definida por $\psi(x) = (f(x),g(x))$. Logo, pela Regra da Cadeia, pela Proposição 46 e pelas considerações do Exemplo 58, temos que σ é diferenciável em a e

$$\begin{aligned}\sigma'(a)h &= \varphi'(\psi(a))\psi'(a)h\\ &= \varphi'(f(a),g(a))(f'(a)h, g'(a)h)\\ &= \varphi(f'(a)h, g(a)) + \varphi(f(a), g'(a)h).\end{aligned}$$

Em particular, se $\alpha, \beta : I \subset \mathbb{R} \to \mathbb{R}^m$ são curvas diferenciáveis no intervalo aberto I, então a função $\sigma(t) = \langle \alpha(t), \beta(t) \rangle$ é diferenciável e, para quaisquer $t \in I$, $h \in \mathbb{R}$, tem-se

$$\sigma'(t)h = \langle \alpha'(t)h, \beta(t) \rangle + \langle \alpha(t), \beta'(t)h \rangle = \big(\langle \alpha'(t), \beta(t) \rangle + \langle \alpha(t), \beta'(t) \rangle \big) h,$$

isto é,

$$\sigma'(t) = \langle \alpha'(t), \beta(t) \rangle + \langle \alpha(t), \beta'(t) \rangle.$$

Exemplo 71 (norma euclidiana). Consideremos a função $f(x) = \|x\|$, $x \in \mathbb{R}^n$, e apliquemos a Regra da Cadeia para verificar que f é diferenciável no aberto $U = \mathbb{R}^n - \{0\}$. Para tanto, basta tomarmos a função $\lambda(t) = \sqrt{t}$, $t \geq 0$, e observarmos que a mesma é diferenciável em $\mathbb{R} - \{0\}$ e satisfaz, para todo $t > 0$, $\lambda'(t) = 1/2\sqrt{t}$. Fazendo-se, então, $g(x) = \|x\|^2$, tem-se $f = \lambda \circ g$. Logo, pela Regra da Cadeia e pelo resultado do Exemplo 56, tem-se, para todo $x \neq 0$, que f é diferenciável em x e

$$f'(x)h = \lambda'(g(x))g'(x)h = \frac{1}{2\sqrt{g(x)}} 2\langle x, h \rangle = \frac{\langle x, h \rangle}{\|x\|} \cdot$$

Por outro lado, para todo vetor unitário $u \in \mathbb{R}^n$, tem-se

$$\lim_{t \to 0_+} \frac{\|tu\|}{t} = 1 \quad \text{e} \quad \lim_{t \to 0_-} \frac{\|tu\|}{t} = -1,$$

donde se infere que a derivada direcional $\frac{\partial f}{\partial u}(0)$ não está definida e, portanto, que a norma euclidiana não é diferenciável em $x = 0$.

Exemplo 72 (norma do máximo). Seja f a norma do máximo em $\mathbb{R}^n$,

$$f(x) = \|x\|_{\max}, \quad x \in \mathbb{R}^n.$$

Dado $i \in \{1, \dots, n\}$, designemos por U_i o subconjunto de $\mathbb{R}^n$ formado pelos pontos $x = (x_1, \dots x_i, \dots, x_n)$, tais que

$$f(x) = |x_i| > |x_j| \ \forall j \in \{1, \dots, i-1, i+1, \dots, n\}.$$

Conforme constatamos no Exemplo 12 do Capítulo 2, cada conjunto U_i é aberto. Assim, dado $x \in U_i$, existe $\delta > 0$, tal que $B(x, \delta) \subset U_i$. Observemos que a função $f|_{B(x,\delta)}$ é positiva e coincide com a restrição a $B(x, \delta)$ da função $y \mapsto |P_i(y)|$, em que $P_i : \mathbb{R}^n \to \mathbb{R}$ é a projeção $P_i(y) = y_i$. Logo, pela Regra da Cadeia, f é diferenciável em $B(x, \delta)$ (em particular, em x) e, para todo $h = (h_1, \dots, h_n) \in \mathbb{R}^n$, tem-se

$$f'(x)h = \frac{1}{2\sqrt{|P_i(x)|^2}} 2P_i(x)P_i'(x)h = \frac{1}{\sqrt{|P_i(x)|^2}} P_i(x)P_i(h) = \frac{x_i h_i}{|x_i|} \cdot$$

Agora, se $y = (y_1, \dots, y_n) \in \mathbb{R}^n - U$, então existem índices distintos $i, j \in \{1, \dots, n\}$, tais que $\|y\|_{\max} = |y_i| = |y_j|$. Desta forma, tomando-se $\epsilon > 0$ e $t \in (-\epsilon, \epsilon)$, tem-se

$$\frac{f(y + te_i) - f(y)}{t} = \frac{\max\{|y_i + t|, |y_j|\} - |y_i|}{t} \cdot$$

Suponhamos, sem perda de generalidade, que $y_i \geq 0$. Neste caso, quando $t > 0$, tem-se $|y_i + t| = y_i + t > y_i = |y_j|$ e, portanto, $f(y + te_i) - f(y) = (y_i + t) - y_i = t$. No entanto, quando $t < 0$, tomando-se ϵ suficientemente pequeno, tem-se

$$|y_i + t| = y_i + t < y_i = |y_j|,$$

donde $f(y+te_i) - f(y) = |y_j| - y_i = 0$. Desta forma,

$$\lim_{t\to 0_+} \frac{f(y+te_i)-f(y)}{t} = 1 \quad \text{e} \quad \lim_{t\to 0_-} \frac{f(y+te_i)-f(y)}{t} = 0,$$

donde f não é diferenciável em y.

Segue-se, então, destas considerações, que a norma do máximo é diferenciável em $x \in \mathbb{R}^n$ se, e somente se, x é um ponto do aberto $U = \bigcup U_i$.

Nas figuras 4 e 5 exibem-se, respectivamente, os gráficos das funções $x \mapsto \|x\|$ e $x \mapsto \|x\|_{\max}$, $x \in \mathbb{R}^2$. Note que, nos pontos destes gráficos que correspondem àqueles em que estas funções não são diferenciáveis (vértice do cone e vértice e arestas da pirâmide) os mesmos não admitem o que, intuitivamente, chamaríamos de plano tangente.

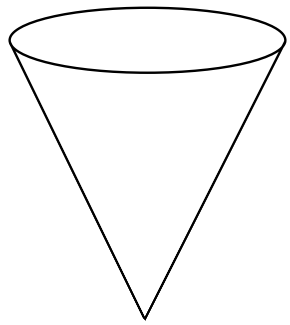

FIGURA 4

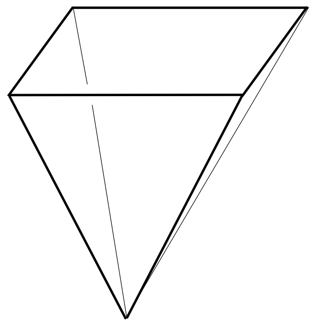

FIGURA 5

EXEMPLO 73 (FUNÇÃO DISTÂNCIA A UM SUBESPAÇO VETORIAL). Sejam $\mathbb{V} \subset \mathbb{R}^n$ um subespaço vetorial próprio de $\mathbb{R}^n$ e $d_{\mathbb{V}}(x) = \|x - P_{\mathbb{V}}\, x\|$, $x \in \mathbb{R}^n$, a função distância a $\mathbb{V}$. Segue-se diretamente da Regra da Cadeia, do Exemplo 71 e da linearidade da projeção ortogonal $P_{\mathbb{V}}$, que $d_{\mathbb{V}}$ é diferenciável em $\mathbb{R}^n - \mathbb{V}$ e

$$d_{\mathbb{V}}'(x)h = \frac{\langle x - P_{\mathbb{V}}\, x, h - P_{\mathbb{V}}\, h\rangle}{\|x - P_{\mathbb{V}}\, x\|} = \left\langle \frac{x - P_{\mathbb{V}}\, x}{\|x - P_{\mathbb{V}}\, x\|}, h \right\rangle \ \ \forall x \in \mathbb{R}^n - \mathbb{V}, h \in \mathbb{R}^n,$$

pois $x - P_{\mathbb{V}}\, x \in \mathbb{V}^{\perp}$ e $P_{\mathbb{V}}\, h \in \mathbb{V}$. Em particular

$$\nabla d_{\mathbb{V}}(x) = \frac{x - P_{\mathbb{V}}\, x}{\|x - P_{\mathbb{V}}\, x\|} \ \ \forall x \in \mathbb{R}^n - \mathbb{V}.$$

Procedendo-se como nos dois exemplos anteriores, conclui-se facilmente que em ponto algum de $\mathbb{V}$ existe qualquer derivada direcional, donde se infere que $d_{\mathbb{V}}$ é diferenciável em $x \in \mathbb{R}^n$ se, e somente se, $x \in \mathbb{R}^n - \mathbb{V}$.

PROPOSIÇÃO 47. *A composta de aplicações de classe C^k é também uma aplicação de classe C^k.*

DEMONSTRAÇÃO. Consideremos abertos $U \subset \mathbb{R}^n$, $V \subset \mathbb{R}^m$ e aplicações diferenciáveis $f : U \to V$ e $g : V \to \mathbb{R}^p$. Pela Regra da Cadeia, $g \circ f$ é diferenciável e, para todo $x \in U$,

$$(g \circ f)'(x) = g'(f(x))f'(x) = (g' \circ f)(x)f'(x).$$

Desta forma, podemos escrever

$$(g \circ f)' = \varphi \circ \psi, \tag{39}$$

em que $\varphi : \mathrm{L}(\mathbb{R}^n, \mathbb{R}^m) \times \mathrm{L}(\mathbb{R}^m, \mathbb{R}^p) \to \mathrm{L}(\mathbb{R}^n, \mathbb{R}^p)$ e $\psi : U \to \mathrm{L}(\mathbb{R}^n, \mathbb{R}^m) \times \mathrm{L}(\mathbb{R}^m, \mathbb{R}^p)$ são dadas por

$$\varphi(T, Z) = ZT \quad \text{e} \quad \psi(x) = (f'(x), (g' \circ f)(x)).$$

Suponhamos que f e g sejam de classe C^1. Neste caso, a aplicação ψ é contínua, já que f' e g', por hipótese, são contínuas, f é contínua e a composta de contínuas é contínua. Além disso, a aplicação φ, por ser bilinear, é de classe C^∞. Segue-se, portanto, da igualdade (39), que $(g \circ f)'$ é contínua, donde o resultado é verdadeiro para $k = 1$.

Supondo-se que o resultado seja verdadeiro para um certo $k \in \mathbb{N}$ e que f e g sejam de classe C^{k+1}, tem-se f' e g' de classe C^k, donde, por hipótese de indução, se conclui que $g' \circ f$ é de clase C^k e, portanto, que ψ é de classe C^k. Logo, $(g \circ f)'$ é de classe C^k e, consequentemente, $g \circ f$ é de classe C^{k+1}.

Segue-se, portanto, do Princípio da Indução, que a composta de aplicações de classe C^k é de classe C^k para todo $k \in \mathbb{N}$. □

6. Difeomorfismos

No capítulo anterior, constatamos que através do conceito de continuidade e seu consequente, o de homeomorfismo, estabelece-se uma relação de equivalência entre subconjuntos de espaços euclidianos. No contexto das aplicações diferenciáveis, o conceito análogo chama-se difeomorfismo, o qual define uma relação de equivalência entre subconjuntos abertos de espaços euclidianos, conforme a definição seguinte.

Definição 25 (difeomorfismo). Dados abertos $U \subset \mathbb{R}^n$, $V \subset \mathbb{R}^m$, uma bijeção $f : U \to V$ é dita um *difeomorfismo* — e U é dito, então, *difeomorfo* a V — quando ambas, f e f^{-1}, são diferenciáveis. Quando f e f^{-1} são de classe C^k, diz-se que f é um *difeomorfismo de classe* C^k.

É imediato que a aplicação identidade de um aberto de $\mathbb{R}^n$ em si mesmo é um difeomorfismo, que a inversa de um difeomorfismo é um difeomorfismo e que a composta de difeomorfismos é um difeomorfismo. Logo, "ser difeomorfo a" é, de fato, uma relação de equivalência.

Segue-se diretamente das considerações do Exemplo 42 do Capítulo 3 que quaisquer duas bolas (euclidianas) abertas de $\mathbb{R}^n$ são difeomorfas.

Verificamos anteriormente que a aplicação

$$\begin{array}{rccc} f: & B & \to & \mathbb{R}^n \\ & x & \mapsto & \frac{x}{1-\|x\|} \end{array}$$

é um homeomorfismo de $B = B(0,1) \subset \mathbb{R}^n$ em $\mathbb{R}^n$ cujo inverso é a aplicação

$$\begin{array}{rccc} g: & \mathbb{R}^n & \to & B \\ & y & \mapsto & \frac{y}{1+\|y\|} \end{array}.$$

Uma vez que a função norma é diferenciável em $\mathbb{R}^n - \{0\}$, temos que f e g são diferenciáveis em $B - \{0\}$ e $\mathbb{R}^n - \{0\}$, respectivamente. Além disso, pode-se verificar facilmente, pelo cálculo das derivadas direcionais, que f e g são diferenciáveis em $x = 0$ e $y = 0$, e que suas respectivas derivadas nestes pontos são, ambas, a aplicação identidade de $\mathbb{R}^n$. Logo, f é um difeomorfismo e, portanto, qualquer bola (euclidiana) aberta de $\mathbb{R}^n$ é difeomorfa a $\mathbb{R}^n$.

Outro exemplo de difeomorfismo, este de classe C^∞, é a inversão de matrizes. Isto decorre do fato de esta aplicação ser bijetiva, de classe C^∞ e ter ela própria como inversa.

PROPOSIÇÃO 48 (DERIVADAS DE DIFEOMORFISMOS SÃO ISOMORFISMOS). *Seja* $f : U \subset \mathbb{R}^n \to V \subset \mathbb{R}^m$ *um difeomorfismo. Então, para todo* $x \in U$, *a derivada* $f'(x) \in \mathrm{L}(\mathbb{R}^n, \mathbb{R}^m)$ *é um isomorfismo. Em particular, tem-se* $m = n$.

DEMONSTRAÇÃO. Com efeito, denotando-se por $g : V \to U$ a inversa de f e tomando-se $x = g(y) \in U$, $y \in V$, temos que $g(f(x)) = x$ e $f(g(y)) = y$. Logo, pela Regra da Cadeia, $g'(f(x))f'(x) = I_n$ e $f'(g(y))g'(y) = I_m$, em que I_n, I_m denotam, respectivamente, as aplicações identidade de $\mathbb{R}^n$ e $\mathbb{R}^m$. Segue-se que $f'(x)$ e $g'(y)$ são isomorfismos, sendo um o inverso do outro. □

Em certos contextos da Análise, faz-se apropriado considerar-se um difeomorfismo $\varphi : U \subset \mathbb{R}^n \to V \subset \mathbb{R}^n$ como uma *mudança de coordenadas*

$$x = (x_1, \dots, x_n) \;\leftrightarrow\; \varphi(x) = (y_1, \dots, y_n).$$

Um exemplo clássico é aquele em que $U = (0, +\infty) \times (-\pi/2, \pi/2) \subset \mathbb{R}^2$, $V = (0, +\infty) \times \mathbb{R} \subset \mathbb{R}^2$ e φ é definida por $\varphi(r, \theta) = (r\cos\theta, r\,\mathrm{sen}\,\theta)$. Verifica-se facilmente que φ é bijetiva, diferenciável e sua inversa, $\varphi^{-1} : V \to U$, dada por $\varphi^{-1}(x, y) = (\sqrt{x^2 + y^2}\,, \mathrm{arctg}\,(y/x))$, é diferenciável, donde φ é um difeomorfismo. Neste caso, as coordenadas (r, θ) são ditas *polares*.

A mudança de coordenadas é uma técnica frequentemente utilizada na solução de problemas em Cálculo e Equações Diferenciais, a qual consiste em exprimir as funções envolvidas com respeito a outras variáveis, que não aquelas que as definem, e que tornem o problema mais simples.

Consideremos, por exemplo, uma função diferenciável $f : V \subset \mathbb{R}^2 \to \mathbb{R}$, V aberto, e um difeomorfismo $\varphi : U \subset \mathbb{R}^2 \to V$, em que

$$\varphi(\xi, \eta) = (x, y) = (x(\xi, \eta), y(\xi, \eta)).$$

Podemos interpretar, desta forma, a composta da função f com o difeomorfismo φ, como a expressão de f com respeito às variáveis ξ e η. Nestas condições, a simplificação que mencionamos anteriormente decorre das relações entre as derivadas parciais de f com respeito às variáveis x, y e aquelas com respeito a ξ e η, que, naturalmente, são obtidas por meio da Regra da Cadeia.

Com efeito, escrevendo-se

$$g(\xi, \eta) = (f \circ \varphi)(\xi, \eta) = f(x(\xi, \eta), y(\xi, \eta)),$$

temos, Pela Regra da Cadeia, que

$$\frac{\partial g}{\partial \xi}(\xi, \eta) = g'(\xi, \eta)e_1 = f'(x, y)\varphi'(\xi, \eta)e_1 = \langle \nabla f(x, y), \varphi'(\xi, \eta)e_1 \rangle.$$

Observando-se que

$$\varphi'(\xi, \eta)e_1 = ((x'(\xi, \eta)e_1, y'(\xi, \eta)e_1) = \left(\frac{\partial x}{\partial \xi}(\xi, \eta), \frac{\partial y}{\partial \xi}(\xi, \eta)\right),$$

obtém-se

$$\frac{\partial g}{\partial \xi}(\xi,\eta) = \frac{\partial f}{\partial x}(\xi,\eta)\frac{\partial x}{\partial \xi}(\xi,\eta) + \frac{\partial f}{\partial y}(\xi,\eta)\frac{\partial y}{\partial \xi}(\xi,\eta). \tag{40}$$

Em geral, no intuito de simplificar a notação, abusa-se da mesma omitindo-se o ponto em que as derivadas são aplicadas, bem como escrevendo-se $g(\xi,\eta) = (f\circ\varphi)(\xi,\eta) = f(\xi,\eta)$. Neste caso, a igualdade (40) assume a forma clássica

$$\frac{\partial f}{\partial \xi} = \frac{\partial f}{\partial x}\frac{\partial x}{\partial \xi} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial \xi}\cdot \tag{41}$$

Evidentemente, vale uma fórmula análoga para $\frac{\partial f}{\partial \eta}$, isto é,

$$\frac{\partial f}{\partial \eta} = \frac{\partial f}{\partial x}\frac{\partial x}{\partial \eta} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial \eta}\cdot \tag{42}$$

7. O Teorema de Motzkin (∗)

Conforme discutimos anteriormente, em $\mathbb{R}^n$, a noção de distância entre dois pontos é naturalmente estendida à de distância entre ponto e conjunto através da igualdade

$$d(x,F) = \inf_{a\in F}\|x-a\|, \tag{43}$$

em que $x\in\mathbb{R}^n$ é um ponto e $F\subset\mathbb{R}^n$ um subconjunto de $\mathbb{R}^n$.

Constatamos, então, que, quando o conjunto F é fechado, a função distância a F,

$$\begin{array}{rccc} d_F: & \mathbb{R}^n & \to & \mathbb{R} \\ & x & \mapsto & d(x,F), \end{array}$$

é uniformemente contínua.

Cabe-nos, pois, indagar se, do ponto de vista da Topologia e da Análise, esta e outras propriedades da função d_F relacionam-se com as propriedades topológicas e métricas do conjunto F.

Neste contexto, um caso simples é aquele em que o conjunto dado é um subespaço vetorial próprio de $\mathbb{R}^n$, $\mathbb{V}$. Ao longo deste capítulo e dos anteriores, verificamos os seguintes fatos:

- Existe uma projeção(iv) $P_{\mathbb{V}}:\mathbb{R}^n\to\mathbb{V}$, tal que $d_{\mathbb{V}}(x) = \|x - P_{\mathbb{V}}x\|$;
- a função distância a $\mathbb{V}$ é diferenciável em $\mathbb{R}^n - \mathbb{V}$;
- $\mathbb{V}$ é convexo.

Em consideração à questão levantada acima, um fato notável é que estas três afirmações não só são equivalentes, mas o são para qualquer conjunto fechado de $\mathbb{R}^n$, isto é, elas não se restringem aos seus subespaços vetoriais. Este resultado, conhecido como Teorema de Motzkin, constitui uma importante caracterização dos conjuntos convexos (fechados) de espaços euclidianos e deve seu nome ao matemático alemão Theodore Motzkin (1908–1970), que o estabeleceu em 1935 (vide [**6**], [**25**]).

(iv)Uma aplicação $P:\mathbb{R}^n\to F\subset\mathbb{R}^n$ é dita uma *projeção* sobre F se $P\circ P = P$.

7.1. Projeções – Conjuntos de Chebyshev. Dado um conjunto fechado $F \subset \mathbb{R}^n$, verificamos que, para todo $x \in \mathbb{R}^n$, existe $a \in F$, tal que $d(x, F) = \|x - a\|$. Entretanto, o ponto $a \in F$ com esta propriedade não é, necessariamente, único. Tomando-se, por exemplo, a esfera $S = S[x, r] \subset \mathbb{R}^n$, temos que todo ponto $a \in S$ cumpre $d(x, S) = \|x - a\| = r$.

Dados, então, um conjunto fechado $F \subset \mathbb{R}^n$ e $x \in \mathbb{R}^n$, designaremos por $\Pi_F(x)$ o conjunto formado por todos os pontos de F que realizam a distância de x a F, isto é,

$$\Pi_F(x) = \{a \in F;\ \|x - a\| = d(x, F)\}.$$

Note que, para todo $x \in \mathbb{R}^n$, $\Pi_F(x)$ é compacto. Com efeito, em virtude da continuidade da função $a \mapsto \|x - a\|$, $\Pi_F(x)$ é fechado. Além disso, para todo $a \in \Pi_F(x)$, tem-se $\|a\| \le \|a - x\| + \|x\| = d(x, F) + \|x\|$, donde $\Pi_F(x)$ é também limitado e, portanto, compacto.

Diz-se que um subconjunto fechado $F \subset \mathbb{R}^n$ é um *conjunto de Chebyshev*[v] se, para todo $x \in \mathbb{R}^n$, $\Pi_F(x)$ contém um único elemento $a \in F$, o qual designamos por $P_F(x)$. Neste caso, fica bem definida a aplicação *projeção sobre* F,

$$\begin{array}{rccc} P_F: & \mathbb{R}^n & \to & F \\ & x & \mapsto & P_F(x), \end{array}$$

a qual, para todo ponto $x \in \mathbb{R}^n$, satisfaz

$$d_F(x) = \|x - P_F(x)\|.$$

PROPOSIÇÃO 49. *A projeção sobre um conjunto de Chebyshev é uma aplicação contínua.*

DEMONSTRAÇÃO. Suponhamos que $F \subset \mathbb{R}^n$ seja um conjunto de Chebyshev. Tomemos a projeção sobre F, $P_F : \mathbb{R}^n \to F$, e uma sequência (x_k) em $\mathbb{R}^n$, tal que $x_k \to x_0 \in \mathbb{R}^n$. Uma vez que, para todo $x \in \mathbb{R}^n$,

$$\|P_F(x)\| \le \|P_F(x) - x\| + \|x\| = d_F(x) + \|x\|,$$

temos que a sequência $(P_F(x_k))$ é limitada, pois, pela continuidade de d_F e da função norma, as sequências $(d_F(x_k))$ e $(\|x_k\|)$ são convergentes, donde, limitadas. Logo, $(P_F(x_k))$ possui uma subsequência convergente, $(P_F(x_{k_i}))$. Supondo-se que $a \in F$ seja o limite desta subsequência, temos que

$$\|x_0 - a\| = \lim_{i\to\infty} \|x_{k_i} - P_F(x_{k_i})\| = \lim_{i\to\infty} d_F(x_{k_i}) = d_F(x_0),$$

donde $a = P_F(x_0)$ e, portanto, P_F é contínua (vide Observação 12 – Capítulo 3). □

7.2. O Teorema de Motzkin. Apresentaremos agora uma demonstração do Teorema de Motzkin que terá como base os dois lemas seguintes, os quais, por sua vez, estabelecem propriedades interessantes da função distância d_F e da projeção P_F, respectivamente.

LEMA 3. *Sejam $F \subset \mathbb{R}^n$ um conjunto fechado e $f : \mathbb{R}^n \to \mathbb{R}$ a função distância a F ao quadrado, isto é, $f(x) = d_F^2(x)$. Então, dado $x \in \mathbb{R}^n$, a função*

$$T(h) = \min_{a \in \Pi_F(x)} 2\langle x - a, h\rangle,\ h \in \mathbb{R}^n,$$

[v] Em consideração ao matemático russo Pafnuty Chebyshev (1821–1894), um dos fundadores da Teoria da Aproximação, cujos temas envolvem questões relativas a projeções e convexidade.

cumpre as seguintes condições:

i) $\displaystyle\lim_{h\to 0} \frac{f(x+h)-f(x)-T(h)}{\|h\|} = 0;$

ii) $\displaystyle\lim_{t\to 0_+} \frac{f(x+th)-f(x)}{t} = T(h).$

Cabe-nos observar que, no lema acima, a função T está bem definida (pois, para todo $x \in F$, $\Pi_F(x)$ é compacto) e é sugerida pela transformação linear $Th = 2\langle x - P_F(x), h\rangle$, que é a derivada de d_F^2 em $x \in \mathbb{R}^n$ quando F é um subespaço vetorial de $\mathbb{R}^n$. Deve-se notar também que a igualdade (i) não assegura que f seja diferenciável em x, a menos que T seja linear.

Demonstração do Lema 3. Tomemos $x \in \mathbb{R}^n$ e observemos que, devido à compacidade de $\Pi_F(x)$ e à continuidade da função $a \mapsto 2\langle x-a, h\rangle$, para cada $h \in \mathbb{R}^n$, o conjunto

$$\Pi_h(x) = \{a \in \Pi_F(x);\, T(h) = 2\langle x-a, h\rangle\}$$

é não-vazio.

Assim, dados $h_0, h \in \mathbb{R}^n$, para quaisquer $a_h \in \Pi_{h_0+h}(x)$ e $a \in \Pi_{h_0}(x)$, tem-se

$$T(h_0+h) - T(h_0) = 2\langle x - a_h, h_0 + h\rangle - 2\langle x - a, h_0\rangle.$$

Porém, $\langle x - a_h, h_0 + h\rangle \le \langle x - a, h_0 + h\rangle$ e $\langle x - a, h_0\rangle \le \langle x - a_h, h_0\rangle$. Logo,

$$2\langle x - a_h, h\rangle \le T(h_0+h) - T(h_0) \le 2\langle x - a, h\rangle. \tag{44}$$

Uma vez que $|\langle x - a_h, h\rangle| \le \|x - a_h\|\,\|h\| = d_F(x)\|h\|$ (note que $a_h \in \Pi_F(x)$), tem-se $\lim\limits_{h\to 0}\langle x - a_h, h\rangle = 0$. Segue-se, portanto, de (44), que

$$\lim_{h\to 0} T(h_0+h) = T(h_0),$$

donde T é contínua.

Agora, tomando-se $h \in \mathbb{R}^n - \{0\}$, $a_h \in \Pi_F(x+h)$ e $a \in \Pi_F(x)$, valem as desigualdades $\|x+h-a_h\| \le \|x+h-a\|$ e $\|x-a\| \le \|x-a_h\|$. Daí, segue-se que

$$2\langle x - a_h, h\rangle + \|h\|^2 \le f(x+h) - f(x) \le 2\langle x - a, h\rangle + \|h\|^2. \tag{45}$$

Assim, tomando-se $a \in \Pi_h(x) \subset \Pi_F(x)$ e observando-se que, para todo real $t > 0$, $T(th) = tT(h)$, obtém-se, de (45), as desigualdades

$$2\left\langle x - a_h, \frac{h}{\|h\|}\right\rangle - T\left(\frac{h}{\|h\|}\right) + \|h\| \le \frac{f(x+h)-f(x)-T(h)}{\|h\|} \le \|h\|. \tag{46}$$

Em particular,

$$2\left\langle x - a_h, \frac{h}{\|h\|}\right\rangle - T\left(\frac{h}{\|h\|}\right) \le 0 \;\forall a_h \in \Pi_F(x+h). \tag{47}$$

Provemos, então, que o primeiro membro de (47) tem limite nulo quando h tende a zero. Para tanto, tomemos uma sequência (h_k) em $\mathbb{R}^n - \{0\}$, tal que $h_k \to 0$. Passando-se a uma subsequência, se necessário, podemos supor que

$$u_k = \frac{h_k}{\|h_k\|} \to u, \;\; \|u\| = 1.$$

Tomando-se uma sequência (a_k) em $\mathbb{R}^n$ satisfazendo, para cada $k \in \mathbb{N}$, $a_k \in \Pi_F(x+h_k)$, tem-se, para todo $b \in F$, que

$$\|a_k\| - \|x+h_k\| \le \|x+h_k-a_k\| \le \|x+h_k-b\|. \tag{48}$$

Considerando-se a primeira e última expressões destas desigualdades, obtém-se

$$\|a_k\| \le \|x+h_k-b\| + \|x+h_k\|.$$

Uma vez que as sequências que têm como termos gerais $\|x+h_k-b\|$ e $\|x+h_k\|$ são limitadas (por serem convergentes), segue-se desta desigualdade que vale o mesmo para a sequência (a_k). Assim, passando-se novamente a uma subsequência, podemos supor que $a_k \to a_0 \in \Pi_F(x)$, pois, tomando-se os limites de ambos os membros na segunda desigualdade (48), obtém-se $\|x-a_0\| \le \|x-b\| \, \forall b \in F$.

Segue-se, portanto, destas últimas considerações, da continuidade de T e de (47), que

$$2\langle x-a_0, u\rangle = \lim_{k\to\infty} 2\langle x-a_k, u_k\rangle \le \lim_{k\to\infty} T(u_k) = T(u) = \min_{a\in\Pi_F(x)} 2\langle x-a, u\rangle,$$

donde $2\langle x-a_0, u\rangle = T(u)$. Desta forma,

$$\lim_{k\to\infty}\left[2\left\langle x-a_k, \frac{h_k}{\|h_k\|}\right\rangle - T\left(\frac{h_k}{\|h_k\|}\right)\right] = 2\langle x-a_0, u\rangle - T(u) = 0.$$

Desta igualdade e da arbitrariedade das sequências (h_k) e (a_k), obtém-se

$$\lim_{h\to 0}\left[2\left\langle x-a_h, \frac{h}{\|h\|}\right\rangle - T\left(\frac{h}{\|h\|}\right)\right] = 0,$$

que, juntamente com a desigualdade (46), nos dá

$$\lim_{h\to 0}\frac{f(x+h)-f(x)-T(h)}{\|h\|} = 0$$

e conclui a demonstração de (i).

Quanto a (ii), basta observarmos que, para todo real $t > 0$,

$$\frac{f(x+th)-f(x)}{t} - T(h) = \frac{f(x+th)-f(x)-T(th)}{\|th\|}\|h\|,$$

donde, por (i),

$$\lim_{t\to 0_+}\frac{f(x+th)-f(x)}{t} = T(h),$$

como desejado. □

Corolário 4. *Nas condições do Lema 3, se F é um conjunto de Chebyshev, então d_F^2 é diferenciável e, para todo $x \in \mathbb{R}^n$,*

$$\nabla d_F^2(x) = 2(x - P_F(x)).$$

Demonstração. Com efeito, sendo F um conjunto de Chebyshev, tem-se, para quaisquer $x, h \in \mathbb{R}^n$, que

$$T(h) = \min_{a\in P_F(x)} 2\langle x-a, h\rangle = 2\langle x - P_F(x), h\rangle,$$

donde T é linear. O resultado segue-se, então, do item (i) do Lema 3. □

No que se segue, dados pontos distintos $x, y \in \mathbb{R}^n$, nos referiremos ao conjunto

$$\sigma = \{(1-t)x + ty \in \mathbb{R}^n;\ t \in [0, +\infty)\}$$

como a *semi-reta* com *origem* em x e que contém y. Note que toda semi-reta σ é conexa, pois é a imagem de $[0, +\infty)$ pela aplicação contínua $t \mapsto (1-t)x + ty$.

LEMA 4. *Sejam $F \subset \mathbb{R}^n$ um conjunto de Chebyshev e $x_0 \in \mathbb{R}^n - F$. Então, a semi-reta σ, com origem em $P_F(x_0)$ e que contém x_0, é tal que*

$$x \in \sigma \ \Rightarrow\ P_F(x) = P_F(x_0).$$

DEMONSTRAÇÃO. Dado $x \in \mathbb{R}^n - F$, todo ponto y do segmento $[P_F(x), x]$ é tal que $P_F(y) = P_F(x)$. Com efeito, neste caso, temos

$$\|x - P_F(x)\| = \|x - y\| + \|y - P_F(x)\|.$$

Logo,

$$\|x - P_F(y)\| \leq \|x - y\| + \|y - P_F(y)\| \leq \|x - y\| + \|y - P_F(x)\| = \|x - P_F(x)\|,$$

donde $P_F(y) = P_F(x)$.

Desta forma, devemos nos ocupar apenas dos pontos da semi-reta $\sigma_0 \subset \sigma$, cuja origem é x_0. Temos, pela continuidade da projeção P_F, que o conjunto

$$\Omega = \{x \in \sigma_0\,;\ P_F(x) \ = \ P_F(x_0)\} \ \subset\ \sigma_0$$

é fechado em σ_0. Vejamos que este conjunto é também aberto em σ_0. Uma vez que $\Omega \neq \emptyset$ (pois $x_0 \in \Omega$), o resultado se seguirá, então, da conexidade de σ_0.

Dado $x \in \Omega \subset \mathbb{R}^n - F$, tomemos $r > 0$, tal que $B[x, r] \cap F = \emptyset$. Façamos, então,

$$\mu = \max_{z \in S[x,r]} d_F(z)$$

e observemos que $\mu > 0$, pois F é fechado e a esfera $S[x, r]$ é disjunta de F. Assim, podemos tomar $\lambda \in \mathbb{R}$, tal que $0 < \lambda < \min\{1, \frac{r^2}{\mu^2}\}$, e definir a função

$$\begin{array}{rcl} \varphi : \mathbb{R}^n & \rightarrow & \mathbb{R} \\ z & \mapsto & d^2_{F_0}(z) - \lambda d^2_F(z), \end{array}$$

em que $F_0 = \{x\}$.

Pelo Corolário 4, φ é diferenciável e satisfaz, para todo $z \in \mathbb{R}^n$,

$$\nabla\varphi(z) = 2((z - x) - \lambda(z - P_F(z))).$$

Em particular, φ é contínua. Sendo assim, a restrição de φ à bola (compacta) $B[x, r]$ assume um valor mínimo em algum ponto $z_0 \in B[x, r]$. Entretanto, $\varphi(x) = -\lambda d^2_F(x) < 0$ e, para todo ponto z da esfera $S[x, r]$, tem-se

$$\varphi(z) = r^2 - \lambda d^2_F(z) \geq r^2 - \lambda\mu^2 > 0.$$

Segue-se que $z_0 \in B(x, r)$ e, portanto, que $\nabla\varphi(z_0) = 0$ (vide Exercício 3), isto é,

$$z_0 - x = \lambda(z_0 - P_F(z_0)). \tag{49}$$

Lembrando-se que $0 < \lambda < 1$, conclui-se de (49) que $x \in (P_F(z_0), z_0)$. Logo, pelas nossas considerações iniciais, $P_F(z_0) = P_F(x) = P_F(x_0)$, ou seja, $x \in (P_F(x_0), z_0)$ (Fig. 6). Fazendo-se, então, $r_0 = \|z_0 - x\|$, tem-se que $I_0 = B(x, r_0) \cap \sigma_0$ é um aberto de σ_0 que contém x e, claramente, está contido em $[x_0, z_0]$. Logo, todo ponto $y \in I_0$ satisfaz $P_F(y) = P_F(x_0)$, donde se infere que $I_0 \subset \Omega$ e, portanto, que Ω é aberto em σ_0. $\square$

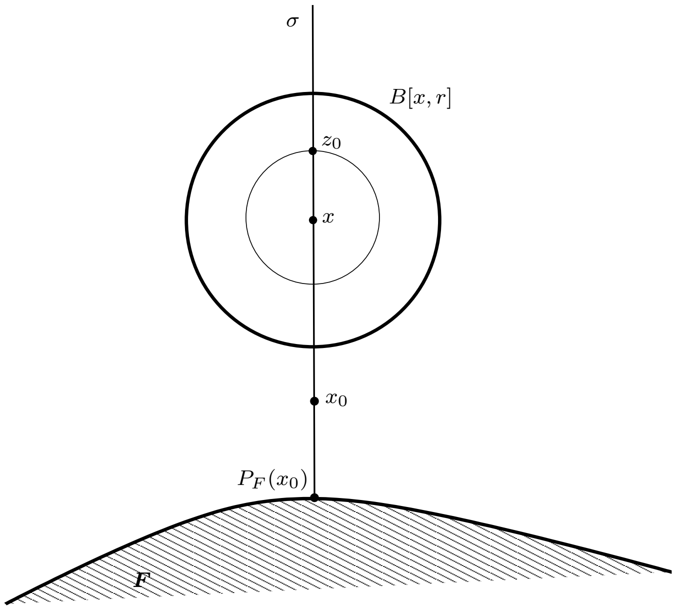

FIGURA 6

TEOREMA DE MOTZKIN. *As seguintes afirmações a respeito de um subconjunto fechado F de $\mathbb{R}^n$ são equivalentes:*

i) d_F *é diferenciável em* $\mathbb{R}^n - F$;

ii) F *é um conjunto de Chebyshev;*

iii) F *é convexo.*

DEMONSTRAÇÃO. Suponhamos que d_F seja diferenciável em $\mathbb{R}^n - F$. Então, vale o mesmo para $f = d_F^2$. Logo, pelo item (ii) do Lema 3, para quaisquer $x \in \mathbb{R}^n - F$ e $h \in \mathbb{R}^n$, tem-se

$$f'(x)h = \frac{\partial f}{\partial h}(x) = \lim_{t\to 0} \frac{f(x+th) - f(x)}{t} = \lim_{t\to 0_+} \frac{f(x+th) - f(x)}{t} = T(h),$$

donde $T = f'(x)$. Em particular,

$$Th = \langle \nabla f(x), h\rangle \quad \forall h \in \mathbb{R}^n.$$

Desta forma, tomando-se $a = x - \frac{\nabla f(x)}{2}$, tem-se, pela definição de T, que

$$\langle x - a, h\rangle \leq \langle x - b, h\rangle \quad \forall h \in \mathbb{R}^n,\ b \in \Pi_F(x).$$

Daí, tomando-se $b \in \Pi_F(x)$ e $h = b - a$, obtém-se $\|b - a\|^2 \leq 0$, ou seja, $b = a$. Logo, F é um conjunto de Chebyshev.

Agora, se $F \subset \mathbb{R}^n$ é um conjunto de Chebyshev, então, pelo Corolário 4, d_F^2 é diferenciável em $\mathbb{R}^n$ e, portanto, $d_F = \sqrt{d_F^2}$ é diferenciável em $\mathbb{R}^n - F$. Desta forma, (i) e (ii) são equivalentes.

Suponhamos agora que F seja convexo. Neste caso, dados $x \in \mathbb{R}^n - F$, $a \in \Pi_F(x)$ e $b \in F$, $b \neq a$, temos que o segmento fechado $[a,b]$ está contido em F. Logo, para todo $x_t = a + t(b-a) \in [a,b]$, $0 \le t \le 1$, tem-se

$$\|x-a\|^2 \le \|x - x_t\|^2 = \|x-a\|^2 - 2t\langle x-a, b-a\rangle + t^2\|b-a\|^2, \tag{50}$$

donde, para todo $t \in (0,1]$, $2\langle x-a, b-a\rangle \le t\|b-a\|^2$, o que nos dá $\langle x-a, b-a\rangle \le 0$. Fazendo-se $t = 1$ na igualdade em (50), obtém-se, então,

$$\|x-b\|^2 - \|x-a\|^2 = -2\langle x-a, b-a\rangle + \|b-a\|^2 > 0,$$

isto é, $\|x-a\| < \|x-b\|$. Segue-se que a é o único elemento de $\Pi_F(x)$ e, portanto, que F é de Chebyshev.

Finalmente, suponhamos que F seja de Chebyshev e, por absurdo, que não seja convexo. Então, existem $a, b \in F$, tais que $[a,b] \not\subset F$. Além disso, podemos supor, sem perda de generalidade, que o segmento aberto (a,b) não intersecta F (Fig. 7). Com efeito, cada componente conexa de $(\mathbb{R}^n - F) \cap (a,b)$ é um aberto e conexo de (a,b) e, portanto, é um segmento da forma (a_0, b_0), em que $a_0, b_0 \in F$ [(vi)].

Fazendo-se $\alpha(t) = a + t(b-a)$, $t \in [0,1]$, tem-se, pelo Corolário 4, que a função $g(t) = d_F^2(\alpha(t))$ é contínua, diferenciável em $(0,1)$ e satisfaz $g(0) = g(1) = 0$. Então, pelo Teorema de Rolle (vide Exercício 3), existe $t_0 \in (0,1)$, tal que

$$0 = g'(t_0) = 2\langle \alpha(t_0) - P_F(\alpha(t_0)), b-a\rangle,$$

donde a semi-reta σ, que contém $x_0 = \alpha(t_0) \in (a,b)$ e tem origem em $P_F(x_0)$, é ortogonal ao segmento $[a,b]$ (Fig. 8).

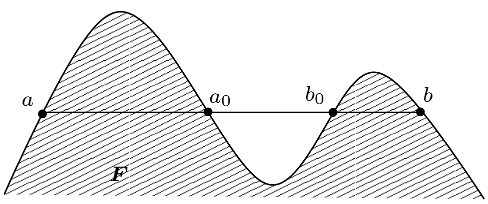

Figura 7

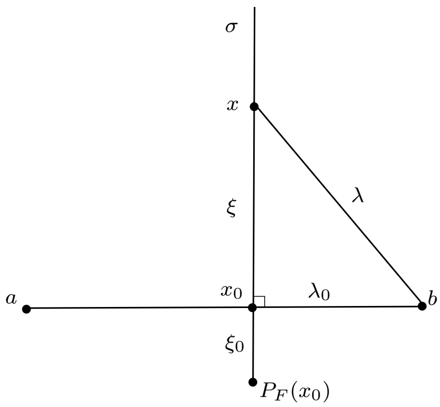

Figura 8

Definamos $u = \frac{x_0 - P_F(x_0)}{\|x_0 - P_F(x_0)\|}$ e consideremos $x = x_0 + \xi u \in \sigma$. Escrevendo-se

$$\xi_0 = \|x_0 - P_F(x_0)\|, \quad \lambda_0 = \|b - x_0\| \quad \text{e} \quad \lambda = \|x-b\|,$$

tem-se

$$d(x_0, F) = \xi_0 < \lambda_0\,, \quad \|x - P_F(x_0)\| = \xi_0 + \xi \quad \text{e} \quad \lambda^2 = \xi^2 + \lambda_0^2\,.$$

Logo, tomando-se $\xi > \frac{\lambda_0^2 - \xi_0^2}{2\xi_0}$, obtém-se

$$\|x - P_F(x_0)\|^2 = (\xi_0 + \xi)^2 = \xi_0^2 + 2\xi_0\xi + \xi^2 > \lambda_0^2 + \xi^2 = \lambda^2 = \|x-b\|^2,$$

[(vi)] Note que o segmento (a,b) é homeomorfo ao intervalo aberto $(0,1) \subset \mathbb{R}$ e que os únicos conjuntos abertos e conexos de $\mathbb{R}$ são os intervalos abertos.

isto é, $\|x-b\| < \|x-P_F(x_0)\|$. Em particular, $P_F(x) \neq P_F(x_0)$, o que contradiz o Lema 4. Segue-se desta contradição que F é convexo e, portanto, que as afirmações (ii) e (iii) são equivalentes. $\square$

8. Exercícios

Seções 1 e 2

1. Em cada item abaixo, mostre que a aplicação dada é diferenciável e determine a sua derivada.
   i) $f:\mathbb{R}^2\to\mathbb{R}^2$, $f(x,y)=(x^2+y, x+y^2)$;
   ii) $f:\mathrm{L}(\mathbb{R}^n)\to\mathrm{L}(\mathbb{R}^n)$, $f(X)=X^3$.

2. Considere a função $f:\mathbb{R}^2\to\mathbb{R}$, definida por

$$f(x,y)=\begin{cases}\left(1-\cos\frac{x^2}{y}\right)\sqrt{x^2+y^2} & \text{se} \quad y\neq 0,\\[2ex] 0 & \text{se} \quad y=0.\end{cases}$$

   Prove que:
   i) A aplicação $T(v)=\frac{\partial f}{\partial v}(0,0)$ está bem definida em $\mathbb{R}^2$ e é linear;
   ii) f não é diferenciável em $(0,0)$.

3. Dados um conjunto $X\subset\mathbb{R}^n$ e uma função $f:X\to\mathbb{R}$, diz-se que $x_0\in X$ é um *ponto de máximo local* de f se existe um aberto (relativo) $V\subset X$, tal que $x_0\in V$ e $f(x_0)\geq f(x)\,\forall x\in V$. De modo análogo, define-se *ponto de mínimo local*.
   i) Seja $f:U\to\mathbb{R}$ uma função definida num aberto U de $\mathbb{R}^n$. Mostre que se $x\in U$ é um ponto de máximo ou de mínimo local de f e f é diferenciável em x, então $f'(x)=0$.
   ii) Prove o *Teorema de Rolle*: Seja U um aberto limitado de $\mathbb{R}^n$ e $f:\overline{U}\to\mathbb{R}$ uma função contínua, diferenciável em U e constante em ∂U. Então, existe $x\in U$, tal que $f'(x)=0$.
   iii) Suponha que $f:\mathbb{R}^n\to\mathbb{R}$ seja diferenciável e que, para cada $u\in S^{n-1}$, valha a desigualdade $f'(u)u>0$. Prove que existe um ponto $x\in\mathbb{R}^n$, tal que $f'(x)=0$.

4. Diz-se que uma função $f:\mathbb{R}^n\to\mathbb{R}$ é *positivamente homogênea* se, para todo $x\in\mathbb{R}^n$ e todo $t>0$, tem-se $f(tx)=tf(x)$. Prove que toda função positivamente homogênea e diferenciável em $x=0$ é, necessariamente, linear. Conclua que a função, $f:\mathbb{R}^2\to\mathbb{R}$, definida por

$$f(x,y)=\begin{cases}\frac{x^3}{x^2+y^2} & \text{se} \quad (x,y)\neq(0,0)\\[2ex] 0 & \text{se} \quad (x,y)=(0,0),\end{cases}$$

   não é diferenciável em $(0,0)$.

5. Seja $f : U \to \mathbb{R}^m$ diferenciável no aberto $U \subset \mathbb{R}^n$. Se, para algum $b \in \mathbb{R}^m$, o conjunto $f^{-1}(\{b\})$ possui um ponto de acumulação $a \in U$, então $f'(a) : \mathbb{R}^n \to \mathbb{R}^m$ não é injetiva.

6. Sejam $U \subset \mathbb{R}^n$ um conjunto aberto e $f : U \to \mathbb{R}^m$ uma aplicação diferenciável que satisfaz, para quaisquer $x, y \in U$, $\|f(x) - f(y)\| \leq \mu\|x - y\|^2$, $\mu > 0$. Prove que, para todo $x \in U$, tem-se $f'(x) = 0$.

7. Considere uma aplicação diferenciável $f : \mathbb{R}^n \to \mathbb{R}^n$, tal que $f(0) = 0$ e $\|f'(0)\| = \mu < 1$. Prove que existe uma bola aberta B de $\mathbb{R}^n$, centrada em 0, satisfazendo $f(B) \subset B$.

8. Dada uma aplicação contínua $f : S^{n-1} \to \mathbb{R}^m$, defina a *extensão radial* de f como a aplicação $\varphi : \mathbb{R}^n \to \mathbb{R}^m$, tal que

$$\varphi(x) = \begin{cases} \|x\|\, f\left(\frac{x}{\|x\|}\right) & x \neq 0, \\ \\ 0 & x = 0. \end{cases}$$

Mostre que φ é diferenciável na origem $0 \in \mathbb{R}^n$ se, e somente se, f é a restrição a S^{n-1} de uma aplicação linear.

9. Prove as afirmações seguintes a respeito da função determinante:
   i) Para toda matriz $H \in \mathrm{M}(n)$, tem-se $\det'(I)H = \text{traço}(H)$, em que I é a matriz identidade de $\mathrm{M}(n)$;
   ii) dada $X \in \mathrm{M}(n)$, $\det'(X) = 0$ se, e somente se, $\text{posto}\,(X) \leq n - 2$.

10. Seja $f : \mathbb{R}^n \to S^n \subset \mathbb{R}^n \times \mathbb{R}$ a inversa da projeção estereográfica, isto é,

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

Mostre que, para todo $x \in \mathbb{R}^n$, f é diferenciável em x e:
   i) $f'(x)$ é um isomorfismo conforme de $\mathbb{R}^n$ sobre $f'(x)(\mathbb{R}^n)$ (vide Exercício 9 – Capítulo 1);
   ii) $f'(x)(\mathbb{R}^n) = \{f(x)\}^{\perp}$.

Seção 3

11. Sejam $U = \mathbb{R}^2 - \{0\} \subset \mathbb{R}^2$ e $f : U \to \mathbb{R}^3$ definida por $f(x, y) = (x^2, y^2, (x+y)^2)$. Prove que, para todo $(x, y) \in U$, $f'(x, y) : \mathbb{R}^2 \to \mathbb{R}^3$ é injetiva.

12. Sejam $U \subset \mathbb{R}^n$ aberto e $f : U \to \mathbb{R}$ diferenciável em $x \in U$. Prove que, se $f'(x) \neq 0$, então, dentre todos os vetores $h \in \mathbb{R}^n$ tais que $\|h\| = 1$, a derivada direcional $\frac{\partial f}{\partial h}(x)$ atinge seu valor máximo quando $h = \nabla f(x)/\|\nabla f(x)\|$.

Seção 4

13. Determine a derivada segunda de cada uma das aplicações do Exercício 1.

14. Seja $f : \mathbb{R}^3 \to \mathbb{R}$ definida por $f(x_1, x_2, x_3) = 3x_1^2 + 2x_2^2 + 2x_1x_3 + x_3^2$. Prove que, para todo $x = (x_1, x_2, x_3) \in \mathbb{R}^3$, a derivada segunda $f''(x)$ é uma forma bilinear positiva definida, isto é, $f''(x)(h, h) > 0 \ \forall h \in \mathbb{R}^3 - \{0\}$.

15. Dados um aberto $U \subset \mathbb{R}^n$ e uma função $f : U \to \mathbb{R}$, diz-se que $x \in U$ é um *ponto crítico* de f se $f'(x) = 0$. Quando f é duas vezes diferenciável num ponto crítico x e $\operatorname{hess} f(x)$ tem determinante não-nulo, x é dito um ponto crítico *não-degenerado*. Prove que:

    i) Todo ponto crítico não-degenerado x de uma função duas vezes diferenciável $f : U \to \mathbb{R}$ é um ponto crítico isolado, isto é, existe um aberto $V \ni x$, tal que $V - \{x\}$ não contém pontos críticos de f;

    ii) o conjunto dos pontos críticos da função $f(x) = 2\|x\|^2 - \|x\|^4$, $x \in \mathbb{R}^{n+1}$, é compacto e contém um único ponto crítico não-degenerado.

Seção 5

16. Determine o maior aberto U de $\mathbb{R}^n$ no qual a função
$$\begin{array}{rccc} f: & \mathbb{R}^n & \to & \mathbb{R} \\ & x & \mapsto & \|x\|_s \end{array}$$
é diferenciável e calcule, para cada $x \in U$, a derivada $f'(x)$.

17. Dados abertos $U \subset \mathbb{R}^n$ e $V \subset \mathbb{R}^m$, suponha que $f : U \to V$ e $\lambda : V \to \mathbb{R}$ sejam diferenciáveis. Dado $x \in U$, prove que
$$\nabla(\lambda \circ f)(x) = (f'(x))^* \nabla\lambda(f(x)).$$

18. Sejam f_1, f_2 as funções-coordenada de uma aplicação holomorfa $f : \mathbb{R}^2 \to \mathbb{R}^2$ e $\alpha_1, \alpha_2 : (0, 1) \to \mathbb{R}^2$ curvas diferenciáveis, tais que, para cada $i \in \{1, 2\}$, a função $f_i \circ \alpha_i$ é constante, isto é, cada α_i é uma *curva de nível* da função f_i. Prove que se $t, s \in (0, 1)$ satisfazem $\alpha_1(t) = \alpha_2(s)$, então $\alpha_1'(t)$ e $\alpha_2'(s)$ são ortogonais.

19. Seja $f : U \to \mathbb{R}^n$ diferenciável no aberto $U \subset \mathbb{R}^n$. Prove que se a função $x \mapsto \|f(x)\|$ é constante em U, então, para todo $x \in U$, o determinante da matriz jacobiana de f em x é identicamente nulo.

20. Diz-se que uma função diferenciável $f : \mathbb{R}^n - \{0\} \to \mathbb{R}$ é *positivamente homogênea* de grau $r \in \mathbb{R}$ se, para quaisquer $t > 0$ e $x \in \mathbb{R}^n - \{0\}$, tem-se $f(tx) = t^r f(x)$. Mostre que se f é positivamente homogênea de grau r, então
$$f'(x)x = rf(x) \, \forall x \in \mathbb{R}^n - \{0\}.$$

21. Sejam $U \subset \mathbb{R}^n$, $V \subset \mathbb{R}^m$ abertos e $f : U \to V$, $g : V \to \mathbb{R}^p$ aplicações duas vezes diferenciáveis. Mostre que, para quaisquer $x \in U$ e $h, k \in \mathbb{R}^n$, tem-se
$$(g \circ f)''(x)(h, k) = g''(f(x))(f'(x)h, f'(x)k) + g'(f(x))f''(x)(h, k).$$

22. O *laplaciano* de uma função duas vezes diferenciável, $f : \mathbb{R}^n \to \mathbb{R}$, é a função $\Delta f : \mathbb{R}^n \to \mathbb{R}$, definida por
$$\Delta f(x) = \text{traço}\,(\operatorname{hess} f(x)) = \sum_{i=1}^{n} \frac{\partial^2 f}{\partial x_i^2}(x).$$

Prove que se $T \in \mathrm{L}(\mathbb{R}^n)$ é um operador ortogonal, então $\Delta(f \circ T) = \Delta f \circ T$.

23. Sejam $I \subset \mathbb{R}$ um intervalo aberto e $A : I \to \mathrm{M}(n)$ uma curva diferenciável em $\mathrm{M}(n)$, tal que $A(t) = (a_{ij}(t))_{n\times n}$, $t \in I$. Suponha que $\det(A(t)) > 0\,\forall t \in I$ e defina $f(t) = \log(\det A(t))$. Prove que
$$f'(t) = \sum_{i,j=1}^{n} a'_{ij}(t) b_{ji}(t),$$
em que $B(t) = (b_{ij}(t))_{n\times n}$ é a inversa de $A(t)$, $t \in I$.

Seção 6

24. Seja $U = \{(x, y) \in \mathbb{R}^2;\, x > |y|\}$. Prove que a aplicação
$$\begin{array}{rccc} f: & U & \to & \mathbb{R}^2 \\ & (x,y) & \mapsto & \left(\log(\frac{x+y}{2}), \log(\frac{x-y}{2})\right). \end{array}$$
é um difeomorfismo.

25. Prove que toda inversão com respeito a uma esfera é um difeomorfismo conforme.

CAPíTULO 5

Teoremas Fundamentais do Cálculo Diferencial

Neste capítulo, concluiremos nossas considerações sobre o cálculo diferencial de aplicações em $\mathbb{R}^n$ estabelecendo seus resultados mais fundamentais, quais sejam, o Teorema do Valor Médio, o Teorema de Schwarz, o Teorema de Taylor, o Teorema da Função Inversa e o Teorema da Função Implícita. A cada um destes teoremas será devotada uma seção e cada uma delas conterá, pelo menos, uma subseção em que se aplica o teorema a ela correspondente. Mantendo o espírito dos capítulos precedentes, encerraremos, então, apresentando um célebre resultado relativo às aplicações diferenciáveis, conhecido como Teorema de Sard.

1. O Teorema do Valor Médio

O Teorema do Valor Médio, atribuído ao matemático francês Joseph-Louis Lagrange (1736–1813), é frequentemente considerado o resultado mais essencial do cálculo das funções reais de uma variável, pois constitui a chave das demonstrações dos teoremas mais importantes desta teoria. Ele afirma que, dada uma função $f : [a,b] \subset \mathbb{R} \to \mathbb{R}$, contínua e diferenciável em (a,b), existe $x \in (a,b)$, tal que

$$f'(x) = \frac{f(b) - f(a)}{b - a}\,, \tag{51}$$

isto é, a reta tangente ao gráfico de f em $(x, f(x))$ é paralela à reta determinada por $(a, f(a))$ e $(b, f(b))$.

Para sua demonstração, basta aplicar o Teorema de Rolle (vide Exercício 3 – Capítulo 4) à função $\varphi : [a,b] \to \mathbb{R}$, dada por

$$\varphi(x) = f(x) - \frac{f(b) - f(a)}{b - a}x.$$

Com efeito, é imediato que φ é contínua, diferenciável em (a,b) e satisfaz $\varphi(a) = \varphi(b)$. Logo, pelo Teorema de Rolle, existe $x \in (a,b)$, tal que $\varphi'(x) = 0$, donde se obtém a igualdade (51).

Fazendo-se $b = a + h$, temos que $x = a + th$ para algum $t \in (0,1)$. Assim, a igualdade (51) assume a forma

$$f(a+h) - f(a) = f'(a+th)h, \tag{52}$$

a qual, convém mencionar, pressupõe apenas a estrutura vetorial de $\mathbb{R}$.

No que se segue, estenderemos o Teorema do Valor Médio para aplicações em espaços euclidianos. Constataremos então, nas seções subsequentes, que o mesmo desempenha um papel crucial nas demonstrações de todos os resultados fundamentais que apresentaremos neste capítulo.

Dados $a, b \in \mathbb{R}^n$, denotaremos respectivamente por $[a,b]$ e (a,b), os segmentos de reta fechado e aberto de extremidades a e b, isto é,

$$[a,b] = \{a + t(b-a);\, 0 \le t \le 1\} \quad \text{e} \quad (a,b) = \{a + t(b-a);\, 0 < t < 1\}.$$

Note que, se $n = 1$, os segmentos $[a,b]$ e (a,b) ficam definidos mesmo quando $a > b$. Neste caso, eles coincidem com os intervalos $[b,a]$ e (b,a).

Teorema do Valor Médio (para Funções). *Seja* $f : U \subset \mathbb{R}^n \to \mathbb{R}$ *contínua no segmento fechado* $[a, a+h] \subset U$ *e diferenciável em* $(a, a+h)$. *Então, existe* $t \in (0,1)$, *tal que*

$$f(a+h) - f(a) = f'(a+th)h.$$

Demonstração. Defina $\varphi : [0,1] \to \mathbb{R}$ por $\varphi(t) = f(a+th)$. Então, φ é contínua e, pela Regra da Cadeia, diferenciável em $(0,1)$. Aplicando-se a φ o Teorema do Valor Médio para funções de uma variável, obtém-se $t \in (0,1)$, tal que $\varphi(1) - \varphi(0) = \varphi'(t)$, o que implica

$$f(a+h) - f(a) = f'(a+th)h,$$

como desejado. □

Quando $m > 1$, não há, em geral, um teorema do valor médio para aplicações $f : U \subset \mathbb{R}^n \to \mathbb{R}^m$ que se traduza através de uma igualdade, como no caso das funções. Considere, por exemplo, a curva $\alpha : [0, 2\pi] \subset \mathbb{R} \to \mathbb{R}^2$, dada por $\alpha(t) = (\cos t, \operatorname{sen} t)$. Tem-se $\alpha'(t) = (-\operatorname{sen} t, \cos t) \ne (0,0)\, \forall t \in (0, 2\pi)$. Por outro lado, $\alpha(0) = \alpha(2\pi)$. Logo, não existe $t_0 \in (0, 2\pi)$, tal que $\alpha(2\pi) - \alpha(0) = \alpha'(t_0)2\pi$.

No entanto, se $\alpha : [a,b] \to \mathbb{R}^n$ é uma curva diferenciável em (a,b) e, para todo $t \in (a,b)$, tem-se $\|\alpha'(t)\| \le \mu$, então vale a desigualdade

$$\|\alpha(b) - \alpha(a)\| \le \mu(b-a). \tag{53}$$

De fato, a função $\varphi : [a,b] \to \mathbb{R}$, dada por

$$\varphi(t) = \langle \alpha(t), \alpha(b) - \alpha(a) \rangle,$$

claramente, cumpre as condições do Teorema do Valor Médio para funções de uma variável. Desta forma, existe $t_0 \in (a,b)$, tal que $\varphi(b) - \varphi(a) = \varphi'(t_0)(b-a)$. Segue-se daí e da desigualdade de Cauchy-Schwarz que

$$|\varphi(b) - \varphi(a)| = |\varphi'(t_0)(b-a)| = |\langle \alpha'(t_0), \alpha(b) - \alpha(a) \rangle|(b-a) \le \mu \|\alpha(b) - \alpha(a)\|(b-a).$$

Logo,

$$\|\alpha(b) - \alpha(a)\|^2 = |\varphi(b) - \varphi(a)| \le \mu \|\alpha(b) - \alpha(a)\|(b-a),$$

donde se obtém (53).

Observemos que, no argumento do parágrafo precedente, usamos fortemente o fato da norma euclidiana ser advinda de um produto interno. Provaremos agora que a desigualdade do valor médio (53) é válida para qualquer norma de $\mathbb{R}^n$, conforme o teorema que se segue.

Teorema do Valor Médio (para Curvas). *Sejam* $\|\ \|$ *uma norma arbitrária em* $\mathbb{R}^n$ *e* $\alpha : [a,b] \subset \mathbb{R} \to \mathbb{R}^n$ *uma curva diferenciável no intervalo* (a,b). *Então, se* $\|\alpha'(t)\| \le \mu$ *para todo* $t \in (a,b)$, *tem-se*

$$\|\alpha(b) - \alpha(a)\| \le \mu(b-a).$$

Demonstração. Dados $\epsilon > 0$ e $b_0 \in (a, b)$, consideremos o conjunto

$$X_\epsilon = \{t \in [a, b_0];\ \|\alpha(b_0) - \alpha(t)\| \leq (\mu + \epsilon)(b_0 - t)\}$$

e observemos que X_ϵ é não-vazio (pois $b_0 \in X_\epsilon$) e limitado. Façamos, então, $t_0 = \inf X_\epsilon$ e provemos que $t_0 = a$.

Dada uma sequência convergente (t_k) em X_ϵ, se $t \in [a, b_0]$ é o seu limite, pela continuidade da função $s \mapsto \|\alpha(b_0) - \alpha(s)\|$, temos

$$\|\alpha(b_0) - \alpha(t)\| = \lim_{k\to\infty} \|\alpha(b_0) - \alpha(t_k)\| \leq \lim_{k\to\infty} (\mu + \epsilon)(b_0 - t_k) = (\mu + \epsilon)(b_0 - t),$$

donde $t \in X_\epsilon$. Segue-se que X_ϵ é fechado e, portanto, que $t_0 \in X_\epsilon$.

Suponhamos, por absurdo, que se tenha $t_0 > a$. Neste caso, $t_0 \in (a, b_0] \subset (a, b)$ e α, desta forma, é diferenciável em t_0. Logo, valem as igualdades

$$\alpha(t_0 + \delta) - \alpha(t_0) = \delta\alpha'(t_0) + r(\delta), \quad \lim_{\delta\to 0} \frac{r(\delta)}{|\delta|} = 0.$$

Podemos, então, tomar $\delta < 0$, de tal forma que $a < t_0 + \delta < t_0$ e $\|r(\delta)\| < \epsilon|\delta|$. Desta forma, teremos

$$\|\alpha(t_0 + \delta) - \alpha(t_0)\| = \|\delta\alpha'(t_0) + r(\delta)\| \leq |\delta|\,\|\alpha'(t_0)\| + \|r(\delta)\| \leq (\mu + \epsilon)|\delta|.$$

Daí, segue-se que

$$\begin{aligned} \|\alpha(b_0) - \alpha(t_0 + \delta)\| &\leq \|\alpha(b_0) - \alpha(t_0)\| + \|\alpha(t_0) - \alpha(t_0 + \delta)\| \\ &\leq (\mu + \epsilon)(b_0 - t_0) + (\mu + \epsilon)|\delta| \\ &= (\mu + \epsilon)(b_0 - (t_0 + \delta)), \end{aligned} \tag{54}$$

donde $t_0 + \delta \in X_\epsilon$. Isto, porém, vai de encontro ao fato de t_0 ser o ínfimo de X_ϵ e prova que $a = \inf X_\epsilon \in X_\epsilon$. Logo, $\|\alpha(b_0) - \alpha(a)\| \leq (\mu + \epsilon)(b_0 - a)$. Uma vez que ϵ foi tomado arbitrariamente, conclui-se que

$$\|\alpha(b_0) - \alpha(a)\| \leq \mu(b_0 - a).$$

Finalmente, tomando-se uma sequência (b_k) em (a, b), tal que $b_k \to b$, tem-se

$$\|\alpha(b) - \alpha(a)\| = \lim_{k\to\infty} \|\alpha(b_k) - \alpha(a)\| \leq \lim_{k\to\infty} \mu(b_k - a) = \mu(b - a),$$

concluindo, assim, a demonstração. □

Teorema do Valor Médio (para Aplicações). *Sejam $U \subset \mathbb{R}^n$ um conjunto aberto e $f : U \to \mathbb{R}^m$ uma aplicação contínua no segmento $[a, a + h] \subset U$ e diferenciável em $(a, a + h)$. Então, se $\|f'(x)\| \leq \mu$ (norma espectral) para todo $x \in (a, a + h)$, vale a desigualdade*

$$\|f(a + h) - f(a)\| \leq \mu\|h\|.$$

Demonstração. Consideremos a aplicação $\alpha : [0, 1] \to \mathbb{R}^m$, definida por $\alpha(t) = f(a+th)$, e observemos que α é contínua (isto é, é uma cuva) e diferenciável em $(0, 1)$. Além disso, para todo $t \in (0, 1)$, tem-se

$$\|\alpha'(t)\| = \|f'(a + th)h\| \leq \|f'(a + th)\|\,\|h\| \leq \mu\|h\|.$$

Segue-se, então, do teorema anterior, que

$$\|f(a + h) - f(a)\| = \|\alpha(1) - \alpha(0)\| \leq \mu\|h\|,$$

como desejado. □

Corolário 5 (aplicações de derivada nula em conexos). *Sejam* $U \subset \mathbb{R}^n$ *aberto e conexo e* $f : U \to \mathbb{R}^m$ *uma aplicação diferenciável, tal que* $f'(x) = 0$ *para todo* $x \in U$. *Então,* f *é constante.*

Demonstração. Tomemos arbitrariamente $a \in U$ e consideremos o conjunto $X_a = \{x \in U;\, f(x) = f(a)\}$. Temos que X_a é não-vazio, pois $a \in X_a$. Além disso, como f é contínua, X_a é fechado em U. Agora, dado $x \in X_a$, uma vez que U é aberto, existe $\delta > 0$, tal que, para todo $h \in \mathbb{R}^n$ satisfazendo $\|h\| < \delta$, tem-se $x + h \in U$. Logo, pelo Teorema do Valor Médio, $\|f(x+h) - f(x)\| = 0$, isto é, $f(x+h) = f(x) = f(a)$. Segue-se que X_a é aberto em U. Como U é conexo e X_a é não vazio, temos que $X_a = U$, donde f é constante. □

1.1. Uma Condição Suficiente para Diferenciabilidade. Vimos, no capítulo anterior, que uma aplicação diferenciável possui todas as suas derivadas direcionais, em particular, suas derivadas parciais. Assinalamos ali também que, em contrapartida, a existência das derivadas direcionais não é uma condição suficiente para a diferenciabilidade. Aplicaremos, então, o Teorema do Valor Médio para mostrar que a existência e continuidade das derivadas parciais de uma aplicação é uma condição suficiente para a diferenciabilidade da mesma.

Teorema 17 (continuidade das derivadas parciais). *Consideremos uma aplicação* $f : U \subset \mathbb{R}^n \to \mathbb{R}^m$, U *aberto, e* $a \in U$, *tais que as derivadas parciais de* f *existem e são contínuas num aberto* $A \ni a$, $A \subset U$. *Então,* f *é diferenciável em* a.

Demonstração. Uma vez que a diferenciabilidade de uma aplicação equivale à de suas coordenadas, basta considerarmos o caso $m = 1$.

Suponhamos, sem perda de generalidade, que $A \subset U$ é uma bola aberta de $\mathbb{R}^n$ centrada em a. Tomemos, então, para cada $i \in \{1, 2, \dots, n\}$, a projeção $P_i : \mathbb{R}^n \to \mathbb{R}$, $P_i(x) = x_i = \langle x, e_i \rangle$, e façamos $I_i = P_i(A)$. Note que cada $I_i \subset \mathbb{R}$ é um intervalo aberto de $\mathbb{R}$, pois P_i leva bolas abertas em bolas abertas (vide Exercício 15 – Capítulo 2).

Escrevamos $a = (a_1, \dots, a_n)$ e consideremos a função $g_1 : I_1 \to \mathbb{R}$, definida por $g_1(x) = f(x, a_2, a_3, \dots, a_n)$. Observemos que g_1 é diferenciável e

$$\begin{aligned} g_1'(x) &= \lim_{t \to 0} \frac{g_1(x+t) - g_1(x)}{t} \\ &= \lim_{t \to 0} \frac{f(x+t, a_2, a_3, \dots, a_n) - f(x, a_2, a_3, \dots, a_n)}{t} \\ &= \frac{\partial f}{\partial x_1}(x, a_2, \dots, a_n). \end{aligned}$$

Tomemos $\delta > 0$ suficientemente pequeno, de tal forma que, para quaisquer $i \in \{2, \dots, n\}$ e $x \in I_i$, se tenha

$$\|h\| < \delta \;\Rightarrow\; (a_1 + h_1, a_2 + h_2, \dots, a_{i-1} + h_{i-1}, x, a_{i+1}, a_{i+2}, \dots, a_n) \in U.$$

Assim, para cada tal i e cada tal h, podemos definir a função

$$g_i : I_i \to \mathbb{R},\;\; g_i(x) = f(a_1 + h_1, a_2 + h_2, \dots, a_{i-1} + h_{i-1}, x, a_{i+1}, a_{i+2}, \dots, a_n).$$

Temos, desta forma, que g_i é diferenciável e, a exemplo de g_1, cumpre

$$g_i'(x) = \frac{\partial f}{\partial x_i}(a_1 + h_1, a_2 + h_2, \dots, a_{i-1} + h_{i-1}, x, a_{i+1}, a_{i+2}, \dots, a_n).$$

Além disso, verifica-se facilmente que, para cada $i \in \{1, 2, \cdots, n\}$, tem-se

$$g_{i+1}(a_{i+1}) = g_i(a_i + h_i).$$

Segue-se destas considerações que

$$\begin{aligned} f(a+h) - f(a) &= f(a_1 + h_1, \ldots, a_n + h_n) - f(a_1, \ldots, a_n) \\ &= g_n(a_n + h_n) - g_1(a_1) \\ &= \sum_{i=1}^{n} g_i(a_i + h_i) - g_i(a_i). \end{aligned}$$

Pelo Teorema do Valor Médio, para cada $i \in \{1, \ldots, n\}$, existe um real $t_i = t_i(h_i) \in (0, 1)$, satisfazendo $g_i(a_i + h_i) - g_i(a_i) = g_i'(a_i + t_i h_i)h_i$. Logo,

$$f(a+h) - f(a) = \sum_{i=1}^{n} g_i(a_i + h_i) - g_i(a_i) = \sum_{i=1}^{n} g_i'(a_i + t_i h_i)h_i = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(v_i)h_i\,,$$

em que $v_i = v_i(h) = (a_1 + h_1, \ldots, a_{i-1} + h_{i-1}, a_i + t_i h_i, a_{i+1}, a_{i+2}, \ldots, a_n)$. Note que

$$\lim_{h \to 0} v_i(h) = a.$$

Naturalmente, nosso candidato a derivada de f em a é a transformação linear

$$T : \mathbb{R}^n \to \mathbb{R}, \quad Th = \langle \nabla f(a), h \rangle = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(a)h_i\,.$$

Esta, por sua vez, determina a função resto

$$r(h) = f(a+h) - f(a) - \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(a)h_i = \sum_{i=1}^{n} \left(\frac{\partial f}{\partial x_i}(v_i) - \frac{\partial f}{\partial x_i}(a) \right) h_i\,.$$

Assim, uma vez que $|h_i|/\|h\| \leq 1$, considerando-se a desigualdade triangular, obtém-se

$$\frac{|r(h)|}{\|h\|} \leq \sum_{i=1}^{n} \left| \frac{\partial f}{\partial x_i}(v_i) - \frac{\partial f}{\partial x_i}(a) \right|.$$

Segue-se, então, da continuidade das derivadas parciais de f em A, que

$$\lim_{h \to 0} \frac{r(h)}{\|h\|} = 0$$

e, portanto, que f é diferenciável em a. $\square$

Dada $f : U \subset \mathbb{R}^n \to \mathbb{R}^m$ diferenciável, observemos que as derivadas parciais de f são as coordenadas de f'. Logo, se as derivadas parciais de f são contínuas em U, f é de classe C^1.

Reciprocamente, se as derivadas parciais de f existem e são contínuas em U, pelo Teorema 17, f é diferenciável em U e, pelo argumento do parágrafo anterior, de classe C^1. Vale, portanto, o resultado seguinte.

Corolário 6. *Uma aplicação $f : U \subset \mathbb{R}^n \to \mathbb{R}^m$, U aberto, é de classe C^1 se, e somente se, suas derivadas parciais existem e são contínuas em U.*

Exemplo 74. Consideremos a função $f : \mathbb{R}^2 \to \mathbb{R}$, definida por

$$f(x,y) = \begin{cases} \frac{xy(x^2-y^2)}{x^2+y^2} & (x,y) \neq (0,0) \\ 0 & (x,y) = (0,0). \end{cases}$$

Temos que

$$\lim_{t\to 0} \frac{f(0+t,0)-f(0,0)}{t} = 0,$$

donde $\frac{\partial f}{\partial x}(0,0) = 0$.

Fazendo-se, para $(x,y) \neq (0,0)$, $g(x,y) = \frac{x^2-y^2}{x^2+y^2}$, tem-se

$$\frac{\partial g}{\partial x}(x,y) = \frac{4xy^2}{(x^2+y^2)^2}\,.$$

Assim, observando-se que, para todo $(x,y) \neq (0,0)$, $f(x,y) = xyg(x,y)$, tem-se, para um tal (x,y),

$$\frac{\partial f}{\partial x}(x,y) = yg(x,y) + xy\frac{\partial g}{\partial x}(x,y) = yg(x,y) + y\left(\frac{2xy}{x^2+y^2}\right)^2, \tag{55}$$

donde se conclui que a função $\partial f/\partial x$ é contínua em $\mathbb{R}^n - \{(0,0)\}$. Além disso, observando-se que

$$g(x,y) < 2 \quad \text{e} \quad \frac{2xy}{x^2+y^2} \leq 1,$$

segue-se de (55) que

$$\lim_{(x,y)\to(0,0)} \frac{\partial f}{\partial x}(x,y) = 0 = \frac{\partial f}{\partial x}(0,0),$$

donde $\partial f/\partial x$ é contínua também em $(0,0)$.

Analogamente, a derivada parcial $\partial f/\partial y$ está bem definida em $\mathbb{R}^2$ e é uma função contínua. Desta forma, pelo Corolário 6, f é uma função de classe C^1.

1.2. Diferenciabilidade e Compacidade. Verificaremos agora, através do Teorema do Valor Médio, que a restrição de uma aplicação de classe C^1 a um subconjunto compacto de seu domínio é lipschitziana, conforme a proposição seguinte.

Proposição 50. *Sejam $U \subset \mathbb{R}^n$ aberto, $f : U \to \mathbb{R}^m$ uma aplicação de classe C^1 e $K \subset U$ compacto. Então, $f|_K$ é lipschitziana.*

Demonstração. Sendo f de classe C^1 e K compacto, temos, pelo Teorema de Weierstrass, que

$$\mu = \sup_{x\in K} \|f'(x)\|$$

está bem definido.

Suponhamos, por absurdo, que, para todo $k \in \mathbb{N}$, existam $x_k, y_k \in K$, tais que

$$\|f(x_k) - f(y_k)\| > k\|x_k - y_k\|. \tag{56}$$

Ainda pela compacidade de K, podemos supor, sem perda de generalidade, que $x_k \to a \in K$ e $y_k \to b \in K$. A função $x \mapsto \|f(x)\|$, $x \in K$, é contínua. Logo, é limitada, pois seu conjunto-imagem é compacto. Em particular, a sequência

$\|f(x_k)-f(y_k)\|$ é limitada. Daí e de (56), infere-se que $\|x_k-y_k\| \to 0$ e, portanto, que $a=b$.

Seja $B \subset U$ uma bola aberta centrada em a e contida em U. Então, existe $k_0 \in \mathbb{N}$, tal que, para todo $k \geq k_0$, tem-se $x_k, y_k \in B$. Uma vez que toda bola é convexa, para todo $k \geq k_0$, o segmento $[x_k, y_k]$ está contido em B e, consequentemente, em U. Assim, podemos fazer

$$\mu_k = \sup_{x \in [x_k, y_k]} \|f'(x)\|, \quad k \geq k_0\,.$$

Além disso, como $[x_k, y_k]$ é compacto e f' é contínua, existe $z_k \in [x_k, y_k]$, tal que $\mu_k = \|f'(z_k)\|$. Observe que z_k não pertence, necessariamente, a K (Fig. 1).

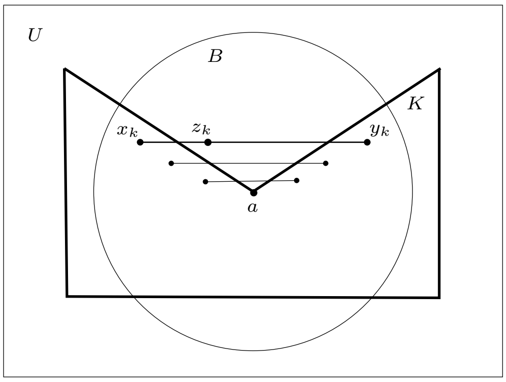

FIGURA 1

Agora, pelo Teorema do Valor Médio, para todo $k \geq k_0$, vale a desigualdade $\|f(x_k)-f(y_k)\| \leq \mu_k\|x_k-y_k\|$, donde, por (56), $\mu_k > k\ \forall k \geq k_0$. Em particular, para todo k suficientemente grande, tem-se $\|f'(z_k)\| = \mu_k > k > 2\mu$. Escrevendo-se, então, $z_k = (1-t_k)x_k + t_k y_k$, $t_k \in [0,1]$, e $a = (1-t_k)a + t_k a$, obtém-se

$$\|z_k - a\| = \|(1-t_k)(x_k-a) + t_k(y_k-a)\| \leq (1-t_k)\|(x_k-a)\| + t_k\|(y_k-a)\|,$$

donde se conclui que $z_k \to a$. Logo, $\|f'(z_k)\| \to \|f'(a)\| \geq 2\mu > \mu$, o que contradiz a definição de μ.

Segue-se que, para algum $k \in \mathbb{N}$,

$$\|f(x)-f(y)\| \leq k\|x-y\| \quad \forall x, y \in K,$$

donde $f|_K$ é lipschitziana. □

2. O Teorema de Schwarz

O Teorema de Schwarz estabelece a simetria da derivada segunda de uma aplicação em todo ponto de seu domínio em que esta é duas vezes diferenciável. Ele é assim designado em consideração ao matemático alemão Hermann Schwarz (1864–1951), a quem, juntamente com o matemático e astrônomo francês Alexis Clairaut (1713–1765)[(i)], atribui-se sua demonstração.

[(i)]Por esta razão, este resultado é também conhecido como *Teorema de Clairaut*.

Segue-se deste teorema que se f é uma aplicação definida num aberto U de $\mathbb{R}^n$ e duas vezes diferenciável em $x \in U$, então

$$\frac{\partial^2 f}{\partial x_i \partial x_j}(x) = \frac{\partial^2 f}{\partial x_j \partial x_i}(x) \quad \forall i, j \in \{1, \dots, n\}.$$

Em particular, a matriz hessiana de f é simétrica e, portanto, o operador em $\mathbb{R}^n$ a ela associado é auto-adjunto.

A simetria da derivada segunda estabelecida pelo Teorema de Schwarz se faz bastante útil em inúmeros contextos, em particular, na resolução de alguns tipos de equações diferenciais parciais, conforme ilustraremos no fim da seção.

Teorema de Schwarz. *Seja $f : U \subset \mathbb{R}^n \to \mathbb{R}^m$ uma aplicação diferenciável no aberto U e duas vezes diferenciável em $x \in U$. Então, $f''(x)$ é uma forma bilinear simétrica, isto é,*

$$f''(x)(h, k) = f''(x)(k, h) \;\; \forall\, h, k \in \mathbb{R}^n.$$

Demonstração. Como na demonstração do teorema precedente, e pelo mesmo motivo, basta provarmos o resultado no caso em que $m = 1$.

Dados, então, $h, k \in \mathbb{R}^n$, tomemos $\epsilon > 0$ suficientemente pequeno, de tal modo que a função

$$\varphi(t) \;=\; f(x + t(h+k)) - f(x + th) - f(x + tk) + f(x), \quad t \in (0, \epsilon),$$

esteja bem definida. Para cada $t \in (0, \epsilon)$, façamos

$$\xi(s) = f(x + sh + tk) - f(x + sh), \;\; s \in [0, t].$$

A função $\xi : [0, t] \to \mathbb{R}$, claramente, é contínua e, pela Regra da Cadeia, é diferenciável em $(0, t)$ e satisfaz

$$\xi'(s) = f'(x + sh + tk)h - f'(x + sh)h \;\; \forall s \in (0, t).$$

Podemos, desta forma, aplicar à mesma o Teorema do Valor Médio e obter $\theta = \theta(t) \in (0, 1)$, tal que

$$\xi(t) - \xi(0) = \xi'(\theta t)t = f'(x + t(\theta h + k))th - f'(x + \theta th)th\,.$$

Observando-se que $\varphi(t) = \xi(t) - \xi(0)$ e fazendo-se $v_1 = v_1(t) = \theta h + k$ e $v_2 = v_2(t) = \theta h$, obtém-se, da última igualdade acima,

$$\varphi(t) = (f'(x + tv_1) - f'(x + tv_2))th. \tag{57}$$

Agora, levando-se em conta que f é duas vezes diferenciável em x, tem-se

$$f'(x + tv_i) = f'(x) + f''(x)(tv_i) + r_i(tv_i), \quad \lim_{t \to 0} \frac{r_i(tv_i)}{\|tv_i\|} = 0, \;\; i = 1, 2.$$

Daí e de (57) conclui-se que

$$\frac{\varphi(t)}{t^2} = f''(x)(v_1, h) - f''(x)(v_2, h) + \left(\frac{r_1(tv_1)}{t} - \frac{r_2(tv_2)}{t}\right) h.$$

No entanto,

$$f''(x)(v_1, h) - f''(x)(v_2, h) = f''(x)(\theta h + k, h) - f''(x)(\theta h, h) = f''(x)(k, h),$$

donde

$$\lim_{t \to 0} \frac{\varphi(t)}{t^2} = f''(x)(k, h).$$

Agora, basta notar que a expressão que define φ permanece inalterada quando permutam-se as variáveis h e k, donde se infere que

$$f''(x)(h,k) = f''(x)(k,h),$$

como desejado. □

Usando-se indução e o Teorema de Schwarz, prova-se facilmente que se f é $k-1$ vezes diferenciável no aberto $U \subset \mathbb{R}^n$ e k vezes diferenciável em $x \in U$, então, a derivada de ordem k de f em x, $f^{(k)}(x)$, é uma aplicação k-linear simétrica.

Reconsideremos a função $f : \mathbb{R}^2 \to \mathbb{R}$ do Exemplo 74, a qual, segundo verificamos, é de classe C^1, cumpre as igualdades

$$\frac{\partial f}{\partial x}(0,0) = \frac{\partial f}{\partial y}(0,0) = 0,$$

e, para $(x,y) \neq (0,0)$, satisfaz

$$\frac{\partial f}{\partial x}(x,y) = yg(x,y) + y\left(\frac{2xy}{x^2+y^2}\right)^2 \quad \text{e} \quad \frac{\partial f}{\partial y}(x,y) = xg(x,y) - x\left(\frac{2xy}{x^2+y^2}\right)^2,$$

em que $g(x,y) = \frac{x^2-y^2}{x^2+y^2}\cdot$ Daí, obtém-se

$$\lim_{t\to 0}\frac{\frac{\partial f}{\partial x}(0,t) - \frac{\partial f}{\partial x}(0,0)}{t} = \lim_{t\to 0} g(0,t) = -1,$$

donde $\frac{\partial^2 f}{\partial y \partial x}(0,0) = -1$, e

$$\lim_{t\to 0}\frac{\frac{\partial f}{\partial y}(t,0) - \frac{\partial f}{\partial y}(0,0)}{t} = \lim_{t\to 0} g(t,0) = 1,$$

isto é, $\frac{\partial^2 f}{\partial x \partial y}(0,0) = 1$. Desta forma,

$$\frac{\partial^2 f}{\partial x \partial y}(0,0) \neq \frac{\partial^2 f}{\partial y \partial x}(0,0).$$

Logo, pelo Teorema de Schwarz, f não é duas vezes diferenciável em $(x,y) = (0,0)$.

2.1. A Equação da Onda. A fim de ilustrar a aplicabilidade do Teorema de Schwarz, apresentaremos o método de d'Alembert[(ii)] para obtenção da solução da equação diferencial parcial

$$\frac{\partial^2 u}{\partial t^2} = \frac{1}{c^2}\frac{\partial^2 u}{\partial x^2}\,, \ u = u(x,t), \ \ c > 0, \tag{58}$$

que satisfaz as *condições de fronteira*

$$u(0,t) = u(L,t) = 0, \ \ t \geq 0, \tag{59}$$

e as *condições iniciais*

$$u(x,0) = f(x) \quad \text{e} \quad \frac{\partial u}{\partial t}(x,0) = 0, \ 0 < x < L. \tag{60}$$

Esta equação, dentre outros fenômenos ondulatórios, modela a propagação da onda produzida numa corda vibrante horizontal (com respeito a um sistema de

[(ii)] Em consideração ao matemático francês Jean-Baptiste le Rond d'Alembert(1717–1783), que o introduziu.

coordenadas xy), de comprimento L e com extremidades fixas (condições de fronteira). A função $u = u(x,t)$ fornece a altura y do ponto da corda correspondente à abscissa x no instante t (Fig. 2), donde as condições iniciais fornecem a posição e a velocidade de cada ponto x no instante $t = 0$, sendo esta última, supostamente, nula. Finalmente, c é uma constante que depende das características físicas da corda.

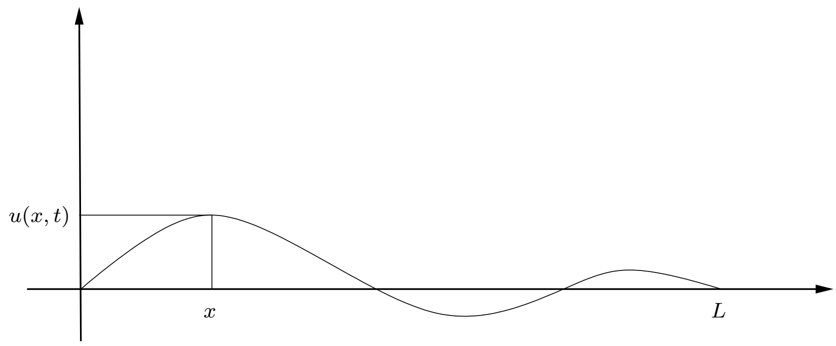

FIGURA 2

Aplicando-se a este fenômeno certos princípios da Física, mais especificamente, as leis de Hooke e a segunda lei de Newton, verifica-se, então, que a função u satisfaz a equação (58), conhecida como *equação da onda unidimensional*.

A ideia do método de d'Alembert é, através de uma mudança de coordenadas conveniente (vide Seção 6 – Capítulo 4), transformar a equação (58) numa equação do tipo

$$\frac{\partial^2 u}{\partial \eta \partial \xi} = 0\,. \tag{61}$$

Antes de fazê-lo, verifiquemos que toda solução de (61) é da forma

$$u(\xi,\eta) = u_1(\xi) + u_2(\eta).$$

Com efeito, sendo $u = u(\xi,\eta)$ uma solução, temos que $\frac{\partial u}{\partial \xi}$ não depende de η, isto é, $\frac{\partial u}{\partial \xi} = u_0(\xi)$. Uma vez que u é, supostamente, duas vezes diferenciável, temos que u_0 é contínua e, portanto, possui uma primitiva $u_1 = u_1(\xi)$, isto é, $u_1' = u_0$. Fazendo-se $u_2(\xi,\eta) = u(\xi,\eta) - u_1(\xi)$, tem-se $\frac{\partial u_2}{\partial \xi} = \frac{\partial u}{\partial \xi} - u_1'(\xi) = 0$, donde $u_2(\xi,\eta) = u_2(\eta)$.

Retomemos a equação (58) e consideremos a mudança de coordenadas

$$\varphi(\xi,\eta) = (x(\xi,\eta), t(\xi,\eta)) = \left(\frac{\xi+\eta}{2}, \frac{\eta-\xi}{2c}\right),$$

cuja inversa é

$$\varphi^{-1}(x,t) = (\xi(x,t), \eta(x,t)) = (x - ct, x + ct).$$

Temos, pela Regra da Cadeia (vide igualdade (41) – Capítulo 4), que

$$\frac{\partial u}{\partial \xi} = \frac{\partial u}{\partial x}\frac{\partial x}{\partial \xi} + \frac{\partial u}{\partial t}\frac{\partial t}{\partial \xi} = \frac{1}{2}\left(\frac{\partial u}{\partial x} - \frac{1}{c}\frac{\partial u}{\partial \xi}\right),$$

donde

$$\begin{aligned}
2\frac{\partial^2 u}{\partial\eta\partial\xi} &= \frac{\partial}{\partial\eta}\left(\frac{\partial u}{\partial x}\right) - \frac{1}{c}\frac{\partial}{\partial\eta}\left(\frac{\partial u}{\partial t}\right)\\
&= \left(\frac{\partial^2 u}{\partial x^2}\frac{\partial x}{\partial\eta} + \frac{\partial^2 u}{\partial t\partial x}\frac{\partial t}{\partial\eta}\right) - \frac{1}{c}\left(\frac{\partial^2 u}{\partial x\partial t}\frac{\partial x}{\partial\eta} + \frac{\partial^2 u}{\partial t^2}\frac{\partial t}{\partial\eta}\right)\\
&= \frac{1}{2}\left(\frac{\partial^2 u}{\partial x^2} - \frac{1}{c^2}\frac{\partial^2 u}{\partial t^2}\right) + \frac{1}{2c}\left(\frac{\partial^2 u}{\partial t\partial x} - \frac{\partial^2 u}{\partial x\partial t}\right).
\end{aligned}$$

Segue-se do Teorema de Schwarz, portanto, que a equação da onda (58) é equivalente à equação (61). Desta forma, temos $u(\xi,\eta) = u_1(\xi) + u_2(\eta)$, isto é,

$$u(x,t) = u_1(x - ct) + u_2(x + ct).$$

Além disso, pela Regra da Cadeia e pelas condições iniciais dadas,

$$f(x) = u(x,0) = u_1(x) + u_2(x) \quad \text{e} \quad 0 = \frac{\partial u}{\partial t}(x,0) = -cu_1'(x) + cu_2'(x).$$

A segunda igualdade implica que $u_2 = u_1 + c_0$, $c_0 \in \mathbb{R}$. Logo, $f = u_1 + u_2 = 2u_1 + c_0$, em que a constante c_0 é determinada pelas condições de fronteira. No caso particular em que $c_0 = 0$, tem-se $u_1 = u_2 = f/2$, donde

$$u(x,t) = \frac{1}{2}(f(x - ct) + f(x + ct))$$

é a solução de (58) que satisfaz as condições (59) e (60).

3. O Teorema de Taylor

Desde o capítulo anterior, vimos estudando o comportamento local das aplicações diferenciáveis valendo-nos do fato de que as mesmas expressam-se, localmente, como aproximações de transformações lineares.

Entretanto, em alguns contextos, a derivada (de ordem 1) não dá informações suficientes sobre a aplicação, fazendo-se necessário, quando pertinente, derivá-la mais de uma vez e, a partir do comportamento das derivadas de ordem superior (no ponto), deduzir o comportamento da aplicação (numa vizinhança do ponto).

A Fórmula de Taylor, que deduziremos a seguir, constitui um forte instrumento em abordagens deste tipo. Ela deve seu nome ao matemático inglês Brook Taylor (1685–1731), que a introduziu em 1712.

Lema 5. *Sejam B uma bola aberta de $\mathbb{R}^n$ com centro na origem e $r : B \to \mathbb{R}^n$ uma aplicação diferenciável e duas vezes diferenciável no ponto 0. Então, se a aplicação r, juntamente com suas derivadas de ordem 1 e 2, se anulam no ponto 0, tem-se*

$$\lim_{h\to 0}\frac{r(h)}{\|h\|^2} = 0.$$

Demonstração. Dado $h \in B$, segue-se da hipótese que

$$r'(h) = r'(0 + h) - r'(0) - r''(0)h$$

e, portanto

$$\lim_{h\to 0}\frac{r'(h)}{\|h\|} = 0.$$

Então, dado $\epsilon > 0$, existe $\delta > 0$, tal que $\|r'(h)\| \le \epsilon\|h\| \ \forall h \in B(0,\delta) \subset B$. Assim, tomando-se $h \in B(0,\delta)$ e $v \in [0,h]$, tem-se $\|r'(v)\| \le \epsilon\|v\| \le \epsilon\|h\|$, donde r' é limitada em $[0,h]$ por $\epsilon\|h\|$.

Desta forma, aplicando-se o Teorema do Valor Médio a r no segmento $[0,h]$, obtém-se $\|r(h)\| \le \epsilon\|h\|^2$, donde

$$\lim_{h\to 0} \frac{r(h)}{\|h\|^2} = 0,$$

como desejado. □

Teorema de Taylor. *Seja $f : U \subset \mathbb{R}^n \to \mathbb{R}^m$ uma aplicação diferenciável no aberto U e duas vezes diferenciável em $x \in U$. Então, a aplicação $r = r(h)$, definida pela igualdade*

$$f(x+h) = f(x) + f'(x)h + \frac{1}{2}f''(x)(h,h) + r(h),$$

satisfaz

$$\lim_{h\to 0} \frac{r(h)}{\|h\|^2} = 0.$$

Demonstração. Seja $\varphi : \mathbb{R}^n \to \mathbb{R}^m$ a aplicação definida por

$$\varphi(h) = \frac{1}{2}f''(x)(h,h).$$

Pela Regra da Cadeia, φ é diferenciável. Além disso, uma vez que $f''(x)$ é uma aplicação bilinear simétrica, tem-se, para todo $v \in \mathbb{R}^n$, que

$$\varphi'(h)v = \frac{1}{2}(f''(x)(v,h) + f''(x)(h,v)) = f''(x)(v,h).$$

Em particular, $\varphi'(0) = 0$.

Observando-se que φ' é linear, tem-se $\varphi''(h) = \varphi' \, \forall h \in \mathbb{R}^n$, donde

$$\varphi''(h)(a,b) = \varphi'(a)b = f''(x)(a,b) \ \ \forall a,b \in \mathbb{R}^n,$$

isto é, $\varphi''(h) = f''(x) \, \forall h \in \mathbb{R}^n$.

Agora, uma vez que $r(h) = f(x+h) - f(x) - f'(x)h - \varphi(h)$, tem-se $r'(h) = f'(x+h) - f'(x) - \varphi'(h)$ e, portanto,

$$r''(h) = f''(x+h) - \varphi''(h) = f''(x+h) - f''(x).$$

Desta forma, $r(0) = r'(0) = r''(0) = 0$ e o resultado segue-se, então, do Lema 5. □

Exemplo 75. Consideremos uma função diferenciável, $f : \mathbb{R}^2 \to \mathbb{R}$, e suponhamo-la duas vezes diferenciável em $(0,0)$. Fazendo-se $h = (x,y)$, tem-se

- $f'(0)h = \langle \nabla f(0,0), h \rangle = \dfrac{\partial f}{\partial x}(0,0)x + \dfrac{\partial f}{\partial y}(0,0)y\,;$
- $f''(0)(h,h) = \dfrac{\partial^2 f}{\partial x^2}(0,0)x^2 + 2\dfrac{\partial^2 f}{\partial xy}(0,0)xy + \dfrac{\partial^2 f}{\partial y^2}(0,0)y^2\,.$

Logo, pelo Teorema de Taylor, numa vizinhança de $(0,0)$, f assume a forma

$$f(x,y) = f(0,0) + \frac{\partial f}{\partial x}x + \frac{\partial f}{\partial y}y + \frac{1}{2}\left(\frac{\partial^2 f}{\partial x^2}x^2 + 2\frac{\partial^2 f}{\partial xy}xy + \frac{\partial^2 f}{\partial y^2}y^2\right) + r(x,y),$$

em que

$$\lim_{(x,y)\to(0,0)} \frac{r(x,y)}{\|(x,y)\|^2} = 0$$

e todas as derivadas indicadas são calculadas no ponto $(0,0)$.

Usando-se indução, prova-se facilmente uma versão do Lema 5, na qual se supõe que a função r é k vezes diferenciável em B e $k+1$ vezes diferenciável no ponto 0, e se conclui que

$$\lim_{h\to 0} \frac{r(h)}{\|h\|^{k+1}} = 0. \tag{62}$$

A partir daí, verifica-se que se $f : U \subset \mathbb{R}^n \to \mathbb{R}^m$ é k vezes diferenciável no aberto U e $k+1$ vezes diferenciável em $x \in U$, então a aplicação $r = r(h)$, definida por

$$f(x+h) = f(x) + \sum_{j=1}^{k+1} \frac{1}{j!} f^{(j)}(x) h^{(j)} + r(h), \tag{63}$$

satisfaz (62). Aqui, $f^{(j)}(x)$ denota a derivada de ordem j de f em x, e $h^{(j)} = (h, \ldots, h)$ é a j-upla de vetores de $\mathbb{R}^n$ cujas entradas são todas iguais a h.

A equação (63) é dita a *fórmula de Taylor de ordem* $k+1$ *com resto infinitesimal*. Ela fornece, numa vizinhança de x, uma aproximação de f por uma aplicação cujas coordenadas são polinômios de grau não superior a $k+1$. Cada um destes polinômios determina uma função resto $r_i = r_i(h)$, $i = 1, \ldots, n$, que é um *infinitésimo de ordem maior que, ou igual a,* $k+1$, significando que $r_i(h)$ converge para zero mais rapidamente que $\|h\|^{k+1}$ quando $h \to 0$.

Uma das aplicações mais comuns do Teorema de Taylor é aquela feita ao estudo dos máximos e mínimos locais de uma função diferenciável, que faremos a seguir.

3.1. Máximos e Mínimos Locais. Dada uma função $f : U \subset \mathbb{R}^n \to \mathbb{R}$, relembremos que $a \in U$ é dito um *máximo local* (respectivamente, *mínimo local*) de f se existe uma vizinhança V de a, tal que $f(x) \leq f(a)$ (respectivamente, $f(x) \geq f(a)$) para todo $x \in U \cap V$. Todo ponto que é um máximo ou mínimo local de uma função é dito um *extremo local* desta.

Se U é aberto, f é diferenciável em $a \in U$ e $f'(a) = 0$, diz-se que a é um *ponto crítico* de f.

Como sabemos (vide Exercício 3 – Capítulo 4), todo extremo local de uma função diferenciável é um ponto crítico. Veremos, na proposição seguinte, condições suficientes para que um ponto crítico seja um extremo local.

Proposição 51. *Sejam* $f : U \subset \mathbb{R}^n \to \mathbb{R}$ *uma função duas vezes diferenciável no aberto* U *e* $a \in U$ *um ponto crítico de* f*. Então:*

i) a *é um mínimo local de* f *se* $f''(a)$ *é positiva definida;*

ii) a *é um máximo local de* f *se* $f''(a)$ *é negativa definida;*

iii) a *não é um extremo local de* f *se* $f''(a)$ *é indefinida.*

Demonstração. A Fórmula de Taylor aplicada a f em $x = a$ nos dá

$$f(a+h) = f(a) + \frac{1}{2}f''(a)(h,h) + r(h), \quad \lim_{h\to 0} \frac{r(h)}{\|h\|^2} = 0. \tag{64}$$

Pela compacidade de S^{n-1} e continuidade de $f''(a)$, temos que

$$\mu = \frac{1}{2}\inf\left\{f''(a)(u,u); u \in S^{n-1}\right\}$$

está bem definido. Além disso, pelo Teorema de Weierstrass, existe $u_0 \in S^{n-1}$, tal que $2\mu = f''(a)(u_0,u_0)$.

Suponhamos que $f''(a)$ seja positiva definida. Neste caso, temos $\mu > 0$. Podemos, então, pelo limite em (64), tomar $\delta > 0$, tal que

$$-\mu < \frac{r(h)}{\|h\|^2} < \mu \ \forall h \in B(0,\delta) - \{0\}.$$

Desta forma,

$$\frac{f(a+h)-f(a)}{\|h\|^2} = \frac{1}{2}f''(a)\left(\frac{h}{\|h\|}, \frac{h}{\|h\|}\right) + \frac{r(h)}{\|h\|^2} > \mu - \mu = 0,$$

isto é, $f(a+h) > f(a)$. Segue-se que, para todo $x = a+h \in B(a,\delta)$, $f(x) \geq f(a)$, donde a é um mínimo local de f. Isto prova (i).

Agora, se $f''(a)$ é negativa definida, fazendo-se $g = -f$, tem-se que $g''(a)$ é positiva definida. Logo, por (i), a é um mínimo local de g e, portanto, um máximo local de f, o que prova (ii).

Finalmente, supondo-se $f''(a)$ indefinida, temos que existem h, k em S^{n-1}, tais que $f''(a)(h,h) > 0$ e $f''(a)(k,k) < 0$. Tomando-se $t > 0$ satisfazendo $|r(th)|/\|th\|^2 = |r(th)|/t^2 < \mu$ e considerando-se (64), como fizemos acima, conclui-se igualmente que $f(a+th) - f(a) > 0$. De modo análogo, obtém-se $s > 0$, tal que $f(a+sk) - f(a) < 0$, isto é, a não é máximo ou mínimo local de f. □

Exemplo 76. Consideremos a função $f : \mathbb{R}^2 \to \mathbb{R}$, definida por

$$f(x,y) = x^2 + y^2 - xy + 2x + 2y + 1.$$

Temos que $\nabla f(x,y) = (2x - y + 2, 2y - x + 2)$ e, portanto, $a = (-2,-2)$ é o único ponto crítico de f, pois é a única solução de $\nabla f(x,y) = (0,0)$. Além disso, dado $h = (h_1,h_2) \in \mathbb{R}^2$, um cálculo direto nos dá

$$f''(a)(h,h) = 2h_1^2 - 2h_1h_2 + 2h_2^2 = (h_1 - h_2)^2 + h_1^2 + h_2^2\,,$$

donde $f''(a)$ é positiva definida. Logo, pela Proposição 51, $a = (-2,-2)$ é um mínimo local de f.

4. O Teorema da Função Inversa

Apresentaremos agora um dos resultados mais importantes da Análise dentre aqueles ligados ao conceito de diferenciabilidade, o Teorema da Função Inversa. Ele estabelece que uma aplicação de classe C^1, $f : U \subset \mathbb{R}^n \to \mathbb{R}^n$, cuja derivada num ponto $a \in U$ é um isomorfismo, quando restrita a um certo aberto $V \ni a$, tem como imagem um aberto $f(V) \ni f(a)$, sendo tal restrição um difeomorfismo de V sobre $f(V)$. A força deste teorema, que é de caráter local, reside no fato de que uma propriedade simples da derivada de uma aplicação num ponto, a saber, a de ser invertível, implica numa propriedade bastante especial da função numa vizinhança deste ponto, a de ser um difeomorfismo.

Constatamos, no capítulo anterior, que a derivada de um difeomorfismo, em todo ponto de seu domínio, é um isomorfismo linear. Por outro lado, considerando-se a aplicação $f : \mathbb{R}^2 \to \mathbb{R}^2$, dada por $f(x,y) = (e^x \cos y, e^x \operatorname{sen} y)$, verifica-se que não vale a recíproca. Com efeito, para todo $(x,y) \in \mathbb{R}^2$, o determinante jacobiano de f em (x,y) é igual a $e^{2x} \neq 0$, donde $f'(x,y)$ é um isomorfismo linear. Entretanto, a aplicação f não é um difeomorfismo, pois, claramente, não é bijetiva.

Pode-se constatar facilmente, porém, que para todo $(x_0,y_0) \in \mathbb{R}^2$, existe um aberto $V \subset \mathbb{R}^2$, tal que $(x_0,y_0) \in V$, $f(V)$ é aberto em $\mathbb{R}^2$ e a restrição $f|_V : V \to f(V)$ é um difeomorfismo. Por exemplo, se $(x_0,y_0) = (0,0)$, basta tomar $V = (-1,1) \times (-\frac{\pi}{4}, \frac{\pi}{4})$ (Fig. 3).

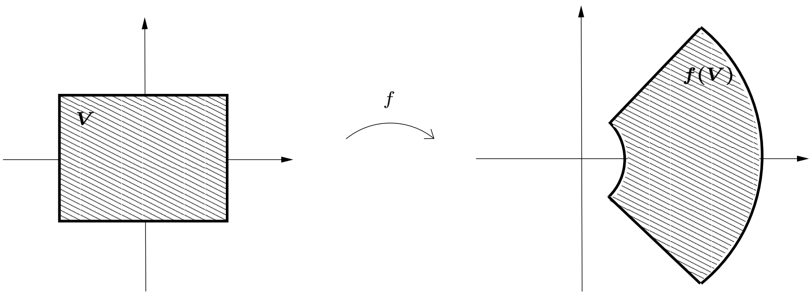

FIGURA 3

Estas considerações motivam a definição seguinte.

DEFINIÇÃO 26 (DIFEOMORFISMO LOCAL). Seja $U \subset \mathbb{R}^n$ um conjunto aberto. Uma aplicação $f : U \to \mathbb{R}^n$ é dita um *difeomorfismo local* (de classe C^k) quando, para todo $x \in U$, existem abertos $V \subset U$ e $W \subset \mathbb{R}^n$, tais que $x \in V$, $f(x) \in W$ e $f|_V : V \to W$ é um difeomorfismo (de classe C^k).

É imediato que todo difeomorfismo local bijetivo é um difeomorfismo. Tem-se, também, que todo difeomorfismo local $f : U \to \mathbb{R}^n$, definido num aberto U de $\mathbb{R}^n$, é uma aplicação aberta. Com efeito, nestas condições, dados um aberto $A \subset U$ e $b \in f(A)$, para todo $a \in f^{-1}(\{b\}) \cap A$, existem abertos $V \ni a$ e $W \ni b$, tais que $V \subset U$ e $f|_V : V \to W$ é um difeomorfismo. Fazendo-se, então, $A_0 = V \cap A$, tem-se que A_0 é aberto, contém a e está contido em V. Assim, uma vez que $f|_V$ é um homeomorfismo, temos que $f|_V(A_0) = f(A_0)$ é aberto e contém $b = f(a)$. Além disso, $f(A_0) = f(A \cap V) = f(A) \cap f(V) \subset f(A)$. Logo, $f(A)$ é aberto e f, desta forma, é uma aplicação aberta.

Observemos que se $f : U \to \mathbb{R}^n$ é um difeomorfismo local, então, para todo $x \in U$, $f'(x)$ é um isomorfismo. O Teorema da Função Inversa, que enunciamos a seguir, nos diz que vale a recíproca quando f é de classe C^1.

TEOREMA DA FUNÇÃO INVERSA. *Sejam $f : U \subset \mathbb{R}^n \to \mathbb{R}^n$ uma aplicação de classe C^1 definida num aberto U de $\mathbb{R}^n$ e $a \in U$. Então, se $f'(a) : \mathbb{R}^n \to \mathbb{R}^n$ é um isomorfismo, existem abertos $V \subset U$ e $W \subset \mathbb{R}^n$, tais que $a \in V$, $f(a) \in W$ e $f|_V : V \to W$ é um difeomorfismo de classe C^1.*

Observemos inicialmente que, nas condições dadas, prova-se facilmente que existe um aberto $V \subset U$, tal que $a \in V$ e a restrição $f|_V : V \to f(V) \subset \mathbb{R}^n$ é bijetiva.

De fato, fazendo-se, por simplicidade de notação, $a = f(a) = 0$ (isto é, considerando-se sistemas de coordenadas em $\mathbb{R}^n$ obtidos pelas respectivas translações da origem aos pontos a e $f(a)$) e $T = f'(0)$, tem-se

$$f(x) = Tx + r(x), \ \lim_{x \to 0} \frac{r(x)}{\|x\|} = 0.$$

Agora, sendo T um isomorfismo, sabemos que existe $\mu > 0$, tal que $\|Tx\| \geq \mu\|x\| \, \forall x \in \mathbb{R}^n$ (vide Exemplo 16 – Capítulo 2). Além disso, a aplicação $r(x) = f(x) - Tx$ é de classe C^1 (pois f e T o são) e $r'(0) = f'(0) - T'(0) = T - T = 0$. Logo, existe $\delta > 0$, tal que $V = B(0, \delta) \subset U$ e $\|r'(x)\| < \frac{\mu}{2} \ \forall x \in V$.

Desta forma, pelo Teorema do Valor Médio, tem-se

$$\|r(x) - r(y)\| \leq \frac{\mu}{2}\|x - y\| \ \ \forall x, y \in V.$$

Tomando-se, então, $x, y \in V$, tem-se $f(x) - f(y) = T(x - y) + r(x) - r(y)$. Logo,

$$\|f(x) - f(y)\| \geq \|T(x - y)\| - \|r(x) - r(y)\| \geq \mu\|x - y\| - \frac{\mu}{2}\|x - y\| = \frac{\mu}{2}\|x - y\|,$$

donde se infere que $f(x) = f(y)$ se, e somente se, $x = y$, isto é, f é injetiva em $V = B(0, \delta)$ e, portanto, $f|_V : V \to f(V)$ é bijetiva, como desejávamos provar.

No entanto, nada garante que $f(V)$ seja um subconjunto aberto de $\mathbb{R}^n$. Este é, justamente, o ponto delicado da demonstração do Teorema da Função Inversa, pois, como veremos adiante, uma vez que este fato esteja estabelecido, a diferenciabilidade da inversa de $f|_V$ será facilmente verificada.

É natural, portanto, que façamos uso de um resultado topológico na demonstração do Teorema da Função Inversa, a saber, o Teorema da Perturbação da Identidade, o qual relembramos (vide Seção 3 – Capítulo 3).

Teorema da Perturbação da Identidade. *Sejam $U \subset \mathbb{R}^n$ um aberto e $\varphi : U \to \mathbb{R}^n$ uma contração. Então, a aplicação $\psi : U \to \psi(U) \subset \mathbb{R}^n$, dada por $\psi(x) = x + \varphi(x)$, é um homeomorfismo e $\psi(U)$ é aberto em $\mathbb{R}^n$.*

Relembremos ainda que $\varphi : U \to \mathbb{R}^n$ é uma *contração* se existe $\lambda \in [0, 1)$, tal que $\|\varphi(x) - \varphi(y)\| \leq \lambda\|x - y\|$ quaisquer que sejam $x, y \in U$.

Demonstração do Teorema da Função Inversa. Retomemos a notação acima, isto é, suponhamos, sem perda de generalidade, que $a = f(a) = 0$ e façamos $T = f'(0)$. Como antes, temos que, numa vizinhança de 0,

$$f(x) = Tx + r(x), \ \lim_{x \to 0} \frac{r(x)}{\|x\|} = 0,$$

r é de classe C^1 e $r'(0) = 0$. Daí, segue-se que, dado $\epsilon > 0$, existe um aberto convexo $V \ni 0$, tal que $\|r'(x)\| < \epsilon$ se $x \in V$ e, então, pelo Teorema do Valor Médio,

$$\|r(x) - r(y)\| \leq \epsilon\|x - y\| \ \ \forall x, y \in V.$$

Consideremos agora a aplicação $\psi : V \to \mathbb{R}^n$, dada por

$$\psi(x) = T^{-1}f(x) = x + T^{-1}r(x),$$

e suponhamos que ϵ seja suficientemente pequeno, de tal forma que $\epsilon\|T^{-1}\| < 1$. Neste caso, $T^{-1}r$ é uma contração, pois dados $x, y \in V$, tem-se

$$\|T^{-1}r(x) - T^{-1}r(y)\| \leq \|T^{-1}\| \, \|r(x) - r(y)\| < \epsilon\|T^{-1}\| \, \|x - y\|.$$

Segue-se, então, do Teorema da Perturbação da Identidade, que ψ é um homeomorfismo de V sobre o aberto $\psi(V) = T^{-1}(f(V)) \subset \mathbb{R}^n$, donde $f(V) = T(\psi(V))$ é aberto em $\mathbb{R}^n$, uma vez que todo isomorfismo linear de $\mathbb{R}^n$ é uma aplicação aberta (vide Exemplo 14 – Capítulo 2). Desta forma, escrevendo-se $f(V) = W$, tem-se que a aplicação

$$T \circ \psi = f|_V : V \to W,$$

por ser uma composta de homeomorfismos, é um homeomorfismo do aberto V sobre o aberto W.

Agora, sendo f de classe C^1, temos que a aplicação $x \mapsto \det f'(x)$, $x \in V$, é contínua. Daí, uma vez que $T = f'(0)$ é um isomorfismo, temos que $\det f'(0) \neq 0$. Logo, pela propriedade de permanência de sinal das funções contínuas, podemos supor que V é suficientemente pequeno, de tal modo que $\det f'(x) \neq 0 \, \forall x \in V$. Assim, temos que, para todo $x \in V$, $f'(x)$ é um isomorfismo linear.

Seja $g : W \to V$ o homeomorfismo inverso de $f|_V$. Dado $y \in W$, façamos $x = g(y)$ e consideremos uma representação de Hadamard de $f|_V$ em x, isto é, uma aplicação $\Phi : V \to \mathrm{L}(\mathbb{R}^n)$, tal que, $\Phi(x) = f'(x)$, Φ é contínua em x e, para todo $x + h \in V$, satisfaz

$$f(x + h) - f(x) = \Phi(x + h)h.$$

Da continuidade de Φ em x e do fato de $\Phi(x) = f'(x)$ ser um isomorfismo, com um argumento análogo ao dado no penúltimo parágrafo acima, podemos supor que, para todo $x + h \in V$, $\Phi(x + h)$ é um isomorfismo. Sendo assim, a aplicação $\Psi : W \to \mathrm{L}(\mathbb{R}^n)$, em que

$$\Psi(y + k) = (\Phi(g(y + k)))^{-1}, \;\; y + k \in W,$$

está bem definida. Além disso, uma vez que g é contínua em y, Φ é contínua em $x = g(y)$ e a inversão de aplicações lineares é contínua, temos que Ψ é contínua em y.

Assim, escrevendo-se $g(y + k) = x + h$, tem-se $h = g(y + k) - g(y)$ e, então,

$$k = f(x + h) - f(x) = \Phi(x + h)h = \Phi(g(y + k))(g(y + k) - g(y)),$$

donde se obtém

$$g(y + k) - g(y) = \Psi(y + k)k.$$

Segue-se que Ψ é uma representação de Hadamard de g em y. Logo, g é diferenciável em y e

$$g'(y) = \Psi(y) = (f'(x))^{-1}.$$

Resta-nos, pois, mostrar que g é de classe C^1. Para isto, basta observarmos que $g' = \varphi \circ f' \circ g$, em que φ é a inversão de transformações lineares. Desta forma, g' é contínua, por ser a composta de aplicações contínuas, e, portanto, a aplicação g é de classe C^1. □

Observação 17. Do argumento do último parágrafo da demonstração do Teorema da Função Inversa, conclui-se que, em seu enunciado, pode-se substituir "C^1" por "C^k, $k \in \mathbb{N}$", pois a inversão de transformações lineares é uma aplicação de classe C^∞ (vide Exemplo 69 – Capítulo 4). No entanto, a hipótese de f ser de classe C^1 é necessária, isto é, ela não pode ser substituída pela hipótese, mais fraca, de f ser diferenciável.

A fim de constatarmos isto, consideremos a função $f : \mathbb{R} \to \mathbb{R}$, definida por

$$f(x) = \begin{cases} \frac{x}{2} + x^2 \text{sen}\,\frac{1}{x} & \text{se} \quad x \neq 0 \\ 0 & \text{se} \quad x = 0. \end{cases}$$

Para $x \neq 0$, tem-se

$$f'(x) = \frac{1}{2} + 2x\,\text{sen}\,\frac{1}{x} - \cos\frac{1}{x}\,,$$

donde f' é descontínua em $x = 0$, pois a sequência $f'(1/k)$ é divergente. Logo, f não é de classe C^1.

Agora, considerando-se as sequências $x_k = (\frac{\pi}{2} + 2k\pi)^{-1}$ e $y_k = (2k\pi)^{-1}$, tem-se $0 < x_k < y_k < \frac{1}{k}$ e $f'(y_k) < 0 < f'(x_k)$. Desta forma, em nenhum intervalo aberto contendo 0 a função f é monótona (se o fosse f' seria não-negativa ou não-positiva nesse intervalo), donde se infere que, a despeito de $f'(0) = 1/2 \neq 0$, não existe um aberto de $\mathbb{R}$ contendo $x = 0$, tal que a restrição de f a este seja invertível.

Exemplo 77 (submersões). Considere uma aplicação $f : U \subset \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^n$, definida num aberto U de $\mathbb{R}^n \times \mathbb{R}^m = \mathbb{R}^{n+m}$, tal que a derivada de f em $a \in U$, $f'(a) : \mathbb{R}^{n+m} \to \mathbb{R}^n$, seja sobrejetiva. Neste caso, diz-se que f é uma *submersão* em a. Suponhamos que f seja de classe C^1 e, através do Teorema da Função Inversa, verifiquemos que existe uma vizinhança V de a em U, tal que $f(V)$ é uma vizinhança de $f(a)$ em $\mathbb{R}^n$.

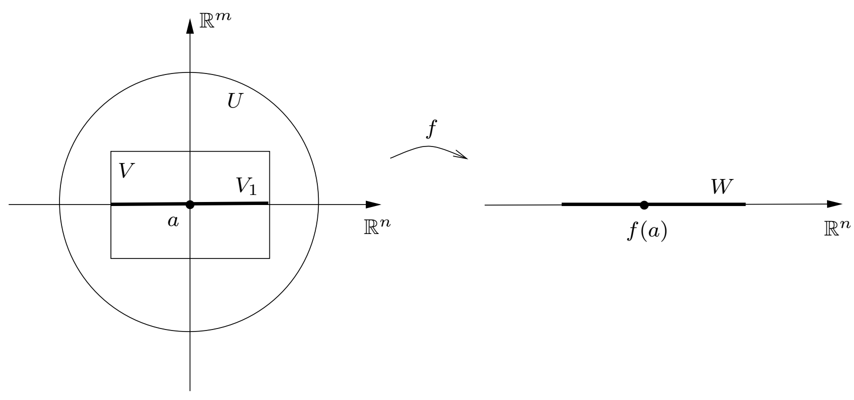

Figura 4

Pelo Teorema do Núcleo e da Imagem, o núcleo de $f'(a)$, $\ker f'(a)$, tem dimensão m. Desta forma, após uma possível mudança de coordenadas, podemos supor que $a = 0$ e $\ker f'(a) = \{0\} \times \mathbb{R}^m \subset \mathbb{R}^n \times \mathbb{R}^m$ (Fig. 4).

Identificando-se, então, o subespaço $\mathbb{R}^n \times \{0\} \subset \mathbb{R}^n \times \mathbb{R}^m$ com $\mathbb{R}^n$, temos que $V_0 = U \cap \mathbb{R}^n$ é um aberto relativo de $\mathbb{R}^n$ que contém a, e a derivada da aplicação $g = f|_{V_0}$ em a, $g'(a) = f'(a)|_{\mathbb{R}^n} : \mathbb{R}^n \to \mathbb{R}^n$, é um isomorfismo. Logo, pelo Teorema da Função Inversa, existem abertos V_1 e W, de $\mathbb{R}^n$, tais que $a \in V_1 \subset U$, $f(a) \in W$ e $g|_{V_1} : V_1 \to W$ é um difeomorfismo de classe C^1. Agora, sendo f

contínua, temos que $V = f^{-1}(W)$ é um aberto de $\mathbb{R}^n$ que contém V_1, donde se infere que $f(V) = W$ e que V, portanto, é a vizinhança desejada.

Dado um aberto U de $\mathbb{R}^n$, uma aplicação diferenciável $f : U \to \mathbb{R}^m$ cuja derivada $f'(x)$ é sobrejetiva para todo $x \in U$ é dita uma *submersão*. Se $\mathbb{V} \subset \mathbb{R}^n$ é um subespaço de $\mathbb{R}^n$, a projeção ortogonal de $\mathbb{R}^n$ sobre $\mathbb{V}$ é um exemplo trivial de submersão[iii]. Note que, pelas considerações acima, toda submersão de classe C^1 é uma aplicação aberta.

4.1. Decomposição em Difeomorfismos Primitivos. Diz-se que um difeomorfismo, definido num aberto U de $\mathbb{R}^n$, é *primitivo* quando é de um dos dois tipos:

i) $x = (x_1, \dots, x_n) \mapsto g(x) = (x_1, \dots, x_{j-1}, \varphi(x), x_{j+1}, \dots, x_n)$;

ii) $x = (x_1, \dots, x_i, \dots, x_j, \dots, x_n) \mapsto Tx = (x_1, \dots, x_j, \dots, x_i, \dots, x_n)$.

Evidentemente, um difeomorfismo $g : U \to V$ é do tipo (i) se, e somente se, para todo $x \in U$, tem-se

$$\langle g(x), e_i \rangle = \langle x, e_i \rangle \ \forall x \in U, \ i \in \{1, \dots, j-1, j+1, \dots, n\}.$$

Deve-se observar também que os difeomorfismos T, do tipo (ii), são restrições a U de isomorfismos lineares, os quais satisfazem $T^{-1} = T$.

No que se segue, estabeleceremos o Teorema de Decomposição Local em Difeomorfismos Primitivos, segundo o qual, localmente e a menos de translações, todo difeomorfismo de classe C^1 em $\mathbb{R}^n$ se exprime como uma composta de $2n$ difeomorfismos primitivos, sendo n do tipo (i) e n do tipo (ii).

Em virtude da simplicidade dos difeomorfismos primitivos, os quais coincidem com a aplicação identidade, exceto por, no máximo, duas funções-coordenada, este teorema de decomposição desempenha um papel importante na demonstração de certos resultados que envolvem mudanças de coordenadas, dentre eles, o teorema fundamental do cálculo integral das funções de várias variáveis, o dito Teorema de Mudança de Coordenadas.

Teorema da Decomposição Local em Difeomorfismos Primitivos. *Seja f um difeomorfismo de classe C^1 definido num aberto U de $\mathbb{R}^n$. Então, para todo $a \in U$, existem difeomorfismos primitivos $g_1, \dots, g_n$, do tipo* (i)*, e $T_1, \dots, T_n$, do tipo* (ii)*, definidos num aberto $V \ni a$, $V \subset U$, tais que*

$$f(x) = f(a) + ((g_n \circ T_n) \circ (g_{n-1} \circ T_{n-1}) \circ \cdots \circ (g_1 \circ T_1))(x - a) \ \ \forall x \in V.$$

Demonstração. Assim como o fizemos na demonstração do Teorema da Função Inversa, translademos os nossos sistemas de coordenadas, de tal modo que tenhamos $a = f(a) = 0$.

Denotemos, então, por ϕ_1 a aplicação f e, para cada $k \in \{1, \dots, n\}$, façamos a seguinte hipótese de indução: Existem uma bola aberta $B_k = B(0, r_k) \subset U$ e um difeomorfismo de classe C^1 definido em B_k, ϕ_k, tal que $\phi_k(0) = 0$. Além disso, quando $2 \leq k \leq n$, tem-se

$$\langle \phi_k(x), e_i \rangle = \langle x, e_i \rangle \ \forall x \in B_k, \ i \in \{1, \dots, k-1\}, \tag{65}$$

[iii]Na Seção 5, veremos que, localmente, toda submersão de classe C^1 se exprime como uma projeção ortogonal (vide Forma Local das Submersões).

isto é, as $k-1$ primeiras funções-coordenada de ϕ_k coincidem com aquelas da aplicação identidade de B_k.

Façamos $T = \phi_k'(0)$ e denotemos por $A = (a_{ij})$ a matriz de T com respeito à base canônica de $\mathbb{R}^n$. Supondo-se $2 \le k \le n$ e diferenciando-se ambos os membros da igualdade (65), obtém-se

$$a_{ij} = \langle e_i, Te_j \rangle = \langle e_i, e_j \rangle \ \forall i \in \{1, \dots, k-1\}.$$

Em particular, as $k-1$ primeiras entradas da k-ésima linha de A são nulas. Desta forma, uma vez A é invertível, para algum $j \in \{k, k+1, \dots, n\}$, tem-se $a_{kj} \neq 0$. Esta última asserção, evidentemente, vale também quando $k = 1$.

Fixemos um tal j e consideremos a projeção ortogonal $P_k = P_{\mathbb{V}_k}$, sobre o subespaço $\mathbb{V}_k = \{x \in \mathbb{R}^n; \langle x, e_k \rangle = 0\}$, e o difeomorfismo primitivo T_k, que faz o intercâmbio das variáveis x_k e x_j, isto é, T_k é a aplicação linear determinada pelas igualdades

$$T_k e_k = e_j\,,\ T_k e_j = e_k\,,\ T_k e_i = e_i \ \ \forall i \neq j, k.$$

Dado $x \in B_k$, façamos

$$g_k(x) = P_k x + \langle \phi_k(T_k x), e_k \rangle e_k$$

e observemos que g_k está bem definida, pois $x \in B_k$ se, e somente se, $T_k x \in B_k$. É imediato que g_k é de classe C^1 e satisfaz $g_k(0) = 0$. Além disso, para todo $h \in \mathbb{R}^n$, tem-se

$$g_k'(0)h = P_k'(0)h + \langle \phi_k'(T_k 0) T_k'(0)h, e_k \rangle e_k = P_k h + \langle TT_k h, e_k \rangle e_k\,.$$

Assim, uma vez que $P_k h$ e e_k são ortogonais, se $g_k'(0)h = 0$, devemos ter $P_k h = 0$ e $\langle TT_k h, e_k \rangle = 0$. Da primeira destas igualdades, segue-se que $h = \mu e_k$, $\mu \in \mathbb{R}$. Então, pela segunda delas,

$$0 = \mu \langle TT_k e_k, e_k \rangle = \mu \langle Te_j, e_k \rangle = \mu a_{kj}\,,$$

donde $\mu = 0$ e, portanto, $h = 0$. Logo, $g_k'(0)$ é um isomorfismo.

Desta forma, pelo Teorema da Função Inversa, existem um aberto de B_k, $U_k \ni 0$, e uma bola aberta $B_{k+1} = B(0, r_{k+1}) \subset U$, tais que $g_k|_{U_k} : U_k \to B_{k+1}$ é um difeomorfismo (Fig. 5).

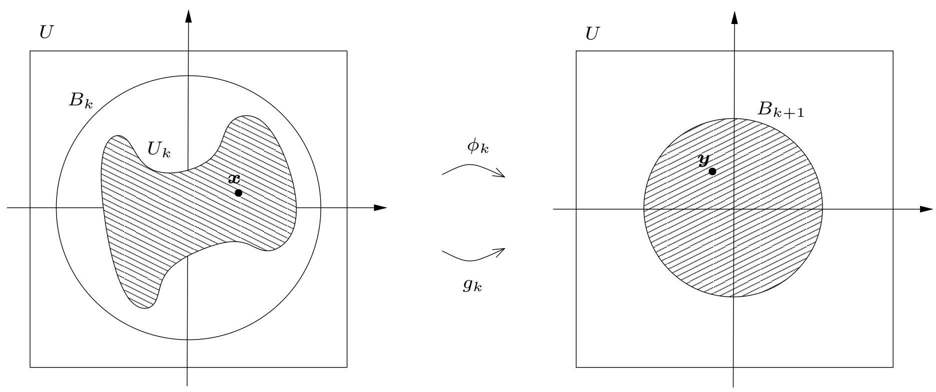

FIGURA 5

Além disso, este difeomorfismo é primitivo do tipo (i), pois, como se pode verificar facilmente, para todo $x \in B_k$,

$$\langle g_k(x), e_i \rangle = \langle x, e_i \rangle \quad \forall i \neq k. \tag{66}$$

Definamos, então,

$$\phi_{k+1}(y) = (\phi_k \circ T_k \circ g_k^{-1})(y), \ y \in B_{k+1}.$$

Sendo uma composta de difeomorfismos, a aplicação $\phi_{k+1} : B_{k+1} \to \phi_{k+1}(B_{k+1})$ é um difeomorfismo. Também, escrevendo-se $x = g_k^{-1}(y) \in U_k$ e lembrando-se que $k \leq j \leq n$, tem-se, para $1 \leq i < k$, $\langle T_k x, e_i \rangle = \langle x, e_i \rangle$. Logo, considerando-se (65) e (66), obtém-se

$$\langle \phi_{k+1}(y), e_i \rangle = \langle \phi_k(T_k x), e_i \rangle = \langle T_k x, e_i \rangle = \langle x, e_i \rangle = \langle g_k(x), e_i \rangle = \langle y, e_i \rangle.$$

Agora, se $k \geq 1$, pelas definições de g_k e ϕ_{k+1}, tem-se, para todo ponto $y = g_k(x)$ em B_{k+1},

$$\langle y, e_k \rangle = \langle g_k(x), e_k \rangle = \langle \phi_k(T_k x), e_k \rangle = \langle \phi_{k+1}(y), e_k \rangle.$$

Segue-se destas considerações que ϕ_{k+1} satisfaz a hipótese de indução, pois $\phi_{k+1}(0) = 0$. Além disso, uma vez que $\phi_{k+1} = \phi_k \circ T_k \circ g_k^{-1}$ e $T_k^{-1} = T_k$, tem-se

$$\phi_k = \phi_{k+1} \circ g_k \circ T_k.$$

Daí, observando-se que ϕ_{n+1} é a aplicação identidade de $B_{n+1} = B(0, r_{n+1})$, fazendo-se $r = \min\{r_1, \dots, r_{n+1}\}$ e $V = B(0, r)$, valem, em V, as igualdades

$$f = \phi_1 = \phi_2 \circ (g_1 \circ T_1) = \phi_3 \circ ((g_2 \circ T_2) \circ (g_1 \circ T_1)) = \cdots = (g_n \circ T_n) \circ \cdots \circ (g_1 \circ T_1),$$

como desejado. $\square$

5. O Teorema da Função Implícita

Nesta seção, através do Teorema da Função Inversa, obteremos o celebrado Teorema da Função Implícita, o qual encontra diversas aplicações em teorias que envolvem equações diferenciais, bem como equações algébricas não lineares, pois garante a existência de funções definidas implicitamente através de uma igualdade.

Mais especificamente, aplicado ao caso particular das funções $f : U \subset \mathbb{R}^2 \to \mathbb{R}$, o Teorema da Função Implícita estabelece condições para que o conjunto de pontos $(x, y) \in U$ que cumprem a igualdade

$$f(x, y) = c, \ c \in \mathbb{R}, \tag{67}$$

seja, numa vizinhança de um ponto $(a, b) \in U$ que satisfaz $f(a, b) = c$, o gráfico de uma função de variável x, $y = y(x)$, ou de variável y, $x = x(y)$. O teorema estabelece que se f é de classe C^1 e $\frac{\partial f}{\partial y}(a, b) \neq 0$, tem-se o primeiro caso, e se $\frac{\partial f}{\partial x}(a, b) \neq 0$, tem-se o segundo.

O conjunto dos pontos (x, y) que satisfazem (67) definem, em muitos casos, o traço de uma curva, razão pela qual é denominado *curva de nível de* f. Assim, decorre do Teorema da Função Implícita que uma curva de nível de uma função de classe C^1 em cujos pontos pelo menos uma das derivadas parciais é diferente de zero é, localmente, o gráfico de uma função de uma única variável.

Considerando-se, por exemplo, a função de classe C^1,

$$f(x, y) = x^2 + y^2, \ (x, y) \in \mathbb{R}^2,$$

tem-se que o conjunto dos pontos $(x,y) \in \mathbb{R}^2$ que cumpre a igualdade $f(x,y) = 1$ é a esfera S^1. Geometricamente, é fácil ver que, para um dado ponto $(a,b) \in S^1$ distinto de $(1,0)$ e $(-1,0)$, existe uma vizinhança do mesmo em S^1 que é o gráfico de uma função $y = y(x)$. Observe-se que $\frac{\partial f}{\partial y}(a,b) = 2b \neq 0$, enquanto $\frac{\partial f}{\partial y}(1,0) = \frac{\partial f}{\partial y}(-1,0) = 0$ (Fig. 6).

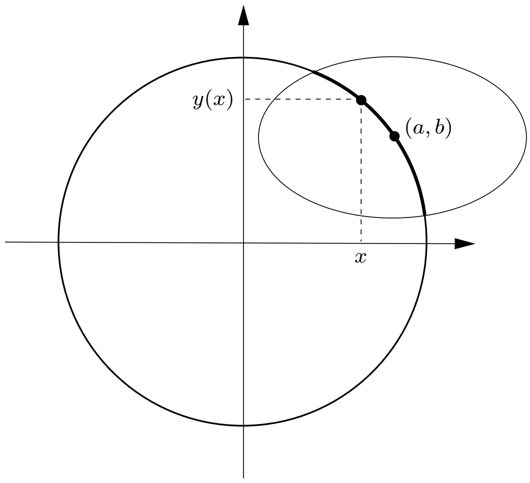

FIGURA 6

TEOREMA DA FUNÇÃO IMPLÍCITA. *Sejam* $U \subset \mathbb{R}^{n+m} = \mathbb{R}^n \times \mathbb{R}^m$ *um conjunto aberto,* $f : U \to \mathbb{R}^m$ *de classe* C^1 *e* $(a,b) \in U$, *tais que* $f(a,b) = c \in \mathbb{R}^m$. *Nestas condições, se* $f'(a,b)|_{\{0\}\times\mathbb{R}^m}$ *é um isomorfismo sobre* $\mathbb{R}^m$, *então existem abertos* $V \ni (a,b)$, $V \subset U$, *e* $A \ni a$, $A \subset \mathbb{R}^n$, *que cumprem as seguintes condições:*

i) *Para todo* $x \in A$, *existe um único* $y = \xi(x) \in \mathbb{R}^m$, *tal que*

$$(x, \xi(x)) \in V \quad e \quad f(x, \xi(x)) = c\,;$$

ii) *a aplicação*

$$\begin{array}{rccc} \xi : & A \subset \mathbb{R}^n & \to & \mathbb{R}^m \\ & x & \to & \xi(x) \end{array}$$

é de classe C^1 *e*

$$\xi'(a) = -T_2^{-1}\, T_1\,,$$

em que $T_1 \in \mathrm{L}(\mathbb{R}^n, \mathbb{R}^m)$ *e* $T_2 \in \mathrm{L}(\mathbb{R}^m, \mathbb{R}^m)$ *são definidas por*

$$T_1(h) = f'(a,b)(h,0) \quad e \quad T_2(k) = f'(a,b)(0,k).$$

Note que, no enunciado do Teorema da Função Implícita, a condição de $f'(a,b)|_{\{0\}\times\mathbb{R}^m}$ ser um isomorfismo equivale à de que seja invertível a matriz $m \times m$ cujos vetores-coluna são os m últimos vetores-coluna da matriz jacobiana de f em (a,b). Cumprida esta condição e sendo f de classe C^1, o teorema assegura a existência de uma vizinhança de (a,b) cuja interseção com $f^{-1}(\{f(a,b)\})$ é o gráfico de uma aplicação ξ, de classe C^1, definida num aberto de $\mathbb{R}^n$ e que toma valores em $\mathbb{R}^m$.

Demonstração do Teorema da Função Implícita. Suponhamos, sem perda de generalidade, que $c=0$ e consideremos a aplicação

$$\begin{array}{rccc} F: & U & \rightarrow & \mathbb{R}^n\times\mathbb{R}^m \\ & (x,y) & \rightarrow & (x,f(x,y)), \end{array}$$

isto é, F tem como coordenadas a projeção $P_1(x,y)=x$ e a aplicação f, donde F é de classe C^1 e

$$F'(a,b)(h,k)=(h,f'(a,b)(h,k)),\ \ (h,k)\in\mathbb{R}^n\times\mathbb{R}^m.$$

Em particular, se $F'(a,b)(h,k)=(0,0)$, tem-se $h=0$ e $0=f'(a,b)(h,k)=f'(a,b)(0,k)$, donde $k=0$, pois, por hipótese, $f'(a,b)|_{\{0\}\times\mathbb{R}^m}$ é um isomorfismo. Logo, $F'(a,b)$ é um isomorfismo de $\mathbb{R}^n\times\mathbb{R}^m$ sobre si mesmo.

Temos, então, pelo Teorema da Função Inversa, que existem abertos $V\ni(a,b)$, $V\subset U$, e $W\ni(a,0)$, $W\subset\mathbb{R}^n\times\mathbb{R}^m$, tais que $F|_V:V\to W$ é um difeomorfismo de classe C^1 (Fig. 7).

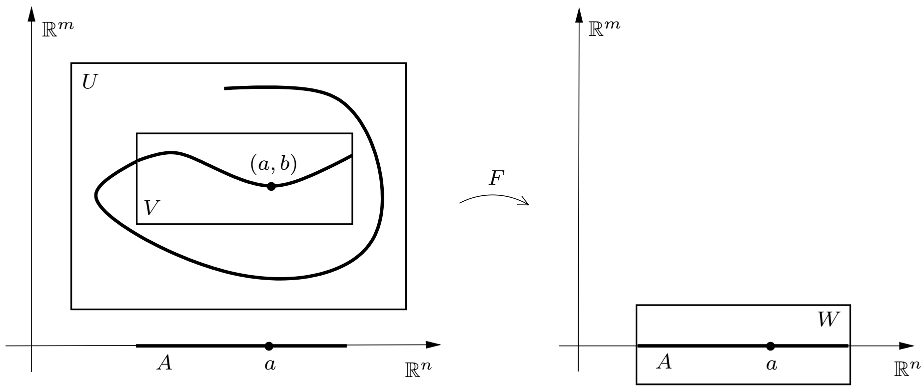

Figura 7

Podemos supor, sem perda de generalidade, que $W=A\times B$, em que A e B são abertos de $\mathbb{R}^n$ e $\mathbb{R}^m$, respectivamente, e $a\in A$. Sendo assim, dado $x\in A$, existe um único $y=\xi(x)\in\mathbb{R}^m$, tal que $(x,y)\in V$ e $F(x,y)=(x,f(x,y))=(x,0)\in W$, pois $F|_V:V\to W$ é bijetiva. Logo, a aplicação ξ, assim definida, satisfaz

$$f(x,\xi(x))=0\ \ \forall x\in A,$$

o que demonstra (i).

Seja $G=(F|_V)^{-1}$. Então, G é de classe C^1 (ainda pelo Teorema da Função Inversa) e $G(x,0)=(x,\xi(x))$. Logo, designando-se por P_2 a projeção $(x,y)\mapsto y$, $(x,y)\in\mathbb{R}^n\times\mathbb{R}^m$, e por ϕ a aplicação $x\mapsto(x,0)$, $x\in\mathbb{R}^n$, tem-se $\xi=P_2\circ G\circ\phi$, donde ξ é de classe C^1.

Finalmente, observemos que, para todo $(h,k)\in\mathbb{R}^n\times\mathbb{R}^m$, tem-se

$$f'(a,b)(h,k)=f'(a,b)((h,0)+(0,k))=T_1h+T_2k.$$

Logo, diferenciando-se ambos os membros de $f(x,\xi(x))=0$ e fazendo-se $x=a$, obtém-se

$$0=f'(a,\xi(a))(h,\xi'(a)h)=f'(a,b)(h,\xi'(a)h)=T_1h+T_2\xi'(a)h\ \ \forall h\in\mathbb{R}^n,$$

donde $\xi'(a)=-T_2^{-1}T_1$. Isto prova (ii) e conclui, desta forma, a demonstração. □

Observação 18. Assim como no Teorema da Função Inversa, e como consequência disto, no enunciado do Teorema da Função Implícita pode-se substituir "C^1" por "C^k, $k \in \mathbb{N}$," isto é, a função ξ obtida é da mesma classe de diferenciabilidade da função dada, f.

Dada uma aplicação $f : U \subset \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^m$, motivados pela terminologia adotada no caso em que $n = m = 1$, para cada $c \in \mathbb{R}^m$, chamaremos o conjunto dos pontos $(x, y) \in U$ que satisfazem $f(x, y) = c$ de *conjunto de nível* de f. Segue-se, portanto, do Teorema da Função Implícita, que se f é uma submersão de classe C^1, então todo conjunto de nível não-vazio de f é, localmente, o gráfico de uma aplicação de classe C^1 definida num aberto de $\mathbb{R}^n$ e tomando valores em $\mathbb{R}^m$.

Reconsideremos a demonstração do Teorema da Função Implícita e dela o difeomorfismo $G : W \to V$, em que $G = (F|_V)^{-1}$ e $F(x, y) = (x, f(x, y))$. Dado $(x, z) \in W$, tem-se $G(x, z) = (x, y)$ se, e somente se, $z = f(x, y)$. Logo,

$$(f \circ G)(x, z) = f(G(x, z)) = f(x, y) = z,$$

isto é, $f \circ G = P_2$. Assim, no aberto $W = A \times B$, cada $z \in B$ determina uma "fibra" $A \times \{z\} \subset W$, cuja imagem pelo difeomorfismo G é o conjunto de nível $f(x, y) = z$ (Fig. 8).

Em suma, vale o resultado seguinte.

Forma Local das Submersões. *Sejam $(a, b) \in U \subset \mathbb{R}^n \times \mathbb{R}^m$, U aberto, e $f : U \to \mathbb{R}^m$ uma aplicação de classe C^1, tais que $f'(a, b)|_{\{0\} \times \mathbb{R}^m}$ é um isomorfismo sobre $\mathbb{R}^m$. Nestas condições, existem abertos $V \subset U$, $A \subset \mathbb{R}^n$, $B \subset \mathbb{R}^m$, e um difeomorfismo $G : A \times B \to V$, tais que*

$$(G \circ f)(x, z) = z \quad \forall (x, z) \in A \times B.$$

Exemplo 78. Consideremos a função $f : \mathbb{R}^3 = \mathbb{R}^2 \times \mathbb{R} \to \mathbb{R}$, definida por

$$f(x, y, z) = x^3 + y^3 + z^3 - 2z(x + y) - 2x + y - 2z + 1.$$

Temos que $f(0, 0, 1) = 0$ e $\frac{\partial f}{\partial z}(0, 0, 1) = 1 \neq 0$. Logo, pelo Teorema da Função Implícita, existem uma vizinhança V de $(0, 0, 1)$ em $\mathbb{R}^3$, uma vizinhança A de $(0, 0)$ em $\mathbb{R}^2$ e uma função $\xi : A \to \mathbb{R}$, tais que a parte do conjunto de nível $f(x, y, z) = 0$ que intersecta V é o gráfico de ξ, isto é, para todo $(x, y) \in A$, tem-se $(x, y, \xi(x, y)) \in V$ e $f(x, y, \xi(x, y)) = 0$. Diferenciando-se esta última igualdade, obtém-se

$$f'(x, y, \xi(x, y))(h, k, \xi'(x, y)(h, k)) = 0 \quad \forall (x, y) \in A, \ (h, k) \in \mathbb{R}^2.$$

Sendo assim, tem-se

$$\langle \nabla f(x, y, \xi(x, y)), (h, k, \xi'(x, y)(h, k)) \rangle = \frac{\partial f}{\partial x} h_1 + \frac{\partial f}{\partial y} h_2 + \frac{\partial f}{\partial z} \xi'(x, y)(h_1, h_2) = 0,$$

em que as derivadas indicadas são todas calculadas no ponto $(x, y, \xi(x, y))$. Aplicando-se esta igualdade aos vetores da base canônica de $\mathbb{R}^2$, obtém-se:

i) $\frac{\partial f}{\partial x}(x, y, \xi(x, y)) + \frac{\partial f}{\partial z}(x, y, \xi(x, y)) \frac{\partial \xi}{\partial x}(x, y) = 0;$

ii) $\frac{\partial f}{\partial y}(x, y, \xi(x, y)) + \frac{\partial f}{\partial z}(x, y, \xi(x, y)) \frac{\partial \xi}{\partial y}(x, y) = 0.$

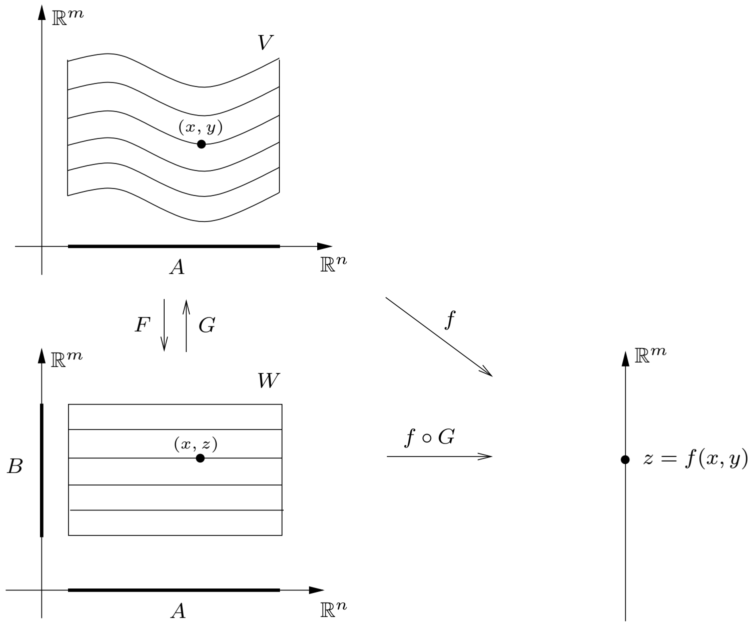

FIGURA 8

Agora, uma vez que f é de classe C^∞, tem-se, em particular, que $\partial f/\partial z$ é contínua, donde podemos supor que $\partial f/\partial z$ não se anula V. Logo, escrevendo-se $f_x = \partial f/\partial x$ e as demais derivadas parciais concomitantemente, segue-se das igualdades (i) e (ii) que

$$\xi_x(x,y) = -\frac{f_x(x,y,\xi(x,y))}{f_z(x,y,\xi(x,y))} \quad \text{e} \quad \xi_y(x,y) = -\frac{f_y(x,y,\xi(x,y))}{f_z(x,y,\xi(x,y))}\,.$$

Em particular, $\xi_x(0,0) = 4$ e $\xi_y(0,0) = 1$.

5.1. Autovalores e Autovetores de Matrizes ao Longo de Curvas. Dada uma curva $t \mapsto A(t)$ em $\mathrm{M}(n)$, pode-se indagar como variam os autovalores e autovetores de $A(t)$ com respeito à variável t. No que se segue, através de uma ideia de Jerry Kazdan (vide [**15**]), aplicaremos o Teorema da Função Implícita para dar uma resposta parcial a esta questão.

Para tanto, nos será conveniente identificar, através do isomorfismo canônico, os espaços $\mathbb{R}^n$ e $\mathrm{M}(n,1)$. Assim, dado um vetor $v \in \mathbb{R}^n$ de coordenadas $x_1, \dots, x_n$ com respeito à base canônica de $\mathbb{R}^n$, escreveremos

$$v = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \quad \text{e} \quad v^* = [x_1 \cdots x_n]\,.$$

Note que, neste caso, dados $v, w \in \mathbb{R}^n$, tem-se $\langle v, w \rangle = w^* v$.

Diz-se que $\lambda \in \mathbb{R}$ é um autovalor *simples* de uma matriz $A \in \mathrm{M}(n)$ se satisfaz a seguinte condição: Dado um autovetor unitário de A associado a λ, u, tem-se

$$v \in \mathbb{R}^n \quad \text{e} \quad (A - \lambda I)^2 v = 0 \quad \Rightarrow \quad v = \mu u,\ \mu \in \mathbb{R},$$

em que I é a matriz identidade de $\mathrm{M}(n)$. Esta condição, convém mencionar, equivale à de λ ser uma raiz simples, isto é, de multiplicidade 1, do polinômio característico de A.

Teorema 18. *Seja $A = A(t), t \in (-\epsilon, \epsilon)$, uma curva de classe C^1 em $\mathrm{M}(n)$, tal que $A(0)$ possui um autovetor unitário u_0 associado a um autovalor simples λ_0. Nestas condições, existem $\delta > 0$, uma função $\lambda : (-\delta, \delta) \to \mathbb{R}$ e uma curva $v : (-\delta, \delta) \to \mathbb{R}^n$, ambas de classe C^1, tais que $\lambda(0) = 0$, $u(0) = u_0$ e, para todo $t \in (-\delta, \delta)$, $\lambda(t)$ é um autovalor simples de $A(t)$ que tem $v(t)$ como autovetor associado, isto é,*

$$A(t)v(t) = \lambda(t)v(t) \quad \forall t \in (-\delta, \delta).$$

Demonstração. Consideremos a aplicação $F : (-\epsilon, \epsilon) \times \mathbb{R} \times \mathbb{R}^n \to \mathbb{R} \times \mathbb{R}^n$, definida por

$$F(t, \lambda, v) = (u_0^* v - 1, (A(t) - \lambda I)v),$$

e observemos que a matriz jacobiana de F em $x = (t, \lambda, v)$, $J_F(x)$, é a matriz $(n+1) \times (n+2)$ que, em blocos, se escreve como

$$\begin{bmatrix} 0 & 0 & u_0^* \\ A'(t)v & -v & A(t) - \lambda I \end{bmatrix}.$$

Assim, identificando-se $\mathbb{R} \times \mathbb{R}^n$ com $\{0\} \times (\mathbb{R} \times \mathbb{R}^n)$, a restrição de $F'(x)$ a este último é o operador linear de $\mathbb{R} \times \mathbb{R}^n = \mathbb{R}^{n+1}$, cuja matriz (com respeito à base canônica de $\mathbb{R}^{n+1}$) é o bloco $(n+1) \times (n+1)$ formado pelas dois últimos blocos-coluna de $J_F(x)$. Em particular, para $x = (0, \lambda_0, u_0)$, esta matriz é

$$M = \begin{bmatrix} 0 & u_0^* \\ -u_0 & A(0) - \lambda_0 I \end{bmatrix}.$$

A fim de aplicar o Teorema da Função Implícita, verifiquemos que o operador linear $T \in \mathrm{L}(\mathbb{R}^{n+1})$ associado a M é um isomorfismo. Para tanto, tomemos $(s, w) \in \mathbb{R} \times \mathbb{R}^n$, tal que $T(s, w) = 0$. Neste caso, tem-se

$$u_0^* w = 0 \quad \text{e} \quad -su_0 + (A(0) - \lambda_0 I)w = 0.$$

A segunda destas igualdades nos dá $(A(0) - \lambda_0 I)w = su_0$, donde

$$(A(0) - \lambda_0 I)^2 w = s(A(0) - \lambda_0 I)u_0 = 0$$

e, portanto, para algum real μ, $w = \mu u_0$, pois λ_0 é um autovalor simples de $A(0)$. Porém, $0 = u_0^* w = \mu u_0^* u_0 = \mu$ e, então, $w = s = 0$, donde se infere que T é um isomorfismo. Logo, pelo Teorema da Função Implícita, existe $\delta > 0$, tal que a função $\lambda = \lambda(t) \in \mathbb{R}$ e a curva $v = v(t) \in \mathbb{R}^n$ estão bem definidas em $(-\delta, \delta)$, são de classe C^1 e satisfazem

$$F(t, \lambda(t), v(t)) = F(0, \lambda_0, u_0) = 0 \ \forall t \in (-\delta, \delta).$$

Desta forma, para todo $t \in (-\delta, \delta)$, tem-se

$$\langle v(t), u_0 \rangle = 1 \quad \text{e} \quad A(t)v(t) = \lambda(t)v(t).$$

Por fim, verifiquemos que, tomando-se $\delta > 0$ suficientemente pequeno, $\lambda(t)$ é um autovalor simples de $A(t)$. Com efeito, se não fosse esse o caso, existiriam sequências (t_k) e (u_k) em $\mathbb{R}$ e S^{n-1}, respectivamente, tais que $t_k \to 0$ $u_k \to u \in S^{n-1}$ e, para cada $k \in \mathbb{N}$, $(A(t_k) - \lambda(t_k))^2 u_k = 0$, sendo u_k e $v_k = v(t_k)$ linearmente independentes. Assim, o núcleo de $(A(t_k) - \lambda(t_k))^2$ teria dimensão maior que 1, donde, em particular, poderíamos tomar v_k e u_k ortogonais. Por continuidade, teríamos, então, $(A(0) - \lambda(0))^2 u = 0$ e u e u_0 ortogonais (pois $v_k \to u_0$), contradizendo o fato de u_0 ser um autovalor simples de $A(0)$. $\square$

Vejamos agora que, no teorema acima, a hipótese de λ_0 ser um autovalor simples de $A(0)$ é, de fato, necessária.

Consideremos em $\mathrm{M}(2)$, por exemplo, a curva

$$A(t) = \begin{bmatrix} 0 & 1 \\ t^2 & 0 \end{bmatrix}, \ t \in \mathbb{R}.$$

Temos que A é de classe C^∞. No entanto, para todo $t \in \mathbb{R}$, as raízes do polinômio característico de $A(t)$ são $\lambda_1(t) = -|t|$ e $\lambda_2(t) = |t|$, isto é, nenhuma das funções λ_1, λ_2 é diferenciável em vizinhança alguma de $t = 0$. Note que $\lambda_0 = 0$ é um autovalor de $A(0)$ que não é simples, pois $A(0)^2 = 0$.

Mesmo quando a curva $A = A(t)$ e a função $\lambda = \lambda(t)$ são diferenciáveis, pode ocorrer da curva de autovetores $u = u(t)$ não ter esta propriedade. A fim de ilustrar este fenômeno, consideremos as funções $\mu, \theta : \mathbb{R} \to \mathbb{R}$, dadas por

$$\mu(t) = \begin{cases} e^{-1/t^2} & \text{se} \quad t \neq 0 \\ \\ 0 & \text{se} \quad t = 0 \end{cases} \qquad \text{e} \qquad \theta(t) = \begin{cases} \frac{1}{t} & \text{se} \quad t \neq 0 \\ \\ 0 & \text{se} \quad t = 0, \end{cases}$$

e definamos a curva

$$A(t) = \mu(t) \begin{bmatrix} \cos 2\theta(t) & \operatorname{sen} 2\theta(t) \\ \\ \operatorname{sen} 2\theta(t) & -\cos 2\theta(t) \end{bmatrix}, \ t \in \mathbb{R} - \{0\}, \ A(0) = 0.$$

Por um cálculo direto, verifica-se que a curva A é de classe C^∞ e, para todo $t \in \mathbb{R}$, $\lambda_1(t) = -\mu(t)$ e $\lambda_2 = \mu(t)$ são os autovalores de $A(t)$ que, quando $t \neq 0$, têm

$$u_1(t) = \begin{bmatrix} \cos \theta(t) \\ \\ \operatorname{sen} \theta(t) \end{bmatrix} \quad \text{e} \quad u_2(t) = \begin{bmatrix} -\operatorname{sen} \theta(t) \\ \\ \cos \theta(t) \end{bmatrix}$$

como respectivos autovetores associados. Uma vez que μ é, sabidamente, de classe C^∞, assim o são λ_1 e λ_2. Porém, os limites de u_1 e u_2 quando t tende a zero não existem, donde se conclui que estas curvas não são, sequer, contínuas em $t = 0$.

5.2. Multiplicadores de Lagrange. O método dos multiplicadores de Lagrange, devido a Joseph-Louis Lagrange (1736–1813), se aplica ao problema de se determinar extremos locais da restrição de uma função diferenciável f, definida num aberto $U \subset \mathbb{R}^n \times \mathbb{R}^m$, a um subconjunto $M \subset U$, que, por sua vez, é um conjunto de nível de uma submersão de classe C^1, $g : U \to \mathbb{R}^m$. Nestas condições, conforme estabeleceremos no teorema seguinte, para que um ponto $p \in M$ seja um extremo local de $f|_M$, é necessário que o gradiente de f em p esteja no espaço gerado pelos gradientes das coordenadas de g, $g_1, \dots, g_m$, em p, isto é, devem existir $\lambda_1, \dots, \lambda_m \in \mathbb{R}$, tais que

$$\nabla f(p) = \lambda_1 \nabla g_1(p) + \cdots + \lambda_m \nabla g_m(p).$$

Neste contexto, os reais $\lambda_1, \dots, \lambda_m$ são chamados de *multiplicadores de Lagrange*.

Desta forma, se M é definida pela igualdade $g(x) = c$, as $n+m$ coordenadas de p, juntamente com os m multiplicadores $\lambda_1, \dots, \lambda_m$, são as variáveis do sistema de equações (vetoriais)

$$\begin{cases} \nabla f(p) & = & \lambda_1 \nabla g_1(p) + \cdots + \lambda_m \nabla g_m(p) \\ \\ g(p) & = & c. \end{cases}$$

Admitindo-se, por hora, este resultado, apliquemo-lo ao seguinte problema: Dados $u \in S^{n-1}$, $n > 1$, e um real $a > 1$, determinar pontos $x_0 \in S^{n-1}$ e $y_0 \in \Pi = \{y \in \mathbb{R}^n; \langle y, u\rangle = a\}$, tais que a distância entre x_0 e y_0 seja a menor possível, isto é,

$$\|x_0 - y_0\| = \inf\{\|x - y\|\,;\, x \in S^{n-1},\ y \in \Pi\}. \tag{68}$$

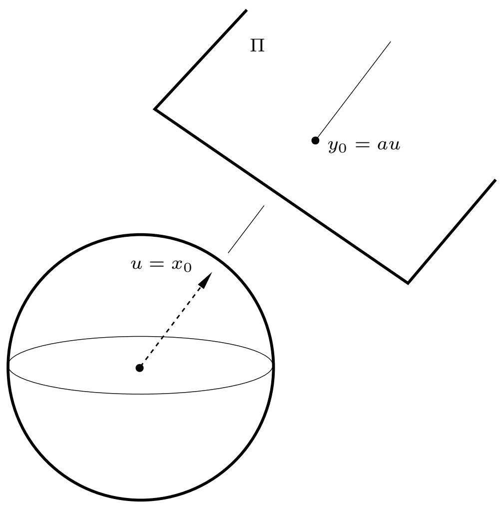

FIGURA 9

Observe que, fazendo-se $\mathbb{V} = \{u\}^{\perp}$, tem-se que Π é o *espaço afim* de $\mathbb{R}^n$ paralelo ao subespaço $\mathbb{V}$ e que contém au, isto é,

$$\Pi = \mathbb{V} + au = \{v + au;\ v \in \mathbb{V}\}.$$

Em particular, dado $w = v + au \in \Pi$, tem-se, pelo Teorema de Pitágoras, que $\|w\|^2 = \|v\|^2 + a^2 \geq a^2 > 1$, donde $\Pi \cap S^{n-1} = \emptyset$.

Façamos $U = \{(x, y) \in \mathbb{R}^n \times \mathbb{R}^n; x \neq 0\}$, $M = S^{n-1} \times \Pi$, e consideremos a função $f : U \to \mathbb{R}$, bem como a aplicação $g : U \to \mathbb{R}^2$, definidas por

$$f(x, y) = \|x - y\|^2 \quad \text{e} \quad g(x, y) = (g_1(x, y), g_2(x, y)) = (\|x\|^2 - 1, \langle y, u\rangle - a).$$

Claramente, f e g são de classe C^1 e $M = g^{-1}(\{0\}) \subset U$. Além disso,

$$\nabla g_1(x, y) = (2x, 0) \qquad \text{e} \qquad \nabla g_2(x, y) = (0, u) \quad \forall (x, y) \in U. \tag{69}$$

Em particular, $\nabla g_1(x, y)$ e $\nabla g_2(x, y)$ são linearmente independentes, donde se conclui que o jacobiano de g em (x, y) tem posto 2 e, portanto, que g é uma submersão.

Nestas condições, podemos aplicar o método dos coeficientes de Lagrange, segundo o qual, a solução do nosso problema será o ponto $p = (x_0, y_0)$ que satisfaz

$$\begin{cases} \nabla f(p) &= \lambda_1 \nabla g_1(p) + \lambda_2 \nabla g_2(p) \\ g(p) &= 0. \end{cases}$$

Uma vez que, para todo par $(x, y) \in U$, tem-se $\nabla f(x, y) = 2(x - y, y - x)$ e $\nabla g_1(x, y)$, $\nabla g_2(x, y)$ são como em (69), este sistema equivale a

$$\begin{cases} x_0 - y_0 &= \lambda_1 x_0 \\ y_0 - x_0 &= \frac{\lambda_2}{2} u \\ \|x_0\| &= 1 \\ \langle y_0, u\rangle &= a. \end{cases}$$

Se tivéssemos $\lambda_1 = 0$, teríamos, pela primeira igualdade, $x_0 = y_0$, o que contradiria o fato de S^{n-1} ter interseção vazia com Π. Logo, devemos ter $\lambda_1 \neq 0$. Neste caso, fazendo-se $\mu = -\frac{\lambda_2}{2\lambda_1}$ e adicionando-se as duas primeiras igualdades, obtém-se $x_0 = \mu u$ e $y_0 = (1 - \lambda_1)\mu u$. Daí e das duas últimas igualdades do sistema, segue-se que $\mu = \pm 1$ e $(1 - \lambda_1)\mu = a$, donde $x_0 = \pm u$ e $y_0 = au$. Observando-se, então, que $\|u - y_0\| = a - 1 < a + 1 = \| - u - y_0\|$ e considerando-se (68), infere-se que $x_0 = u$.

Por fim, dados $x \in S^{n-1}$ e $y \in \Pi$, sejam $v, v' \in \mathbb{V}$ e $t \in \mathbb{R}$, tais que $x = v + tu$ e $y = v' + au$, donde $x - y = (v - v') + (t - a)u$. Uma vez que $|t| \leq \|x\| = 1$, tem-se $\|x - y\| \geq |t - a| \geq |a| - |t| \geq a - 1 = \|u - au\| = \|x_0 - y_0\|$, donde se conclui que $x_0 = u$ e $y_0 = au$ são os pontos desejados (Fig. 9).

TEOREMA 19 (MULTIPLICADORES DE LAGRANGE). *Sejam $U \subset \mathbb{R}^n \times \mathbb{R}^m$ um conjunto aberto, $f : U \to \mathbb{R}$ uma função diferenciável e $M \subset U$ um conjunto de nível de uma submersão de classe C^1, $g = (g_1, \ldots, g_m) : U \to \mathbb{R}^m$. Então, se $p \in M$ é um extremo local de $f|_M$, existem $\lambda_1, \ldots, \lambda_m \in \mathbb{R}$, tais que*

$$\nabla f(p) = \lambda_1 \nabla g_1(p) + \cdots + \lambda_m \nabla g_m(p).$$

DEMONSTRAÇÃO. Seja $p = (x_0, y_0) \in M$ um extremo local de $f|_M$. Uma vez que g é uma submersão, temos que $g'(p) \in \mathrm{L}(\mathbb{R}^{n+m}, \mathbb{R}^m)$ é sobrejetiva, donde podemos supor, sem perda de generalidade, que $g'(p)|_{\{0\}\times\mathbb{R}^m}$ é um isomorfismo sobre $\mathbb{R}^m$. Logo, pelo Teorema da Função Implícita, existem um aberto de $\mathbb{R}^n$, $A \ni x_0$, e uma aplicação de classe C^1, $\xi : A \to \mathbb{R}^m$, tais que o gráfico de ξ está contido em M, isto é, $(x, \xi(x)) \in M \; \forall x \in A$.

Consideremos o gráfico de $\xi'(x_0)$,

$$\mathbb{V} = \{(h, \xi'(x_0)h); h \in \mathbb{R}^n\} \subset \mathbb{R}^n \times \mathbb{R}^m,$$

e observemos que este é um subespaço vetorial de $\mathbb{R}^n \times \mathbb{R}^m$ de dimensão n (de fato, $\{(e_1, \xi'(x_0)e_1), \ldots, (e_n, \xi'(x_0)e_n)\}$ é uma base de $\mathbb{V}$). Sendo M um conjunto de nível de g, temos que $g(x, \xi(x))$ é constante. Em particular, dado $h \in \mathbb{R}^n$, tem-se $g'((x_0, \xi(x_0))(h, \xi'(x_0)h) = 0$, isto é, $g'(p)(\mathbb{V}) = 0$. Logo, para todo vetor $v = (h, \xi'(x_0)h) \in \mathbb{V}$, tem-se

$$0 = g'(p)v = (g_1'(p)v, \ldots, g_m'(p)v) = (\langle \nabla g_1(p), v\rangle, \ldots, \langle \nabla g_m(p), v\rangle),$$

donde se infere que, para cada $j \in \{1, \ldots, m\}$, $\nabla g_j(p)$ é ortogonal a $\mathbb{V}$. Porém, os vetores-linha da matriz jacobiana de g em p são, justamente, $\nabla g_1(p), \ldots, \nabla g_m(p)$ e, portanto, são linearmente independentes, já que $g'(p)$ tem posto m. Segue-se, então, destas considerações, que o complemento ortogonal de $\mathbb{V}$ em $\mathbb{R}^n \times \mathbb{R}^m$ é o subespaço gerado pelos vetores $\nabla g_1(p), \dots, \nabla g_m(p)$.

Agora, uma vez que p é um extremo local de $f|_M$, temos que x_0 é um extremo local de $f(x, \xi(x))$, $x \in A$. Dado, então, $v = (h, \xi'(x_0)h) \in \mathbb{V}$, tem-se

$$0 = f'(x_0, \xi(x_0))(h, \xi'(x_0)h) = \langle \nabla f(p), v \rangle,$$

donde $\nabla f(p) \in \mathbb{V}^\perp$. Logo, existem reais $\lambda_1, \dots, \lambda_m$, tais que

$$\nabla f(p) = \lambda_1 \nabla g_1(p) + \dots + \lambda_m \nabla g_m(p),$$

como desejávamos provar. □

Exemplo 79 (desigualdade de hadamard). Dada uma matriz $A \in \mathrm{M}(n)$, sejam $v_1, \dots, v_n \in \mathbb{R}^n$ seus vetores-coluna. Verifiquemos, através do método dos multiplicadores de Lagrange, a *desigualdade de Hadamard*,

$$|\det A| = |\det(v_1, \dots, v_n)| \leq \|v_1\| \dots \|v_n\|,$$

em que a igualdade ocorre se, e somente se, os vetores $v_1, \dots, v_n$ são ortogonais entre si.

Para tanto, consideremos a esfera $S = \{A \in \mathrm{M}(n)\,;\, \|A\|_e^2 = n\}$, em que $\|A\|_e = \sqrt{\text{traço}(AA^*)}$ é a norma euclidiana de A, bem como as funções $f, g : \mathrm{M}(n) - \{0\} \to \mathbb{R}$, dadas por

$$f(A) = \det A \quad \text{e} \quad g(A) = \|A\|_e^2 .$$

Claramente, a aplicação g é de classe C^1 e, para toda matriz $A \in M(n) - \{0\}$, $\nabla g(A) = 2A \neq 0$, donde g é uma submersão. Logo, uma vez que $S = g^{-1}(\{n\})$, pelo método dos multiplicadores de Lagrange, para todo extremo A_0 de $f|_S$, valem as igualdades:

i) $\nabla f(A_0) = \lambda \nabla g(A_0) = 2\lambda A_0$, $\lambda \in \mathbb{R}$;

ii) $\|A_0\|_e^2 = n$.

Lembrando que $A_0(\nabla f(A_0))^* = (\det A_0)I$ (vide Exemplo 64 – Capítulo 4), segue-se de (i) que $2\lambda A_0 A_0^* = (\det A_0)I$, donde $2\lambda\, \text{traço}(A_0 A_0^*) = n \det A_0$. Daí e de (ii), obtém-se $2\lambda = \det A_0$ e, portanto, $A_0 A_0^* = I$, isto é, A_0 é ortogonal[(iv)]. Assim, o valor máximo de $f|_S$ é 1, enquanto o mínimo é -1, isto é, $f(S) = [-1, 1]$ (pois f é contínua e S é compacto e conexo) e $|\, f|_S(A_0)\,| = 1$ se, e somente se, A_0 é ortogonal.

Agora, dada uma matriz $A \in \mathrm{M}(n)$, se um de seus vetores-coluna $v_1, \dots, v_n$ for nulo, a desigualdade de Hadamard será trivialmente satisfeita por A. Suponhamos, então, que cada um dos vetores v_j seja não-nulo. Neste caso, a matriz cujos vetores-coluna são $v_1/\|v_1\|, \dots, v_n/\|v_n\|$ pertence a S. Logo,

$$|\det A| = |\det(v_1, \dots, v_n)| = \|v_1\| \dots \|v_n\| \left| \det\left(\frac{v_1}{\|v_1\|}, \dots, \frac{v_n}{\|v_n\|} \right) \right| \leq \|v_1\| \dots \|v_n\|,$$

em que a igualdade ocorre se, e somente se, $v_1, \dots, v_n$ são ortogonais.

[(iv)] Diz-se que uma matriz quadrada A é *ortogonal* quando $AA^* = I$, o que equivale aos vetores-coluna de A formarem uma base ortonormal de $\mathbb{R}^n$. Note-se que, da igualdade $\det A = \det A^*$, segue-se que o determinante de uma matriz ortogonal é 1 ou -1.

6. O Teorema de Sard $(*)$

Em Cálculo, quando do estudo da integral de funções de duas variáveis reais, verifica-se que toda função contínua definida num conjunto limitado $\Omega \subset \mathbb{R}^2$ é integrável, isto é, continuidade é uma condição suficiente para integrabilidade. No entanto, ela não é necessária. Verifica-se, por exemplo, que se Ω é uma bola fechada de $\mathbb{R}^2$ e a função em questão é descontínua apenas ao longo de um segmento de Ω, então ela é integrável. Considerando-se as noções de área e volume associadas ao conceito de integral, constata-se facilmente que esta propriedade decorre do fato de um segmento de reta do plano ter área nula. Considerando-se, então, a área como uma maneira de medir conjuntos em $\mathbb{R}^2$, pode-se dizer que, desse ponto de vista, um segmento do plano tem *medida nula*.

Mesmo destituídos de um conceito generalizado de área, podemos, a partir da definição natural de área de um retângulo como o produto dos comprimentos dos seus lados, caracterizar a medida nula de um conjunto em $\mathbb{R}^2$ como uma noção de "ausência" de área. Para isto, observamos que um segmento arbitrário X de $\mathbb{R}^2$ está contido em um retângulo que tem área tão pequena quanto se queira, isto é, dado $\epsilon > 0$, existe um retângulo K de $\mathbb{R}^2$ que contém X e cuja área é menor que ϵ. Com efeito, basta tomar K com base X e altura menor que ϵ/ρ, em que ρ é o comprimento de X. Motivados por este e outros exemplos simples, dizemos que um subconjunto X de $\mathbb{R}^2$ tem medida nula se, para qualquer $\epsilon > 0$, pode-se obter uma família enumerável de retângulos que o cobrem, cuja soma das áreas não excede ϵ.

Esta ideia, devida ao matemático francês Henri Lebesgue (1875–1941), pode ser facilmente generalizada para se introduzir conjuntos de medida nula em $\mathbb{R}^n$ (vide Definição 27, abaixo). Daí, verifica-se que uma função definida num conjunto limitado de $\mathbb{R}^n$ é integrável se, e somente se, o conjunto dos seus pontos de descontinuidade tem medida nula em $\mathbb{R}^n$.

Ocorrem muito comumente, em Topologia e Análise, fenômenos como o descrito, em que certas funções ideais (contínuas, no caso considerado) têm uma certa propriedade (integrabilidade, idem) que permanece quando estas deixam de ser ideais em subconjuntos de medida nula.

O Teorema de Sard, que apresentaremos, ilustra um desses fenômenos. Nele, consideram-se as funções diferenciáveis definidas em abertos de $\mathbb{R}^n$ e que tomam valores neste espaço. Neste contexto, em vista do Teorema da Função Inversa, uma tal função é ideal se a sua derivada, em cada ponto, é um isomorfismo de $\mathbb{R}^n$. O Teorema de Sard estabelece, então, que, se a função em questão for de classe C^1, a imagem do conjunto dos pontos de seu domínio cuja derivada não é um isomorfismo — ditos, singulares — tem medida nula.

6.1. Conjuntos de Medida Nula. Relembremos que um *paralelepípedo* K em $\mathbb{R}^n$ é um conjunto dado pelo produto cartesiano de n intervalos fechados de $\mathbb{R}$, isto é,

$$K = [a_1, b_1] \times [a_2, b_2] \times \cdots \times [a_n, b_n].$$

A *diagonal* de K, ρ, é definida por

$$\rho = \left(\sum_{i=1}^{n} (b_i - a_i)^2 \right)^{1/2},$$

donde, para quaisquer $x, y \in K$, tem-se $\|x - y\| \leq \rho$ (vide Seção 4 – Capítulo 2).

De um paralelepípedo $K = [a_1, b_1] \times \cdots \times [a_n, b_n] \subset \mathbb{R}^n$ diz-se também que é n-dimensional e define-se o seu volume por

$$\operatorname{vol} K = \prod_{i=1}^{n} (b_i - a_i) = (b_1 - a_1)(b_2 - a_2) \cdots (b_n - a_n).$$

Um paralelepípedo $C = [a_1, b_1] \times \cdots \times [a_n, b_n] \subset \mathbb{R}^n$ é dito um *cubo* quando os intervalos que o compõem têm comprimentos iguais, isto é,

$$(b_i - a_i) = (b_j - a_j) \ \ \forall i, j \in \{1, \ldots, n\}.$$

Neste caso, fazendo-se $\lambda = b_1 - a_1$, tem-se, evidentemente, $\operatorname{vol} C = \lambda^n$.

Deve-se notar também que, para quaisquer $a \in \mathbb{R}^n$ e $r > 0$, a bola

$$B_{\max}[a, r] = \{x \in \mathbb{R}^n ; \ \|x - a\|_{\max} \leq r\} \subset \mathbb{R}^n$$

é um cubo de $\mathbb{R}^n$ cujo volume é $(2r)^n$.

DEFINIÇÃO 27 (MEDIDA NULA). Diz-se que um conjunto $X \subset \mathbb{R}^n$ tem *medida nula* em $\mathbb{R}^n$ se, para todo $\epsilon > 0$, existe uma cobertura de X por uma família enumerável de paralelepípedos de $\mathbb{R}^n$, $\{K_1, \ldots, K_k, \ldots\}$, tais que

$$\sum_{k=1}^{\infty} \operatorname{vol} K_k < \epsilon.$$

Todo conjunto enumerável $X = \{x_1, \ldots, x_k, \ldots\} \subset \mathbb{R}^n$ tem medida nula, pois, dado $\epsilon > 0$, fazendo-se $C_k = B_{\max}[x_k, r_k]$, em que $0 < 2r_k < \sqrt[n]{\epsilon/2^k}$, tem-se $X \subset \bigcup C_k$ e $\operatorname{vol} C_k < (\epsilon/2^k)$. Logo,

$$\sum_{k=1}^{\infty} \operatorname{vol} C_k < \sum_{k=1}^{\infty} \frac{\epsilon}{2^k} = \epsilon.$$

Vale, na verdade, o resultado seguinte.

PROPOSIÇÃO 52. *Toda reunião enumerável de conjuntos de medida nula é um conjunto de medida nula.*

DEMONSTRAÇÃO. Seja $\{X_1, \ldots, X_k, \ldots\}$ uma família enumerável de conjuntos de medida nula em $\mathbb{R}^n$. Dado, então, $\epsilon > 0$, para cada $k \in \mathbb{N}$, existe uma família enumerável de paralelepípedos de $\mathbb{R}^n$, $\{K_{1k}, \ldots, K_{ik}, \ldots\}$, que cobre X_k e satisfaz

$$\sum_{i=1}^{\infty} \operatorname{vol} K_{ik} < \frac{\epsilon}{2^k} \cdot$$

Fazendo-se $X = \bigcup X_k$, temos que a família (enumerável) $\{K_{ik}; i, k \in \mathbb{N}\}$ constitui, certamente, uma cobertura de X. Além disso,

$$\sum_{i,k=1}^{\infty} \operatorname{vol} K_{ik} = \sum_{k=1}^{\infty} \sum_{i=1}^{\infty} \operatorname{vol} K_{ik} < \sum_{k=1}^{\infty} \frac{\epsilon}{2^k} = \epsilon,$$

donde X tem medida nula em $\mathbb{R}^n$. □

Vejamos, agora, uma interessante propriedade dos conjuntos abertos de $\mathbb{R}^n$, a qual aplicaremos na demonstração do Teorema de Sard.

PROPOSIÇÃO 53. *Em $\mathbb{R}^n$, todo conjunto aberto e não-vazio se exprime como uma união enumerável de cubos.*

DEMONSTRAÇÃO. Seja $\mathbb{Q} = \{q_1\,, \dots, q_k\,, \dots\}$ uma enumeração de $\mathbb{Q}$. Dados $i, j \in \mathbb{N},\ i \neq j,$ façamos $I_{ij} = [q_i, q_j]$. Assim, cada $2n$-upla

$$(i_1\,, j_1\,, i_2\,, j_2 \dots, i_n\,, j_n) \in \mathbb{N}^{2n},\ \ i_k \neq j_k\, \forall k \in \{1, \dots, n\},$$

determina um paralelepípedo $K = I_{i_1j_1} \times \cdots \times I_{i_nj_n}$ e reciprocamente. Logo, uma vez que $\mathbb{N}^{2n}$ é enumerável, o conjunto Γ, formado por todos os paralelepípedos de $\mathbb{R}^n$ cujos extremos dos intervalos que os compõem são números racionais, é enumerável.

Consideremos, agora, um conjunto aberto e não-vazio $U \subset \mathbb{R}^n$. Dado, então, um ponto $a = (a_1, \dots, a_n) \in U$, tomemos $r > 0$, tal que $B(a, r) \subset U$. A projeção ortogonal de $B(a, r)$ em cada eixo coordenado de $\mathbb{R}^n$ é um intervalo aberto $I_i = (a_i - r, a_i + r)$. Logo, devido à densidade de $\mathbb{Q}$ em $\mathbb{R}$, para cada $i \in \{1, \dots, n\}$, podemos tomar racionais $q_i, q_i' \in I_i$, tais que $q_i < a_i < q_i'$. Além disso, podemos escolhê-los de tal forma que $q_i' - q_i = q_j' - q_j\ \forall i, j \in \{1, \dots, n\}$ (Fig. 10).

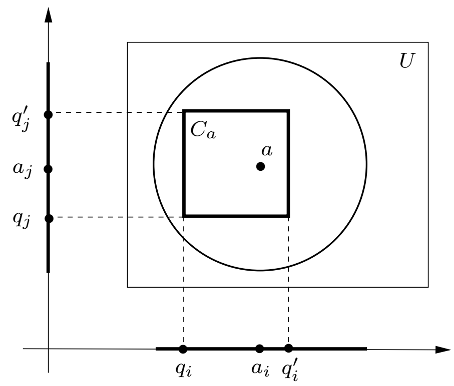

FIGURA 10

Segue-se destas considerações que, para cada $a \in U$, o cubo

$$C_a = [q_1, q_1'] \times \cdots \times [q_n, q_n'],$$

assim obtido, satisfaz $C_a \subset B(a, r) \subset U$. Evidentemente, $U = \bigcup_{a \in U} C_a$ e a família $\{C_a, a \in U\}$, por ser uma subfamília infinita de Γ, é enumerável. $\square$

6.2. O Teorema de Sard.

DEFINIÇÃO 28 (PONTO SINGULAR – PONTO REGULAR). Dados um aberto U de $\mathbb{R}^n$ e uma aplicação diferenciável $f : U \to \mathbb{R}^n$, diz-se que $x \in U$ é um *ponto singular* de f, ou, equivalentemente, uma *singularidade* de f, se $f'(x)$ não é um isomorfismo linear de $\mathbb{R}^n$, isto é, se $\det f'(x) = 0$. Um ponto de U que não é singular é dito *regular*.

Nas vizinhanças de pontos regulares, conforme estabelecido pelo Teorema da Função Inversa, as aplicações diferenciáveis de classe C^1 (definidas em abertos de $\mathbb{R}^n$ e tomando valores nesse espaço) tem uma certa regularidade de comportamento (daí a terminologia), pois são difeomorfismos locais. Esta regularidade, entretanto, deixa de ocorrer em vizinhanças de pontos singulares. Neste contexto, o célebre Teorema de Sard, devido ao matemático americano Arthur Sard (1909–1980) (vide [**30**]), estabelece um resultado extremamente forte, que, dentre outras virtudes, desempenha um papel fundamental em teorias que envolvem singularidades de aplicações, como a Topologia Diferencial e a Teoria de Morse.

Convém mencionar que o conceito de ponto singular se estende ao contexto das aplicações diferenciáveis entre espaços euclidianos de dimensões distintas, e que o Teorema de Sard é válido, também, neste caso. No que se segue, apresentaremos uma versão mais simples do mesmo, na qual os espaços euclidianos envolvidos têm mesma dimensão.

Denotemos por $\operatorname{sing} f$ o conjunto dos pontos singulares de uma aplicação diferenciável f.

Teorema de Sard. *Seja U um conjunto aberto de $\mathbb{R}^n$. Se $f : U \to \mathbb{R}^n$ é uma aplicação de classe C^1, então $f(\operatorname{sing} f)$ tem medida nula em $\mathbb{R}^n$.*

Demonstração. O resultado é trivial quando $\operatorname{sing} f = \emptyset$. Suponhamos, então, que $\operatorname{sing} f$ seja não-vazio e consideremos um cubo $K \subset U$, tal que $\operatorname{sing} f \cap K \neq \emptyset$. Tomando-se $1 < q \in \mathbb{N}$ e uma partição de cada intervalo que define K em q intervalos de comprimentos iguais, obtém-se uma decomposição de K em q^n cubos

$$C_1, C_2, \dots, C_{q^n},$$

cada um com diagonal ρ/q, em que ρ é a diagonal de K. Em particular,

$$\|x - y\| \le \rho/q \ \ \forall x, y \in C_k, \ k = 1, \dots, q^n. \tag{70}$$

Uma vez que f é de classe C^1, a aplicação $x \mapsto f'(x)$ é contínua em K. Logo, é uniformemente contínua, pois K é compacto. Dado, então, $\epsilon > 0$, existe $\delta > 0$, tal que

$$x, y \in K \quad \text{e} \quad \|x - y\| < \delta \quad \Rightarrow \quad \|f'(x) - f'(y)\| < \epsilon.$$

Assim, tomando-se q suficientemente grande (de tal modo que $\rho/q < \delta$), tem-se, para todo $k \in \{1, \dots, q^n\}$,

$$\|f'(x) - f'(y)\| < \epsilon \ \ \forall x, y \in C_k. \tag{71}$$

Fixemos um cubo $C = C_k$ desta partição cuja interseção com $\operatorname{sing} f$ seja não-vazia e tomemos $a \in \operatorname{sing} f \cap C$. Então, $f'(a) \in \mathrm{L}(\mathbb{R}^n)$ não é um isomorfismo. Logo, existe um subespaço $\mathbb{V}$ de $\mathbb{R}^n$, de dimensão $n-1$, que contém o conjunto-imagem de $f'(a)$. Em particular, para todo $x \in C$, o vetor $w = f(a) + f'(a)(x-a)$ pertence ao *espaço afim*

$$\mathbb{W} = f(a) + \mathbb{V} = \{w \in \mathbb{R}^n;\ w = f(a) + v,\ v \in \mathbb{V}\}.$$

Agora, fazendo-se $\varphi(x) = f(x) - f'(a)x,\ x \in C$, tem-se $\varphi'(x) = f'(x) - f'(a)$. Logo, pela desigualdade (71), $\|\varphi'(x)\| < \epsilon\ \forall x \in C$. Segue-se, então, do Teorema do Valor Médio (note que C é convexo), bem como da desigualdade (70), que

$$\|\varphi(x) - \varphi(a)\| \le \epsilon \|x - a\| \le \epsilon \frac{\rho}{q} \cdot$$

Dado, então, $x \in C$, fazendo-se $w = f(a) + f'(a)(x-a)$, tem-se

$$\|f(x) - w\| = \|f(x) - (f(a) + f'(a)(x-a))\| = \|\varphi(x) - \varphi(a)\| \le \epsilon \frac{\rho}{q}\,,$$

donde se conclui que

$$d(f(x), \mathbb{W}) = \min_{w \in \mathbb{W}} \|f(x) - w\| \le \epsilon \frac{\rho}{q} \ \ \forall x \in C. \tag{72}$$

Também, escrevendo-se $\mu = \max\{\|f'(x)\|;\, x \in K\}$ e aplicando-se novamente o Teorema de Valor Médio, obtém-se

$$\|f(x) - f(a)\| \le \mu\|x - a\| \le \mu\frac{\rho}{q} \quad \forall x \in C. \tag{73}$$

Considerando-se, momentaneamente, um sistema de coordenadas $(y_1, \dots, y_n)$ em $\mathbb{R}^n$, em que $f(a)$ seja a origem e $\mathbb{W}$ coincida com o *subespaço coordenado*(v)

$$\{(y_1, \dots, y_n) \in \mathbb{R}^n\,;\ y_n = 0\},$$

vê-se, pela desigualdade (73), que a projeção ortogonal de $f(x)$ sobre $\mathbb{W}$ está contida na bola de $\mathbb{W}$ (aqui identificado com $\mathbb{R}^{n-1}$) com centro em $f(a)$ e raio $\mu\rho/q$, que, por sua vez, está contida no cubo $(n-1)$-dimensional

$$C_0 = B_{\max}[f(a), \mu\rho/q] \cap \mathbb{W}.$$

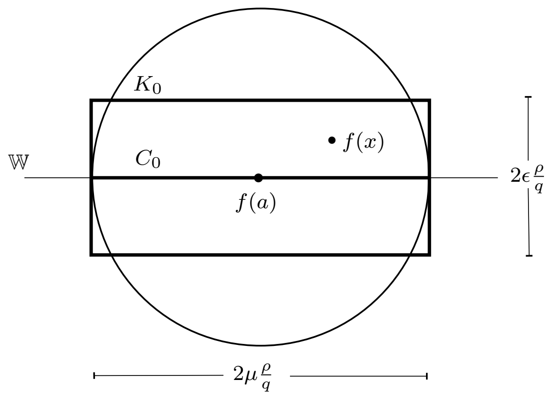

FIGURA 11

Daí e de (72), segue-se que, para todo $x \in C$, $f(x)$ pertence ao paralelepípedo $K_0 = C_0 \times [-\epsilon\rho/q, \epsilon\rho/q] \subset \mathbb{R}^n$ (Fig. 11), para o qual tem-se

$$\operatorname{vol} K_0 = \left(2\epsilon\frac{\rho}{q}\right)\operatorname{vol} C_0 = \left(2\epsilon\frac{\rho}{q}\right)\left(\frac{2\rho\mu}{q}\right)^{n-1} = \frac{(2\rho)^n\mu^{n-1}}{q^n}\epsilon.$$

Adotando-se este procedimento em cada um dos cubos da partição que intersectam $\operatorname{sing} f$, concluímos que $f(\operatorname{sing} f \cap K)$ está contido numa união de, no máximo, q^n paralelepípedos, cada um deles com volume igual a $((2\rho)^n\mu^{n-1}/q^n)\epsilon$. Logo, a soma de seus volumes é menor que, ou igual a,

$$q^n\frac{(2\rho)^n\mu^{n-1}}{q^n}\epsilon = (2\rho)^n\mu^{n-1}\epsilon,$$

donde se conclui que $f(\operatorname{sing} f \cap K)$ tem medida nula em $\mathbb{R}^n$.

Agora, pela Proposição 53,

$$U = K_1 \cup \cdots \cup K_k \cup \cdots\,,$$

(v)Note que a aplicação que faz esta mudança de coordenadas é uma isometria $T - f(a)$, em que T é um operador ortogonal de $\mathbb{R}^n$ que leva $\mathbb{V}$ no subespaço coordenado $\{x \in \mathbb{R}^n;\, x_n = 0\}$.

em que, para cada $k \in \mathbb{N}$, K_k é um cubo de $\mathbb{R}^n$. Além disso, para cada $k \in \mathbb{N}$, conforme o concluído acima, $f(\operatorname{sing} f \cap K_k)$ tem medida nula em $\mathbb{R}^n$.

Uma vez que

$$\operatorname{sing} f = \operatorname{sing} f \cap U = \operatorname{sing} f \cap \left(\bigcup K_k\right) = \bigcup(\operatorname{sing} f \cap K_k),$$

tem-se

$$f(\operatorname{sing} f) = f\left(\bigcup(\operatorname{sing} f \cap K_k)\right) = \bigcup f(\operatorname{sing} f \cap K_k),$$

donde se infere que $f(\operatorname{sing} f)$ tem medida nula em $\mathbb{R}^n$, pois, pela Proposição 52, toda reunião enumerável de conjuntos de medida nula tem medida nula. □

7. Exercícios

Seção 1

1. Seja $U \subset \mathbb{R}^n$ aberto e conexo. Prove que se $f : U \to \mathbb{R}^m$ é diferenciável e $f'(x) = T$ para todo $x \in U$, então existe $a \in \mathbb{R}^m$, tal que $f(x) = Tx + a$.

2. Seja $f : \mathbb{R} \to \mathbb{R}$ uma função real de classe C^1. Defina $F : \mathbb{R}^2 \to \mathbb{R}$ pondo $F(x,t) = \frac{f(x+t)-f(x)}{t}$ quando $t \neq 0$ e $F(x,0) = f'(x)$. Demonstre que F é uma função contínua.

3. Seja $g : \mathbb{R}^n \to \mathbb{R}^n$ uma aplicação diferenciável, tal que, para todo $x \in \mathbb{R}^n$, $\|g'(x)\| \leq \lambda < 1$. Considere o Exercício 13 do Capítulo 3 e prove que a aplicação $f : \mathbb{R}^n \to \mathbb{R}^n$, dada por $f(x) = x + g(x)$, é um homeomorfismo.

4. Seja $f : \mathbb{R}^n \to \mathbb{R}^m$ uma aplicação diferenciável em $\mathbb{R}^n - \{0\}$, contínua em $x = 0$ e que satisfaz
$$\lim_{x \to 0} \frac{\partial f}{\partial x_i}(x) = 0\,.$$
Prove que f é diferenciável.

Seção 2

5. Sejam $f, g : U \subset \mathbb{R}^n \to \mathbb{R}^m$ aplicações de classe C^2 no aberto U. Suponha que, para quaisquer $x \in U$ e $i \in \{1, \dots, n\}$, tenha-se $\langle f(x), \frac{\partial g}{\partial x_i}(x)\rangle = 0$. Prove que, para todo $x \in U$, a matriz $A(x) = (a_{ij}(x))$, em que $a_{ij}(x) = \langle \frac{\partial f}{\partial x_i}(x), \frac{\partial g}{\partial x_j}(x)\rangle$, é simétrica.

6. Seja $f : \mathbb{R}^2 \to \mathbb{R}$ uma função duas vezes diferenciável. Suponha que, para quaisquer $x, y, t \in \mathbb{R}$, tenha-se
$$f(x+t, y+t) + f(x-t, y-t) - f(x+t, y-t) - f(x-t, y+t) = 0.$$
Prove que $\dfrac{\partial^2 f}{\partial x \partial y}(x,y) = 0 \ \ \forall (x,y) \in \mathbb{R}^2.$

7. Diz-se que uma função duas vezes diferenciável é *harmônica* quando seu laplaciano (vide Exercício 7 – Capítulo 4) é uma função identicamente nula. Prove que as coordenadas de uma aplicação holomorfa (vide Exemplo 65 – Capítulo 4), $f : \mathbb{R}^2 \to \mathbb{R}^2$, são funções harmônicas.

8. Seja $f : \mathbb{R}^n \to \mathbb{R}^n$ uma aplicação duas vezes diferenciável, tal que, para cada $x \in \mathbb{R}^n$, $f'(x) : \mathbb{R}^n \to \mathbb{R}^n$ é uma isometria, isto é, uma transformação linear ortogonal. Prove que a derivada segunda de f é identicamente nula e conclua, daí e do Exercício 1, que f é uma isometria de $\mathbb{R}^n$.

Seção 3

9. Seja $f : \mathbb{R}^n \to \mathbb{R}$ uma função duas vezes diferenciável cujo laplaciano, em todos os pontos de $\mathbb{R}^n$, é positivo. Mostre que f não possui máximo local.

10. Seja $f : \mathbb{R}^n \to \mathbb{R}$ uma função duas vezes diferenciável cuja derivada segunda $f''(x)$ é positiva definida em todo ponto $x \in \mathbb{R}^n$. Prove que f tem, no máximo, um ponto crítico.

Seção 4

11. Sejam f e g as aplicações do Exercício 3. Prove que se g é de classe C^1, então $f : \mathbb{R}^n \to \mathbb{R}^n$ é um difeomorfismo de classe C^1.

12. Sejam $f : \mathbb{R} \to \mathbb{R}$ de classe C^1 e $\varphi : \mathbb{R}^2 \to \mathbb{R}^2$, tais que

$$\varphi(x, y) = (u(x, y), v(x, y)) = (f(x), xf(x) - y).$$

Dado $(x_0, y_0) \in \mathbb{R}^2$, suponha que $f'(x_0) \neq 0$. Prove que, em uma vizinhança de (x_0, y_0), φ é invertível e satisfaz $\varphi^{-1}(u, v) = (g(u), ug(u) - v)$, em que g é uma função de classe C^1.

13. Mostre que a aplicação $f : \mathrm{L}(\mathbb{R}^n) \to \mathrm{L}(\mathbb{R}^n)$, dada por $f(X) = X^2$, é um difeomorfismo de uma vizinhança da identidade sobre outra vizinhança da identidade mas não é um difeomorfismo local.

14. Prove que existe $\delta > 0$, tal que, se $A \in \mathrm{M}(n)$ satisfaz $\|A\| < \delta$, então existe uma matriz $X \in \mathrm{M}(n)$, tal que $X^2 + X^* = A$.

15. Seja $f : \mathbb{R}^n \to \mathbb{R}^n$ uma aplicação de classe C^1, tal que a matriz jacobiana de f em cada ponto x de $\mathbb{R}^n$ é invertível. Mostre que se f é própria (vide Exercício 21 – Capítulo 3), então $f(\mathbb{R}^n) = \mathbb{R}^n$.

16. Seja $f : \mathbb{R}^n \to \mathbb{R}^n$ uma aplicação de classe C^1, tal que

$$\|x - y\| \leq \|f(x) - f(y)\| \ \ \forall x, y \in \mathbb{R}^n.$$

Prove que f é um difeomorfismo de classe C^1.

17. Mostre que se $U \subset \mathbb{R}^n$ é aberto e $f : U \to \mathbb{R}^m$ é uma submersão, então $\|f(x)\|$ não assume valor máximo.

18. Mostre que a função determinante é uma submersão em $X \in \mathrm{M}(n)$ se, e somente se, $\mathrm{posto}(X) \geq n - 1$.

19. Prove a *Forma Local das Imersões*: Sejam $U \subset \mathbb{R}^n$ um conjunto aberto e $f : U \to \mathbb{R}^n \times \mathbb{R}^m = \mathbb{R}^{n+m}$ uma aplicação de classe C^1, tal que f é uma

imersão em $a \in U$, isto é, $f'(a) \in \mathrm{L}(\mathbb{R}^n, \mathbb{R}^{n+m})$ é injetiva[(vi)]. Então, existem abertos $A \ni a$, de U, e $W \ni f(a)$, de $\mathbb{R}^{n+m}$, bem como um difeomorfismo $G : W \to G(W) \subset \mathbb{R}^{n+m}$, tais que a composta $G \circ f : A \to \mathbb{R}^{n+m}$ está bem definida e, para todo $x \in A$, vale a igualdade $(G \circ f)(x) = (x, 0)$ (Fig. 12).

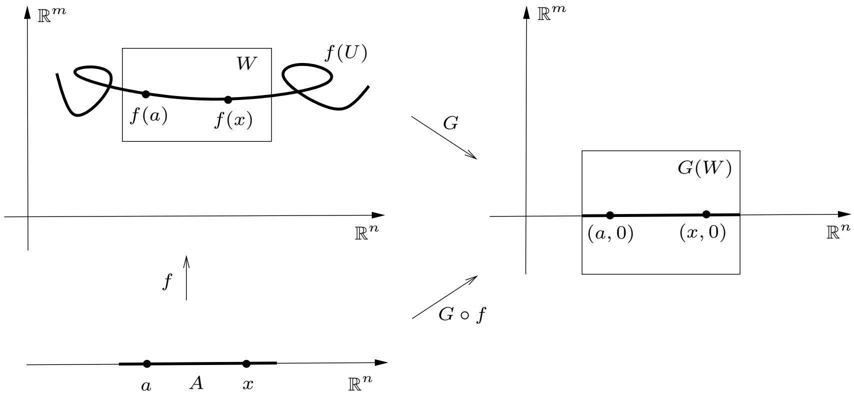

FIGURA 12

Seção 5

20. Sejam $f : \mathbb{R}^2 \to \mathbb{R}$ e $\xi : A \subset \mathbb{R}^2 \to \mathbb{R}$ como no Exemplo 78. Prove que ξ é de classe C^∞ e

$$\frac{\partial^2 \xi}{\partial x^2}(x, y) = \xi_{xx}(x, y) = \frac{f_x(f_{xz} + f_{zz}\xi_x) - f_z(f_{xx} + f_{xz}\xi_x)}{f_z^2},$$

em que as aplicações f_x, f_z, f_{xx}, f_{xz} e f_{zz} são calculadas no ponto $(x, y, \xi(x, y))$ e ξ_x é calculada em (x, y). Obtenha relações análogas para ξ_{xy} e ξ_{yy} e determine a fórmula de Taylor de segunda ordem de ξ numa vizinhança de $(0, 0)$.

21. Prove que existem um aberto de $\mathbb{R}^2$, $A \ni (1, -1)$, e um intervalo $(-\epsilon, \epsilon) \subset \mathbb{R}$, tais que, para todo ponto $(x, y) \in A$, a equação de variável real t,

$$xt^2 + e^{2t} + y = 0,$$

admite uma única solução $t = t(x, y)$ em $(-\epsilon, \epsilon)$, e a função $t \mapsto t(x, y)$, assim definida, é de classe C^∞.

22. Sejam $f, g : \mathbb{R} \to \mathbb{R}$ funções de classe C^1. Suponha que $f(0) = 0$ e $f'(0) \neq 0$. Mostre que existem $\delta > 0$ e uma função diferenciável $x = x(t)$, definida em $(-\delta, \delta)$, tais que $f(x(t)) = tg(x(t))$ $\forall t \in (-\delta, \delta)$.

23. Dada uma matriz $B \in \mathrm{M}(n)$, prove que existe uma curva de classe C^∞ $A : (-\epsilon, \epsilon) \to \mathrm{M}(n)$, tal que $A(0) = I$, $A'(0) = -\frac{1}{2}B$ e, para todo $t \in (-\epsilon, \epsilon)$, tem-se $A(t)^2 + tBA(t) = I$.

[(vi)]Quando a derivada $f'(x)$ é injetiva para todo $x \in U$, f é dita, simplesmente, uma *imersão*.

24. Dado $a = (a_0, \dots, a_n) \in \mathbb{R}^{n+1}$, denotemos por p_a o polinômio de variável real t, dado por $a_0 + a_1 t + \cdots + a_n t^n$. Suponha que $t_0 \in \mathbb{R}$ seja uma raiz simples de p_a, isto é, existe um polinômio $q = q(t)$, tal que $p_a(t) = (t - t_0)q(t)$ e $q(t_0) \neq 0$. Mostre que existem uma bola aberta $B(a, r) \subset \mathbb{R}^{n+1}$ e um intervalo aberto $(-\epsilon, \epsilon) \subset \mathbb{R}$, tais que, para todo $x \in B(a, r)$, o polinômio p_x tem um único zero $t = t(x) \in (-\epsilon, \epsilon)$, o qual é simples. Além disso, a função $x \mapsto t(x)$, $x \in B(a, r)$, é de classe C^∞ e satisfaz $t(a) = t_0$.

25. Sejam p e q números reais positivos, tais que $\frac{1}{p} + \frac{1}{q} = 1$. Prove, pelo método dos multiplicadores de Lagrange, a desigualdade
$$\frac{1}{p}x^p + \frac{1}{q}y^q \geq 1 \;\; \forall x, y > 0, \;\; xy = 1.$$
Conclua, então, que
$$\frac{1}{p}x^p + \frac{1}{q}y^q \geq xy \;\; \forall x, y > 0,$$
e, daí, a validez da *desigualdade de Hölder*,
$$\sum_{i=1}^n x_i y_i \leq \left(\sum_{i=1}^n x_i^p\right)^{\frac{1}{p}} \left(\sum_{i=1}^n y_i^q\right)^{\frac{1}{q}} \quad \forall x_i, y_i > 0, \;\; i = 1, \dots, n.$$

APÊNDICE A

Variedades Diferenciáveis

As superfícies constituem, desde a antiguidade, objetos geométricos por excelência, às quais vêm sendo devotadas teorias que as estudam e generalizam sob os mais diversos aspectos. Neste processo histórico, o advento do Cálculo, no século XVII, tornou possível abordar as superfícies através do conceito de derivada, o que deu origem à teoria conhecida como *Geometria Diferencial*, brilhantemente estabelecida no início do século XIX por Karl Friedrich Gauss (1777–1855).

Subsequentes generalizações do conceito de superfície regular, a qual é estudada pela Geometria Diferencial, conduziram naturalmente ao conceito mais geral de *variedade diferenciável*, o qual é atribuído a uma certa classe de espaços topológicos que se caracterizam pelo fato de se poder associar a cada um de seus pontos um espaço vetorial, dito *tangente*. A partir daí, torna-se possível estender às variedades diferenciáveis a teoria do cálculo diferencial e, em particular, o conceito de difeomorfismo, que estabelece, então, uma relação de equivalência na classe das variedades diferenciáveis.

No que concerne a esta classificação, há um célebre resultado, devido ao matemático americano Hassler Whitney (1907–1989), segundo o qual toda variedade diferenciável é difeomorfa a uma subvariedade de um espaço euclidiano $\mathbb{R}^n$ — isto é, um subespaço topológico deste que é uma variedade diferenciável. Tendo-o em consideração, não se perde generalidade quando se reduz o estudo das variedades diferenciáveis àquele das subvariedades do espaço $\mathbb{R}^n$, um procedimento que, por simplicidade, adotaremos.

No que se segue, a fim de ilustrar a força e efetividade da teoria discutida nos capítulos precedentes, introduziremos, a partir dela, o conceito de variedade diferenciável, bem como estenderemos o conceito de derivada às aplicações entre variedades diferenciáveis.

Cabe-nos, por fim, mencionar que as variedades diferenciáveis são objetos de intensa investigação, as quais estão intimamente relacionadas com uma das questões matemáticas mais relevantes dos últimos 100 anos, conhecida como *conjectura de Poincaré*[i], que, em 2002, foi demonstrada pelo matemático russo Grigori Yakovlevich Perelman (1966–).

0.1. Definições – Exemplos. Dados conjuntos $X \subset \mathbb{R}^n$ e $Y \subset \mathbb{R}^m$, diz-se que uma aplicação $f : X \to Y$ é *diferenciável em* $x \in X$ se existe uma extensão local diferenciável de f a um aberto $U \subset \mathbb{R}^n$ que contém x, isto é, uma aplicação $F : U \to \mathbb{R}^m$, diferenciável em $x \in U$, e tal que $F|_{U \cap X} = f|_{U \cap X}$. Quando f é diferenciável em todos os pontos de X, diz-se, simplesmente, que f é *diferenciável*. Diz-se, ainda, que f é um *difeomorfismo* se for um homeomorfismo e as aplicações f e f^{-1} forem, ambas, diferenciáveis.

[i]Em consideração ao matemático francês Henri Poincaré (1854–1912), que a levantou.

Observe que estes conceitos de diferenciabilidade e difeomorfismo são extensões daqueles introduzidos anteriormente para aplicações definidas em subconjuntos abertos de espaços euclidianos.

A *aplicação antípoda* de S^n, por exemplo,

$$\begin{array}{rccc} f: & S^n & \to & S^n \\ & x & \mapsto & -x, \end{array}$$

é um difeomorfismo, pois a aplicação $F(x) = -x$, $x \in \mathbb{R}^{n+1}$, é um difeomorfismo de $\mathbb{R}^{n+1}$ em si mesmo que cumpre $F|_{S^n} = f$.

Definição 29 (variedade diferenciável). Um conjunto $M \subset \mathbb{R}^m$ é dito uma *variedade diferenciável* de *dimensão* n se, para cada $x \in M$, existe um difeomorfismo $\varphi : U \subset \mathbb{R}^n \to V \subset M$, em que U é um aberto de $\mathbb{R}^n$ e V é um aberto (relativo) de M que contém x. A aplicação φ é dita, então, uma *parametrização local* de M em x (Fig. 1).

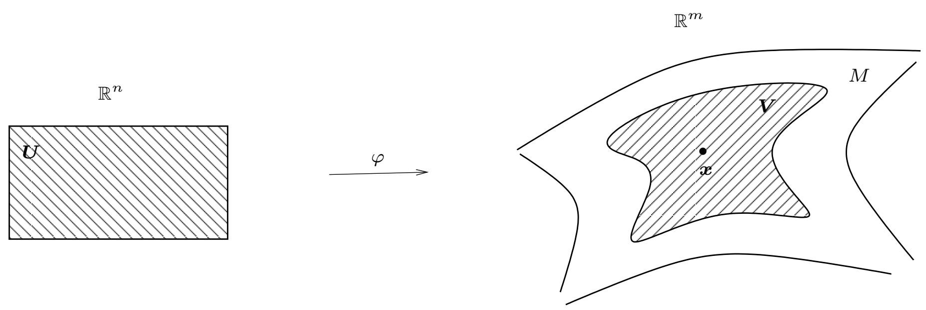

Figura 1

Quando se deseja indicar a dimensão n de uma variedade diferenciável M, costuma-se denotá-la por M^n.

Exemplo 80 (espaços euclidianos – espaços afins). O espaço euclidiano $\mathbb{R}^n$ é uma variedade de dimensão n, pois, para todo $x \in \mathbb{R}^n$, a aplicação identidade de $\mathbb{R}^n$ é uma parametrização de $\mathbb{R}^n$ em x. Também, se $\mathbb{V} \subset \mathbb{R}^m$ é um subespaço vetorial de $\mathbb{R}^m$ cuja dimensão é $n > 0$ e $a \in \mathbb{R}^n$, então o *subespaço afim*

$$\mathbb{V} + a = \{v + a\,;\, v \in \mathbb{V}\}$$

é uma variedade diferenciável de dimensão n. Com efeito, nestas condições, existe um isomorfismo linear $T : \mathbb{R}^n \to \mathbb{V}$. Logo, $\varphi = T + a$ é um homeomorfismo diferenciável de $\mathbb{R}^n$ em $\mathbb{V}+a$. Além disso, tomando-se uma extensão linear qualquer de T^{-1} a $\mathbb{R}^m$ e denotando-a por τ, tem-se que $\tau - a$ é uma extensão diferenciável de φ^{-1} a $\mathbb{R}^m$, donde φ^{-1} é diferenciável e, portanto, φ é uma parametrização de $\mathbb{V} + a$ em todos os seus pontos.

Exemplo 81 (a esfera S^n). Conforme discutimos no Exemplo 46 do Capítulo 3, considerando-se a decomposição $\mathbb{R}^{n+1} = \mathbb{R}^n \times \mathbb{R}$ e fazendo-se $p = (0, 1) \in S^n$, a projeção estereográfica $f : S^n - \{p\} \to \mathbb{R}^n$, definida por

$$f(x, s) = \frac{x}{1 - s}\,,$$

é um homeomorfismo. Além disso, sua inversa, $\varphi = f^{-1}$, define-se por

$$\varphi(y) = f^{-1}(y) = \left(\frac{2y}{\|y\|^2+1}, \frac{\|y\|^2-1}{\|y\|^2+1}\right) \in S^n \subset \mathbb{R}^n \times \mathbb{R}.$$

É imediato que φ é uma aplicação diferenciável e que $\varphi^{-1} = f$ estende-se diferenciavelmente ao aberto $\{(x,s) \in \mathbb{R}^{n+1};\, s \neq 1\}$, donde, para todo $x \in S^n - \{p\}$, φ é uma parametrização de S^n em x.

Analogamente, tomando-se a restrição da aplicação antípoda a $S^n - \{-p\}$,

$$\begin{array}{rccc} g: & S^n - \{-p\} & \to & S^n - \{p\} \\ & x & \mapsto & -x, \end{array}$$

verifica-se que, para todo $x \in S^n - \{-p\}$, a aplicação $f \circ g : S^n - \{-p\} \to \mathbb{R}^n$ é um homeomorfismo cujo inverso é uma parametrização de S^n em x. Logo, S^n é uma variedade diferenciável de dimensão n.

Exemplo 82 (gráficos de aplicações diferenciáveis). Dada uma aplicação diferenciável $f : U \subset \mathbb{R}^n \to \mathbb{R}^m$, definida num aberto U de $\mathbb{R}^n$, consideremos o seu gráfico,

$$\operatorname{graf} f = \{(x, f(x))\,;\, x \in U\} \subset \mathbb{R}^n \times \mathbb{R}^m.$$

Conforme constatamos anteriormente (vide Exemplo 45 – Capítulo 3), a aplicação $\varphi : U \to \operatorname{graf} f$, $\varphi(x) = (x, f(x))$, é um homeomorfismo cujo inverso é a restrição, ao gráfico de f, da projeção ortogonal P de $\mathbb{R}^n \times \mathbb{R}^m$ sobre $\mathbb{R}^n$. Daí, uma vez que f e P são diferenciáveis, infere-se que φ e φ^{-1} são diferenciáveis, donde, para todo $(x, f(x)) \in \operatorname{graf} f$, φ é uma parametrização de $\operatorname{graf} f$ em $(x, f(x))$. Logo, $\operatorname{graf} f$ é uma variedade diferenciável de dimensão n.

Em particular, a *sela*, gráfico da função $f(x,y) = x^2 - y^2$, $(x,y) \in \mathbb{R}^2$, é uma variedade diferenciável de dimensão 2 (Fig. 2).

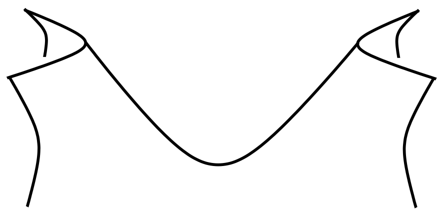

Figura 2

Exemplo 83 (matrizes ortogonais). Identifiquemos o espaço $\mathrm{M}(n)$, das matrizes quadradas de ordem $n > 1$, com $\mathbb{R}^{n^2}$, e o espaço das matrizes simétricas de $\mathrm{M}(n)$ com $\mathbb{R}^{\frac{n(n+1)}{2}}$. Definindo-se $f : \mathbb{R}^{n^2} \to \mathbb{R}^{\frac{n(n+1)}{2}}$ por $f(A) = AA^*$, temos que f é de classe C^∞ e $f^{-1}(\{I\})$ é o conjunto das matrizes ortogonais de $\mathrm{M}(n)$, $\mathrm{O}(n)$. Além disso, para quaisquer $A, H \in \mathrm{M}(n)$, tem-se $f'(A)H = AH^* + HA^*$. Em particular, se A é ortogonal e $B \in \mathrm{M}(n)$ é simétrica, a matriz $H = \frac{1}{2}BA$, claramente, satisfaz $f'(A)H = B$, donde se infere que $f'(A)$ é sobrejetiva e, portanto, que f é uma submersão em todo ponto $A \in O(n)$. Logo, pelo Teorema da Função Implícita, para cada $A \in O(n)$, existem uma bola aberta $B(A,r)$ de $\mathrm{M}(n)$ e um aberto U de $\mathbb{R}^{\frac{n(n-1)}{2}}$, tais que $B(A,r) \cap O(n)$ é o gráfico de uma aplicação de classe C^∞, $\xi : U \to \mathbb{R}^{\frac{n(n+1)}{2}}$. Daí e das considerações do exemplo anterior, segue-se que $O(n)$ é uma variedade diferenciável de dimensão $n(n-1)/2$.

De modo análogo, estabelece-se o fato geral de que todo conjunto de nível de uma submersão de classe C^1, $f : U \subset \mathbb{R}^{n+m} \to \mathbb{R}^m$, é uma variedade diferenciável de dimensão n.

Exemplo 84 (produtos cartesianos). Sejam $M^m \subset \mathbb{R}^q$ e $N^n \subset \mathbb{R}^k$ variedades diferenciáveis de dimensões m e n, respectivamente. Neste caso, dado $(x,y) \in M \times N$, existem parametrizações $\varphi : U \subset \mathbb{R}^m \to M$, de M em x, e $\psi : V \subset \mathbb{R}^n \to N$, de N em y. Daí, verifica-se facilmente que a aplicação $\xi : U \times V \to M \times N$, $\xi(x,y) = (\varphi(x), \psi(y))$, é uma parametrização de $M \times N$ em (x,y) e, portanto, que $M^m \times N^n$ é uma variedade diferenciável de dimensão $m+n$.

Em particular, o *cilindro* $S^n \times \mathbb{R} \subset \mathbb{R}^{n+2}$ é uma variedade diferenciável de dimensão $n+1$ (Fig. 3) e o *toro* $S^n \times S^n$ é uma variedade diferenciável de dimensão $2n$ (Fig. 4).

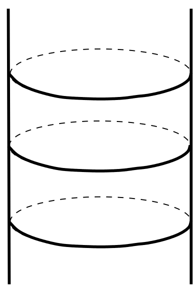

Figura 3. $S^1 \times \mathbb{R}$

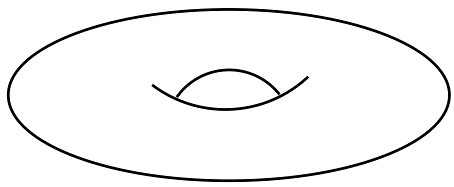

Figura 4. $S^1 \times S^1$

0.2. Cálculo em Variedades. Introduziremos, agora, a noção de derivada de aplicações entre variedades associando a cada ponto x de uma variedade diferenciável M um espaço vetorial, T_xM, o qual chamaremos *espaço tangente* a M em x. Dada, então, uma aplicação diferenciável $f : M \to N$, entre variedades M e N, definiremos a *derivada* de f em $x \in M$ como uma transformação linear $f'(x) : T_xM \to T_{f(x)}N$.

0.2.1. *O Espaço Tangente.* Sejam $\varphi : U \subset \mathbb{R}^n \to M$ e $\psi : V \subset \mathbb{R}^n \to M$ parametrizações distintas de uma variedade diferenciável M num ponto $x \in M$. Uma vez que φ e ψ são homeomorfismos, tem-se que:

i) $O = \varphi(U) \cap \varphi(V)$ é um aberto de M que contém x;

ii) $U_0 = \varphi^{-1}(O) \subset U$ e $V_0 = \psi^{-1}(O) \subset V$ são abertos de $\mathbb{R}^n$.

Assim, escrevendo-se $\varphi_0 = \varphi|_{U_0}$ e $\psi_0 = \psi|_{V_0}$, a aplicação composta $\xi = \psi_0^{-1} \circ \varphi_0 : U_0 \to V_0$ é um homeomorfismo, o qual denomina-se *mudança de parâmetros* (Fig. 5).

Verifiquemos que ξ é um difeomorfismo. Para tanto, tomemos os abertos U e V suficientemente pequenos, de tal modo que existam extensões diferenciáveis de φ_0^{-1} e ψ_0^{-1} a um aberto de $\mathbb{R}^n$ que contém O. Designando-as, respectivamente,

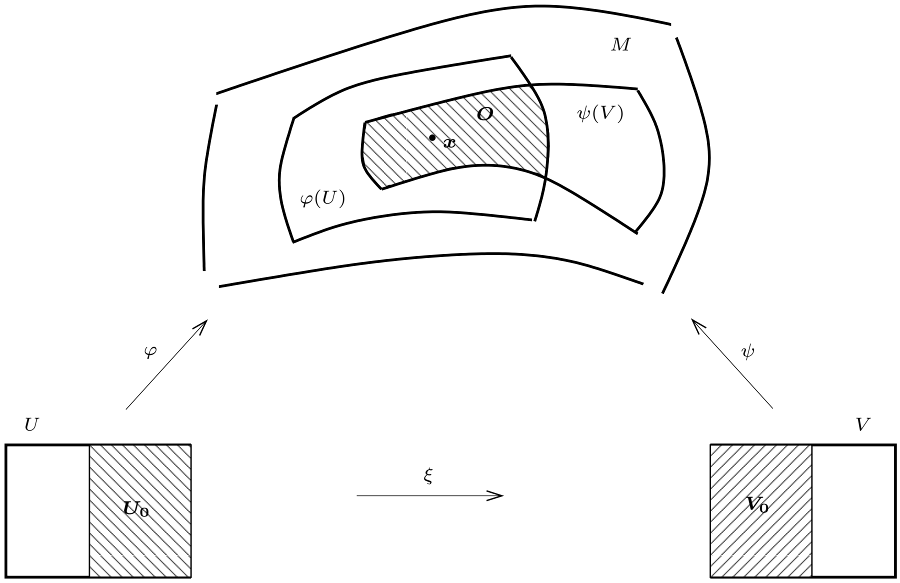

Figura 5

por f e g, tem-se

$$\xi = \psi_0^{-1} \circ \varphi_0 = g \circ \varphi_0 \quad \text{e} \quad \xi^{-1} = \varphi_0^{-1} \circ \psi_0 = f \circ \psi_0 \,.$$

Logo, pela Regra da Cadeia, ξ e ξ^{-1} são diferenciáveis, donde ξ é um difeomorfismo.

Desta forma, fazendo-se $p = \varphi^{-1}(x)$ e $q = \psi^{-1}(x)$, tem-se, em particular, que $\xi'(p)$ é um isomorfismo. Logo, uma vez que $\psi_0 \circ \xi = \varphi_0$, $\varphi_0'(p) = \varphi'(p)$ e $\psi_0'(q) = \psi'(q)$, segue-se da Regra da Cadeia que $\psi'(q)\xi'(p) = \varphi'(p)$, donde $\varphi'(p)(\mathbb{R}^n) = \psi'(q)(\xi'(p)(\mathbb{R}^n)) = \psi'(q)(\mathbb{R}^n)$, isto é, $\varphi'(p)$ e $\psi'(q)$ têm o mesmo conjunto-imagem. Assim, o conjunto

$$T_xM = \varphi'(p)(\mathbb{R}^n), \; p = \varphi^{-1}(x),$$

ao qual chamamos *espaço tangente a* M *em* x, está bem definido, isto é, independe da parametrização φ.

Finalmente, considerando-se a igualdade $\xi = g \circ \varphi_0$ e aplicando-se, uma vez mais, a Regra da Cadeia, tem-se $\xi'(p) = g'(x)\varphi'(p)$, donde se infere que $\varphi'(p)$ é injetiva, já que $\xi'(p)$, por ser um isomorfismo, o é. Daí, segue-se que

$$\varphi'(p) : \mathbb{R}^n \to T_xM$$

é um isomorfismo linear e, portanto, que T_xM é um subespaço vetorial de $\mathbb{R}^m$ de dimensão n.

Dado um ponto x de uma variedade diferenciável $M^n \subset \mathbb{R}^m$, diz-se que um vetor $v \in \mathbb{R}^m$ é *tangente* a M em x, quando existe uma curva diferenciável $\gamma : (-\epsilon, \epsilon) \to M$, $\epsilon > 0$, tal que $\gamma(0) = x$ e $\gamma'(0) = v$.

Esta terminologia justifica-se pela igualdade

$$T_xM = \{v \in \mathbb{R}^3; \, v \text{ é tangente a } M \text{ em } x\}.$$

Para estabelecê-la, consideremos uma parametrização de M em x, $\varphi : U \subset \mathbb{R}^n \to M$, e $p = \varphi^{-1}(x)$. Dado $v \in \mathbb{R}^3$, tangente a M em x, seja $\gamma : (-\epsilon, \epsilon) \to M$, tal que $\gamma(0) = x$ e $\gamma'(0) = v$. Tomemos $\epsilon > 0$ suficientemente pequeno, de tal modo que o traço de γ esteja contido em $V = \varphi(U)$. Neste caso, se f é uma extensão diferenciável de φ^{-1} a um aberto $O \ni x$ de $\mathbb{R}^m$ cuja interseção com M está contida em V, tem-se que $\sigma = f \circ \gamma = \varphi^{-1} \circ \gamma$ é uma curva diferenciável. Daí, da igualdade $\varphi \circ \sigma = \gamma$ e da Regra da Cadeia, obtém-se $\varphi'(\sigma(0))\sigma'(0) = \gamma'(0)$, isto é, $v = \varphi'(p)\sigma'(0) \in T_xM$.

Reciprocamente, dado $v \in T_xM$, uma vez que $\varphi'(p) : \mathbb{R}^n \to T_xM$ é um isomorfismo, existe um único $u \in \mathbb{R}^n$ satisfazendo $\varphi'(p)u = v$. Tomando-se, então, $\epsilon > 0$ suficientemente pequeno, tem-se, para todo $t \in (-\epsilon, \epsilon)$, $p + tu \in U$. Logo, a curva $\gamma(t) = \varphi(p + tu)$ está bem definida, é diferenciável e cumpre as igualdades $\gamma(0) = \varphi(p) = x$ e $\gamma'(0) = \varphi'(p)u = v$, donde v é tangente a M em x (Fig. 6).

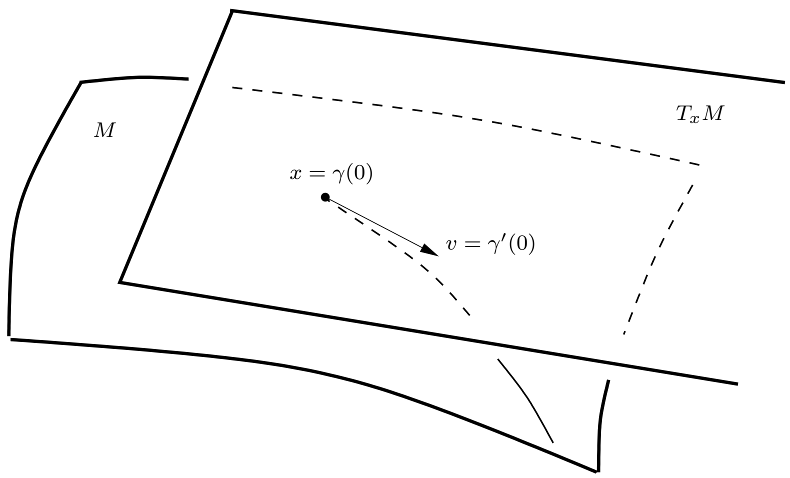

FIGURA 6

Reconsideremos os exemplos de variedades diferenciáveis dados acima e vejamos como seus espaços tangentes podem ser caracterizados.

No caso do espaço $\mathbb{R}^n$, é imediato que, para todo $x \in \mathbb{R}^n$, $T_x\mathbb{R}^n = \mathbb{R}^n$. Agora, se $M = \mathbb{V} + a$ é um espaço afim de $\mathbb{R}^m$ e $T : \mathbb{R}^n \to \mathbb{V}$ é um isomorfismo linear, constatamos que, para todo $x \in M$, a aplicação $\varphi = T + a$ é uma parametrização de M em x. Fazendo-se, então, $p = \varphi^{-1}(x)$, tem-se $\varphi'(p) = T$, donde

$$T_xM = \varphi'(p)(\mathbb{R}^n) = T(\mathbb{R}^n) = \mathbb{V}.$$

Passemos à esfera $S^n \subset \mathbb{R}^{n+1}$. Dado $v \in T_xS^n$, tomemos uma curva diferenciável $\gamma : (-\epsilon, \epsilon) \to S^n$, tal que $\gamma(0) = x$ e $\gamma'(0) = v$. Em particular, para todo $t \in (-\epsilon, \epsilon)$, tem-se $\langle \gamma(t), \gamma(t) \rangle = 1$. Diferenciando-se, então, com respeito a t e fazendo-se $t = 0$, obtém-se $\langle v, x \rangle = 0$, isto é, $T_xS^n \subset \{x\}^\perp$. Logo, $T_xS^n = \{x\}^\perp$, pois $\dim T_xS^n = n = \dim\{x\}^\perp$.

Suponhamos, agora, que $M \subset \mathbb{R}^{n+m}$ seja um conjunto de nível de uma submersão $f : U \subset \mathbb{R}^{n+m} \to \mathbb{R}^m$. Dados, então, $x \in M$ e $v \in T_xM$, seja $\gamma : (-\epsilon, \epsilon) \to M$ uma curva diferenciável satisfazendo $\gamma(0) = x$ e $\gamma'(0) = v$. Nestas condições, uma vez que $f \circ \gamma$ é constante, devemos ter $0 = f'(\gamma(0))\gamma'(0) = f'(x)v$, isto é, T_xM coincide com o núcleo de $f'(x)$, pois as dimensões desses dois subespaços são iguais. No caso em que $M = O(n) = f^{-1}(\{I\})$, $f(A) = AA^*$, tem-se

$$T_AM = \{H \in \mathrm{M}(n); AH^* + HA^* = 0\}.$$

Em particular, T_IM é o subespaço de $M(n)$ formado pelas matrizes $H \in \mathrm{M}(n)$ que satisfazem $H + H^* = 0$, ditas *anti-simétricas*.

Por fim, segue-se diretamente das considerações dos exemplos 82 e 84 e da definição de espaço tangente que:

- Se M é o gráfico de uma aplicação diferenciável $f : U \subset \mathbb{R}^n \to \mathbb{R}^m$ e $x \in M$, então $T_xM = \operatorname{graf} f'(x) \subset \mathbb{R}^n \times \mathbb{R}^m$;
- dadas variedades diferenciáveis, $M \subset \mathbb{R}^q$ e $N \subset \mathbb{R}^k$, o espaço tangente a $M \times N$ num ponto $(x, y) \in M \times N$ é $T_xM \times T_yN \subset \mathbb{R}^{q+k}$.

0.2.2. *A Derivada de Aplicações entre Variedades.* Sejam $M \subset \mathbb{R}^q$ e $N \subset \mathbb{R}^k$ variedades diferenciáveis, $f : M \to N$ uma aplicação diferenciável e x um ponto de M. Dados, então, $v \in T_xM$ e extensões diferenciáveis de f a um aberto $O \subset \mathbb{R}^q$ que contém x, $F, G : O \to \mathbb{R}^k$, temos que

$$F'(x)v = G'(x)v \in T_{f(x)}N.$$

De fato, dada uma curva diferenciável $\gamma : (-\epsilon, \epsilon) \to O \cap M$ satisfazendo $\gamma(0) = x$ e $\gamma'(0) = v$, tem-se, para todo $t \in (-\epsilon, \epsilon)$, $F(\gamma(t)) = G(\gamma(t))$, donde, diferenciando-se com respeito a t e fazendo-se $t = 0$, obtém-se $F'(x)v = G'(x)v$. Além disso, a curva $\sigma(t) = f(\gamma(t)) = F(\gamma(t))$ tem traço contido em N, é diferenciável e satisfaz $\sigma(0) = f(x)$, donde $F'(x)v = \sigma'(0) \in T_{f(x)}N$.

Logo, a *derivada* de f *em* x,

$$\begin{array}{rccc} f'(x): & T_xM & \to & T_{f(x)}N \\ & v & \mapsto & F'(x)v, \end{array}$$

está bem definida, isto é, independe da extensão diferenciável F. Note que, neste caso, a derivada de f em x nada mais é que a restrição de $F'(x)$ a T_xM.

Vale salientar que a derivada de aplicações entre variedades, como introduzida acima, tem essencialmente as mesmas propriedades que a derivada de aplicações definidas em abertos de espaços euclidianos, sendo aquelas facilmente estabelecidas a partir destas. Incluem-se aí as propriedades operatórias, a Regra da Cadeia e o Teorema da Função Inversa.

Exemplo 85 (conformalidade da projeção estereográfica). Tomemos, uma vez mais, a projeção estereográfica $f : S^n - \{p\} \subset \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}^n$, $p = (0, 1)$, dada por

$$f(x, s) = \frac{x}{1 - s} \cdot$$

Façamos $\mu(x, s) = \frac{1}{1-s}$ e verifiquemos que f cumpre

$$\|f'(x, s)(v, t)\| = \mu(x, s)\|(v, t)\| \;\; \forall (x, s) \in S^n - \{p\}, \;\; (v, t) \in T_{(x,s)}S^n,$$

donde se conclui que a projeção estereográfica é um *difeomorfismo conforme* (vide Exercício 9 – Capítulo 1).

Temos que $\nabla\mu(x,s) = (0, \frac{\partial\mu}{\partial s}(x,s)) = (0, 1/(1-s)^2)$, donde

$$\mu'(x,s)(v,t) = \langle \nabla\mu(x,s), (v,t)\rangle = \frac{t}{(1-s)^2}\cdot$$

Logo, uma vez que $f(x,s) = \mu(x,s)P(x,s)$, $P(x,s) = x$, tem-se

$$f'(x,s)(v,t) = (\mu'(x,s)(v,t))P(x,s) + \mu(x,s)P'(x,s)(v,t) = \frac{tx}{(1-s)^2} + \frac{v}{1-s}\,,$$

isto é,

$$f'(x,s)(v,t) = \mu(t\mu x + v), \quad \mu = \mu(x,s).$$

Daí e das igualdades

$$\|x\|^2 + s^2 = 1 \quad \text{e} \quad \langle x, v\rangle + ts = 0,$$

que são válidas por termos $(x,s) \in S^n$ e $(v,t) \in T_{(x,s)}S^n = \{(x,s)\}^\perp$, conclui-se facilmente que

$$\|f'(x,s)(v,t)\| = \mu(x,s)\|(v,t)\|,$$

como desejado.

Em conclusão, aplicaremos o conceito de derivada de aplicações entre variedades para estabelecer um resultado clássico da Álgebra Linear.

Teorema Espectral. *Para todo operador auto-adjunto* $A : \mathbb{R}^n \to \mathbb{R}^n$, *existe uma base ortonormal de* $\mathbb{R}^n$ *formada por autovetores de* A.

Demonstração. O resultado é imediato para $n = 1$. Suponhamo-lo, então, verdadeiro para $n - 1 \geq 1$ e provemo-lo verdadeiro para n. Para tanto, consideremos a função

$$\begin{array}{rccc} f: & S^{n-1} & \to & \mathbb{R} \\ & x & \mapsto & \langle Ax, x\rangle \end{array}$$

e observemos que f é diferenciável e satisfaz

$$f'(x)v = \langle Av, x\rangle + \langle Ax, v\rangle = 2\langle Ax, v\rangle \quad \forall x \in S^{n-1},\ v \in T_xS^{n-1}. \tag{74}$$

Logo, $f'(x) = 0$ se, e somente se,

$$Ax \in (T_xS^{n-1})^\perp = \{x\}^{\perp\perp} = \{\mu x; \mu \in \mathbb{R}\},$$

isto é, se, e somente se, x for um autovetor de A. No caso afirmativo, tem-se, então, para algum $\mu \in \mathbb{R}$, $f(x) = \langle Ax, x\rangle = \langle \mu x, x\rangle = \mu$ e, portanto, $Ax = f(x)x$.

Agora, uma vez que a função f é contínua e S^{n-1} é compacto, existe um ponto $x_1 \in S^{n-1}$, no qual f assume seu valor máximo. Dados, então, $v \in T_{x_1}S^{n-1}$ e uma curva diferenciável $\gamma : (-\epsilon, \epsilon) \to S^{n-1}$ satisfazendo $\gamma(0) = x_1$, $\gamma'(0) = v$, tem-se que a função $g(t) = f(\gamma(t))$ assume um máximo em $t = 0$, donde

$$0 = g'(0) = f'(\gamma(0))\gamma'(0) = f'(x_1)v,$$

isto é, $f'(x_1) = 0$. Logo, x_1 é um autovetor de A e $Ax_1 = f(x_1)x_1$.

Seja $\{x_1, y_2, \dots, y_n\}$ uma base ortonormal de $\mathbb{R}^n$ cujo primeiro vetor é o autovetor x_1. Então, para todo $i \in \{2, \dots, n\}$, tem-se

$$0 = f(x_1)\langle x_1, y_i\rangle = \langle f(x_1)x_1, y_i\rangle = \langle Ax_1, y_i\rangle = \langle x_1, Ay_i\rangle.$$

Desta forma, o subespaço $\mathbb{V} = \{x_1\}^\perp \subset \mathbb{R}^n$, gerado pelos vetores $y_2, \dots, y_n$, satisfaz $A(\mathbb{V}) \subset \mathbb{V}$, donde o operador (auto-adjunto) $A|_{\mathbb{V}} : \mathbb{V} \to \mathbb{V}$ está bem definido. Uma vez que $\dim \mathbb{V} = n-1$, por hipótese de indução, existe uma base ortonormal de $\mathbb{V}$, $\{x_2, \dots, x_n\}$, formada por autovetores de $A|_{\mathbb{V}}$. Logo, $\{x_1, x_2, \dots, x_n\}$ é uma base ortonormal de $\mathbb{R}^n$ formada por autovetores de A, concluindo, assim, a demonstração. $\square$

APÊNDICE B

Soluções dos Exercícios

Capítulo 1

1. Fazendo-se $\mu = \|1\| > 0$, dado $x \in \mathbb{R}$, tem-se $\|x\| = \|x.1\| = |x|\,\|1\| = \mu|x|$.

2. Aplicando-se a desigualdade triangular $\|z+y\| \le \|z\| + \|y\|$ a $z = x - y$, obtém-se $\|x\| - \|y\| \le \|x-y\|$. Analogamente, $\|y\| - \|x\| \le \|y - x\|$. Logo, $|\,\|x\| - \|y\|\,| \le \|x-y\|$.

3. A desigualdade $\langle x,y\rangle(\|x\|+\|y\|) \le \|x\|\,\|y\|\,\|x+y\|$ é evidente quando $\langle x,y\rangle \le 0$. Suponhamos, então, que $\langle x,y\rangle > 0$. Considerando-se a igualdade $\|x+y\|^2 = \langle x+y, x+y\rangle$, bem como a bilinearidade do produto interno, obtém-se $\langle x,y\rangle^2(\|x\|+\|y\|)^2 - \|x\|^2\|y\|^2\|x+y\|^2 = (\langle x,y\rangle^2 - \|x\|^2\|y\|^2)(\|x\|^2+\|y\|^2) + 2\|x\|\,\|y\|\langle x,y\rangle(\langle x,y\rangle - \|x\|\,\|y\|)$, donde, pela desigualdade de Cauchy-Schwarz, tem-se que o segundo membro desta equação é menor que, ou igual a, zero. Logo, $\langle x,y\rangle^2(\|x\|+\|y\|)^2 \le \|x\|^2\|y\|^2\|x+y\|^2$. Tomando-se, então, a raiz quadrada de ambos os membros desta última desigualdade, obtém-se o desejado.

   Considerando-se um vetor não-nulo $x \in \mathbb{R}^n$ e fazendo-se $y = -x$, tem-se $|\langle x,y\rangle|(\|x\|+\|y\|) = 2\|x\|^3 > 0 = \|x\|\,\|y\|\,\|x+y\|$, donde se conclui que, na desigualdade dada, não se pode substituir $\langle x,y\rangle$ por $|\langle x,y\rangle|$.

4. Eleve ambos os membros da igualdade $\|x+y\| = \|x\|+\|y\|$ ao quadrado e use o resultado do Teorema 1. Fazendo-se $x = e_1$ e $y = e_2$, temos que $\|e_1+e_2\|_s = \|e_1\|_s + \|e_2\|_s$. No entanto, e_1 e e_2 não são múltiplos um do outro. Para a norma do máximo, considere $x = (1,-1,0,\dots,0)$ e $y = (1,1,0,\dots,0)$.

5. A identidade $y = (1-t)x + tz$ é satisfeita para $t = 0$, se $x = y$, e para para $t = 1$, se $y = z$. Logo, podemos supor que $v = x-y$ e $w = y-z$ são não-nulos. Da hipótese, segue-se que $\|v+w\| = \|v\|+\|w\|$. Logo, pelo exercício anterior, existe $\lambda \ne 0$, tal que $w = \lambda v$. Desta forma, $\|(1+\lambda)v\| = (1+|\lambda|)\|v\|$, isto é, $|1+\lambda| = 1+|\lambda|$. Segue-se que $1+\lambda \ne 0$ e que (eleve ambos os membros ao quadrado) $\lambda = |\lambda| > 0$. Usando-se, uma vez mais, a igualdade $w = \lambda v$ e as definições de v e w, obtém-se $y = \frac{\lambda}{1+\lambda}x + \frac{1}{1+\lambda}z$. Tomando-se agora $t = \frac{1}{1+\lambda}$ conclui-se o desejado.

   A interpretação geométrica é: Se $x,y,z \in \mathbb{R}^n$ são como no enunciado, então (os pontos) x, y e z são colineares e, quando dois a dois distintos, y está entre x e z.

6. A identidade do paralelogramo segue-se diretamente das igualdades $\|x+y\|^2 = \langle x+y, x+y\rangle$ e $\|x-y\|^2 = \langle x-y, x-y\rangle$, bem como da bilinearidade do produto interno.

Em $\mathbb{R}^n$, a identidade do paralelogramo não é satisfeita pela norma do máximo ou da soma quando se tomam $x = e_1$ e $y = e_2$. Logo, nenhuma destas normas provém de um produto interno.

7. Dado $x \in \mathbb{R}^n$, escrevendo-se $x = x_1e_1 + \cdots + x_ne_n$, temos, pela linearidade de T, que $Tx = x_1Te_1 + \cdots + x_nTe_n = \langle x, a\rangle$, em que $a = (Te_1, \dots, Te_n)$. Agora, se $b \in \mathbb{R}^n$ satisfaz $Tx = \langle x, b\rangle \, \forall x \in \mathbb{R}^n$, tem-se $\langle x, a\rangle = \langle x, b\rangle \, \forall x \in \mathbb{R}^n$, donde $a = b$.

8. Designemos, por abuso de notação, as respectivas bases canônicas de $\mathbb{R}^n$ e $\mathbb{R}^m$ por $\{e_1, \dots, e_n\}$ e $\{e_1, \dots, e_m\}$. Fazendo-se, então, $x = (x_1, \dots, x_n)$ e $y = (y_1, \dots, y_m)$, segue-se da bilinearidade de f que
$$f(x, y) = \sum_{i,j=1}^{n,m} x_iy_jf(e_i, e_j).$$
Logo, escrevendo-se $\mu_0 = \max\{\|f(e_i, e_j)\|;\ 1 \leq i \leq n,\ 1 \leq j \leq m\}$, tem-se
$$\|f(x, y)\| \leq \sum_{i,j=1}^{n,m} |x_i|\,|y_j|\,\|f(e_i, e_j)\| \leq mn\mu_0\|x\|_{\max}\,\|y\|_{\max} \leq \mu\|x\|\,\|y\|,$$
em que $\mu = \mu_0mn$.

Agora, seja $f : \mathbb{R}^{j_1} \times \cdots \times \mathbb{R}^{j_n} \to \mathbb{R}^m$ uma aplicação n-linear. Dado $x_i \in \mathbb{R}^{j_i}$, $i \in \{1, \dots, n\}$, escrevamos
$$x_i = (x_{i1}, x_{i2}, \dots, x_{ij_i}) = \sum_{k_i=1}^{j_i} x_{ik_i}e_{k_i}.$$
Temos, então,
$$\begin{aligned} f(x_1, \dots, x_n) &= f\left(\sum_{k_1=1}^{j_1} x_{1k_1}e_{k_1}, \dots, \sum_{k_n=1}^{j_n} x_{nk_n}e_{k_n}\right) \\ &= \sum_{k_1,\dots,k_n=1}^{j_1,\dots,j_n} x_{1k_1} \dots x_{nk_n} f(e_{k_1}, \dots, e_{k_n}). \end{aligned}$$
Fazendo-se $\mu_0 = \max\{\|f(e_{k_1}, \dots, e_{k_n})\|;\ 1 \leq k_i \leq j_i\,,\ 1 \leq i \leq n\}$ e procedendo-se como no caso $n = 2$, obtém-se facilmente uma constante $\mu > 0$, tal que $\|f(x_1, \dots, x_n)\| \leq \mu\|x_1\| \dots \|x_n\|$.

9. (i) $\Rightarrow$ (ii) Dados vetores não-nulos e ortogonais, $x, y \in \mathbb{R}^n$, tem-se que Tx e Ty são não-nulos (pois T é um isomorfismo) e ortogonais (pois, por hipótese, T preserva ângulo). Suponhamos, então, que x, y sejam vetores não-nulos e que não sejam ortogonais. Neste caso, existem um único real não-nulo λ e um único vetor $z \in \{x\}^\perp$, tais que $y = \lambda x + z$. Logo,
$$\lambda\frac{\|x\|}{\|y\|} = \frac{\langle x, y\rangle}{\|x\|\|y\|} = \frac{\langle Tx, Ty\rangle}{\|Tx\|\|Ty\|} = \frac{\langle Tx, T(\lambda x + z)\rangle}{\|Tx\|\|Ty\|} = \lambda\frac{\langle Tx, Tx\rangle}{\|Tx\|\|Ty\|} = \lambda\frac{\|Tx\|}{\|Ty\|},$$
donde $\|Tx\|/\|x\| = \|Ty\|/\|y\|$, isto é, a função $x \mapsto \|Tx\|/\|x\|$ é constante em $\mathbb{R}^n - \{0\}$. Desta forma, existe $\mu > 0$, tal que, para todo $x \neq 0$, $\|Tx\|/\|x\| = \mu$, donde $\|Tx\| = \mu\|x\| \, \forall x \in \mathbb{R}^n - \{0\}$. No entanto, esta última igualdade é trivialmente verdadeira para $x = 0$ e vale, portanto, para todo $x \in \mathbb{R}^n$.

(ii) $\Rightarrow$ (iii) Para quaisquer $x, y \in \mathbb{R}^n$, tem-se, por hipótese, que $\|T(x+y)\|^2 = \mu^2\|x+y\|^2$, isto é, $\langle Tx+Ty, Tx+Ty\rangle = \mu^2\langle x+y, x+y\rangle$. Daí, da bilinearidade do produto interno e das igualdades $\|Tx\|^2 = \mu^2\|x\|^2$ e $\|Ty\|^2 = \mu^2\|y\|^2$ segue-se que $\langle Tx, Ty\rangle = \lambda\langle x, y\rangle$, em que $\lambda = \mu^2$.

(iii) $\Rightarrow$ (i) Dados vetores não-nulos $x, y \in \mathbb{R}^n$, tem-se

$$\|Tx\|\,\|Ty\| = \sqrt{\langle Tx, Tx\rangle\langle Ty, Ty\rangle} = \sqrt{\lambda^2\langle x, x\rangle\langle y, y\rangle} = \lambda\|x\|\,\|y\|.$$

Logo, $\measuredangle(Tx, Ty) = \frac{\langle Tx, Ty\rangle}{\|Tx\|\,\|Ty\|} = \frac{\lambda\langle x, y\rangle}{\lambda\|x\|\,\|y\|} = \frac{\langle x, y\rangle}{\|x\|\,\|y\|} = \measuredangle(x, y)$.

10. Temos que as matrizes de T e T^* são, respectivamente,

$$A = \begin{pmatrix} \sqrt{5} & -2 \\ 4 & 2\sqrt{5} \end{pmatrix} \qquad \text{e} \qquad A^* = \begin{pmatrix} \sqrt{5} & 4 \\ -2 & 2\sqrt{5} \end{pmatrix}.$$

Consequentemente, a matriz de TT^* é $AA^* = \begin{pmatrix} 9 & 0 \\ 0 & 36 \end{pmatrix}$, cujos autovalores são 9 e 36. Segue-se que $\|T\| = \sqrt{36} = 6$.

11. O operador Z, por ser ortogonal, é um isomorfismo de $\mathbb{R}^n$ que preserva norma. Assim, para todo $x \in \mathbb{R}^n$, existe um único $y \in \mathbb{R}^n$, tal que $Zx = y$ e reciprocamente. Além disso, $\|x\| = 1$ se, e somente se, $\|y\| = 1$. Logo, $\{Tx \in \mathbb{R}^n;\, \|x\| = 1\} = \{(TZ)y \in \mathbb{R}^n;\, \|y\| = 1\}$, donde $\|TZ\| = \|T\|$. Da ortogonalidade de Z, segue-se também a igualdade $\{\|Tx\| \in \mathbb{R}^n;\, \|x\| = 1\} = \{\|(ZT)x\| \in \mathbb{R}^n;\, \|x\| = 1\}$, o que nos dá $\|ZT\| = \|T\|$.

12. Seja $C = \{\mu \in \mathbb{R}\,;\, \|Tx\| \le \mu\|x\|\, \forall x \in \mathbb{R}^n\}$. Dado $x \in \mathbb{R}^n$, temos que $\|Tx\| \le \|T\|\,\|x\|$, isto é, $\|T\| \in C$. Por outro lado, dado $\mu \in C$, para todo vetor unitário $x \in \mathbb{R}^n$, tem-se $\mu \ge \|Tx\|$. Logo, μ é uma cota superior do conjunto $\{\|Tx\|;\, \|x\| = 1\}$, o que nos dá, $\mu \ge \|T\|$, ou seja, $\|T\|$ é uma cota inferior de C. Desta forma, $\|T\| = \inf C$ (pois $\|T\| \in C$).

13. Suponhamos, por absurdo, que T não seja invertível. Neste caso, o núcleo de T é não-trivial, isto é, existe um vetor unitário $u \in \mathbb{R}^n$, tal que $Tu = 0$, donde $\|(T-I)u\| = \|-u\| = 1 > \|T - I\| = \sup\{\|(T-I)x\|\,;\, \|x\| = 1\}$, o que, claramente, é uma contradição. Logo, T é invertível.

14. Da desigualdade $\|Tx\| \ge \mu\|x\|\, \forall x \in \mathbb{R}^n$, segue-se que o núcleo de T é trivial e, portanto, que T é invertível. Agora, dado um vetor unitário $y \in \mathbb{R}^n$, seja $x = T^{-1}y$. Então, $\|T^{-1}y\| = \|x\| \le \frac{1}{\mu}\|Tx\| = \frac{1}{\mu}\|y\| = \frac{1}{\mu}\cdot$ Logo, $\|T^{-1}\| \le \frac{1}{\mu}\cdot$

15. Seja $\lambda \ne 0$ um autovalor de ZT. Então, existe $v \in \mathbb{R}^n - \{0\}$, tal que $(ZT)v = \lambda v$, donde $Tv \ne 0$ e $T(ZT)v = \lambda Tv$, isto é, Tv é um autovetor de TZ associado a λ. Segue-se que todos os autovalores não-nulos de ZT são autovalores de TZ. Analogamente, verifica-se que os autovalores não-nulos de TZ são também autovalores de ZT.

Agora, se todos os autovalores do operador (auto-adjunto) T^*T são nulos, então $T^*T = 0$, donde $\|T\| = \sqrt{\|T^*T\|} = 0$, isto é, $T = T^* = 0$ e, portanto, $\|T\| = \|T^*\|$. Caso contrário, pelas considerações acima, tem-se $\|T^*T\| = \|TT^*\|$ e, então, $\|T\| = \sqrt{\|T^*T\|} = \sqrt{\|TT^*\|} = \|T^*\|$.

16. (i) Dado $a = (a_1, \dots, a_n) \in \mathbb{R}^n$, uma vez que $\mathbb{Q}$ é denso em $\mathbb{R}$, para cada $k \in \mathbb{N}$, existem $x_{k1}, \dots, x_{kn} \in \mathbb{Q}$, tais que $x_{ki} \in (a_i - 1/k, a_i + 1/k)$. Assim, todos os termos da sequência (x_k), em que $x_k = (x_{k1}, \dots, x_{kn})$, são pontos de $\mathbb{Q}^n$. Além disso, $\|x_k - a\|_{\max} = \max\{|x_{k1} - a_1|, \dots, |x_{kn} - a_n|\} < 1/k$ para todo $k \in \mathbb{N}$, donde $\|x_k - a\|_{\max} \to 0$ e, portanto, $x_k \to a$.

(ii) Uma vez que X é denso em $\mathbb{R}^n$, existe uma sequência $(y_i)_{i \in \mathbb{N}}$ em X, tal que $y_i \to a$. Além disso, por hipótese, para cada $i \in \mathbb{N}$, $\lim_{k\to\infty} \|x_k - y_i\| = \|a - y_i\|$. Dado, então, $\epsilon > 0$, existem $i_0, k_0 \in \mathbb{N}$, tais que $\|y_i - a\| < \epsilon \ \forall i \geq i_0$ e $|\,\|x_k - y_{i_0}\| - \|a - y_{i_0}\|\,| < \epsilon \ \forall k \geq k_0$. Dessa última desigualdade tem-se, em particular, $\|x_k - y_{i_0}\| < \epsilon + \|a - y_{i_0}\| \ \forall k \geq k_0$. Logo, $\|x_k - a\| \leq \|x_k - y_{i_0}\| + \|y_{i_0} - a\| < \epsilon + \|a - y_{i_0}\| + \|y_{i_0} - a\| < 3\epsilon$, donde $x_k \to a$.

17. Segue-se das considerações do Exemplo 6 que, para algum $a \in \mathbb{R}^n$, $T(x_k) \to Ta$, já que o conjunto-imagem de T é um subespaço de $\mathbb{R}^m$ e a sequência (Tx_k), por hipótese, é convergente. Além disso, a aplicação $Z : \mathbb{R}^n \to T(\mathbb{R}^n)$, em que $Zx = Tx$, é um isomorfismo linear, pois T, também por hipótese, é injetiva. Logo, pela Proposição 8, $Z^{-1}(Tx_k) \to Z^{-1}(Ta)$, isto é, $x_k \to a$.

18. Considerando-se as propriedades da transformação adjunta, o Exercício 15 e a convergência $T_k \to T$, tem-se $\|T_k^* - T^*\| = \|(T_k - T)^*\| = \|T_k - T\| \to 0$, donde $T_k^* \to T^*$. Agora, se (T_k) é uma sequência de operadores ortogonais em $\mathrm{L}(\mathbb{R}^n)$, então $\|T_k\| = 1 \, \forall k \in \mathbb{N}$. Em particular, (T_k) é limitada. Logo, pelo Teorema de Bolzano-Weierstrass, existe uma subsequência (T_{k_i}), de (T_k), que converge para um operador $T \in \mathrm{L}(\mathbb{R}^n)$, donde $T_{k_i}^* \to T^*$. Daí e do resultado do Exemplo 8, obtém-se $TT^* = \lim T_{k_i} \lim T_{k_i}^* = \lim(T_{k_i} T_{k_i}^*) = \lim I = I$, implicando que T é ortogonal.

19. Uma vez que (A_k^{-1}) é limitada, pelo Teorema de Bolzano-Weierstrass, esta sequência possui uma subsequência convergente $(A_{k_i}^{-1})$. Supondo-se que B seja o seu limite, teremos, pelo Exemplo 8, $I = A_{k_i} A_{k_i}^{-1} \to AB$, donde A é invertível e $B = A^{-1}$.

20. Dada uma matriz $A = [x_1 \; x_2 \; \cdots \; x_{n-1} \; x_n] \in \mathrm{M}(n)$ (isto é, $x_1, \dots, x_n$ são os vetores-coluna de A), suponhamos que $\det A \neq 0$. Para cada $k \in \mathbb{N}$, façamos $\lambda_k = (1 + 1/k)$ e consideremos a sequência (A_k), em $\mathrm{M}(n)$, dada por $A_k = [x_1 \; x_2 \; \cdots \; x_{n-1} \; \lambda_k x_n]$. É imediato que $A_k \to A$ e que, para cada $k \in \mathbb{N}$, $A_k \neq A$. Além disso, tem-se $\det A_k = \lambda_k \det[x_1 \; x_2 \; \cdots \; x_n] = \lambda_k \det A \neq 0$, donde cada matriz A_k é invertível. Suponhamos agora que $\det A = 0$. Se $A = 0$, basta fazermos $A_k = \frac{1}{k} I$, em que I é a matriz identidade de $\mathbb{R}^n$. Caso contrário, denotando-se por $\mathbb{W} \subset \mathbb{R}^n$ o subespaço gerado pelos vetores-coluna de A, $x_1, \dots, x_n$, tem-se $0 < \dim \mathbb{W} < n$. Suponhamos, então, sem perda de generalidade, que $\{x_1, \dots, x_m\}$ seja uma base de $\mathbb{W}$. Tomando-se vetores $y_{m+1}, \dots, y_n$, tais que $\{x_1, \dots, x_m, y_{m+1}, \dots, y_n\}$ seja uma base de $\mathbb{R}^n$, definimos, para cada $k \in \mathbb{N}$, a matriz $A_k = \left[x_1 \quad x_2 \quad \cdots \quad x_m \quad x_{m+1} + \frac{1}{k} y_{m+1} \quad \cdots \quad x_n + \frac{1}{k} y_n\right]$. Como anteriormente, temos que $A_k \to A$ e que, para cada $k \in \mathbb{N}$, $A_k \neq A$. Uma vez que os vetores $x_1, \dots, x_m, \dots x_n$ são linearmente dependentes e os vetores $x_1, \dots, x_m, \dots, y_{m+1}, \dots, y_n$ são linearmente independentes, segue-se da n-linearidade do determinante que $\det A_k = \frac{1}{k^{n-m}} \det[x_1 \; \dots \; x_m \; y_{m+1} \; \dots \; y_n]$. Logo, cada matriz A_k tem determinante não-nulo e é, portanto, invertível.

21. (i) Dado $v \in \mathbb{V} \cap \mathbb{W}$, temos que $P_{\mathbb{V}} v = P_{\mathbb{W}} v = v$, donde $(P_{\mathbb{V}} P_{\mathbb{W}} P_{\mathbb{V}})v = v$, isto é, $\mathbb{V} \cap \mathbb{W} \subset \mathbb{V}_0 = \{v \in \mathbb{R}^n; (P_{\mathbb{V}} P_{\mathbb{W}} P_{\mathbb{V}})v = v\}$. Agora, dado $v \in \mathbb{V}_0$, tem-se $v = (P_{\mathbb{V}} P_{\mathbb{W}} P_{\mathbb{V}})v = P_{\mathbb{V}}(P_{\mathbb{W}} P_{\mathbb{V}} v)$, donde $\mathbb{V}_0 \subset P_{\mathbb{V}}(\mathbb{R}^n) = \mathbb{V}$. Então, para todo $v \in \mathbb{V}_0$, tem-se $(P_{\mathbb{V}} P_{\mathbb{W}})v = v$ e, portanto, $P_{\mathbb{V}}(P_{\mathbb{W}} v - v) = 0$. Segue-se que $P_{\mathbb{W}} v - v \in \mathbb{V}^{\perp}$. Porém, $P_{\mathbb{W}}(P_{\mathbb{W}} v - v) = 0$, o que implica $P_{\mathbb{W}} v - v \in \mathbb{W}^{\perp}$. Assim, uma vez que $v \in \mathbb{V}_0 \subset \mathbb{V}$ e $P_{\mathbb{W}} v \in \mathbb{W}$, tem-se $\langle v - P_{\mathbb{W}} v, v\rangle = 0 = \langle P_{\mathbb{W}} v - v, P_{\mathbb{W}} v\rangle$, donde $\|P_{\mathbb{W}} v - v\|^2 = 0$. Segue-se que $P_{\mathbb{W}} v = v$ e, portanto, que $v \in \mathbb{W}$, isto é, $\mathbb{V}_0 \subset \mathbb{W}$. Desta forma, $\mathbb{V}_0 \subset \mathbb{V} \cap \mathbb{W}$, donde $\mathbb{V}_0 = \mathbb{V} \cap \mathbb{W}$.

    (ii) Escrevamos $T = P_{\mathbb{V}} P_{\mathbb{W}} P_{\mathbb{V}}$ e observemos que T é um operador auto-adjunto, pois $T^* = (P_{\mathbb{V}} P_{\mathbb{W}} P_{\mathbb{V}})^* = P_{\mathbb{V}}^* P_{\mathbb{W}}^* P_{\mathbb{V}}^* = P_{\mathbb{V}} P_{\mathbb{W}} P_{\mathbb{V}} = T$. Se λ é um autovalor de T associado a um autovetor unitário $u \in \mathbb{R}^n$, tem-se $\lambda = \langle Tu, u\rangle = \langle (P_{\mathbb{V}} P_{\mathbb{W}} P_{\mathbb{V}})u, u\rangle = \langle P_{\mathbb{W}}(P_{\mathbb{V}} u), P_{\mathbb{V}} u\rangle \geq 0$ (vide Exemplo 2). Além disso, $\|T\| = \|P_{\mathbb{V}} P_{\mathbb{W}} P_{\mathbb{V}}\| \leq \|P_{\mathbb{V}}\|^2 \|P_{\mathbb{W}}\| = 1$. Logo, existe uma base ortonormal de $\mathbb{R}^n$ formada por autovetores de T, $\mathfrak{B} = \{u_1, \ldots, u_n\}$, cujos respectivos autovalores, $\lambda_1, \ldots, \lambda_n$, satisfazem $0 \leq \lambda_i \leq 1 \, \forall i = 1, \ldots, n$. Temos, pelo item (i), que $\lambda_i = 1$ se, e somente se, $u_i \in \mathbb{V} \cap \mathbb{W}$. Assim, uma vez que $T^k(u_i) = \lambda_i^k u_i$, tem-se $T^k(u_i) \to u_i$ se $u_i \in \mathbb{V} \cap \mathbb{W}$. Caso contrário, tem-se $0 \leq \lambda_i < 1$ e, então, $T^k(u_i) \to 0$. Tomando-se $v = x_1 u_1 + \cdots + x_n u_n \in \mathbb{R}^n$, tem-se $T^k(v) = x_1 T^k(u_1) + \cdots + x_n T^k(u_n)$. Se $\mathbb{V} \cap \mathbb{W} = \{0\}$, então todos os autovalores de T são menores que 1, donde, neste caso, tem-se $T^k(v) \to 0 = P_{\mathbb{V} \cap \mathbb{W}}(v)$. Agora se $0 < \dim(\mathbb{V} \cap \mathbb{W}) = m \leq n$, podemos supor, sem perda de generalidade, que $\{u_1, \ldots, u_m\}$ é uma base de $\mathbb{V} \cap \mathbb{W}$, donde $T^k(v) \to x_1 u_1 + \cdots + x_m u_m = P_{\mathbb{V} \cap \mathbb{W}}(v)$. Logo, pela Proposição 10, $(P_{\mathbb{V}} P_{\mathbb{W}} P_{\mathbb{V}})^k \to P_{\mathbb{V} \cap \mathbb{W}}$.

22. Seja $\| \ \|$ uma norma em $\mathbb{V}$ e $d : \mathbb{V} \times \mathbb{V} \to \mathbb{R}$ a métrica $d(x, y) = \|x - y\|$. Então, $d(x + a, y + a) = \|x + a - (y + a)\| = \|x - y\| = d(x, y)$ e $d(\lambda x, \lambda y) = \|\lambda x - \lambda y\| = |\lambda| d(x, y)$ $\forall x, y, a \in \mathbb{R}^n$ e $\lambda \in \mathbb{R}$. Reciprocamente, se d cumpre as condições dadas, definimos $\|x\| = d(x, 0)$, $x \in \mathbb{V}$. É imediato que $\|x\| \geq 0$ e $\|x\| = 0 \Leftrightarrow x = 0$. Além disso, dado $\lambda \in \mathbb{R}$, tem-se $\|\lambda x\| = d(\lambda x, 0) = d(\lambda x, \lambda 0) = |\lambda| d(x, 0) = |\lambda| \|x\|$. Finalmente, dados $x, y \in \mathbb{V}$, vale $\|x + y\| = d(x + y, 0) = d(x + y, -y + y) = d(x, -y) \leq d(x, 0) + d(-y, 0) = \|x\| + \| - y\| = \|x\| + \|y\|$. Logo, $\| \ \|$ define uma norma em $\mathbb{V}$. Ademais, $\|x - y\| = d(x - y, 0) = d(x - y, y - y) = d(x, y)$, isto é, d é proveniente da norma $\| \ \|$.

    Dados $x, y, z \in \mathbb{R}^n$, temos que $d'(x, y) = \sqrt{d(x, y)}$. Logo, $d'(x, y) \geq 0$ e $d'(x, y) = 0$ se, e somente se, $x = y$. É fácil ver também que d' é simétrica. Além disso, $d'(x, y) = \sqrt{d(x, y)} \leq \sqrt{d(x, z) + d(z, y)} \leq \sqrt{d(x, z)} + \sqrt{d(z, y)} = d'(x, z) + d'(z, y)$. Logo, d' é uma métrica em $\mathbb{R}^n$. No entanto, $d'(\lambda x, \lambda y) = \sqrt{\|\lambda x - \lambda y\|} = \sqrt{|\lambda|} d'(x, y)$, isto é, $d'(\lambda x, \lambda y) \neq |\lambda| d'(x, y)$ para todo λ não-nulo cujo valor absoluto é diferente de 1. Desta forma, d' não provém de um produto interno de $\mathbb{R}^n$.

23. Fazendo-se, para cada $k \in \mathbb{N}$, $x_k = 1/k$, temos que (x_k), claramente, é uma sequência de Cauchy em (M, d). Porém, (x_k) é convergente em $(\mathbb{R}, d)$ e seu limite é $0 \notin M$. Logo, (x_k) é divergente em (M, d), donde (M, d) não é um espaço métrico completo.

    Observando-se que, para quaisquer $x, y \in M$, $d'(x, y) = d(1/x, 1/y)$, conclui-se facilmente que a função d' é não-negativa, simétrica e satisfaz $d'(x, y) = 0 \Leftrightarrow$

$x = y$. Além disso, para quaisquer $x, y, z \in M$, tem-se $d'(x,z) = d(1/x, 1/z) \leq d(1/x, 1/y) + d(1/y, 1/z) = d'(x,y) + d'(y,z)$, donde d' satisfaz a desigualdade triangular. Segue-se que d' é uma métrica em M. A fim de verificar que (M, d') é completo, tomemos uma sequência de Cauchy, (x_k), em (M, d'). Uma vez que, para quaisquer $x, y \in M$, tem-se $d(x,y) = xyd'(x,y) \leq d'(x,y) = d(1/x, 1/y)$, conclui-se que (x_k) e $(1/x_k)$ são sequências de Cauchy em $(\mathbb{R}, d)$. Sendo este completo, ambas são, então, convergentes em $(\mathbb{R}, d)$. Em particular, o limite de (x_k) em $(\mathbb{R}, d)$ é um real a não-nulo (donde $a \in (0,1] = M$) e o limite de $(1/x_k)$ em $(\mathbb{R}, d)$ é $1/a$. Logo, $d'(x_k, a) = d(1/x_k, 1/a) \to 0$, donde se infere que $x_k \to a$ em (M, d') e, portanto, que (M, d') é completo.

24. Temos que P é uma isometria se, e somente se, para quaisquer $t_1, t_2 \in \mathbb{R}$, $d(P(t_1, at_1), P(t_2, at_2)) = d_{\max}((t_1, at_1), (t_2, at_2))$. Porém, esta igualdade é válida se, e somente se, $d(at_1, at_2) = \|(t_1 - t_2, a(t_1 - t_2)\|_{\max}$, isto é, $|a|\,|t_1 - t_2| = \max\{|t_1 - t_2|, |a|\,|t_1 - t_2|\}$, o que ocorre se, e somente se, $|a| \geq 1$.

25. Dados $x, y \in S^n$, $x \neq y$, sejam $\{x_2, \dots, x_n\}$ e $\{y_2, \dots, y_n\}$ bases ortonormais de $\{x\}^\perp$ e $\{y\}^\perp$, respectivamente. Então, $\{x, x_2, \dots, x_n\}$ e $\{y, y_2, \dots, y_n\}$ são bases ortonormais de $\mathbb{R}^n$. Existe, portanto, uma única transformação linear $T \in \mathrm{L}(\mathbb{R}^n)$, tal que $T(x) = y$ e $T(x_i) = y_i \, \forall i \in \{2, \dots, n\}$. Além disso, T é ortogonal, pois tem uma base ortonormal como imagem de uma base ortonormal. Em particular, T preserva norma. Logo, a aplicação $\varphi = T|_S : S \to S$ está bem definida, leva x em y e preserva distância, pois T é uma isometria de $\mathbb{R}^n$. Agora, uma vez que $T^{-1} \in \mathrm{L}(\mathbb{R}^n)$ é também ortogonal, dado $b \in S$, temos que $a = T^{-1}(b) \in S$, donde $\varphi(a) = b$. Segue-se que a aplicação φ é sobrejetiva e, portanto, é uma isometria de (S, d).

Capítulo 2

1. Seja $a \in \operatorname{int}(X)$. Então, existe $r > 0$ satisfazendo $B(a,r) \subset X$. A bola $B(a,r)$, porém, é aberta. Sendo assim, dado $x \in B(a,r)$, existe $\delta > 0$, tal que $B(x,\delta) \subset B(a,r) \subset X$. Logo, $x \in \operatorname{int}(X)$. Segue-se que $B(a,r) \subset \operatorname{int}(X)$ e, portanto, que $\operatorname{int}(X)$ é aberto. Agora, se $A \subset X$ é um conjunto aberto, dado $a \in A$, existe $r > 0$, tal que $B(a,r) \subset A \subset X$, donde $a \in \operatorname{int} X$, isto é, $A \subset \operatorname{int} X$.

2. (i) Seja $a \in \operatorname{int}(X \cap Y)$. Então, existe $r > 0$, tal que $B(a,r) \subset X \cap Y$, donde $B(a,r) \subset X$ e $B(a,r) \subset Y$. Logo, $a \in \operatorname{int}(X) \cap \operatorname{int}(Y)$ e, portanto, $\operatorname{int}(X \cap Y) \subset \operatorname{int}(X) \cap \operatorname{int}(Y)$. Agora, se $a \in \operatorname{int}(X) \cap \operatorname{int}(Y)$, existem $r_1, r_2 > 0$, tais que $B(a, r_1) \subset X$ e $B(a, r_2) \subset Y$. Logo, tomando-se $r = \min\{r_1, r_2\}$, tem-se $B(a,r) \subset X \cap Y$, isto é, $a \in \operatorname{int}(X \cap Y)$. Segue-se que $\operatorname{int}(X) \cap \operatorname{int}(Y) \subset \operatorname{int}(X \cap Y)$, ou seja, $\operatorname{int}(X) \cap \operatorname{int}(Y) = \operatorname{int}(X \cap Y)$.

   (ii) Tem-se, $a \in \operatorname{int}(X) \cup \operatorname{int}(Y) \Rightarrow \exists\, r > 0$, tal que $B(a,r) \subset X$ ou $B(a,r) \subset Y \Rightarrow B(a,r) \subset X \cup Y \Rightarrow a \in \operatorname{int}(X \cup Y)$. Desta forma, $\operatorname{int}(X) \cup \operatorname{int}(Y) \subset \operatorname{int}(X \cup Y)$.

   (iii) Segue-se da Proposição 12, item (ii), e do item (i), acima, que $\overline{X \cup Y} = \mathbb{R}^n - \operatorname{int}(\mathbb{R}^n - (X \cup Y)) = \mathbb{R}^n - \operatorname{int}((\mathbb{R}^n - X) \cap (\mathbb{R}^n - Y)) = \mathbb{R}^n - (\operatorname{int}(\mathbb{R}^n - X) \cap \operatorname{int}(\mathbb{R}^n - Y)) = (\mathbb{R}^n - \operatorname{int}(\mathbb{R}^n - X)) \cup (\mathbb{R}^n - \operatorname{int}(\mathbb{R}^n - Y)) = \overline{X} \cup \overline{Y}$.

(iv) Novamente pela Proposição 12, item (ii), e pelo item (ii), acima, tem-se $\overline{X \cap Y} = \mathbb{R}^n - \operatorname{int}(\mathbb{R}^n - (X \cap Y)) = \mathbb{R}^n - \operatorname{int}((\mathbb{R}^n - X) \cup (\mathbb{R}^n - Y)) \subset \mathbb{R}^n - (\operatorname{int}(\mathbb{R}^n - X) \cup \operatorname{int}(\mathbb{R}^n - Y)) = (\mathbb{R}^n - \operatorname{int}(\mathbb{R}^n - X)) \cap (\mathbb{R}^n - \operatorname{int}(\mathbb{R}^n - Y)) = \overline{X} \cap \overline{Y}$.

Fazendo-se $X = [0,1)$ e $Y = [1,2]$, obtém-se $\operatorname{int}(X) \cup \operatorname{int}(Y) = (0,1) \cup (1,2) \subsetneq (0,2) = \operatorname{int}(X \cup Y)$. Além disso, $\overline{X \cap Y} = \emptyset \subsetneq \{1\} = \overline{X} \cap \overline{Y}$.

3. Suponha que $A \subset \mathbb{R}^n$ seja aberto e tome $a \in A \cap \overline{X}$ (se $A \cap \overline{X} = \emptyset$ não há nada a ser feito). Existem, então, uma sequência (x_k) em X, tal que $x_k \to a$, e $\epsilon > 0$, tal que $B(a, \epsilon) \subset A$. No entanto, para este ϵ, existe $k_0 \in \mathbb{N}$, tal que $x_k \in B(a, \epsilon)\, \forall k \geq k_0$. Desta forma, a subsequência $(x_k)_{k \geq k_0}$ é uma sequência em $A \cap X$, o que nos dá $a \in \overline{A \cap X}$, isto é, $A \cap \overline{X} \subset \overline{A \cap X}$. Reciprocamente, suponhamos que, para todo $X \subset \mathbb{R}^n$, $A \cap \overline{X} \subset \overline{A \cap X}$. Fazendo-se, então, $X = \mathbb{R}^n - A$, tem-se $A \cap \overline{\mathbb{R}^n - A} \subset \overline{A \cap (\mathbb{R}^n - A)} = \emptyset$, o que implica $\overline{\mathbb{R}^n - A} \subset \mathbb{R}^n - A$. Segue-se que $\mathbb{R}^n - A$ é fechado e, portanto, que A é aberto.

4. Sejam $x_0 \in \overline{X}$ e (x_k) uma sequência em X, tais que $x_k \to x_0$. Então, para cada $k \in \mathbb{N}$, existe $a_k \in F$, tal que $\|x_k - a_k\| = r$. Porém, $\|a_k\| \leq \|x_k - a_k\| + \|x_k\| = r + \|x_k\|$, donde (a_k) é limitada, pois (x_k), sendo convergente, é limitada. Logo, existe uma subsequência (a_{k_i}), de (a_k), tal que $a_{k_i} \to a \in F$, pois F é fechado. Daí, tem-se $\|x_0 - a\| = \|\lim(x_{k_i} - a_{k_i})\| = \lim \|x_{k_i} - a_{k_i}\| = r$, donde $x_0 \in X$. Segue-se que $\overline{X} = X$ e, portanto, que X é fechado.

5. Seja $\mathbb{V} \subset \mathbb{R}^n$ um subespaço vetorial próprio de $\mathbb{R}^n$. Segue-se do Exemplo 6 do Capítulo 1 que $\overline{\mathbb{V}} = \mathbb{V}$, donde $\mathbb{V}$ é fechado. Pode-se verificar esta propriedade, também, da seguinte forma. Considere $x \in \mathbb{R}^n - \mathbb{V}$. Então, fazendo-se $r = \|x - P_{\mathbb{V}} x\|/2$, tem-se que a bola $B(x, r)$ é disjunta de $\mathbb{V}$ (vide desigualdade (5) – Capítulo 1), donde $B(x, r) \subset (\mathbb{R}^n - \mathbb{V})$. Sendo assim, $\mathbb{R}^n - \mathbb{V}$ é aberto e, portanto, $\mathbb{V}$ é fechado. Agora, dado $a \in \mathbb{V}$, para todo $\epsilon > 0$, existe $x \in B(a, \epsilon) - \mathbb{V}$. Basta fazer $x = a + h$, em que h é um vetor arbitrário de $\mathbb{V}^{\perp}$ que satisfaz $0 < \|h\| < \epsilon$. Logo, $\mathbb{V}$ tem interior vazio em $\mathbb{R}^n$.

6. (i) Consideremos $T \in \mathrm{L}(\mathbb{R}^n, \mathbb{R}^m)$ e $A \subset \mathbb{R}^m$ um conjunto aberto. Se T é identicamente nula, então $T^{-1}(A) = \mathbb{R}^n$ (o que ocorre se $0 \in A$) ou $T^{-1}(A) = \emptyset$ (o que ocorre se $0 \notin A$). Em qualquer dos casos, $T^{-1}(A)$ é aberto em $\mathbb{R}^n$. Suponhamos, então, que $T \neq 0$ e $T^{-1}(A) \neq \emptyset$. Dado $a \in T^{-1}(A)$, existe $r > 0$, tal que $B(Ta, r) \subset A$, pois A é aberto. Fazendo-se, então, $r_0 = r/\|T\|$ e tomando-se $x \in B(a, r_0)$, tem-se $\|Tx - Ta\| = \|T(x - a)\| \leq \|T\|\, \|x - a\| < \|T\| r_0 = r$, isto é, $Tx \in B(Ta, r) \subset A$. Logo, $x \in T^{-1}(A)$, donde $B(a, r_0) \subset T^{-1}(A)$ e, portanto, $T^{-1}(A)$ é aberto.

   (ii) Dado um conjunto fechado $F \subset \mathbb{R}^m$, temos que $\mathbb{R}^m - F$ é aberto em $\mathbb{R}^m$. Logo, pelo item (i), $T^{-1}(\mathbb{R}^m - F)$ é aberto em $\mathbb{R}^n$. Porém, $T^{-1}(\mathbb{R}^m - F) = T^{-1}(\mathbb{R}^m) - T^{-1}(F) = \mathbb{R}^n - T^{-1}(F)$, donde $T^{-1}(F)$ é fechado em $\mathbb{R}^n$.

7. (i) Dada uma transformação linear injetiva $T : \mathbb{R}^n \to \mathbb{R}^{n+m}$, tem-se que $T(\mathbb{R}^n)$ é um subespaço n-dimensional de $\mathbb{R}^{n+m}$. Desta forma, identificando-se o espaço $\mathbb{R}^{n+m}$ com $\mathbb{R}^n \times \mathbb{R}^m$, podemos supor, sem perda de generalidade, que $T(\mathbb{R}^n) = \mathbb{R}^n \times \{0\} \subset \mathbb{R}^n \times \mathbb{R}^m$. Feito isto, definamos $\Phi : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^n \times \mathbb{R}^m$

por $\Phi(x,y) = Tx+y$. Claramente, Φ é linear. Além disso, se para um dado par $(x,y) \in \mathbb{R}^n \times \mathbb{R}^m$ tem-se $\Phi(x,y) = 0$, então $Tx + y = 0$, o que nos dá, $Tx = 0$ (e, portanto, $x = 0$) e $y = 0$, pois $Tx \in \mathbb{R}^n \times \{0\}$ e $y \in \{0\} \times \mathbb{R}^m$. Logo, Φ é um isomorfismo. Tomando-se, então, $F \subset \mathbb{R}^n$ fechado, tem-se, obviamente, que $F \times \{0\}$ é um subconjunto fechado de $\mathbb{R}^n \times \mathbb{R}^m$. Daí, sendo Φ um isomorfismo (vide Exemplo 17), segue-se que $T(F) = \Phi(F \times \{0\})$ é um subconjunto fechado de $\mathbb{R}^n \times \mathbb{R}^m$.

(ii) Seja $T : \mathbb{R}^{n+m} \to \mathbb{R}^n$ uma transformação linear sobrejetiva. Mantenhamos as identificações feitas acima e suponhamos, sem perda de generalidade, que o núcleo de T é o subespaço $\{0\} \times \mathbb{R}^m \subset \mathbb{R}^n \times \mathbb{R}^m$. Sendo assim, definindo-se $\Psi : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^n \times \mathbb{R}^m$ por $\Psi(x,y) = (T(x,y), y)$, tem-se que Ψ é linear. Agora, a igualdade $\Psi(x,y) = (0,0)$ implica $T(x,y) = 0$ e $y = 0$, o que nos dá $T(x,0) = 0$ e, então, $x = 0$. Segue-se que Ψ é um isomorfismo e, em particular, uma aplicação aberta. Denotando-se por P a projeção ortogonal $(x,y) \mapsto x$, tem-se $T = P \circ \Psi$, donde se infere que T é aberta, por ser uma composta de aplicações abertas (vide Exemplo 15).

8. (i) Considere as projeções ortogonais $P_1 : \mathbb{R}^{n+m} \to \mathbb{R}^n$, $P_2 : \mathbb{R}^{n+m} \to \mathbb{R}^m$, em que $P_1(x,y) = x$, $P_2(x,y) = y$. Observando-se que $X = P_1(X \times Y)$, $Y = P_2(X \times Y)$ e $X \times Y = P_1^{-1}(X) \cap P_2^{-1}(Y)$, conclui-se do Exemplo 15, do Exercício 6-(i) e do fato de a interseção finita de abertos ser aberta, que $X \times Y$ é aberto se, e somente se, X e Y são abertos (note que P_1 e P_2 são aplicações lineares).

   (ii) A igualdade $X \times Y = P_1^{-1}(X) \cap P_2^{-1}(Y)$ e o Exercício 6-(ii) implicam que $X \times Y$ é fechado se X e Y são fechados, pois toda interseção de fechados é fechada. Suponhamos agora que $X \times Y$ seja fechado e tomemos $a \in \overline{X}$. Neste caso, existe uma sequência (x_k) em X, tal que $x_k \to a$. Tomando-se arbitrariamente $b \in Y$, temos que a sequência (z_k) em $X \times Y$, em que $z_k = (x_k, b)$, converge para (a,b). Uma vez que $X \times Y$ é fechado, devemos ter $(a,b) \in X \times Y$, donde $a \in X$. Logo, $\overline{X} = X$ e, portanto, X é fechado. De modo análogo, prova-se que Y é também fechado.

9. Suponhamos, por absurdo, que A_1 não seja aberto. Então, existem um transformação linear injetiva $T \in \mathrm{L}(\mathbb{R}^n, \mathbb{R}^m)$ e uma sequência (H_k) em $\mathrm{L}(\mathbb{R}^n, \mathbb{R}^m)$, tais que $H_k \to 0$ e, para cada $k \in \mathbb{N}$, $T + H_k$ não é injetiva. Assim, para cada $k \in \mathbb{N}$, existe um vetor unitário $u_k \in S^{n-1}$, tal que $(T + H_k)u_k = 0$. Uma vez que a sequência (u_k) é limitada, pelo Teorema de Bolzano-Weierstrass, podemos passar a uma subsequência, se necessário, e supor que $u_k \to u \in S^{n-1}$, donde (vide Proposição 8 – Capítulo 1) $Tu_k \to Tu$. Além disso, $\|H_k u_k\| \le \|H_k\|\,\|u_k\| = \|H_k\| \to 0$, isto é, $H_k u_k \to 0$. Daí, tem-se $0 = \lim(T + H_k)u_k = \lim(Tu_k + H_k u_k) = \lim Tu_k + \lim H_k u_k = Tu$, o que contradiz a injetividade de T. Segue-se que A_1 é aberto.

   Seja $T \in \mathrm{L}(\mathbb{R}^n, \mathbb{R}^m)$ sobrejetiva. Neste caso, devemos ter $n \ge m$. Tomemos, então, uma base $\{v_1, \dots, v_m, \dots, v_n\}$ de $\mathbb{R}^n$, tal que $\{Tv_1, \dots, Tv_m\}$ seja uma base de $\mathbb{R}^m$ (isto é, $\{v_{m+1}, \dots, v_n\}$ é uma base do núcleo de T) e denotemos por $\mathbb{V}$ o subespaço de $\mathbb{R}^n$ gerado por $\{v_1, \dots, v_m\}$. Sendo assim, a restrição $T_0 = T|_{\mathbb{V}} : \mathbb{V} \to \mathbb{R}^m$ é um isomorfismo linear e, em particular, injetiva. Logo, pela primeira parte do exercício, existe $r > 0$, tal que, para toda $H_0 \in \mathrm{L}(\mathbb{V}, \mathbb{R}^m)$

satisfazendo $\|H_0\| < r$, tem-se $T_0 + H_0 \in \mathrm{L}(\mathbb{V}, \mathbb{R}^m)$ injetiva e, portanto, sobrejetiva, já que $\dim \mathbb{V} = m$. Logo, tomando-se $H \in \mathrm{L}(\mathbb{R}^n, \mathbb{R}^m)$, tal que $\|H\| < r$, e fazendo-se $H_0 = H|_{\mathbb{V}}$, tem-se $\|H_0\| \le \|H\| < r$, donde $T_0 + H_0$ é sobrejetiva. Uma vez que $T_0 + H_0 = (T + H)|_{\mathbb{V}}$, segue-se que $T + H$ é sobrejetiva, donde se infere que a bola aberta $B(T, r)$ de $\mathrm{L}(\mathbb{R}^n, \mathbb{R}^m)$ está contida em A_2. Logo, A_2 é um aberto de $\mathrm{L}(\mathbb{R}^n, \mathbb{R}^m)$.

10. Seja $A \subset \mathbb{R}^n$ aberto e suponhamos, por absurdo, que exista $c \in \operatorname{int}(\partial A)$. Neste caso, existe $r > 0$, tal que $B(c, r) \subset \partial A$. No entanto, $c \in \partial A$. Logo, existe $x \in B(c, r) \cap A$, donde $x \in A \cap \partial A$. Isto, porém, contradiz a hipótese, pois, sendo A aberto, devemos ter $A \cap \partial A = \emptyset$.

11. Sejam $r = \inf\{\|x - a\|; x \in F - \{a\}\}$ e (x_k) uma sequência em $F - \{a\}$, tais que $\|x_k - a\| \to r$. Uma vez que a é um ponto isolado de F, devemos ter $r > 0$ e, pela definição de r, $B(a, r) \cap F = \{a\}$. Além disso, temos que (x_k) é limitada, pois $\|x_k\| \le \|x_k - a\| + \|a\|$. Logo, (x_k) possui uma subsequência convergente (x_{k_i}) cujo limite é um ponto $b \in F$, pois F é fechado. Assim, $\|b - a\| = \|\lim(x_{k_i} - a)\| = \lim \|x_{k_i} - a\| = r$, donde $b \in S[a, r] \cap F$.

    Consideremos o conjunto (não-fechado) $X = \{0\} \cup (1, 2] \subset \mathbb{R}$ e observemos que 0 é um ponto isolado de X. No entanto, como se vê facilmente, não existe $r > 0$, tal que $B(0, r) \cap X = \{0\}$ e $S[0, r] \cap X \neq \emptyset$.

12. Dado $a \in S^n$, tomemos $b \in S^n$, de tal forma que a e b sejam linearmente independentes, isto é, $a \neq \pm b$. Escrevendo-se, para cada $k \in \mathbb{N}$, $x_k = (1/k)b + (1 - 1/k)a$, temos que $x_k \to a$. Além disso, da independência linear entre a e b, tem-se, para todo $k \in \mathbb{N}$, que x_k e a são linearmente independentes. Em particular, $x_k \neq 0 \, \forall k \in \mathbb{N}$. Logo, a sequência (y_k) em S^n, dada por $y_k = x_k / \|x_k\|$, está bem definida e, claramente, converge para a. Por fim, segue-se da independência linear entre x_k e a que, para todo $k \in \mathbb{N}$, $y_k \neq a$, donde se infere que a é um ponto de acumulação da esfera S^n.

13. Denotemos por τ a família formada pelos abertos relativos de X, isto é, $\tau = \{A \subset X;\ A = U \cap X,\ U \subset \mathbb{R}^n \text{ aberto}\}$. Temos que $X = \mathbb{R}^n \cap X$ e $\emptyset = \emptyset \cap X$, donde $\emptyset, X \in \tau$, já que $\emptyset$ e $\mathbb{R}^n$ são abertos de $\mathbb{R}^n$. Seja $\{A_\lambda\}_{\lambda \in \Lambda}$ uma subfamília de τ. Então, para cada $\lambda \in \Lambda$, existe um aberto U_λ de $\mathbb{R}^n$, tal que $A_\lambda = U_\lambda \cap X$. Se Λ é um conjunto finito, então, pelo Teorema 4, $\bigcap U_\lambda$ é aberto. Daí, uma vez que $\bigcap A_\lambda = \bigcap (U_\lambda \cap X) = (\bigcap U_\lambda) \cap X$, conclui-se que $\bigcap A_\lambda \in \tau$. Agora, se Λ é um conjunto arbitrário de índices, tem-se, novamente pelo Teorema 4, que $\bigcup U_\lambda$ é um aberto de $\mathbb{R}^n$. Como $\bigcup A_\lambda = \bigcup (U_\lambda \cap X) = \bigcup (U_\lambda) \cap X$, tem-se, então, $\bigcup A_\lambda \in \tau$. Segue-se, destas considerações, que a família τ define uma topologia em X.

14. Dado $a \in \mathbb{Z}^n$, temos que a bola aberta $B_a = B(a, 1/2)$ é tal que $B_a \cap \mathbb{Z}^n = \{a\}$, isto é, $\mathbb{Z}^n$ é um subconjunto discreto de $\mathbb{R}^n$. Tomando-se, então, $X \subset \mathbb{Z}^n$ e fazendo-se $U = \bigcup_{a \in X} B_a$, temos que U é aberto e

$$U \cap \mathbb{Z}^n = (\bigcup_{a \in X} B_a) \cap \mathbb{Z}^n = \bigcup_{a \in X} (B_a \cap \mathbb{Z}^n) = \bigcup_{a \in X} \{a\} = X,$$

    isto é, X é aberto. Uma vez que X foi tomado arbitrariamente, tem-se, em particular, que $\mathbb{Z}^n - X$ é aberto em $\mathbb{Z}^n$, donde X é fechado em $\mathbb{Z}^n$.

15. Se F é fechado em X, temos, pela Proposição 15, que $F = \overline{F} \cap X$, o que prova a parte "somente se" da afirmação, pois $\overline{F}$ é fechado em $\mathbb{R}^n$. Reciprocamente, suponhamos que $G \subset \mathbb{R}^n$ seja um fechado de $\mathbb{R}^n$, tal que $F = G \cap X$. Neste caso, temos $X - F = X - G = (\mathbb{R}^n - G) \cap X$, donde $X - F$ é aberto em X, pois $\mathbb{R}^n - G$ é aberto em $\mathbb{R}^n$. Logo, F é fechado em X.

16. Provemos, inicialmente, que $X \subset \mathbb{R}^n$ e $Y \subset \mathbb{R}^m$ são limitados se, e somente se, $X \times Y \subset \mathbb{R}^{n+m}$ é limitado. De fato, considerando-se a norma do máximo nos espaços envolvidos, temos que se X e Y são limitados, então existem constantes $\mu_1, \mu_2 > 0$, tais que $\|x\|_{\max} < \mu_1 \, \forall x \in X$ e $\|y\|_{\max} < \mu_2 \, \forall y \in Y$. Então, $\|(x,y)\|_{\max} = \max\{\|x\|_{\max}, \|y\|_{\max}\} < \mu \, \forall (x,y) \in X \times Y$, em que $\mu = \max\{\mu_1, \mu_2\}$. Logo, $X \times Y$ é limitado. Reciprocamente, se $X \times Y$ é limitado, existe $\mu > 0$, tal que $\|(x,y)\|_{\max} = \max\{\|x\|_{\max}, \|y\|_{\max}\} < \mu \, \forall (x,y) \in X \times Y$, donde $\|x\|_{\max} < \mu \, \forall x \in X$ e $\|y\|_{\max} < \mu \, \forall y \in Y$, isto é, X e Y são limitados. Agora, pelo Exercício 8-(ii), X e Y são fechados se, e somente se, $X \times Y$ é fechado em $\mathbb{R}^{n+m}$. O resultado segue-se, então, do Teorema de Heine-Borel.

17. Seja $X \subset \mathbb{R}^n$ um conjunto limitado. Então, existe $r > 0$, tal que $X \subset B(0,r)$, donde $\overline{X} \subset \overline{B(0,r)} = B[0,r]$, isto é, o fecho de X é limitado. Assim, uma vez que $\partial X = \overline{X} \cap \overline{\mathbb{R}^n - X} \subset \overline{X}$, tem-se que ∂X é fechado e limitado. Logo, pelo Teorema de Heine-Borel, ∂X é compacto.

18. Seja (Tx_k), $x_k \in K$, uma sequência em $T(K)$. Temos que K é sequencialmente compacto, por ser compacto. Logo, a sequência (x_k) possui uma subsequência (x_{k_i}), tal que $x_{k_i} \to a \in K$. Então, pela Proposição 8 do Capítulo 1, $Tx_{k_i} \to Ta \in T(K)$, donde $T(K)$ é sequencialmente compacto e, portanto, compacto.

19. Dado $a \in \overline{X}$, seja (x_k) uma sequência em X que converge para a. Então, dado $\epsilon > 0$, existe $k_0 \in \mathbb{N}$, tal que $k \geq k_0 \Rightarrow x_k \in B(a, \epsilon)$. Em particular, a subsequência $(x_k)_{k \geq k_0}$ é uma sequência no conjunto $B[a,\epsilon] \cap X$, o qual, por hipótese, é compacto, já que a bola $B[a,\epsilon]$ é compacta. Então, pelo Teorema de Heine-Borel, $B[a,\epsilon] \cap X$ é fechado, donde $a \in B[a,\epsilon] \cap X$. Logo, $a \in X$ e, portanto, $\overline{X} = X$, isto é, X é fechado.

20. Consideremos, para cada $x \in K$, um real positivo r_x, tal que a bola aberta $B(x, 2r_x)$ esteja contida em algum aberto de $\mathscr{A}$. A família $\{B(x, r_x)\}_{x \in K}$, claramente, constitui uma cobertura aberta de K, da qual podemos extrair uma subcobertura finita $B(x_1, r_1), \dots, B(x_k, r_k)$, $r_i = r_{x_i}$, pois K é compacto. Daí, segue-se que $\epsilon = \min\{r_1, \cdots, r_k\}$ é um número de Lebesgue de $\mathscr{A}$. De fato, dado $x \in K$, temos que $x \in B(x_i, r_i)$ para algum $i \in \{1, \cdots, k\}$. Logo, se $y \in B(x, \epsilon)$, tem-se $\|y - x_i\| \leq \|y - x\| + \|x - x_i\| < \epsilon + r_i \leq r_i + r_i = 2r_i$, isto é, $B(x, \epsilon) \subset B(x_i, 2r_i) \subset A_\lambda$ para algum $\lambda \in \Lambda$.

21. Sejam $\{A_k\}_{k \in \mathbb{N}}$ uma família enumerável de abertos densos de $\mathbb{R}^n$ e $X = \bigcap A_k$. Tomemos $a \in \mathbb{R}^n$ juntamente com uma bola aberta $B = B(a, r)$ e provemos que $X \cap B \neq \emptyset$, donde se concluirá que X é denso em $\mathbb{R}^n$. Uma vez que A_1 é denso em $\mathbb{R}^n$, existe $a_1 \in A_1 \cap B$. Porém, A_1 e B são abertos, donde $A_1 \cap B$ é aberto. Logo, existe $r_1 > 0$, tal que $B[a_1, r_1] \subset A_1 \cap B$. Analogamente,

existem $a_2 \in A_2 \cap B(a_1, r_1)$ e $r_2 > 0$, tais que $B[a_2, r_2] \subset A_2 \cap B(a_1, r_1)$, donde $B[a_2, r_2] \subset (A_1 \cap A_2) \cap B$ e $B[a_2, r_2] \subset B[a_1, r_1]$. Procedendo-se indutivamente, obtém-se uma família enumerável de bolas fechadas e encaixadas

$$B[a_1, r_1] \supset B[a_2, r_2] \supset \cdots \supset B[a_k, r_k] \supset \cdots,$$

tal que, para cada $k \in \mathbb{N}$, $B[a_k, r_k] \subset (A_1 \cap A_2 \cap \cdots \cap A_k) \cap B$. Como toda bola fechada é compacta, pelo Teorema dos Compactos Encaixados, tem-se $\bigcap B[a_k, r_k] \neq \emptyset$. Daí, uma vez que $\bigcap B[a_k, r_k] \subset (\bigcap A_k) \cap B = X \cap B$, infere-se que $X \cap B \neq \emptyset$.

Para a conclusão, observemos inicialmente que, pela Proposição 12-(ii), um conjunto $X \subset \mathbb{R}^n$ é denso em $\mathbb{R}^n$ se, e somente se, seu complementar tem interior vazio. Desta forma, se $\{F_k\}_{k \in \mathbb{N}}$ é uma família enumerável de fechados de $\mathbb{R}^n$ de interior vazio, fazendo-se, para cada $k \in \mathbb{N}$, $A_k = \mathbb{R}^n - F_k$, temos que cada $A_k \subset \mathbb{R}^n$ será um aberto denso em $\mathbb{R}^n$, donde, pelo Teorema de Baire, $\bigcap A_k$ é denso em $\mathbb{R}^n$. Assim, uma vez que $\bigcup F_k = \bigcup(\mathbb{R}^n - A_k) = \mathbb{R}^n - \bigcap A_k$, temos que $\bigcup F_k$ tem interior vazio.

22. Dada uma cisão de Y, $Y = A \cup B$, temos que $X = (X \cap A) \cup (X \cap B)$ é uma cisão de X. Como X é conexo, devemos ter $X \cap A = \emptyset$ ou $X \cap B = \emptyset$. Suponhamos que ocorra o primeiro caso. Teremos, então, $X = X \cap B$, donde $X \subset B$ e, portanto, $\overline{X} \subset \overline{B}$. Como $A \cap \overline{B} = \emptyset$ (pois $A \cup B$ é uma cisão), tem-se $\overline{X} \cap A = \emptyset$. Porém, $Y \subset \overline{X}$, o que nos dá $A = Y \cap A = \emptyset$. Logo, Y é conexo. Em particular, $\overline{X}$ é conexo se X for conexo, pois $X \subset \overline{X} \subset \overline{X}$. Daí, segue-se que se C é uma componente conexa de um subconjunto X de $\mathbb{R}^n$, então C é fechado em X. De fato, sendo C conexo, $\overline{C}$ é conexo. Logo, uma vez que $C \subset X \cap \overline{C} \subset \overline{C}$, temos que $X \cap \overline{C}$ é conexo, donde $C = X \cap \overline{C}$, pois nenhum subconjunto conexo de X pode conter C propriamente. Segue-se, então, da Proposição 15, que C é fechado em X.

23. Temos que $A \in \mathrm{I}(n)$ se, e somente se, $\det A \neq 0$. Assim, $\mathrm{I}(n) = \mathrm{D}_+(n) \cup \mathrm{D}_-(n)$, em que $\mathrm{D}_+(n) = \{A \in \mathrm{I}(n); \det A > 0\}$ e $\mathrm{D}_-(n) = \{A \in \mathrm{I}(n); \det A < 0\}$. Dada, $A \in \overline{\mathrm{D}_+(n)}$, seja (A_k) uma sequência em $\mathrm{D}_+(n)$ cujo limite é A. Então (vide Proposição 9 – Capítulo 1), $\det A_k \to \det A$, donde $\det A \geq 0$, pois um real negativo não pode ser limite de uma sequência de reais positivos. Em particular, $A \notin \mathrm{D}_-(n)$, isto é, $\overline{\mathrm{D}_+(n)} \cap \mathrm{D}_-(n) = \emptyset$. Analogamente, verifica-se que $\overline{\mathrm{D}_-(n)} \cap \mathrm{D}_+(n) = \emptyset$. Uma vez que, claramente, $\mathrm{D}_+(n)$ e $\mathrm{D}_-(n)$ são não-vazios, segue-se que a decomposição $\mathrm{I}(n) = \mathrm{D}_+(n) \cup \mathrm{D}_-(n)$ é uma cisão não-trivial de $\mathrm{I}(n)$ e, portanto, que $\mathrm{I}(n)$ não é conexo.

24. Considerando-se a igualdade $\mathbb{R}^n = \operatorname{int} X \cup \operatorname{int}(\mathbb{R}^n - X) \cup \partial X$ e lembrando-se que esta união é disjunta, tem-se que $\partial X = \emptyset$ se, e somente se, $\mathbb{R}^n = \operatorname{int} X \cup \operatorname{int}(\mathbb{R}^n - X)$. Esta igualdade, por sua vez, ocorre se, e somente se, $\operatorname{int} X = \mathbb{R}^n$ e $\operatorname{int}(\mathbb{R}^n - X) = \emptyset$ ou $\operatorname{int} X = \emptyset$ e $\operatorname{int}(\mathbb{R}^n - X) = \mathbb{R}^n$, pois $\mathbb{R}^n$ é conexo e os conjuntos $\operatorname{int} X$ e $\operatorname{int}(\mathbb{R}^n - X)$ são, ambos, abertos (vide Exercício 1). Porém, esta última condição equivale a $X = \mathbb{R}^n$, pois $X \neq \emptyset$.

25. Dada uma cisão $X \cup Y = A \cup B$, de $X \cup Y$, tem-se que as decomposições $X = (A \cap X) \cup (B \cap X)$ e $Y = (A \cap Y) \cup (B \cap Y)$ são cisões de X e Y, respectivamente. Sendo estes conexos, ambas estas cisões são triviais, donde se

conclui que cada um dos conjuntos X e Y está contido em A ou em B. Porém, se $X \subset A$, por exemplo, temos que $\partial X \subset \overline{X} \subset \overline{A}$, o que implica $Y \subset A$, pois $\partial X \subset Y$ (por hipótese) e $\overline{A} \cap B = \emptyset$ (pois $A \cup B$ é uma cisão). Segue-se que $X \cup Y = A$ e, portanto, que $X \cup Y$ é conexo.

26. Seja $X = A \cup B$ uma cisão de $X = \bigcup X_k$. Então, para cada $k \in \mathbb{N}$, $X_k = (A \cap X_k) \cup (B \cap X_k)$ é uma cisão de X_k. Como cada X_k é conexo, devemos ter $A \cap X_k = \emptyset$ ou $B \cap X_k = \emptyset$, donde $X_k \subset A$ ou $X_k \subset B$. Suponhamos, sem perda de generalidade, que $X_1 \subset A$. Uma vez que A e B são disjuntos e, para todo $k \in \mathbb{N}$, $X_k \cap X_{k+1} \neq \emptyset$, se $X_k \subset A$, então $X_{k+1} \subset A$. Logo, pelo princípio da indução, $X_k \subset A \, \forall k \in \mathbb{N}$, donde se conclui que $X = A$ e, portanto, que X é conexo.

27. Sejam $A = X \cap C$ e $B = (\mathbb{R}^n - X) \cap C$. Por hipótese, A e B são não-vazios e, claramente, disjuntos. Como $C = A \cup B$ e C é conexo, existe $a \in \overline{A} \cap B$ ou $b \in A \cap \overline{B}$. No primeiro caso, temos $a \in \overline{A} = \overline{X \cap C} \subset \overline{X}$. Desta forma, toda bola aberta centrada em a intersecta X (pois $a \in \overline{X}$) e $\mathbb{R}^n - X$ (pois $a \in B \subset \mathbb{R}^n - X$). Logo, $a \in \partial X \cap C$. Analogamente, prova-se que, no segundo caso, $b \in \partial X \cap C$.

28. Temos, por hipótese, que $\emptyset \in \tau$. Além disso, como o conjunto vazio é finito, temos que $\mathbb{R}^n = \mathbb{R}^n - \emptyset \in \tau$. Seja $\{A_\lambda\}_{\lambda \in \Lambda}$ uma subfamília de conjuntos não-vazios de τ. Então, para cada $\lambda \in \Lambda$, existe um subconjunto finito de $\mathbb{R}^n$, Ω_λ, tal que $A_\lambda = \mathbb{R}^n - \Omega_\lambda$. Se $\Lambda = \{1, \dots, k\} \subset \mathbb{N}$, então, $A_1 \cap \cdots \cap A_k = (\mathbb{R}^n - \Omega_1) \cap \cdots \cap (\mathbb{R}^n - \Omega_k) = \mathbb{R}^n - (\Omega_1 \cup \cdots \cup \Omega_k) \in \tau$, pois qualquer união finita de conjuntos finitos é um conjunto finito. Por fim, se Λ é um conjunto arbitrário de índices, temos $\bigcup A_\lambda = \bigcup (\mathbb{R}^n - \Omega_\lambda) = \mathbb{R}^n - \bigcap \Omega_\lambda \in \tau$, pois a interseção de uma família qualquer de conjuntos finitos é um conjunto finito. Logo, τ é uma topologia em $\mathbb{R}^n$.

    Sejam A e B abertos não-vazios de τ. Então, $\mathbb{R}^n - A$ e $\mathbb{R}^n - B$ são finitos. Logo, $\mathbb{R}^n - (A \cap B) = (\mathbb{R}^n - A) \cup (\mathbb{R}^n - B) \neq \mathbb{R}^n$, donde $A \cap B \neq \emptyset$, isto é, quaisquer dois abertos não-vazios de τ se intersectam. Desta forma, $(\mathbb{R}^n, \tau)$ não é de Hausdorff e, portanto, não é metrizável.

    Sejam $X \subset \mathbb{R}^n$ um subespaço de $(\mathbb{R}^n, \tau)$ e $\mathscr{A} = \{A_\lambda\}$ uma família de abertos relativos de X que o cobrem. Tomando-se $\lambda_0 \in \Lambda$ arbitrariamente, tem-se que existe um conjunto finito $\Omega_{\lambda_0} \subset \mathbb{R}^n$, tal que $A_{\lambda_0} = (\mathbb{R}^n - \Omega_{\lambda_0}) \cap X = X - \Omega_{\lambda_0}$. Agora, uma vez que $X \cap \Omega_{\lambda_0}$ é finito, existe um número finito de abertos de $\mathscr{A}$ que o cobrem. Estes, juntamente com A_{λ_0}, formam, então, uma subcobertura finita de $\mathscr{A}$, donde X é compacto.

    Suponhamos agora que $X \subset \mathbb{R}^n$ seja infinito e consideremos uma cisão $X = A \cup B$. Temos que $A = U \cap X$, em que $U \subset \mathbb{R}^n$ é um aberto de $(\mathbb{R}^n, \tau)$, isto é, $\mathbb{R}^n - U$ é finito. Daí, uma vez que $B = X - A = X - (U \cap X) = (\mathbb{R}^n - U) \cap X$, temos que $B = X$ (o que ocorre se $U = \emptyset$) ou B é finito. Analogamente, $A = X$ ou A é finito. Como X é infinito, não podemos ter A e B finitos, o que implica que um deles é igual a X e outro é vazio, donde X é conexo.

29. Dados $a, r \in \mathbb{R}_+$, escrevamos $B(a, r) = \{x \in \mathbb{R}\,;\, d(x, a) < r\}$ e $B'(a, r) = \{x \in \mathbb{R}_+\,;\, d'(x, a) < r\}$. Tomemos $r' > 0$, tal que $r' < \min\{\frac{r}{a(a+r)}, \frac{1}{a}\}$

e provemos que $B'(a,r') \subset B(a,r)$. De fato, dado $x \in B'(a,r')$, tem-se $\frac{|x-a|}{ax} = d'(x,a) < r'$, donde $|x-a| < axr'$ e, então, $x < \frac{a}{1-ar'}\cdot$ Considerando-se as duas últimas desigualdades e a escolha de r', obtém-se $d(x,a) = |x-a| < axr' < \frac{ar}{(1-ar')(a+r)} < r$, pois a desigualdade $r' < \frac{r}{a(a+r)}$ equivale a $\frac{a}{(1-ar')(a+r)} < 1$. Logo, $x \in B(a,r)$ e, portanto, $B'(a,r') \subset B(a,r)$. Analogamente, dados $a, r' > 0$, tomando-se $r > 0$, tal que $r < \min\{\frac{a^2r'}{1+ar'}, a\}$, e observando-se que a desigualdade $r < \frac{a^2r'}{1+ar'}$ equivale a $\frac{r}{a(a-r)} < r'$, conclui-se facilmente que $B(a,r) \subset B'(a,r')$. Desta forma, cada bola aberta de $(\mathbb{R}_+, d)$ contém uma bola aberta de $(\mathbb{R}_+, d')$ e vice-versa, donde se conclui que as topologias de $\mathbb{R}_+$ advindas de d e d' coincidem, isto é, denotando-as respectivamente por τ_d e $\tau_{d'}$, tem-se $A \in \tau_d$ se, e somente se, $A \in \tau_{d'}$.

Fazendo-se $x_k = 1/k\,, k \in \mathbb{N}$, tem-se que (x_k) é de Cauchy em $(\mathbb{R}, d)$, por ser convergente. Logo, (x_k) é de Cauchy em $(\mathbb{R}_+, d)$. Por outro lado, para quaisquer $k, p \in \mathbb{N}$, $d'(x_{k+p}, x_k) = p$, donde se conclui que o limite $\lim_{p\to\infty} d'(x_{k+p}, x_k)$ não existe e, portanto, que (x_k) não é de Cauchy em $(\mathbb{R}_+, d')$.

30. Seja $A \subset X$ um aberto relativo de X. Então, existe um aberto $U \subset \mathbb{R}^n$, tal que $A = U \cap X$, donde $\varphi(A) = \varphi(U) \cap \varphi(X)$, pois φ é, em particular, bijetiva. Como φ é um homeomorfismo, temos que $\varphi(U)$ é aberto em $\mathbb{R}^n$ e, portanto, $\varphi(A) = \varphi|_X(A)$ é aberto em $\varphi(X)$. Logo, $\varphi|_X$ é uma aplicação aberta. De modo análogo, verifica-se que $(\varphi|_X)^{-1} : \varphi(X) \to X$, igualmente, é aberta, donde se conclui que $\varphi|_X : X \to \varphi(X)$ é um homeomorfismo.

    Dados $a, b, c > 0$, a aplicação $T : \mathbb{R}^3 \to \mathbb{R}^3$, $T(x,y,z) = (x/\sqrt{a}, y/\sqrt{b}, z/\sqrt{c})$ é, claramente, um isomorfismo linear. Logo, T é um homeomorfismo (vide Exemplo 14) e, portanto, $\varphi|_S : S \to \varphi(S)$ é um homeomorfismo. No entanto, $\varphi(S) = E$, donde S e E são homeomorfos.

Capítulo 3

1. Tome $0 < \epsilon < \frac{\|f(a)-g(a)\|}{2}$ e considere as bolas (disjuntas) $B_1 = B(f(a), \epsilon)$ e $B_2 = B(g(a), \epsilon)$. Sendo f e g contínuas em a, existem $\delta_1, \delta_2 > 0$ satisfazendo $f(X \cap B(a, \delta_1)) \subset B_1$ e $g(X \cap B(a, \delta_2)) \subset B_2$. Fazendo-se $r = \min\{\delta_1, \delta_2\}$ e tomando-se $B = B(a, r)$, tem-se, então, $f(X \cap B) \subset B_1$ e $g(X \cap B) \subset B_2$. Logo, $f(x) \neq g(x)$ sempre que $x \in X \cap B$.

2. Suponhamos que $f : \mathbb{R}^n \to \mathbb{R}^m$ seja contínua e consideremos um subconjunto Y de $\mathbb{R}^m$. A inclusão $\partial f^{-1}(Y) \subset f^{-1}(\partial Y)$ é trivialmente satisfeita quando $\partial f^{-1}(Y) = \emptyset$. Suponhamos, então, que $\partial f^{-1}(Y)$ seja não-vazio e tomemos $a \in \partial f^{-1}(Y)$. Dado $\epsilon > 0$, segue-se da continuidade de f que existe $\delta > 0$, tal que $f(B(a,\delta)) \subset B(f(a), \epsilon)$ e, pela escolha de a, existem x_0, x_1 em $\mathbb{R}^n$, tais que $x_0 \in B(a,\delta) \cap f^{-1}(Y)$ e $x_1 \in B(a,\delta) \cap (\mathbb{R}^n - f^{-1}(Y))$. Logo, $f(x_0) \in f(B(a,\delta) \cap f^{-1}(Y)) \subset f(B(a,\delta)) \cap f(f^{-1}(Y)) \subset B(f(a), \epsilon) \cap Y$ e, analogamente, $f(x_1) \in B(f(a), \epsilon) \cap (\mathbb{R}^m - Y)$. Segue-se que $f(a) \in \partial Y$, e, portanto, que $a \in f^{-1}(\partial Y)$, donde se infere que $\partial f^{-1}(Y) \subset f^{-1}(\partial Y)$.

    Reciprocamente, suponhamos que, para todo $Y \subset \mathbb{R}^m$, tenha-se $\partial f^{-1}(Y) \subset f^{-1}(\partial Y)$. Se Y for aberto, teremos $Y \cap \partial Y = \emptyset$ e, então, $f^{-1}(Y) \cap \partial f^{-1}(Y) \subset$

$f^{-1}(Y) \cap f^{-1}(\partial Y) = f^{-1}(Y \cap \partial Y) = \emptyset$, donde $f^{-1}(Y)$ é aberto. Segue-se, portanto, do Teorema 12, que f é contínua.

3. Sejam X um subconjunto não-vazio de $\mathbb{R}^n$ e $b \in f(\overline{X})$. Então, existe $a \in \overline{X}$, tal que $f(a) = b$. Tomando-se uma sequência (x_k) em X, tal que $x_k \to a$, tem-se que $(f(x_k))$ é uma sequência em $f(X)$. Além disso, como f é contínua, tem-se $f(x_k) \to f(a) = b$, donde $b \in \overline{f(X)}$. Logo, $f(\overline{X}) \subset \overline{f(X)}$.

   Suponhamos agora que f seja fechada. Neste caso, uma vez que $\overline{X}$ é fechado, $f(\overline{X})$ é fechado. Daí e da inclusão $f(X) \subset f(\overline{X})$ (a qual decorre de $X \subset \overline{X}$), obtém-se $\overline{f(X)} \subset \overline{f(\overline{X})} = f(\overline{X})$, donde $f(\overline{X}) = \overline{f(X)}$.

4. Seja $X \subset F$ limitado. Suponhamos, por absurdo, que $f(X)$ seja ilimitado. Neste caso, para cada $k \in \mathbb{N}$, existe $y_k = f(x_k) \in f(X)$, tal que $\|y_k\| > k$. Como X é limitado, a sequência (x_k) é limitada em X e, portanto, possui uma subsequência (x_{k_i}) que converge para um ponto $a \in F$, pois F é fechado. No entanto, $f(x_{k_i}) = y_{k_i}$ é ilimitada, não sendo, desta forma, convergente. Isto contraria o fato de f ser contínua e prova que $f(X) \subset \mathbb{R}^n$ é limitado.

   O intervalo aberto $(0,1) \subset \mathbb{R} - \{0\}$ é limitado e a função $f : \mathbb{R} - \{0\} \to \mathbb{R}$, $f(x) = 1/x$, é contínua. Porém, $f(\,(0,1)\,) = (0, +\infty)$ é ilimitado. Isto mostra que, no enunciado da proposição deste exercício, a hipótese de F ser fechado é, de fato, necessária.

5. Seja $b \in \mathbb{R}^m$. Como f é sobrejetiva, existe $a \in \mathbb{R}^n$, tal que $f(a) = b$. Porém, X é denso em $\mathbb{R}^n$. Assim, existe uma sequência (x_k), em X, satisfazendo $x_k \to a$. Pela continuidade de f, tem-se $f(x_k) \to b$, isto é, $b \in \overline{f(X)}$. Logo, $\mathbb{R}^m = \overline{f(X)}$.

6. Devemos provar que, quaisquer que sejam $\lambda \in \mathbb{R}$ e $x \in \mathbb{R}^n$, vale a igualdade $T(\lambda x) = \lambda T(x)$. Temos que $T(0) = T(0+0) = T(0) + T(0)$, donde $T(0) = 0$. Logo, $0 = T(0) = T(x + (-x)) = T(x) + T(-x)$, isto é, $T(-x) = -T(x)$ para todo $x \in \mathbb{R}^n$. Podemos, portanto, nos restringir ao caso $\lambda > 0$. Por indução, verifica-se facilmente que $T(x_1 + \cdots + x_q) = T(x_1) + \cdots + T(x_q)$ quaisquer que sejam $x_1, \dots, x_q \in \mathbb{R}^n$. Assim, para quaisquer $q \in \mathbb{N}$ e $x \in \mathbb{R}^n$, tem-se $T(qx) = qT(x)$. Daí, temos que, para todo $q \in \mathbb{N}$ e $x \in \mathbb{R}^n$, $T(x) = T(\frac{q}{q}x) = qT(\frac{1}{q}x)$, isto é, $T(\frac{1}{q}x) = \frac{1}{q}T(x)$. Logo, $T(\frac{p}{q}x) = pT(\frac{1}{q}x) = \frac{p}{q}T(x)\,\forall p, q \in \mathbb{N}$, donde se conclui que a igualdade $T(\lambda x) = \lambda T(x)$ se verifica para quaisquer $x \in \mathbb{R}^n$ e $\lambda \in \mathbb{Q}$. Agora, uma vez que $\mathbb{Q}$ é denso em $\mathbb{R}$, dado $\lambda \in \mathbb{R}$, existe uma sequência convergente (λ_k), em $\mathbb{Q}$, cujo limite é λ. Desta forma, pela continuidade de T, tem-se $T(\lambda x) = \lim T(\lambda_k x) = \lim(\lambda_k T(x)) = \lambda T(x)$, como desejado.

7. Observemos, inicialmente, que a sequência $(\,\|x_k\|\,)$ é limitada e estritamente decrescente em $\mathbb{R}$, pois, para todo $k \in \mathbb{N}$, $0 \le \|x_{k+1}\| = \|f(x_k)\| < \|x_k\| < 1$. Logo, $(\,\|x_k\|\,)$ é convergente e $\lim \|x_k\| < 1$. Temos também que a sequência (x_k) é limitada e, portanto, possui uma subsequência convergente, (x_{k_i}). Fazendo-se $a = \lim x_{k_i}$, tem-se $\|a\| = \lim \|x_{k_i}\| < 1$, donde $a \in B$ e, portanto, $f(x_{k_i}) \to f(a)$, já que f é contínua. Porém, $\|f(a)\| = \lim \|f(x_{k_i})\| =$

$\lim \|x_{k_i+1}\| = \|a\|$. Como $\|f(x)\| < \|x\| \, \forall x \in B - \{0\}$, temos que $a = 0$. Isto implica que $\|x_k\| \to 0$ e, portanto, que $x_k \to 0$.

8. Observemos que as funções μ e λ se escrevem como $\mu(x) = \frac{f(x)+g(x)-\|f(x)-g(x)\|}{2}$ e $\lambda(x) = \frac{f(x)+g(x)+\|f(x)-g(x)\|}{2}$. Segue-se, então, das propriedades operatórias das aplicações contínuas, da continuidade da norma e do fato de a composta de aplicações contínuas ser contínua, que μ e λ são contínuas.

9. Uma vez que f e g são limitadas, existe $\mu > 0$, tal que, para todo $x \in X$, $\|f(x)\| \le \mu$ e $\|g(x)\| \le \mu$. Além disso, pela continuidade uniforme destas funções, dado $\epsilon > 0$, existe $\delta > 0$, tal que $|f(x) - f(y)| < \epsilon$ e $|g(x) - g(y)| < \epsilon$ sempre que $\|x - y\| < \delta$. Sendo assim, para quaisquer $x, y \in X$ satisfazendo $\|x - y\| < \delta$, tem-se $|(fg)(x) - (fg)(y)| = |f(x)g(x) - f(y)g(y)| = |f(x)(g(x) - g(y)) + g(y)(f(x) - f(y)| \le |f(x)|\,|g(x) - g(y)| + |g(y)|\,|f(x) - f(y)| \le 2\mu\epsilon$, donde se infere que fg é uniformemente contínua.

   A função f, dada, é tal que $f = f_1 f_2 \dots f_n$, em que $f_i(x) = \langle \frac{x}{\|x\|}, e_i \rangle$, $x \in \mathbb{R}^n - B(0,r)$. Para cada $i \in \{1, \dots, n\}$, a função f_i é a composta das funções $x \mapsto x/\|x\|$ e $x \mapsto \langle x, e_i \rangle$, que são uniformemente contínuas em $\mathbb{R}^n - B(0,r)$, donde f_i é uniformemente contínua. Além disso, pela desigualdade de Cauchy-Schwarz, cada f_i é limitada. Segue-se, então, do resultado deste exercício (e do princípio da indução) que f é uniformemente contínua.

10. Uma vez que os conjuntos F_1 e F_2 são fechados e disjuntos, temos que a função $x \mapsto d(x, F_1) + d(x, F_2)$ é (uniformemente) contínua e nunca se anula. Logo, a função $f(x) = \frac{d(x,F_1)}{d(x,F_1)+d(x,F_2)}$, $x \in \mathbb{R}^n$, claramente, tem as propriedades desejadas.

11. Sendo U um subconjunto próprio de $\mathbb{R}^n$, não-vazio e aberto, temos que $\partial U \neq \emptyset$ (vide Exercício 24–Capítulo 2) e $U \cap \partial U = \emptyset$. Logo, para todo $x \in U$, tem-se $d(x, \partial U) > 0$, pois ∂U é um subconjunto fechado de $\mathbb{R}^n$. Definamos, então, $f : \mathbb{R}^n \to \mathbb{R}$, em que $f(x) = -d(x, \partial U)$ para $x \in U \cap \mathbb{Q}^n$ e $f(x) = d(x, \partial U)$ para $x \in \mathbb{R}^n - (U \cap \mathbb{Q}^n)$. Sejam $x \in U \cap \mathbb{Q}^n$ e $y \in U - \mathbb{Q}^n \subset \mathbb{R}^n - \mathbb{Q}^n$. Uma vez que $\mathbb{Q}^n$ e $\mathbb{R}^n - \mathbb{Q}^n$ são, ambos, densos em $\mathbb{R}^n$, podemos tomar sequências (x_k), em $U - \mathbb{Q}^n$, e (y_k), em $U \cap \mathbb{Q}^n$, tais que $x_k \to x$ e $y_k \to y$. Então, $f(x_k) = d(x_k, \partial U) \to d(x, \partial U) = -f(x) \neq f(x)$ (pois f nunca se anula em U) e $f(y_k) = -d(y_k, \partial U) \to -d(y, \partial U) = -f(y) \neq f(y)$. Segue-se que f é descontínua em x e em y, donde f é descontínua em U. Por outro lado, fora de U, f coincide com a função distância a ∂U. Logo, $f|_{\mathbb{R}^n - U}$ é uniformemente contínua e não-constante.

12. i) O cone C é o gráfico da função contínua $f : \mathbb{R}^2 \to \mathbb{R}$, dada por $f(x,y) = \sqrt{x^2 + y^2}$. Logo, C é homeomorfo a $\mathbb{R}^2$ (vide Exemplo 45).

    ii) Considere a aplicação $g : \mathbb{R}^{n+1} - \{0\} \to S^n \times \mathbb{R} \subset \mathbb{R}^{n+2}$, definida por $g(x) = (\frac{x}{\|x\|}, \log \|x\|)$, e note que g é contínua, bijetiva e $g^{-1}(u,t) = e^t u$ é contínua. Logo, g é um homeomorfismo.

13. Suponhamos, por absurdo, que U esteja contido propriamente em $\mathbb{R}^n$. Neste caso, $\partial U \neq \emptyset$ (vide Exercício 24 – Capítulo 2). Tomemos $a \in \partial U$ e uma

sequência (x_k), em U, tais que $x_k \to a$. Sendo convergente, (x_k) é de Cauchy. Logo, pelo Corolário 2, $(f(x_k))$ é de Cauchy e, portanto, convergente. Seja $b = \lim f(x_k)$. Uma vez que f é um homeomorfismo, temos que f^{-1} é contínua. Sendo assim, $f^{-1}(b) = f^{-1}(\lim f(x_k)) = \lim f^{-1}(f(x_k)) = \lim x_k = a$, donde $a \in \partial U \cap U$. Isto, porém, contradiz o fato de U ser aberto. Segue-se que $\partial U = \emptyset$ e, portanto, que $U = \mathbb{R}^n$.

14. Procedendo-se como no Exemplo 42, prova-se que uma bola aberta de $\mathbb{R}^n$ com respeito a uma norma qualquer é homeomorfa à bola unitária aberta com respeito à mesma norma e com centro na origem. Agora, procedendo-se como no Exemplo 44, prova-se que esta última é homeomorfa à bola unitária aberta com respeito a qualquer outra norma e também com centro na origem. Segue-se, portanto, da transitividade da relação de homeomorfismo, que, em $\mathbb{R}^n$, duas bolas abertas quaisquer, isto é, relativas a quaisquer normas, são homeomorfas. Analogamente, prova-se que vale o mesmo para bolas fechadas.

15. (i) Dados $a, b \in (-1,1) \subset \mathbb{R}$, tomando-se, em $\mathbb{R}^2$, os pontos $p_1 = (-1,-1)$, $p_2 = (1,1)$ e $p = (a,b)$, obtém-se um homeomorfismo $\lambda : (-1,1) \to (-1,1)$ considerando-se a função cujo gráfico é o conjunto $X - \{p_1, p_2\} \subset \mathbb{R}^2$, em que X é a união dos segmentos de reta $[p_1, p]$ e $[p, p_2]$. Mais precisamente, escrevemos

$$\lambda(x) = \begin{cases} \frac{1+b}{1+a}x + \frac{b-a}{1+a} & \text{se} \quad x \in (-1,a] \\ \\ \frac{1-b}{1-a}x + \frac{b-a}{1-a} & \text{se} \quad x \in (a,1). \end{cases}$$

Claramente, $\lambda(a) = b$ e λ é contínua, pois cada uma das expressões que a definem representam funções contínuas que coincidem em $x = a$. Ademais, por um cálculo direto, verifica-se que λ é bijetiva e sua inversa é

$$\lambda^{-1}(x) = \begin{cases} \frac{1+a}{1+b}x + \frac{a-b}{1+b} & \text{se} \quad x \in (-1,b] \\ \\ \frac{1-a}{1-b}x + \frac{a-b}{1-b} & \text{se} \quad x \in (b,1), \end{cases}$$

que é também contínua. Logo λ é um homeomorfismo, donde se conclui que o intervalo $(-1,1) \subset \mathbb{R}$ é topologicamente homogêneo.

(ii) Sejam $a = (a_1, \dots, a_n), b = (b_1, \dots, b_n) \in B = B_{\max}(0,1) \subset \mathbb{R}^n$. Pelo resultado do item (i), para cada $i \in \{1, \dots, n\}$, existe um homeomorfismo $\lambda_i : (-1,1) \to (-1,1)$, tal que $\lambda_i(a_i) = b_i$. Definindo-se $\varphi : B \to B$ por $\varphi(x_1, \dots, x_n) = (\lambda_1(x_1), \dots, \lambda_n(x_n))$, vê-se facilmente que esta aplicação é um homeomorfismo que leva a em b, cujo inverso é $\varphi^{-1}(y_1, \dots, y_n) = (\lambda^{-1}(y_1), \dots, \lambda^{-1}(y_n))$, donde a bola aberta B é topologicamente homogênea.

Agora, se $\| \;\|_0$ é uma norma qualquer em $\mathbb{R}^n$ e $B_0 \subset \mathbb{R}^n$ é uma bola aberta com respeito a esta norma, temos, pelo exercício anterior, que existe um homeomorfismo $\psi : B_0 \to B$. Assim, dados $a_0, b_0 \in B_0$, tomando-se um homeomorfismo $\varphi : B \to B$, tal que $\varphi(\psi(a_0)) = \psi(b_0)$, tem-se que a aplicação $\psi^{-1} \circ \varphi \circ \psi : B_0 \to B_0$ é um homeomorfismo que leva a_0 em b_0, donde B_0 é topologicamente homogênea.

16. (i) Seja $\lambda : (-1,1) \to (-1,1)$, tal que

$$\lambda(x) = \begin{cases} \frac{2}{3}x - \frac{1}{3} & \text{se} \quad x \in (-1, 1/2] \\ \\ 2x - 1 & \text{se} \quad x \in (1/2, 1). \end{cases}$$

A exemplo da função λ da resolução do exercício anterior, verifica-se facilmente que λ é um homeomorfismo cujo inverso é

$$\lambda^{-1}(x) = \begin{cases} \frac{3}{2}x + \frac{1}{2} & \text{se} \quad x \in (-1, 0] \\ \\ \frac{1}{2}x + \frac{1}{2} & \text{se} \quad x \in (0, 1). \end{cases}$$

Além disso, a equação $\lambda(x) = x$ não admite solução em $(-1,1)$, donde λ é um homeomorfismo sem pontos fixos.

(ii) Pelo item (i), existe um homeomorfismo $\lambda : (-1,1) \to (-1,1)$ sem pontos fixos. Façamos $B = B_{\max}(0,1) \subset \mathbb{R}^n$ e consideremos a decomposição $\mathbb{R}^n = \mathbb{R} \times \mathbb{R}^{n-1}$. Dado $(t,x) \in B$, definimos $\varphi(t,x) = (\lambda(t), x) \in B$. A aplicação $\varphi : B \to B$, então definida, é um homeomorfismo cujo inverso é a aplicação $\varphi^{-1}(s,y) = (\lambda^{-1}(s), y)$. Além disso, a equação $\varphi(t,x) = (t,x)$ nos dá $\lambda(t) = t$, que não tem solução em $(-1,1)$, donde φ não tem pontos fixos.

Considerando-se agora uma bola aberta B_0 com respeito a uma norma $\| \, \|_0$ de $\mathbb{R}^n$, um homeomorfismo $\psi : B_0 \to B$ e um homeomorfismo $\varphi : B \to B$ sem pontos fixos, temos que a aplicação $\varphi_0 = \psi^{-1} \circ \varphi \circ \psi : B_0 \to B_0$ é um homeomorfismo sem pontos fixos. Com efeito, uma vez que $\psi \circ \varphi_0 = \varphi \circ \psi$, se $x \in B_0$ fosse um ponto fixo de φ_0, $\psi(x)$ seria um ponto fixo de φ.

17. A função contínua $x \to \|x\|$, definida em K, é, por hipótese, limitada. Logo, K é limitado. Agora, dado $a \in \overline{K}$, devemos ter $a \in K$. Caso contrário, ficaria bem definida a função $x \to \frac{1}{\|x-a\|}$, $x \in K$, a qual seria contínua e ilimitada. Desta forma, K é também fechado e, portanto, compacto.

18. Suponhamos, por absurdo, que exista $\epsilon > 0$, tal que, para todo $k \in \mathbb{N}$, tenha-se $x_k, y_k \in K$ satisfazendo a desigualdade

$$(*) \qquad \|f(x_k) - f(y_k)\| > k\|x_k - y_k\| + \epsilon.$$

Uma vez que f é contínua e K é compacto, temos que $f(K)$ é compacto. Em particular, fazendo-se $z_k = f(x_k) - f(y_k)$, temos que a sequência (z_k) é limitada. Agora, pela desigualdade $(*)$, $\|x_k - y_k\| < \frac{\|f(x_k)-f(y_k)\|-\epsilon}{k}$, donde $x_k - y_k \to 0$. Porém, f, sendo contínua e definida num compacto, é uniformemente contínua. Logo, pela Proposição 30, devemos ter $f(x_k) - f(y_k) \to 0$, o que, claramente, contradiz $(*)$. Segue-se desta contradição que, para todo $\epsilon > 0$, existe $\mu > 0$, tal que $\|f(x) - f(y)\| \leq \mu\|x - y\| + \epsilon \ \forall x, y \in K$, como desejado.

19. Suponhamos que $f : K \to \mathbb{R}^n$ seja contínua. Então, pela Proposição 45, $\operatorname{graf}(f)$ é homeomorfo à K, donde $\operatorname{graf}(f)$ é compacto. Reciprocamente, suponhamos que o gráfico de $f : K \to \mathbb{R}^n$ seja compacto e consideremos a aplicação $\varphi : \operatorname{graf}(f) \to K$, dada por $\varphi(x, f(x)) = x$. Temos que φ é contínua, pois é a restrição, ao gráfico de f, da projeção $(x,y) \mapsto x$, $(x,y) \in \mathbb{R}^n \times \mathbb{R}^m$. Uma vez

que φ é, claramente, bijetiva, tem-se pela Proposição 33, que φ é um homeomorfismo. Em particular, sua inversa $\varphi^{-1}(x) = (x, f(x))$, $x \in K$, é contínua. Logo, f, sendo uma função-coordenada de φ^{-1}, é contínua.

20. Uma vez que $f : X \to f(X)$ é uma isometria e a composta de isometrias é uma isometria, para todo $k \in \mathbb{N}$, a aplicação $f^k : X \to f^k(X)$ é uma isometria. Assim, dados $k, l \in \mathbb{N}$, tem-se

$$\|x_{k+l} - x_k\| = \|f^{k+l}(x_0) - f^k(x_0)\| = \|f^l(x_0) - x_0\| = \|f(f^{l-1}(x_0)) - x_0\|,$$

donde $\|x_{k+l} - x_k\| \geq \inf\{\|f(x) - x_0\|; x \in X\} = d(x_0, f(X))$.

Suponhamos, agora, que X seja compacto. Sendo uma isometria, a aplicação f é, em particular, contínua e injetiva. Logo, pela Proposição 33, f será um homeomorfismo se for sobrejetiva. Porém, se f não fosse sobrejetiva, haveria $x_0 \in X - f(X)$. Uma vez que, pelo Teorema 14, $f(X)$ é compacto e, portanto, fechado, tem-se $d(x_0, f(X)) > 0$. Procedendo-se, então, como acima, obteríamos uma sequência (x_k) em X, tal que $\|x_{k+l} - x_k\| \geq d(x_0, f(X)) > 0$. Em particular, toda subsequência de (x_k) seria divergente, por não ser de Cauchy, o que contradiria o fato de X ser compacto. Segue-se que f é sobrejetiva e, portanto, um homeomorfismo.

21. (i) Consideremos $F \subset \mathbb{R}^n$ fechado e $b \in \overline{f(F)}$. Existe, então, uma sequência (x_k) em F, tal que $f(x_k) \to b$. Assim, tomando-se $\epsilon > 0$ e a bola $K = B[b, \epsilon]$, existe $k_0 \in \mathbb{N}$, tal que $f(x_k) \in K \, \forall k \geq k_0$. Em particular, $x_k \in f^{-1}(K) \cap F \; \forall k \geq k_0$. Agora, uma vez que K é compacto e f é própria, tem-se $f^{-1}(K)$ compacto, donde $f^{-1}(K) \cap F$, por ser um subconjunto fechado de um compacto, é compacto. Logo, passando-se a uma subsequência, se necessário, podemos supor que $x_k \to a \in f^{-1}(K) \cap F$. Daí e da continuidade de f, segue-se que $f(x_k) \to f(a) = b \in f(F)$. Desta forma, $\overline{f(F)} \subset f(F)$ e, portanto, $f(F)$ é fechado.

(ii) Seja $K \subset \mathbb{R}^m$ compacto e $X = f^{-1}(K) \subset \mathbb{R}^n$ não-vazio (se $X = \emptyset$ não há nada a ser feito). Uma vez que f é contínua, X é fechado. Portanto, $f(X) \subset K$ é fechado, pois f, por hipótese, é fechada. Além disso, a aplicação $g = f|_X : X \to f(X) \subset K$ é contínua, bijetiva e fechada. Logo, sua inversa, g^{-1}, é contínua. Como $f(X)$ é compacto (por ser um subconjunto fechado de um compacto), pela Proposição 33, g é um homeomorfismo. Desta forma, X é compacto, donde f é própria.

22. Dada $A \in \mathrm{O}(n)$, tem-se $AA^* = I$. Logo, $1 = \det(AA^*) = \det A \det A^* = (\det A)^2$, donde $\det A = \pm 1$. Além disso, para todo $n \in \mathbb{N}$, existem $A, B \in \mathrm{O}(n)$, tais que $\det A = 1$ e $\det B = -1$. Basta tomar A, por exemplo, como a matriz identidade e B a matriz obtida desta substituindo-se uma coluna qualquer, que corresponde a um vetor e_j da base canônica de $\mathbb{R}^n$, por $-e_j$. Destas considerações, conclui-se que a imagem de $\mathrm{O}(n)$ pela função (contínua) det é o conjunto (desconexo) $\{-1, 1\}$. Logo, $\mathrm{O}(n)$ é desconexo.

23. Se X é conexo e $f : X \to \{0, 1\}$ é contínua, então $f(X) \subset \{0, 1\}$ é conexo. No entanto, os únicos subconjuntos conexos e não-vazios de $\{0, 1\}$ são $\{0\}$ e $\{1\}$. Logo, f é constante. Se X não é conexo, existe uma cisão não-trivial de X, $X = A \cup B$. Assim, a função $f : X \to \{0, 1\}$, tal que $f(x) = 0 \, \forall x \in A$

e $f(x) = 1\, \forall x \in B$ é contínua, pois a imagem inversa, por f, de qualquer subconjunto de $\{0,1\}$ é um aberto de X. Uma vez que f não é constante, o resultado segue-se por contraposição. Isto prova a primeira parte do exercício.

i) Seja $\{X_\lambda\}_{\lambda\in\Lambda}$ uma família de conjuntos conexos de $\mathbb{R}^n$, tal que $\bigcap X_\lambda \neq \emptyset$. Tomemos, então, $a \in \bigcap X_\lambda$ e façamos $X = \bigcup X_\lambda$. Assim, dada uma função contínua, $f : X \to \{0,1\}$, temos, para cada $\lambda \in \Lambda$, que $f|_{X_\lambda}$ é contínua, donde constante. Agora, dado $x \in X$, existe $\lambda \in \Lambda$, tal que $x \in X_\lambda$. Desta forma, $f(x) = f|_{X_\lambda}(x) = f|_{X_\lambda}(a) = f(a)$, donde se conclui que f é constante. Logo, X é conexo.

ii) Sejam $X \subset \mathbb{R}^n$ um conjunto conexo e $f : \overline{X} \to \{0,1\}$ uma função contínua. Temos, então, que $f|_X$ é constante. Agora, dados $x_0 \in X$ e $a \in \overline{X} - X$ (se $\overline{X} = X$ não há nada a ser feito), existe uma sequência (x_k) em X cujo limite é a. Daí e da continuidade de f, tem-se $f(a) = f(\lim x_k) = \lim f(x_k) = \lim f(x_0) = f(x_0)$. Logo, f é constante e, portanto, $\overline{X}$ é conexo.

24. A implicação (i) é verdadeira, pois se $f : \mathbb{R}^n \to \mathbb{R}$ é contínua, o seu gráfico é homeomorfo a $\mathbb{R}^n$ (vide Exemplo 45), que é conexo. No entanto, a implicação (ii) é falsa. Um contra-exemplo é a função $f : \mathbb{R} \to \mathbb{R}$, definida por $f(x) = \cos(1/x)$, $x \neq 0$, e $f(0) = 0$. Fazendo-se $f_1 = f|_{(0,+\infty)}$ e $f_2 = f|_{(-\infty,0)}$, temos que f_1 e f_2 são contínuas e seus domínios são conexos. Logo, seus respectivos gráficos são conexos. Além disso, a sequência (x_k), em $(0,+\infty)$, dada por $x_k = \frac{2}{(2k-1)\pi}$, é tal que $x_k \to 0$ e $f_1(x_k) \to 0$, donde $(0,0) \in \overline{\mathrm{graf}(f_1)}$. Desta forma, $G_1 = \mathrm{graf}(f_1) \cup \{(0,0)\}$ é conexo, pois $\mathrm{graf}(f_1) \subset G_1 \subset \overline{\mathrm{graf}(f_1)}$ (vide Exercício 22 – Capítulo 2). Analogamente, $G_2 = \mathrm{graf}(f_2) \cup \{(0,0)\}$ é conexo. Agora, $\mathrm{graf}(f) = G_1 \cup G_2$ e $G_1 \cap G_2 = \{(0,0)\} \neq \emptyset$, donde $\mathrm{graf}(f)$ é conexo. No entanto, f não é contínua em $(0,0)$, pois a sequência (y_k), em que $y_k = \frac{1}{2k\pi}$, satisfaz $y_k \to 0$ e $f(y_k) \to 1 \neq f(0)$.

25. Conforme verificamos, para todo $p \in S^2$, $S^2 - \{p, -p\}$ é conexo por caminhos e, portanto, conexo. Por outro lado, dados pontos distintos $q, q' \in S^1$, tomando-se em $\mathbb{R}^2$ um sistema de coordenadas no qual $q = (0,1)$ seja o polo-norte de S^1, tem-se que a projeção estereográfica $\rho : S^1 - \{q\} \to \mathbb{R}$ é um homeomorfismo. Logo, $S^1 - \{q, q'\}$ é homeomorfo a $\mathbb{R} - \{\rho(q')\}$, que é desconexo, donde $S^1 - \{q, q'\}$ é desconexo. Assim, se S^2 fosse homeomorfo a S^1, o conexo $S^2 - \{p, -p\}$ seria homeomorfo ao desconexo $S^1 - \{f(p), f(-p)\}$. Logo, S^2 e S^1 não são homeomorfos.

26. (i) Uma vez que as projeções $(x,y) \mapsto x \in X$, $(x,y) \mapsto y \in Y$, $(x,y) \in X \times Y$, são contínuas e aplicações contínuas levam conexos por caminhos em conexos por caminhos (Proposição 37), tem-se que X e Y são conexos por caminhos se $X \times Y$ o for. Reciprocamente, se X e Y são conexos por caminhos, dados $p = (x_1, y_1)$, $q = (x_2, y_2) \in X \times Y$, existem curvas $\alpha : [0,1] \to X$ e $\beta : [0,1] \to Y$, tais que $\alpha(0) = x_1$, $\alpha(1) = x_2$ e $\beta(0) = y_1$, $\beta(1) = y_2$. Desta forma, a aplicação $\gamma : [0,1] \to X \times Y$, $\gamma(t) = (\alpha(t), \beta(t))$, é uma curva em $X \times Y$, tal que $\gamma(0) = p$ e $\gamma(1) = q$, donde $X \times Y$ é conexo por caminhos.

(ii) Façamos $X = \bigcup X_\lambda$ e tomemos $a \in \bigcap X_\lambda$. Dados $x, y \in X$, existem $\lambda_1, \lambda_2 \in \Lambda$, tais que $x \in X_{\lambda_1}$ e $y \in X_{\lambda_2}$. Uma vez que cada X_λ é conexo por caminhos, temos que existem curvas $\alpha : [0,1] \to X_{\lambda_1}$ e $\beta : [0,1] \to X_{\lambda_2}$, tais que $\alpha(0) = x, \alpha(1) = a$ e $\beta(0) = a, \beta(1) = y$, donde a justaposição $\alpha \vee \beta : [0,1] \to X$ é uma curva em X que liga x a y. Logo, X é conexo por caminhos.

(iii) Seja $U \subset \mathbb{R}^n$ um conjunto aberto, conexo e não-vazio. Dado $a \in U$, considere o conjunto $A \subset U$, formado por todos os pontos $x \in U$, tais que existe um caminho $\alpha_x : [0,1] \to U$ satisfazendo $\alpha_x(0) = x$ e $\alpha_x(1) = a$. Temos que $A \neq \emptyset$, pois $a \in A$. Seja $x_0 \in \overline{A} \cap U$. Uma vez que U é aberto, existe $r > 0$, tal que $B(x_0, r) \subset U$. Existe, também, $x \in B(x_0, r) \cap A$, pois x_0 é aderente ao conjunto A. Porém, a bola $B(x_0, r)$ é conexa por caminhos. Logo, existe um caminho β, em $B(x_0, r)$, que liga x_0 a x, donde $\beta \vee \alpha_x$ é um caminho em U que liga x_0 ao ponto a. Segue-se que $\overline{A} \cap U \subset A$, isto é, A é fechado em U. Agora, dados $x \in A$ e uma bola aberta $B(x, r) \subset U$, procedendo-se de forma análoga, obtém-se, para todo $y \in B(x, r)$, um caminho que liga y ao ponto a. Logo A é aberto em U e, portanto, $A = U$, pois U é conexo. Assim, temos que $U = \bigcup_{x \in U} \alpha_x[0,1]$. Além disso, $a \in \bigcap_{x \in U} \alpha_x[0,1]$ e, para todo $x \in U$, $\alpha_x[0,1]$ é conexo por caminhos, pois é a imagem de um conexo por caminhos por uma aplicação contínua. Logo, pelo resultado do item anterior, U é conexo por caminhos.

27. Seja (x_k) uma sequência em $X - \{a\}$, tal que $x_k \to a$. Da hipótese e da Proposição 41, segue-se que $\lim f(x_k) = x_0$ e $\lim \Phi(x_k) = T$. Uma vez que, para todo $k \in \mathbb{N}$, $\Phi(x_k)f(x_k) = \Phi(x_k)(f(x_k) - x_0) + \Phi(x_k)x_0$, bem como $\Phi(x_k)(f(x_k) - x_0) \to 0$ (pois $(\Phi(x_k))$ é limitada), tem-se, pela Proposição 5 do Capítulo 1, que $\Phi(x_k)f(x_k) \to Tx_0$. Segue-se, então, da Proposição 41, que $\lim\limits_{x \to a} \Phi(x)f(x) = Tx_0$.

28. Tome $0 < \epsilon < (\lambda_2 - \lambda_1)/2$ e note que os intervalos $I_1 = (\lambda_1 - \epsilon, \lambda_1 + \epsilon)$ e $I_2 = (\lambda_2 - \epsilon, \lambda_2 + \epsilon)$ são tais que $t_1 < t_2$ sempre que $t_1 \in I_1$ e $t_2 \in I_2$. Da hipótese, temos que existem reais positivos, δ_1, δ_2, os quais satisfazem $f(X \cap B(a, \delta_1) - \{a\}) \subset I_1$ e $g(X \cap B(a, \delta_2) - \{a\}) \subset I_2$. Tomando-se, então, $\delta = \min\{\delta_1, \delta_2\}$ e $V = B(a, \delta)$, tem-se, para todo $x \in V - \{a\}$, $f(x) \in I_1$ e $g(x) \in I_2$. Logo, $f(x) < g(x) \, \forall x \in V - \{a\}$.

29. Suponhamos que X seja limitado e, por absurdo, que $f(X)$ seja ilimitado. Neste caso, para todo $k \in \mathbb{N}$, existe $x_k \in X$, tal que $\|f(x_k)\| > k$. Em particular, a sequência $(f(x_k))$ não possui subsequências convergentes. Agora, como X é limitado, a sequência (x_k) possui uma subsequência (x_{k_i}) que converge para algum $a \in \mathbb{R}^n$. Porém, $(f(x_{k_i}))$ é divergente, o que contradiz a hipótese de existência do limite de $f(x)$ quando x tende a a.

30. i) Sejam $a \in X'$ e (x_k) uma sequência em $X - \{a\}$, tal que $x_k \to a$. Uma vez que f é uniformemente contínua, pelo Corolário 2, a sequência $(f(x_k))$ é de Cauchy em $\mathbb{R}^m$ e, portanto, converge para um ponto $b \in \mathbb{R}^m$. Se (y_k) é uma outra sequência em $X - \{a\}$, tal que $y_k \to a$, temos que $x_k - y_k \to 0$. Logo, ainda pela continuidade uniforme de f, devemos ter $f(x_k) - f(y_k) \to 0$. Daí e da desigualdade $\|f(y_k) - b\| \leq \|f(x_k) - f(y_k)\| + \|f(x_k) - b\|$, tem-se, então, $f(y_k) \to b$. Logo, pela Proposição 41, $\lim\limits_{x \to a} f(x) = b$.

ii) Pelo resultado do item (i), temos que a aplicação $\varphi : \mathbb{R}^n \to \mathbb{R}^m$, dada por $\varphi(x_0) = \lim_{x \to x_0} f(x)$, está bem definida. Além disso, para todo $x_0 \in X$, $\varphi(x_0) = \lim_{x \to x_0} f(x) = f(x_0)$, pois f é contínua em X. Dado, então, $x_0 \in \mathbb{R}^n - X$, seja (x_k) uma sequência em $X - \{x_0\}$, tal que $x_k \to x_0$ (uma tal sequência existe devido à densidade de X em $\mathbb{R}^n$). Neste caso, tem-se $\lim \varphi(x_k) = \lim f(x_k) = \lim_{x \to x_0} f(x) = \varphi(x_0)$ donde φ é contínua. Se $\psi : \mathbb{R}^n \to \mathbb{R}^m$ é uma aplicação contínua, tal que $\psi|_X = f$, tem-se ainda $\psi(x_0) = \lim \psi(x_k) = \lim f(x_k) = \lim_{x \to x_0} f(x) = \varphi(x_0)$, isto é, φ é única. Agora, dado $\epsilon > 0$, existe $\delta > 0$, tal que, para quaisquer $x, y \in X$ satisfazendo $\|x - y\| < \delta$, tem-se $\|f(x) - f(y)\| < \epsilon$. Além disso, dados $x_0, y_0 \in \mathbb{R}^n$, existem $\delta_1 = \delta_1(x_0) > 0$ e $\delta_2 = \delta_2(y_0) > 0$, tais que $\|\varphi(x_0) - f(x)\| < \epsilon$ e $\|\varphi(y_0) - f(y)\| < \epsilon$ para quaisquer $x, y \in X$ satisfazendo $0 < \|x - x_0\| < \delta_1$ e $0 < \|y - y_0\| < \delta_2$, pois $\varphi(x_0) = \lim_{x \to x_0} f(x)$ e $\varphi(y_0) = \lim_{y \to y_0} f(y)$. Tomando-se, então, $x_0, y_0 \in \mathbb{R}^n$, tais que $\|x_0 - y_0\| < \delta$, e $x, y \in [x_0, y_0]$, tais que $0 < \|x - x_0\| < \delta_1(x_0)$ e $0 < \|y - y_0\| < \delta_2(y_0)$, tem-se $\|x - y\| < \delta$ e, portanto, $\|\varphi(x_0) - \varphi(y_0)\| \leq \|\varphi(x_0) - f(x)\| + \|f(x) - f(y)\| + \|f(y) - \varphi(y_0)\| < 3\epsilon$, donde φ é uniformemente contínua.

Capítulo 4

1. i) Dados $(x, y), (h, k) \in \mathbb{R}^2$, temos que
$$f(x + h, y + k) - f(x, y) = (2xh + h^2 + k, h + 2yk + k^2)\,.$$
Logo, $f(x+h, y+k) - f(x, y) = T(h, k) + r(h, k)$, em que T é a transformação linear $T(h, k) = (2xh+k, h+2yk)$ e r é a aplicação $r(h, k) = (h^2, k^2)$. Tomando-se a norma do máximo em $\mathbb{R}^2$, obtemos
$$\frac{\|r(h, k)\|_{\max}}{\|(h, k)\|_{\max}} = \frac{\|(h^2, k^2)\|_{\max}}{\|(h, k)\|_{\max}} = \max\{\|h\|, \|k\|\}\,.$$
Assim,
$$\lim_{(h,k) \to 0} \frac{\|r(h, k)\|_{\max}}{\|(h, k)\|_{\max}} = 0,$$
donde f é diferenciável e $f'(x, y)(h, k) = T(h, k) = (2xh + k, h + 2yk)\,.$

   ii) Tomando-se $X, H \in \mathrm{L}(\mathbb{R}^n)$ e desenvolvendo-se $(X + H)^3 - X^3$, obtém-se $(X + H)^3 - X^3 = XHX + X^2H + XH^2 + HX^2 + H^2X + HXH + H^3$. Assim, fazendo-se $TH = XHX + X^2H + HX^2$ e $R(H) = XH^2 + H^2X + HXH + H^3$, tem-se $f(X + H) - f(X) = TH + R(H)$ e $\|R(H)\| \leq \|XH^2\| + \|H^2X\| + \|HXH\| + \|H^3\| \leq 3\|X\|\,\|H\|^2 + \|H\|^3$. Desta forma, $\lim_{H \to 0} \frac{\|R(H)\|}{\|H\|} = 0$, isto é, f é diferenciável e $f'(X)H = XHX + X^2H + HX^2 \;\; \forall X, H \in \mathrm{L}(\mathbb{R}^n)$.

2. Verifiquemos, inicialmente, que as derivadas direcionais de f em $(0, 0)$ existem e são todas nulas. De fato, dado $v = (h, k) \in \mathbb{R}^2$, $k \neq 0$, tem-se
$$\lim_{t \to 0} \frac{f((0, 0) + t(h, k)) - f(0, 0)}{t} = \lim_{t \to 0} \left(1 - \cos(th^2/k)\right) \sqrt{h^2 + k^2}\frac{|t|}{t} = 0,$$
pois $t \to \frac{|t|}{t}$ é uma função limitada. O mesmo vale, trivialmente, para os vetores da forma $(h, 0)$, $h \in \mathbb{R}$. Logo, a aplicação $v \to \frac{\partial f}{\partial v}(0, 0)$, $v \in \mathbb{R}^n$, é identicamente nula e, em particular, linear. Isto prova (i). Quanto a (ii), se f fosse diferenciável em $(0, 0)$, a derivada de f neste ponto seria a aplicação

linear nula. Esta, por sua vez, definiria a função resto $r(h,k) = f(h,k)$ e, daí, teríamos $r(h,k)/\|(h,k)\| = 1 - \cos\frac{h^2}{k}$, $k \neq 0$. Tomando-se as sequências $(h_i,k_i) = (0, 1/i)$ e $(h_j,k_j) = (\sqrt{\pi/2j}, 1/j)$, tem-se $(h_i,k_i) \to 0$ e $(h_j,k_j) \to 0$. No entanto, $\lim(1 - \cos\frac{h_i^2}{k_i}) = 0$ e $\lim(1 - \cos\frac{h_j^2}{k_j}) = 1$, donde se conclui que o limite $\lim_{(h,k)\to(0,0)} \frac{r(h,k)}{\|(h,k)\|}$ não existe e, portanto, que f não é diferenciável em $(0,0)$.

3. (i) Suponha que x_0 seja um ponto de máximo local de f e considere um aberto $V \subset U$, tal que $f(x_0) \geq f(x)\,\forall x \in V$. Então, dado $h \in \mathbb{R}^n$, tomando-se $\epsilon > 0$ suficientemente pequeno, tem-se $x_0 + th \in V\,\forall t \in (-\epsilon, \epsilon)$, donde, para tais valores de t, $f(x_0 + th) \leq f(x_0)$. Logo, tomando-se limites laterais, tem-se

$$\lim_{t\to 0^+} \frac{f(x_0+th) - f(x_0)}{t} \leq 0 \quad \text{e} \quad \lim_{t\to 0^-} \frac{f(x_0+th) - f(x_0)}{t} \geq 0,$$

donde $f'(x_0)h = \frac{\partial f}{\partial h}(x_0) = 0$. Um argumento análogo se aplica ao caso em que x_0 é um ponto de mínimo local de f.

(ii) Uma vez que $\overline{U}$ é compacto e f é contínua, pelo Teorema de Weierstrass, f atinge seus valores máximo e mínimo, respectivamente, em pontos $x_0, y_0 \in \overline{U}$. Porém, f é constante em ∂U. Desta forma, a menos que f seja constante em $\overline{U}$, o que implicaria $f'(x) = 0\,\forall x \in U$, devemos ter $x_0 \in U$ ou $y_0 \in U$, donde, pelo resultado do item (i), se conclui que $f'(x_0) = 0$ ou $f'(y_0) = 0$.

(iii) Dado $u \in S^{n-1}$, temos que

$$0 < f'(u)u = \frac{\partial f}{\partial u}(u) = \lim_{t\to 0} \frac{f((1+t)u) - f(u)}{t} \cdot$$

Assim (vide Exercício 28 – Capítulo 3), existe $\epsilon \in (0,1)$, tal que

$$\frac{f((1+t)u) - f(u)}{t} > 0 \;\; \forall t \in (-\epsilon, \epsilon) - \{0\}.$$

Tomando-se $t \in (-\epsilon, 0)$, tem-se, então, $f((1+t)u) - f(u) < 0$. Uma vez que $\|(1+t)u\| = 1+t < 1$, tem-se, em particular, que o valor mínimo da restrição de f à bola (compacta) $B[0,1]$ não é atingido em nenhum ponto de S^{n-1}. Logo, $f|_{B[0,1]}$ tem um ponto de mínimo x no aberto $B(0,1)$, donde, pelo item (i), a derivada de f em x é nula.

4. Seja $f : \mathbb{R}^n \to \mathbb{R}$ uma função positivamente homogênea. Dado $x \neq 0$, tem-se, pela continuidade de f em $x = 0$ e pelo fato de f preservar multiplicação por escalar (positivo), que

$$f(0) = f\left(\lim_{k\to\infty} \frac{x}{k}\right) = \lim_{k\to\infty} f\left(\frac{x}{k}\right) = \lim_{k\to\infty}\left(\frac{1}{k} f(x)\right) = 0.$$

Além disso, para $t > 0$ e $x \in \mathbb{R}^n$, segue-se da diferenciabilidade de f em $x = 0$ a validez das igualdades

$$tf(x) = f(tx) = f(0) + f'(0)tx + r(t), \quad \lim_{t\to 0_+} \frac{r(t)}{t} = 0.$$

Logo, dividindo-se por t e passando-se ao limite com $t \to 0_+$, obtém-se $f(x) = f'(0)x$, donde f é linear. A função $f : \mathbb{R}^2 \to \mathbb{R}$ dada, claramente, satisfaz as condições do exercício, porém, não é linear. Logo, não pode ser diferenciável em $(0,0)$.

5. Seja (x_k) uma sequência em $f^{-1}(\{b\})$, tal que $x_k \to a$, $x_k \neq a$. Como f é contínua e $a \in U$, tem-se $f(a) = b$. Fazendo-se $h_k = x_k - a \neq 0$, temos que $h_k \to 0$ e $0 = f(a + h_k) - f(a) = f'(a)h_k + r(h_k)$, $\lim \frac{r(h_k)}{\|h_k\|} = 0$. Passando-se a uma subsequência, se necessário, podemos supor que $h_k/\|h_k\| \to u \in S^{n-1}$. Dividindo-se, então, ambos os membros da igualdade acima por $\|h_k\|$ e tomando-se o limite quando $k \to \infty$, obtém-se $f'(a)u = 0$. Logo, $f'(a)$ não é injetiva.

6. Dado $h \in \mathbb{R}^n$, tem-se $f(x + th) - f(x) = f'(x)th + r(t)$, $\lim_{t\to 0} \frac{r(t)}{t} = 0$. Logo, $\|f'(x)h + \frac{r(t)}{t}\| = \frac{\|f'(x)th + r(t)\|}{|t|} = \frac{\|f(x+th) - f(x)\|}{|t|} \leq \mu|t|\|h\|^2$, em que, na última desigualdade, usamos a hipótese sobre f. Tomando-se, então, o limite com $t \to 0$, obtém-se $\|f'(x)h\| = 0$, donde $f'(x) = 0$.

7. Tome $0 < \epsilon < 1 - \mu$. Uma vez que f é diferenciável em 0 e $f(0) = 0$, tem-se $f(x) = f'(0)x + r(x)$, em que $\lim_{x\to 0} r(x)/\|x\| = 0$. Assim, para o ϵ tomado, existe $\delta > 0$ satisfazendo $\|x\| < \delta \Rightarrow \|r(x)\| < \epsilon\|x\|$. Logo, $\|f(x)\| \leq \|f'(0)\|\,\|x\| + \|r(x)\| = \mu\|x\| + \|r(x)\| < (\mu + \epsilon)\|x\| < \|x\| < \delta$. Desta forma, a bola aberta $B = B(0,\delta)$ é tal que $f(B) \subset B$.

8. Dado $u \in S^{n-1}$, façamos $h_k = \frac{1}{k}u$, $k \in \mathbb{N}$. Então, supondo-se φ diferenciável na origem, tem-se $\varphi(h_k) = \varphi'(0)h_k + r(h_k)$, em que $\lim r(h_k)/\|h_k\| = 0$. Logo, $\|h_k\| f(h_k/\|h_k\|) = \varphi'(0)h_k + r(h_k)$. Dividindo-se ambos os membros dessa igualdade por $\|h_k\|$ e passando-se ao limite quando $k \to \infty$, obtém-se $f(u) = \varphi'(0)u$. Reciprocamente, suponhamos que exista uma transformação linear $T : \mathbb{R}^n \to \mathbb{R}^m$, tal que $T|_{S^{n-1}} = f$. Neste caso, para todo $x \neq 0$, tem-se $\varphi(x) = \|x\| f(x/\|x\|) = \|x\| T(x/\|x\|) = Tx$, donde φ é diferenciável no ponto 0 e $\varphi'(0) = T$.

9. (i) Pelas considerações do Exemplo 64, temos que $\nabla \det(I) = I$. Assim, dada uma matriz $H \in M(n)$, tem-se $\det'(I)H = \langle \nabla \det(I), H\rangle = \langle I, H\rangle = \text{traço}(HI^*)$ $= \text{traço}(H)$.

   (ii) Temos que $\det'(X) = 0$ se, e somente se, $\nabla \det(X) = 0$. Esta igualdade, por sua vez (vide Exemplo 64), equivale a $\det X_{ij} = 0\ \forall i, j \in \{1, \dots, n\}$, o que ocorre se, e somente se, o posto de X é menor que $n - 1$.

10. Escrevendo-se $\mu(x) = 1/(\|x\|^2 + 1)$, tem-se que as coordenadas de f são a aplicação $f_1(x) = 2\mu(x)x$ e a função $f_2(x) = \mu(x)(\|x^2\| - 1) = 1 - 2\mu(x)$ (observe que $\mu(x)\|x\|^2 = 1 - \mu(x)$). Uma vez que a função $x \mapsto \|x\|^2$ é diferenciável (Exemplo 56), segue-se das propriedades operatórias das aplicações diferenciáveis que μ, f_1 e f_2 são diferenciáveis. Assim, pela Proposição 46, f é diferenciável e, para quaisquer $x, h \in \mathbb{R}^n$, tem-se $f'(x)h = (f_1'(x)h, f_2'(x)h)$. Diferenciando-se a função μ, obtém-se, $\mu'(x)h = -2(\mu(x))^2\langle x, h\rangle$. Desta forma, $f_1'(x)h = 2(\mu'(x)h)x + 2\mu(x)h = 2\mu(x)(h - 2\mu(x)\langle x, h\rangle x)$ e $f_2'(x)h = -2\mu'(x)h =$

$4(\mu(x))^2\langle x,h\rangle$. Escrevendo-se $\mu=\mu(x)$, tem-se $|f_2'(x)h|^2=16\mu^4\langle x,h\rangle^2$ e

$$\begin{aligned}\|f_1'(x)h\|^2 &= 4\mu^2\langle h-2\mu\langle x,h\rangle x, h-2\mu\langle x,h\rangle x\rangle\\ &= 4\mu^2(\|h\|^2-4\mu\langle x,h\rangle^2+4\mu^2\langle x,h\rangle^2\|x\|^2)\\ &= 4\mu^2(\|h\|^2-4\mu\langle x,h\rangle^2(1-\mu\|x\|^2)\\ &= 4\mu^2(\|h\|^2-4\mu^2\langle x,h\rangle^2).\end{aligned}$$

Logo, $\|f'(x)h\|^2=\|f_1'(x)h\|^2+\|f_2'(x)h\|^2=4\mu^2\|h\|^2$, donde $\|f'(x)h\|=2\mu\|h\|$. Daí, segue-se que $f'(x):\mathbb{R}^n\to f'(x)(\mathbb{R}^n)$ é um isomorfismo conforme. Além disso, temos que $\langle f'(x)h,f(x)\rangle=\langle 2\mu(h-2\mu\langle x,h\rangle x),2\mu x\rangle+4\mu^2\langle x,h\rangle(1-2\mu)=8\mu^2\langle x,h\rangle(1-\mu\|x\|^2-\mu)=0$, donde se infere que $f'(x)(\mathbb{R}^n)\subset\{f(x)\}^\perp$ e, daí, que $f'(x)(\mathbb{R}^n)=\{f(x)\}^\perp$, pois as dimensões de $f'(x)(\mathbb{R}^n)$ e $\{f(x)\}^\perp$ são, ambas, iguais a n.

11. Basta observar que os vetores-coluna da matriz jacobiana de f num ponto $(x,y)\neq(0,0)$ são $2(x,0,x+y)$ e $2(0,y,x+y)$ e que estes são linearmente independentes. Logo, para todo $(x,y)\in U$, a matriz jacobiana de f em (x,y) tem posto 2, donde $f'(x,y):\mathbb{R}^2\to\mathbb{R}^3$ é injetiva.

12. Dado $h\in\mathbb{R}^n$, $\|h\|=1$, temos, pela definição de gradiente e pelo Teorema 1 do Capítulo 1 que

$$\frac{\partial f}{\partial h}(x)=f'(x)h=\langle\nabla f(x),h\rangle\leq\|\nabla f(x)\|\,\|h\|=\|\nabla f(x)\|,$$

em que a igualdade ocorre se, e somente se, $h=\frac{\nabla f(x)}{\|\nabla f(x)\|}$.

13. (i) Procedendo-se como no Exemplo 66, página 123, obtém-se facilmente

$$f''(x,y)[(h_1,h_2),(k_1,k_2)]=(2h_1k_1,2h_2k_2)\ \forall(x,y),(h_1,h_2),(k_1,k_2)\in\mathbb{R}^2.$$

(ii) Vimos anteriormente que, dado $X\in\mathrm{L}(\mathbb{R}^n)$, tem-se

$$f'(X)H=XHX+X^2H+HX^2,\ H\in\mathrm{L}(\mathbb{R}^n).$$

Fixando-se $H\in\mathrm{L}(\mathbb{R}^n)$ e considerando-se $g_1:\mathrm{L}(\mathbb{R}^n)\to\mathrm{L}(\mathbb{R}^n)$, dada por $g_1(X)H=XHX$, tem-se

$$(g_1(X+K)-g_1(X))H=XHK+KHX+KHK.$$

Definindo-se, então, $T_1(K),R_1(K)\in\mathrm{L}(\mathbb{R}^n)$ por

$$T_1(K)H=XHK+KHX\quad\text{e}\quad R_1(K)H=KHK,$$

tem-se que T_1 é linear com respeito à variável K e $\|R_1(K)\|\leq\|K\|^2$, isto é, $\|R_1(K)\|/\|K\|\leq\|K\|$. Logo, $\lim_{K\to0}R_1(K)/\|K\|=0$, donde g_1 é diferenciável e $g_1'(X)K=T_1(K)$.

De forma análoga, deduz-se que as aplicações $g_2,g_3:\mathrm{L}(\mathbb{R}^n)\to\mathrm{L}(\mathbb{R}^n)$, dadas por $g_2(X)H=X^2H$ e $g_3(X)H=HX^2$, cumprem $g_2'(X)K=T_2(K)$ e $g_3'(X)K=T_3(K)$, em que $T_2(K)H=XKH+KXH$ e $T_3(K)H=HXK+HKX$. Logo,

$$\begin{aligned}f''(X)(K,H) &= (g_1'(X)K+g_2'(X)K+g_3'(X)K)H\\ &= XHK+KHX+XKH+KXH+HXK+HKX.\end{aligned}$$

14. Dado $x \in \mathbb{R}^3$, um cálculo direto nos dá

$$\operatorname{hess} f(x) = \left(\frac{\partial^2 f}{\partial x_i \partial x_j}(x)\right) = \begin{pmatrix} 6 & 0 & 2 \\ 0 & 4 & 0 \\ 2 & 0 & 2 \end{pmatrix}.$$

Daí, fazendo-se $h = (h_1, h_2, h_3)$ e $k = (k_1, k_2, k_3)$, obtém-se

$$f''(x)(h,k) = \langle Th, k\rangle = 6h_1k_1 + 2h_3k_1 + 4h_2k_2 + 2h_1k_3 + 2h_3k_3,$$

em que T é o operador em $\mathbb{R}^3$ associado a $\operatorname{hess} f(x)$. Logo, para todo $h \neq 0$,

$$f''(x)(h,h) = 2(3h_1^2 + 2h_1h_3 + 2h_2^2 + h_3^2) = 2(2h_1^2 + (h_1 + h_3)^2 + 2h_2^2) > 0.$$

15. (i) Seja $x \in U$ um ponto crítico não-degenerado de f e suponhamos, por absurdo, que x não seja um ponto crítico isolado, isto é, que x seja o limite de uma sequência (x_k) em $U - \{x\}$, cujos termos são pontos críticos de f. Fazendo-se $h_k = x_k - x \neq 0$, temos que $h_k \to 0$ e

$$0 = f'(x_k) = f'(x + h_k) = f''(x)h_k + r(h_k), \ \lim \frac{r(h_k)}{\|h_k\|} = 0.$$

Passando-se a uma subsequência, se necessário, podemos supor que $\frac{h_k}{\|h_k\|} \to u$, $\|u\| = 1$. Daí, obtém-se

$$f''(x)u = \lim \left(f''(x)\frac{h_k}{\|h_k\|} + \frac{r(h_k)}{\|h_k\|}\right) = 0,$$

contrariando, assim, o fato de $\operatorname{hess} f(x)$ ter determinante não-nulo.

(ii) Conforme constatamos no Exemplo 56, a função $\varphi(x) = \|x\|^2$ é diferenciável em $\mathbb{R}^n$ e sua derivada num ponto $x \in \mathbb{R}^n$ é a aplicação linear $h \to 2\langle x, h\rangle$. Segue-se, portanto, das propriedades operatórias da derivada, que a função $f(x) = 2\|x\|^2 - \|x\|^4 = 2\varphi(x) - (\varphi(x))^2$ é diferenciável e, para quaisquer $x, h \in \mathbb{R}^n$, tem-se $f'(x)h = 2\varphi'(x)h - 2\varphi(x)\varphi'(x)h = 4\langle x, h\rangle(1 - \|x\|^2)$. Daí, conclui-se facilmente que o conjunto dos pontos críticos de f é o compacto $K = S^n \cup \{0\}$. Pelo estabelecido em (i), nenhum ponto de S^n é um ponto crítico não-degenerado de f, pois nenhum ponto de S^n é isolado. Agora, observando-se que $f'(x) = 2(1 - \varphi(x))\varphi'(x)$ e que $\varphi' : U \to \mathrm{L}(\mathbb{R}^n, \mathbb{R})$ é linear, tem-se que f é duas vezes diferenciável e, para quaisquer $x, h \in \mathbb{R}^n$, vale $f''(x)h = 2(-(\varphi'(x)h)\varphi'(x) + (1 - \varphi(x))\varphi''(x)h) = 2(-(\varphi'(x)h)\varphi'(x) + (1 - \varphi(x))\varphi'(h))$, donde $f''(x)(h,k) = 2(-(\varphi'(x)h)(\varphi'(x)k) + (1 - \varphi(x))(\varphi'(h)k)$, isto é, para quaisquer $x, h, k \in \mathbb{R}^n$, tem-se

$$f''(x)(h,k) = 2(-4\langle x, h\rangle\langle x, k\rangle + 2(1 - \|x\|^2)\langle h, k\rangle).$$

Em particular, $\frac{\partial^2 f}{\partial x_i \partial x_j}(0) = f''(0)(e_i, e_j) = 4\langle e_i, e_j\rangle$. Segue-se que $\operatorname{hess} f(0) = 4I$, em que I é a matriz identidade de $\mathbb{R}^n$. Logo, 0 é um ponto crítico não-degenerado de f.

16. Consideremos, para cada $i \in \{1, \ldots n\}$, a função $\varphi_i(x) = |x_i| = |\langle x, e_i\rangle|$, $x \in \mathbb{R}^n$, e o aberto $U_i = \{x \in \mathbb{R}^n; \langle x, e_i\rangle \neq 0\}$. Conforme verificamos no Exemplo 72, tem-se, para quaisquer $x \in U_i$ e $h \in \mathbb{R}^n$, $\varphi_i'(x)h = \frac{x_ih_i}{|x_i|} \cdot$ Daí,

segue-se que $f = \varphi_1 + \cdots + \varphi_n$ é diferenciável no aberto $U = U_1 \cap \ldots \cap U_n$ e

$$(*) \quad f'(x)h = \sum_{i=1}^{n} \varphi_i'(x)h = \sum_{i=1}^{n} \frac{x_i h_i}{|x_i|} \quad \forall x \in U,\ h \in \mathbb{R}^n.$$

Agora, dado $x \in \mathbb{R}^n - U$, tem-se $x_i = \langle x, e_i \rangle = 0$ para algum $i \in \{1, \ldots, n\}$. Neste caso, para todo $t \neq 0$, tem-se

$$\frac{f(x + te_i) - f(x)}{t} = \frac{|t|}{t},$$

donde se conclui facilmente que a derivada parcial $\frac{\partial f}{\partial x_i}(x)$ não existe e, portanto, que f não é diferenciável em x. Segue-se que f é diferenciável em $x \in \mathbb{R}^n$ se, e somente se, $x \in U$, sendo a derivada de f em U dada pela igualdade $(*)$.

17. Dados $x \in U$ e $h \in \mathbb{R}^n$, tem-se, pela definição de gradiente e pela Regra da Cadeia, que $\langle \nabla(\lambda \circ f)(x), h \rangle = (\lambda \circ f)'(x)h = \lambda'(f(x))f'(x)h = \langle \nabla\lambda(f(x)), f'(x)h \rangle$ $= \langle (f'(x))^* \nabla\lambda(f(x)), h \rangle$. Logo, $\nabla(\lambda \circ f)(x) = (f'(x))^* \nabla(\lambda(f(x))$.

18. Uma vez que a aplicação $f : \mathbb{R}^2 \to \mathbb{R}^2$ é holomorfa, suas coordenadas, f_1 e f_2, satisfazem as equações de Cauchy-Riemann,

$$\frac{\partial f_1}{\partial x_1}(x) = \frac{\partial f_2}{\partial x_2}(x) \qquad \text{e} \qquad \frac{\partial f_1}{\partial x_2}(x) = -\frac{\partial f_2}{\partial x_1}(x),$$

donde se infere que, para todo $x \in \mathbb{R}^2$, os gradientes de f_1 e f_2 em x são ortogonais. Temos também que $f_1 \circ \alpha_1$ e $f_2 \circ \alpha_2$ são constantes. Logo, pela Regra da Cadeia, devemos ter $f_1'(\alpha_1(t))\alpha_1'(t) = f_2'(\alpha_2(s))\alpha_2'(s) = 0$, isto é, fazendo-se $x = \alpha_1(t) = \alpha_2(s)$, tem-se $\langle \nabla f_1(x), \alpha_1'(t) \rangle = \langle \nabla f_2(x), \alpha_2'(s) \rangle = 0$. Daí, segue-se que existem λ, $\mu \in \mathbb{R}$, tais que, $\alpha_1'(t) = \lambda \nabla f_2(x)$ e $\alpha_2'(s) = \mu \nabla f_1(x)$, donde $\alpha_1'(t)$ e $\alpha_2'(s)$ são ortogonais.

19. Por hipótese, a função $g(x) = \langle f(x), f(x) \rangle$, $x \in U$, é constante. Se g é identicamente nula, então f é identicamente nula e o resultado é imediato. Suponhamos, então, que, para todo $x \in U$, tenhamos $f(x) \neq 0$. Neste caso, para quaisquer $x \in U$ e $h \in \mathbb{R}^n$, tem-se (veja o Exemplo 70) $0 = g'(x)h = 2\langle f'(x)h, f(x) \rangle$. Segue-se que o conjunto imagem de $f'(x)$ está contido no complemento ortogonal de $f(x)$ em $\mathbb{R}^n$, donde $f'(x)$ não é sobrejetiva e, portanto, tem determinante nulo.

20. Derivando-se, com respeito à variável t, a identidade $f(tx) = t^r f(x)$, tem-se $f'(tx)x = rt^{r-1}f(x)$, donde, fazendo-se $t = 1$, obtém-se $f'(x)x = rf(x)$.

21. Temos, pela Regra da Cadeia, que $(g \circ f)' = \varphi \circ \psi$, em que φ e ψ são as aplicações definidas na demonstração da Proposição 47. Ainda pela Regra da Cadeia, temos $(g \circ f)''(x) = (\varphi \circ \psi)'(x) = \varphi'(\psi(x))\psi'(x)$. Segue-se daí e da bilinearidade de φ que, para todo $h \in \mathbb{R}^n$,

$$\begin{aligned} (g \circ f)''(x)h &= \varphi'(f'(x), g'(f(x))(f''(x)h, g''(f(x))f'(x)h) \\ &= \varphi(f'(x), g''(f(x))f'(x)h) + \varphi(f''(x)h, g'(f(x))) \\ &= (g''(f(x))f'(x)h)f'(x) + g'(f(x))f''(x)h. \end{aligned}$$

Logo, como forma bilinear, $(g \circ f)''(x)$ assume a forma

$$(g \circ f)''(x)(h, k) = g''(f(x))(f'(x)h, f'(x)k) + g'(f(x))f''(x)(h, k).$$

22. Escrevendo-se $\varphi(x) = (f \circ T)(x) = f(Tx), x \in \mathbb{R}^n$, tem-se, pela Regra da Cadeia, que φ é duas vezes diferenciável, pois f e T o são. Além disso,
$$\varphi'(x) = f'(Tx)T'(x) = f'(Tx)T,$$
donde $\varphi''(x)h = (f''(Tx)Th)T$ e, portanto,
$$\varphi''(x)(h,k) = f''(Tx)(Th, Tk)\ \forall h, k \in \mathbb{R}^n.$$
Seja $Z \in \mathrm{L}(\mathbb{R}^n)$ o operador em $\mathbb{R}^n$ cuja matriz com respeito à base canônica de $\mathbb{R}^n$ é
$$(\operatorname{hess} f)(Tx) = \left(\frac{\partial^2 f}{\partial x_i \partial x_j}(Tx)\right).$$
Lembrando que T^*T é e aplicação identidade de $\mathbb{R}^n$ e que, para quaisquer operadores T_1, $T_2 \in \mathrm{L}(\mathbb{R}^n)$, vale a igualdade $\text{traço}(T_1T_2) = \text{traço}(T_2T_1)$, tem-se
$$\begin{aligned}\Delta(f \circ T)(x) &= \Delta\varphi(x) = \sum_{i=1}^{n} \varphi''(x)(e_i, e_i) = \sum_{i=1}^{n} f''(Tx)(Te_i, Te_i)\\ &= \sum_{i=1}^{n} \langle ZTe_i, Te_i\rangle = \sum_{i=1}^{n} \langle T^*ZTe_i, e_i\rangle = \text{traço}(T^*ZT)\\ &= \text{traço}(Z) = \Delta f(Tx),\end{aligned}$$
isto é, $\Delta(f \circ T) = \Delta f \circ T$.

23. Façamos $A(t) = [x_1(t) \ldots x_n(t)]$, em que $x_j(t)$ é o j-ésimo vetor-coluna de $A(t)$, isto é, $x_j(t) = \sum_{i=1}^n a_{ij}(t)e_i$, donde $x_j'(t) = \sum_{i=1}^n a_{ij}'(t)e_i$. Sendo $B(t) = (b_{ij}(t))$ a inversa de $A(t)$, temos, pela fórmula dos cofatores, que as entradas de $B(t)$ satisfazem
$$b_{ji}(t) = \frac{\det[x_1(t) \ldots x_{j-1}(t)\ e_i\ x_{j+1}(t) \ldots x_n(t)]}{\det A(t)}.$$
Agora, pela Regra da Cadeia e pela n-linearidade da função determinante, a derivada de $f(t) = \log(\det A(t))$ é
$$\begin{aligned} f'(t) &= \frac{1}{\det A(t)} \det{}'(A(t))A'(t)\\ &= \frac{1}{\det A(t)} \det{}'[x_1(t) \ldots x_n(t)][x_1'(t) \ldots x_n'(t)]\\ &= \frac{1}{\det A(t)} \sum_{j=1}^{n} \det[x_1(t) \ldots x_j'(t) \ldots x_n(t)]\\ &= \frac{1}{\det A(t)} \sum_{j=1}^{n} \det[x_1(t) \ldots \sum_{i=1}^{n} a_{ij}'(t)e_i \ldots x_n(t)]\\ &= \sum_{j=1}^{n}\sum_{i=1}^{n} a_{ij}'(t) \frac{\det[x_1(t) \ldots e_i \ldots x_n(t)]}{\det A(t)}\\ &= \sum_{i,j=1}^{n} a_{ij}'(t) b_{ji}(t).\end{aligned}$$

24. Verifica-se facilmente que, dado $(\xi, \eta) \in \mathbb{R}^2$,
$$(x, y) = (e^\xi + e^\eta, e^\xi - e^\eta)$$
é o único ponto de U que satisfaz a igualdade $f(x, y) = (\xi, \eta)$, donde f é bijetiva e $f^{-1}(\xi, \eta) = (x, y)$. Claramente, as coordenadas de f e f^{-1} são diferenciáveis, donde f é um difeomorfismo.

25. Seja
$$\begin{array}{rccc} f: & \mathbb{R}^n - \{a\} & \to & \mathbb{R}^n - \{a\} \\ & x & \mapsto & r^2 \frac{x-a}{\|x-a\|^2} + a \end{array}$$
a inversão de $\mathbb{R}^n - \{a\}$ com respeito à esfera $S[a, r] \subset \mathbb{R}^n$. Constatamos anteriormente que f é um homeomorfismo cujo inverso é a própria aplicação f. Além disso, é imediato que f é diferenciável, donde se conclui que f é um difeomorfismo. Fazendo-se, então, $\mu(x) = \frac{1}{\|x-a\|^2}$, $x \in \mathbb{R}^n - \{a\}$, e $u = u(x) = \frac{x-a}{\|x-a\|}$, tem-se $f(x) = r^2\mu(x)(x - a) + a$ e
$$\mu'(x)h = -2\frac{\langle x - a, h\rangle}{\|x - a\|^4} = -\frac{2\mu(x)}{\|x - a\|}\langle u, h\rangle.$$
Desta forma,
$$f'(x)h = r^2((\mu'(x)h)(x - a) + \mu(x)h) = r^2\mu(x)(h - 2\langle u, h\rangle u).$$
Daí, um cálculo direto nos dá $\|f'(x)h\| = r^2\mu(x)\|h\|$, donde f é conforme.

Capítulo 5

1. A aplicação $g(x) = f(x) - Tx$, $x \in U$, é diferenciável e, para todo $x \in U$, $g'(x) = f'(x) - T = 0$. Uma vez que U é conexo, segue-se do Corolário 5 que g é constante, isto é, existe $a \in \mathbb{R}^n$, tal que $g(x) = a\,\forall x \in U$, donde $f(x) = Tx + a\,\forall x \in U$.

2. É evidente que F é contínua nos pontos (x, t), $t \neq 0$. Agora, dado $(x, 0) \in \mathbb{R}^2$, seja (x_k, t_k) uma sequência em $\mathbb{R}^2$, tal que $(x_k, t_k) \to (x, 0)$ e $t_k \neq 0$. Então, pelo Teorema do Valor Médio, para cada $k \in \mathbb{N}$, existe $\theta_k \in (0, 1)$, tal que $F(x_k, t_k) = \frac{f(x_k+t_k)-f(x_k)}{t_k} = f'(x_k + \theta_k t_k)$. Uma vez que f' é contínua, segue-se que $\lim F(x_k, t_k) = f'(x) = F(x, 0)$, donde F é contínua em $(x, 0)$.

3. Dados $x, y \in \mathbb{R}^n$, segue-se do Teorema do Valor Médio que $\|g(x) - g(y)\| \leq \lambda\|x - y\|$, donde g é uma contração. Então, pelo Teorema da Perturbação da Identidade, f é um homeomorfismo de $\mathbb{R}^n$ sobre o aberto $U = f(\mathbb{R}^n)$. Dados $y_1, y_2 \in U$, fazendo-se $x_1 = f^{-1}(y_1)$ e $x_2 = f^{-1}(y_2)$, tem-se $y_1 = f(x_1) = x_1 + g(x_1)$ e $y_2 = f(x_2) = x_2 + g(x_2)$, isto é, $x_1 = y_1 - g(x_1)$ e $x_2 = y_2 - g(x_2)$. Sendo assim, temos
$$\begin{aligned} \|f^{-1}(y_1) - f^{-1}(y_2)\| &= \|x_1 - x_2\| \\ &= \|y_1 - g(x_1) - (y_2 - g(x_2)\| \\ &\leq \|y_1 - y_2\| + \|g(x_1) - g(x_2)\| \\ &\leq \|y_1 - y_2\| + \lambda\|x_1 - x_2\| \\ &= \|y_1 - y_2\| + \lambda\|f^{-1}(y_1) - f^{-1}(y_2)\|, \end{aligned}$$

donde $\|f^{-1}(y_1)-f^{-1}(y_2)\| \le \frac{1}{1-\lambda}\|y_1-y_2\|$. Logo, f^{-1} é lipschtziana e, em particular, uniformemente contínua. Segue-se, então, do Exercício 13 do Capítulo 3, que $U=\mathbb{R}^n$ e, portanto, que f é um homeomorfismo.

4. Designando-se por $J_f(x)$ a matriz jacobiana de f em x, temos, pela igualdade $\lim_{x\to 0}\frac{\partial f}{\partial x_i}(x)=0$, que $\lim_{x\to 0}J_f(x)=0$, donde $\lim_{x\to 0}f'(x)=0$. Logo, existe $\delta>0$, tal que f' é limitada em $V=B(0,\delta)-\{0\}$. Além disso, fazendo-se, para cada $x\in V$, $\mu(x)=\sup_{y\in(0,x)}\|f'(y)\|$, tem-se $\lim_{x\to 0}\mu(x)=0$. Agora, uma vez que f é contínua em $x=0$ e diferenciável em $\mathbb{R}^n-\{0\}$, pelo Teorema do Valor Médio, todo $x\in V$ cumpre $\|f(x)-f(0)\|\le \mu(x)\|x\|$. Logo, $\lim_{x\to 0}\frac{\|f(x)-f(0)\|}{\|x\|}=0$, donde f é diferenciável em $x=0$ e $f'(0)=0$.

5. Derivando-se ambos os membros de $\langle f(x),\frac{\partial g}{\partial x_i}(x)\rangle=0$ com respeito à variável x_j, obtém-se
$$\left\langle \frac{\partial f}{\partial x_j}(x),\frac{\partial g}{\partial x_i}(x)\right\rangle+\left\langle f(x),\frac{\partial^2 g}{\partial x_j\partial x_i}(x)\right\rangle=0\ \forall i,j\in\{1,\dots,n\}.$$
Desta igualdade, obtém-se uma segunda intercambiando-se as variáveis i e j. Subtraindo-se, membro a membro, uma da outra e considerando-se o Teorema de Schwarz, chega-se finalmente a
$$\left\langle \frac{\partial f}{\partial x_j}(x),\frac{\partial g}{\partial x_i}(x)\right\rangle-\left\langle \frac{\partial f}{\partial x_i}(x),\frac{\partial g}{\partial x_j}(x)\right\rangle=0,$$
donde se infere que a matriz A, dada, é simétrica.

6. Fixemos $(x,y)\in\mathbb{R}^2$ e, dado $t\in\mathbb{R}$, façamos
$$F(t)=f(x+t,y+t)+f(x-t,y-t)-f(x+t,y-t)-f(x-t,y+t).$$
Derivando-se a função F duas vezes no ponto $t=0$, obtém-se $F''(0)=2f''(x,y)(v,v)-2f''(x,y)(w,w)$, em que $v=(1,1)$ e $w=(-1,1)$. Porém, F é, por hipótese, identicamente nula. Logo, para todo $t\in\mathbb{R}$, tem-se $F''(t)=0$, o que nos dá $f''(x,y)(v,v)=f''(x,y)(w,w)$. Observando-se, então, que $e_1=\frac{v-w}{2}$ e $e_2=\frac{v+w}{2}$, segue-se daí e do Teorema de Schwarz que
$$\frac{\partial^2 f}{\partial x\partial y}=f''(x,y)(e_1,e_2)=\frac{1}{2}f''(x,y)(v-w,v+w)=0.$$

7. Temos que as coordenadas de uma aplicação holomorfa $f:\mathbb{R}^2\to\mathbb{R}^2$, f_1 e f_2, satisfazem, para todo $x\in\mathbb{R}^2$, as equações de Cauchy-Riemann
$$\frac{\partial f_1}{\partial x_1}(x)=\frac{\partial f_2}{\partial x_2}(x)\qquad \text{e}\qquad \frac{\partial f_1}{\partial x_2}(x)=-\frac{\partial f_2}{\partial x_1}(x)\,.$$
Derivando-se a primeira delas com relação a x_1, a segunda com relação a x_2, e adicionando-se membro a membro, obtém-se
$$\Delta f_1(x)=\frac{\partial^2 f_1}{\partial x_1^2}(x)+\frac{\partial^2 f_1}{\partial x_2^2}(x)=\frac{\partial^2 f_2}{\partial x_1\partial x_2}(x)-\frac{\partial^2 f_2}{\partial x_2\partial x_1}(x).$$
Logo, pelo Teorema de Schwarz, para todo $x\in\mathbb{R}^2$, $\Delta f_1(x)=0$, isto é, f_1 é harmônica. De modo análogo verifica-se que f_2, igualmente, é uma função harmônica.

8. Dados $i, j \in \{1, \dots, n\}$, façamos, por simplicidade de notação, $u_i = u_i(x) = \frac{\partial f}{\partial x_i}(x)$ e $u_{ij} = u_{ij}(x) = \frac{\partial^2 f}{\partial x_i \partial x_j}(x)$. Temos, pelo Teorema de Schwarz, que $u_{ij} = u_{ji}$. Além disso, os vetores u_i são os vetores-coluna da matriz jacobiana de f em x. Uma vez que $f'(x)$ é ortogonal, temos que $\mathfrak{B} = \{u_1, \dots, u_n\}$ é uma base ortonormal de $\mathbb{R}^n$, isto é, para quaisquer $i, j \in \{1, \dots, n\}$, tem-se
$$\langle u_i, u_j \rangle = \delta_{ij} = \begin{cases} 1 & \text{se} \quad i = j \\ 0 & \text{se} \quad i \neq j. \end{cases}$$
Derivando-se esta igualdade com respeito a $k \in \{1, \dots, n\}$, obtém-se
$$\langle u_{ki}, u_j \rangle + \langle u_i, u_{kj} \rangle = 0$$
e, intercambiando-se k e i,
$$\langle u_{ik}, u_j \rangle + \langle u_k, u_{ij} \rangle = 0.$$
Subtraindo-se membro a membro estas duas últimas igualdades, obtém-se
$$\langle u_{ij}, u_k \rangle - \langle u_{kj}, u_i \rangle = 0.$$
Agora, derivando-se, com respeito a j, a igualdade $\langle u_i, u_k \rangle = \delta_{ik}$, chega-se a
$$\langle u_{ji}, u_k \rangle + \langle u_i, u_{jk} \rangle = 0.$$
Finalmente, adicionando-se, membro a membro, estas duas últimas igualdades, obtém-se, para quaisquer $i, j, k \in \{1, \dots, n\}$, $\langle u_{ij}, u_k \rangle = 0$. Logo, u_{ij} é ortogonal a cada um dos vetores da base $\mathfrak{B}$, donde $u_{ij} = 0$. Desta forma, para todo $x \in \mathbb{R}^n$, $f''(x) = 0$. Sendo $\mathbb{R}^n$ conexo, tem-se, em particular, que f' é constante, isto é, $f'(x) = T \ \forall x \in \mathbb{R}^n$, em que T é um operador ortogonal em $\mathbb{R}^n$. Segue-se, então, do Exercício 1, que, para todo $x \in \mathbb{R}^n$, $f(x) = Tx + a$, $a \in \mathbb{R}^n$, e, portanto, que f é uma isometria de $\mathbb{R}^n$ (vide Teorema 3 – Capítulo 1).

9. Suponhamos, por absurdo, que f possua um máximo local $a \in \mathbb{R}^n$. Neste caso, existe $\delta > 0$, tal que, para todo $h \in \mathbb{R}^n$ satisfazendo $\|h\| < \delta$, tem-se $f(a+h) - f(a) \leq 0$. Considerando-se isto e a Fórmula de Taylor de f em a (note que $f'(a) = 0$), obtém-se
$$0 \geq f(a+h) - f(a) = \frac{1}{2} f''(a)(h,h) + r(h), \quad \lim_{h \to 0} \frac{r(h)}{\|h\|^2} = 0.$$
Fazendo-se, então, $h = te_i$, $|t| < \delta$, tem-se
$$0 \geq \frac{f(a+te_i) - f(a)}{t^2} = \frac{1}{2} f''(a)(e_i, e_i) + \frac{r(te_i)}{t^2},$$
donde
$$0 \geq \sum_{i=1}^{n} \left(\frac{1}{2} f''(a)(e_i, e_i) + \frac{r(te_i)}{t^2} \right) = \frac{1}{2} \Delta f(a) + \sum_{i=1}^{n} \frac{r(te_i)}{t^2} \cdot$$
Agora, uma vez que $\Delta f(a) > 0$ e $\lim_{t \to 0} \frac{r(te_i)}{t^2} = 0$, podemos tomar δ suficientemente pequeno, de tal forma que, para todo $t \in (-\delta, \delta)$ e todo $i \in \{1, \dots, n\}$, valha $-\frac{\Delta f(a)}{2n} < \frac{r(te_i)}{t^2} < \frac{\Delta f(a)}{2n} \cdot$ Assim, teremos
$$0 \geq \frac{1}{2} \Delta f(a) + \sum_{i=1}^{n} \frac{r(te_i)}{t^2} > \frac{1}{2} \Delta f(a) - \frac{1}{2} \Delta f(a) = 0.$$

Segue-se desta contradição que f não possui máximos locais.

10. Suponha que $x \in \mathbb{R}^n$ seja um ponto crítico de f. Dado, então, $y \in \mathbb{R}^n$, $y \neq x$, seja $\alpha : \mathbb{R} \to \mathbb{R}^n$ a curva cujo traço é a reta que contém x e y, isto é, $\alpha(t) = x + t(y - x)$. Fazendo-se $v = y - x$ e $g(t) = f(\alpha(t))$, tem-se $g'(t) = f'(\alpha(t))\alpha'(t) = f'(\alpha(t))v$. Logo, $g''(t) = f''(\alpha(t))(v, v) > 0$, donde g' é crescente. Sendo assim, $0 = f'(x)v = g'(0) < g'(1) = f'(y)v$ e, portanto, y não é ponto crítico de f. Segue-se que f possui, no máximo, um ponto crítico.

11. Uma vez que g e a aplicação identidade de $\mathbb{R}^n$ são de classe C^1, temos que f é de classe C^1. Além disso, para todo $h \in \mathbb{R}^n$, $f'(x)h = h + g'(x)h$. Logo, $\|f'(x)h\| \geq \|h\| - \|g'(x)h\| \geq (1 - \lambda)\|h\|$. Como $1 - \lambda > 0$, temos que, para todo $x \in \mathbb{R}^n$, o núcleo de $f'(x)$ é trivial. Desta forma, $f'(x)$ é um isomorfismo, donde, pelo Teorema da Função Inversa, $f : \mathbb{R}^n \to \mathbb{R}^n$ é um difeomorfismo local. Porém, f é um homeomorfismo e, em particular, uma bijeção. Logo, f é um difeomorfismo.

12. Pelo Teorema da Função Inversa, existem intervalos abertos $I \ni x_0$ e $J \ni f(x_0)$, tais que $f|_I : I \to J = f(I)$ é um difeomorfismo de classe C^1. Seja $g = g(u)$ a inversa de $f|_I$, a qual, ainda pelo Teorema da Função Inversa, é de classe C^1. Tomando-se o aberto $V = I \times \mathbb{R} \subset \mathbb{R}^2$, o qual contém (x_0, y_0), e $(u, v) = \varphi(x, y) = (f(x), xf(x) - y) \in \varphi(V)$, tem-se $x \in I$ e $u = f(x)$, donde $x = g(u)$. Além disso, $v = xf(x) - y$, e, portanto, $y = v - ug(u)$. Segue-se, então, destas considerações, que a restrição $\varphi|_V : V \to \varphi(V)$ é invertível e $\varphi^{-1}(u, v) = (g(u), ug(u) - v)$.

13. Conforme constatamos anteriormente, para quaisquer $X, H \in \mathrm{L}(\mathbb{R}^n)$, tem-se $f'(X)H = XH + HX$. Em particular, $f'(I)H = 2H$, isto é, $f'(I) = 2I$, em que I é a aplicação identidade de $\mathbb{R}^n$. Logo, $f'(I)$ é um isomorfismo, donde, pelo Teorema da Função Inversa, f é um difeomorfismo de uma vizinhança de I em uma vizinhança de $f(I) = I$. A aplicação f, no entanto, não é um difeomorfismo local, pois $f'(0)$ é identicamente nula.

14. Seja $f : \mathrm{M}(n) \to \mathrm{M}(n)$ a aplicação definida por $f(X) = X^2 + X^*$. Temos, para quaisquer $X, H \in \mathrm{M}(n)$, que $f'(X)H = XH + HX + H^*$. Logo, $f'(0)H = H^*$, donde $f'(0)$ é um isomorfismo. Então, pelo Teorema da Função Inversa, f é um difeomorfismo de um aberto $V \ni 0$ em um aberto $W \ni f(0) = 0$. Tomando-se $\delta > 0$, tal que $B(0, \delta) \subset W$, temos que se $A \in \mathrm{M}(n)$ e $\|A\| < \delta$, então existe um único X em V, tal que $f(X) = A$, como desejado.

15. Segue-se diretamente da hipótese e do Teorema da Função Inversa que f é um difeomorfismo local. Em particular, f é aberta. Além disso, sendo f própria, pelo Exercício 21 do Capítulo 3, f é fechada. Logo, $f(\mathbb{R}^n)$ é fechado e aberto em $\mathbb{R}^n$. Uma vez que $\mathbb{R}^n$ é conexo e $f(\mathbb{R}^n) \neq \emptyset$, tem-se $f(\mathbb{R}^n) = \mathbb{R}^n$.

16. Dados $x, h \in \mathbb{R}^n$ e $t \in \mathbb{R}$, tem-se $\|f(x + th) - f(x)\| \geq \|th\|$. Logo,
$$\|f'(x)h\| = \left\| \lim_{t \to 0} \frac{f(x + th) - f(x)}{t} \right\| = \lim_{t \to 0} \left\| \frac{f(x + th) - f(x)}{t} \right\| \geq \|h\|.$$
Segue-se que o núcleo de $f'(x)$ é trivial e, portanto, que $f'(x)$ é um isomorfismo. Pelo Teorema da Função Inversa, f é um difeomorfismo local, sendo, desta

forma, uma aplicação aberta. Além disso, f é injetiva, pois, por hipótese, $\|f(x)-f(y)\| \geq \|x-y\| \, \forall x, y \in \mathbb{R}^n$. Provemos que f é também sobrejetiva. Para isto, basta mostrarmos que f é fechada (veja a solução do exercício anterior). Tomemos, pois, $F \subset \mathbb{R}^n$ fechado e uma sequência convergente $(f(x_k))$ em $f(F)$, isto é, $x_k \in F \, \forall k \in \mathbb{N}$. A sequência $(f(x_k))$ é, em particular, de Cauchy. Daí e da desigualdade $\|f(x_k)-f(x_l)\| \geq \|x_k - x_l\|$, válida para quaisquer $k, l \in \mathbb{N}$, segue-se que (x_k) é de Cauchy e, portanto, converge para um ponto $a \in F$. Uma vez que f é contínua, temos que $\lim f(x_k) = f(a) \in f(F)$, donde $f(F)$ é fechado. Segue-se que f é fechada e, desta forma, sobrejetiva. Desta forma, por ser um difeomorfismo local bijetivo, a aplicação f é um difeomorfismo.

17. Suponhamos, por absurdo, que x_0 seja um ponto de máximo da função $x \mapsto \|f(x)\|$, $x \in U$. Então, $f(x_0) \neq 0$. Caso contrário, f seria identicamente nula e não seria, desta forma, uma submersão. Temos, então, que x_0 é um ponto de máximo da função (diferenciável) $\varphi(x) = \|f(x)\|^2$ e, portanto, é um ponto crítico de φ. Assim, para todo $h \in \mathbb{R}^n$, $0 = \varphi'(x_0)h = 2\langle f'(x_0)h, f(x_0)\rangle$. Segue-se que o conjunto-imagem de $f'(x_0)$ está contido no complemento ortogonal de $f(x_0) \neq 0$ em $\mathbb{R}^n$. Em particular, $f'(x_0)$ não é sobrejetiva, o que contraria o fato de f ser uma submersão.

18. Observemos, inicialmente, que uma função f definida num aberto U de $\mathbb{R}^n$ é uma submersão em $x \in U$ se, e somente se, $\nabla f(x) \neq 0$. Com efeito, esta última condição equivale a $f'(x)e_i = \frac{\partial f}{\partial x_i}(x) \neq 0$ para algum $i \in \{1, \dots, n\}$, o que ocorre se, e somente se, $f'(x)$ é sobrejetiva. Agora, conforme verificamos anteriormente (Exemplo 64 – Capítulo 4), o gradiente da função determinante em $X \in \mathrm{M}(n)$ é a matriz dos cofatores de X. Assim, a função det é uma submersão em X se, e somente se, a sua matriz dos cofatores é não-nula, o que ocorre se, e somente se, o posto de X é maior que, ou igual a, $n-1$.

19. Uma vez que $f'(a)$ é injetiva, podemos supor, sem perda de generalidade, que $f'(a)(\mathbb{R}^n) = \mathbb{R}^n \times \{0\} \subset \mathbb{R}^{n+m}$. Definamos $F : U \times \mathbb{R}^m \subset \mathbb{R}^{n+m} \to \mathbb{R}^{n+m}$ por

$$F(x,y) = f(x) + (0,y) = f(P_1(x,y)) + P_2(x,y),$$

em que $P_1 : \mathbb{R}^{n+m} \to \mathbb{R}^n \times \{0\}$ e $P_2 : \mathbb{R}^{n+m} \to \{0\} \times \mathbb{R}^m$ são as projeções ortogonais de $\mathbb{R}^{n+m}$ sobre $\mathbb{R}^n \times \{0\}$ e $\{0\} \times \mathbb{R}^m$, respectivamente, e, por abuso de notação, identificam-se U e $U \times \{0\}$. Dado, então, $(h,k) \in \mathbb{R}^n \times \mathbb{R}^m$, tem-se

$$F'(a,0)(h,k) = f'(P_1(a,0))P_1'(a,0)(h,k) + P_2'(a,0)(h,k) = f'(a)h + (0,k),$$

donde se infere que o conjunto-imagem de $F'(a,0)$ é $\mathbb{R}^n \times \{0\} \oplus \{0\} \times \mathbb{R}^m = \mathbb{R}^{n+m}$. Segue-se que $F'(a,0)$ é sobrejetiva e, portanto, um isomorfismo. Logo, pelo Teorema da Função Inversa, existem abertos de $\mathbb{R}^{n+m}$, $V \ni (a,0)$ e $W \ni F(a,0) = f(a)$, tais que $F|_V : V \to W$ é um difeomorfismo. Desta forma, fazendo-se $G = (F|_V)^{-1}$ e $A = P_1(V) \subset U$, tem-se que A é aberto e, para todo $x \in A$,

$$(G \circ f)(x) = G(f(x)) = G(f(x) + (0,0)) = G(F(x,0)) = (x,0).$$

20. A função f dada, por ser polinomial, é de classe C^∞. Logo, pelo Teorema da Função Implícita, ξ também o é (vide Observação 18). No Exemplo 78, vimos

que

$$\xi_x = -\frac{f_x(x,y,\xi(x,y))}{f_z(x,y,\xi(x,y))} \quad \text{e} \quad \xi_y = -\frac{f_y(x,y,\xi(x,y))}{f_z(x,y,\xi(x,y))}\,.$$

Aplicando-se, então, a fórmula de derivação do quociente de funções e lembrando que, para uma dada função diferenciável $p = p(x,y,z)$, tem-se

$$\frac{\partial}{\partial x}(p(x,y,\xi(x,y)) = p_x(x,y,\xi(x,y)) + p_z(x,y,\xi(x,y))\xi_x(x,y)$$

e

$$\frac{\partial}{\partial y}(p(x,y,\xi(x,y)) = p_y(x,y,\xi(x,y)) + p_z(x,y,\xi(x,y))\xi_y(x,y),$$

obtém-se facilmente as igualdades

$$\xi_{xx} = \frac{f_x(f_{xz} + f_{zz}\xi_x) - f_z(f_{xx} + f_{xz}\xi_x)}{f_z^2}\,,$$

$$\xi_{xy} = \frac{f_x(f_{yz} + f_{zz}\xi_y) - f_z(f_{xy} + f_{xz}\xi_y)}{f_z^2}$$

e

$$\xi_{yy} = \frac{f_y(f_{yz} + f_{zz}\xi_y) - f_z(f_{yy} + f_{yz}\xi_y)}{f_z^2}\,.$$

Daí, segue-se que

$$\xi_{xx}(0,0) = -80, \quad \xi_{xy}(0,0) = -14 \quad \text{e} \quad \xi_{yy}(0,0) = -2.$$

Logo, a fórmula de Taylor de segunda ordem de ξ numa vizinhança de $(0,0)$ é (vide Exemplo 75)

$$\xi(x,y) = -40x^2 - 14xy - y^2 + 4x + y + 1 + r(x,y), \quad \lim_{(x,y)\to(0,0)} \frac{r(x,y)}{\|(x,y)\|^2} = 0.$$

21. Dados $x,y,t \in \mathbb{R}$, façamos $f(x,y,t) = xt^2 + e^{2t} + y$ e $p = (1,-1,0)$. A função $f : \mathbb{R}^3 \to \mathbb{R}$, assim definida, é de classe C^∞ e cumpre as igualdades $f(p) = 0$ e $\frac{\partial f}{\partial t}(p) = 1$. Logo, pelo Teorema da Função Implícita, existem abertos $A \ni (1,-1)$, de $\mathbb{R}^2$, e uma bola aberta $B(p,\epsilon) \subset \mathbb{R}^3$, tais que, para todo ponto $(x,y) \in A$, existe um único $t = t(x,y) \in (-\epsilon,\epsilon)$ que satisfaz $f(x,y,t) = 0$. Além disso, uma vez que f é de classe C^∞, o mesmo vale para a função $t = t(x,y)$.

22. Consideremos a função $F : \mathbb{R}^2 \to \mathbb{R}$, definida por $F(t,x) = f(x) - tg(x)$. Temos que $F(0,0) = 0$ e $\frac{\partial F}{\partial x}(0,0) = f'(0) \neq 0$. Logo, pelo Teorema da Função Implícita, existem $\delta > 0$ e uma função $x = x(t)$, definida em $(-\delta,\delta)$, tais que $F(t,x(t)) = 0$, isto é, $f(x(t)) - tg(x(t)) = 0$.

23. Seja $f : \mathbb{R} \times \mathrm{M}(n) \to \mathrm{M}(n)$ a aplicação definida por $f(t,A) = A^2 + tBA - I$. Então, f é de classe C^∞, $f(0,I) = 0$ e $\frac{\partial f}{\partial A}(t,A)$ é a aplicação linear $H \mapsto AH + HA + tBH$, $H \in \mathrm{M}(n)$. Em particular, $\frac{\partial f}{\partial A}(0,I)$ é o isomorfismo $H \mapsto 2H$. Segue-se, portanto, do Teorema da Função Implícita, que existe um intervalo aberto $(-\epsilon,\epsilon)$, tal que, para cada $t \in (-\epsilon,\epsilon)$, existe uma matriz $A = A(t)$ satisfazendo $f(t,A(t)) = 0$. Além disso, a curva $A(t)$, assim definida, é de classe C^∞ e cumpre $A(0) = I$. Por fim, derivando-se a igualdade $A(t)^2 + tBA(t) = I$ com respeito a t, obtém-se

$$A(t)A'(t) + A'(t)A(t) + BA(t) + tBA'(t) = 0,$$

donde $2A'(0) + B = 0$.

24. Definamos $f : \mathbb{R}^{n+1} \times \mathbb{R} \to \mathbb{R}$ por $f(x,t) = p_x(t) = x_0 + x_1 t + \cdots + x_n t^n$, em que $x = (x_0, \ldots, x_n) \in \mathbb{R}^{n+1}$. Temos que f é de classe C^∞, $f(a, t_0) = 0$ e
$$\frac{\partial f}{\partial t}(a, t_0) = \frac{\partial p_a}{\partial t}(t_0) = q(t_0) \neq 0.$$
Logo, fazendo-se $z = (a, t_0)$, pelo Teorema da Função Implícita, existem bolas abertas $B(a,r) \subset \mathbb{R}^{n+1}$ e $B(z,\epsilon) \subset \mathbb{R}^{n+1} \times \mathbb{R}$, tais que, para todo $x \in B(a,r)$, existe um único $t = t(x) \in \mathbb{R}$ que satisfaz $(x, t(x)) \in B(z,\epsilon)$ e $f(x,t(x)) = 0$. Uma vez que a projeção de $B(z,\epsilon)$ sobre $\{0\} \times \mathbb{R}$ é o intervalo $(t_0 - \epsilon, t_0 + \epsilon)$, isto significa que $t(x)$ é a única raiz de p_x neste intervalo. A função $t = t(x)$ é C^∞, pois f o é, e, claramente, satisfaz $t(a) = t_0$. Também, observando-se que uma raiz t de um polinômio p é simples se, e somente se, $p'(t) \neq 0$, tem-se que, se r é suficientemente pequeno, então, para todo $x \in B(a,r)$, $t(x)$ é uma raiz simples de p_x. Com efeito, uma vez que $\frac{\partial f}{\partial t}(a, t(a)) \neq 0$, pela continuidade de f', bem como da aplicação $x \mapsto (x, t(x))$, existe $r > 0$, tal que, para todo $x \in B(a,r)$, tem-se
$$0 \neq \frac{\partial f}{\partial t}(x, t(x)) = \frac{\partial p_x}{\partial t}(t(x)).$$

25. Sejam $U = \{(x,y) \in \mathbb{R}^2; x, y > 0\}$ e $f, g : U \to \mathbb{R}$ as funções definidas por $f(x,y) = \frac{1}{p}x^p + \frac{1}{q}y^q$ e $g(x,y) = xy$. Devemos, então, provar que o valor mínimo de $f|_M$ é 1, em que $M = g^{-1}(\{1\})$. Temos que f e g são de classe C^∞, $\nabla f(x,y) = (x^{p-1}, y^{q-1})$ e $\nabla g(x,y) = (y,x) \neq (0,0)$. Em particular, g é uma submersão. Assim, pelo método dos multiplicadores de Lagrange, os pontos críticos (x,y) de $f|_M$ são tais que, para algum $\lambda \in \mathbb{R}$,
$$\begin{cases} x^{p-1} &= \lambda y \\ y^{q-1} &= \lambda x \\ xy &= 1, \end{cases}$$
donde se obtém $\lambda = x^p = y^q$ e, então, $x^{p+q} = 1$. Uma vez que $p + q = pq \neq 0$, devemos ter $x = 1$ e, portanto, $y = 1$, donde se conclui que o valor mínimo de $f|_M$ é $f(1,1) = \frac{1}{p} + \frac{1}{q} = 1$, como desejado (note que $f|_M$ é ilimitada superiormente, pois a sequência $(f(k, 1/k))_{k \in \mathbb{N}}$ é ilimitada superiormente).

Dados reais $x, y > 0$, façamos $x_0 = \sqrt[q]{x}/\sqrt[p]{y}$, $y_0 = \sqrt[p]{y}/\sqrt[q]{x}$, e observemos que $xyx_0^p = x^p$ e $xyy_0^q = y^q$. Uma vez que $x_0 y_0 = 1$, pelas considerações do parágrafo anterior, temos que $1 \leq \frac{1}{p}x_0^p + \frac{1}{q}y_0^q$. Assim, multiplicando-se ambos os membros dessa desigualdade por xy, obtém-se $xy \leq \frac{1}{p}x^p + \frac{1}{q}y^q$.

Consideremos agora $2n$ números reais positivos, $x_1, \ldots, x_n, y_1, \ldots, y_n$, e façamos
$$\sigma = \left(\sum_{i=1}^{n} x_i^p\right)^{\frac{1}{p}}, \quad \mu = \left(\sum_{i=1}^{n} y_i^q\right)^{\frac{1}{q}}.$$
Dado, então, $i \in \{1, \ldots, n\}$, escrevendo-se $x = x_i/\sigma$, $y = y_i/\mu$, e considerando-se a última desigualdade do parágrafo anterior, obtém-se
$$\frac{x_i y_i}{\sigma\mu} \leq \frac{1}{p\sigma^p}x_i^p + \frac{1}{q\mu^q}y_i^q, \quad i = 1, \ldots, n.$$

Somando-se, então, membro a membro, estas n desigualdades, chega-se a

$$\frac{1}{\sigma\mu}\sum_{i=1}^{n} x_i y_i \leq \frac{1}{p\sigma^p}\sum_{i=1}^{n} x_i^p + \frac{1}{q\mu^q}\sum_{i=1}^{n} y_i^q = \frac{1}{p} + \frac{1}{q} = 1,$$

isto é,

$$\sum_{i=1}^{n} x_i y_i \leq \sigma\mu = \left(\sum_{i=1}^{n} x_i^p\right)^{\frac{1}{p}} \left(\sum_{i=1}^{n} y_i^q\right)^{\frac{1}{q}}.$$

Bibliografia

[1] M. Aigner, G. Ziegler. *Proofs from The Book* (third edition). Springer (2004).

[2] K. Borsuk. *Drei Sätze über die n-dimensionale euklidische Sphäre*, Fund. Math., 20 (1933), 177–190.

[3] D. Bridges. *Foundations of Real and Abstract Analysis*. Springer (1998).

[4] R. Courant. *Differential and Integral Calculus,* Volume II. John Willey, inc. (1968).

[5] R. Courant, H. Robbins. *What is Mathematics* (second edition). Oxford University Press (1996).

[6] R. F. de Lima. *Uma nova demonstração do Teorema de Motzkin*. A ser publicado em Revista Matemática Universitária (SBM).

[7] P. N. de Souza, J.N. Silva. *Berkeley Problems in Mathematics* (third edition). Springer (2004).

[8] J. J. Duistermaat, J. A. Kolk. *Multidimensional Real Analysis I*. Cambridge University Press (2004).

[9] G. Flory. *Ejercicios de Topología y de Análisis,* Tomo III. Editorial Reverté, S.A. (1981).

[10] H. Furstenberg. *On the Infinitude of Primes*. Amer. Math. Mont. 62 (1955), 353.

[11] B. J. Gardner, M. Jackson. *The Kuratowski Closure-Complement Theorem*. New Zealand Journal of Mathemaics (2008), 9–44.

[12] P. R. Halmos. *Von Neumann on Measure and Ergodic Theory*. Bull. Am. Math. Soc. 64 (1958), 86–94.

[13] ——— *Finite-Dimensional Vector Spaces* (third edition). Springer-Verlag New York (1987).

[14] K. Hoffman. *Analysis in Euclidian Space*. Prentice-Hall (1975).

[15] J. L. Kazdan. *Matrices A(t) depending on a Parameter t*. http://www.math.upenn.edu/ kazdan/509S07/eigenv5b.pdf (acesso em 11/04/2013).

[16] C. Kuratowski. *Sur l'opération* $\overline{A}$ *de l'Analysis Situs*, Fund. Math. 3 (1922), 182–199.

[17] S. Lang. *Analysis I*, Addison-Wesley Publishing Company (1968).

[18] E.L. Lima. *Curso de Análise*, Vol. 1 (quarta edição). Projeto Euclides, IMPA (1982).

[19] ——— *Espaços Métricos* (segunda edição). Projeto Euclides, IMPA (1983).

[20] ——— *Curso de Análise*, Vol. 2 (terceira edição). Projeto Euclides, IMPA (1989).

[21] ——— *Análise no Espaço* $\mathbb{R}^n$. Coleção Matemática Universitária, IMPA (2001).

[22] ——— *Análise Real*, Vol. 2. Coleção Matemática Universitária, IMPA (2004).

[23] ——— *Álgebra Linear*. Coleção Matemática Universitária, IMPA (2009).

[24] J. Milnor. *Topology from the Differentiable Viewpoint*. Princeton University Press (1965).

[25] T. Motzkin. *Sur quelques propriétés caractéristiques des ensembles convexes*, Atti Accad. Naz. Lincei. Rend. 8 vol. 21 (1935), 773–779.

[26] J. Munkres. *Topology* (second edition). Prentice Hall (2000).

[27] T. C. Ribeiro. $\mathbb{R}^n$ *não é homeomorfo a* $\mathbb{R}^n - \{0\}$: *Uma prova elementar*. Mat. Univ. – 37 (2004), 33–34.

[28] T. Roque. *História da Matemática: Uma visão crítica, desfazendo mitos e lendas*. Zahar (2012).

[29] W. Rudin. *Principles of Mathematical Analysis* (second edition). McGraw-Hill (1964).

[30] A. Sard. *The measure of the critical values of differentiable maps*. Bull. Amer. Math. Soc. 48 (1942), 883–890.

[31] D. Sherman. *Variations on Kuratowski's 14-Set Theorem*. Amer. Math. Mont. – vol. 117, no. 2 (2010), 113–123.

[32] M. Spivak. *Calculus on Manifolds*. Addison-Wesley (1965).

Índice