



**UNIVERSIDADE FEDERAL DO CEARÁ**  
**CENTRO DE CIÊNCIAS**  
**DEPARTAMENTO DE MATEMÁTICA**  
**PROGRAMA DE PÓS-GRADUAÇÃO EM MATEMÁTICA EM REDE NACIONAL**  
**MESTRADO PROFISSIONAL EM MATEMÁTICA**

**ANDRÉ PINHEIRO DA SILVA ZACARIAS**

**O TEOREMA DA APLICAÇÃO INVERSA E APLICAÇÕES**

**FORTALEZA**

**2022**

ANDRÉ PINHEIRO DA SILVA ZACARIAS

O TEOREMA DA APLICAÇÃO INVERSA E APLICAÇÕES

Dissertação apresentada ao Curso de Mestrado Profissional em Matemática do Programa de Pós-Graduação em Matemática em Rede Nacional do Centro de Ciências da Universidade Federal do Ceará, como requisito parcial à obtenção do título de mestre em Matemática. Área de Concentração: Ensino de Matemática.

Orientador: Prof. Dr. Marcos Ferreira de Melo.

FORTALEZA

2022

Dados Internacionais de Catalogação na Publicação  
Universidade Federal do Ceará  
Sistema de Bibliotecas  
Gerada automaticamente pelo módulo Catalog, mediante os dados fornecidos pelo(a) autor(a)

---

- Z15t Zacarias, André Pinheiro da Silva.  
O teorema da aplicação inversa e aplicações / André Pinheiro da Silva Zacarias. – 2022.  
62 f. : il.
- Dissertação (mestrado) – Universidade Federal do Ceará, Centro de Ciências, Departamento de Matemática, Programa de Pós-Graduação em Matemática em Rede Nacional, Fortaleza, 2022.  
Orientação: Prof. Dr. Marcos Ferreira de Melo.
1. Teorema da aplicação inversa. 2. Teorema da função implícita. 3. Multiplicadores de Lagrange. 4. Forma local das imersões. 5. Forma local das submersões. I. Título.

CDD 510

---

ANDRÉ PINHEIRO DA SILVA ZACARIAS

O TEOREMA DA APLICAÇÃO INVERSA E APLICAÇÕES

Dissertação apresentada ao Curso de Mestrado Profissional em Matemática do Programa de Pós-Graduação em Matemática em Rede Nacional do Centro de Ciências da Universidade Federal do Ceará, como requisito parcial à obtenção do título de mestre em Matemática. Área de Concentração: Ensino de Matemática.

Aprovada em: 26/12/2022.

BANCA EXAMINADORA

---

Prof. Dr. Marcos Ferreira de Melo (Orientador)  
Universidade Federal do Ceará (UFC)

---

Prof. Dr. Marcelo Ferreira de Melo  
Universidade Federal do Ceará (UFC)

---

Prof. Dr. Carlos Augusto David Ribeiro  
Universidade Federal do Delta do Parnaíba (UFDPAr)

Dedico este trabalho à minha esposa Nilian  
Maia.

## AGRADECIMENTOS

Agradeço, primeiramente, e o mais importante de todos, ao Soberano Senhor Jeová Deus, pela vida e saúde concedida, sem as quais seria impossível realizar esse trabalho.

Aos meus pais, Luiz e Sandra, que contribuíram de diversas formas e, assim, eu consegui chegar até aqui.

A minha esposa, Nilian Maia, que sempre esteve ao meu lado e com sua positividade pude chegar até o fim dessa jornada.

Ao meu orientador, Marcos Melo, por aceitar me ajudar com essa tarefa, e pelas excelentes observações sempre se mostrando solícito.

Aos professores Marcelo Melo e Carlos Augusto, por aceitarem fazer parte da banca examinadora dessa dissertação.

Ao meu amigo, Alexsandro Belém, pelas dicas e conselhos pertinentes.

À profissional de Letras, Nirla Maia Freire, pelo trabalho de revisão desse texto.

A todos os professores que contribuíram com suas aulas, a saber, Marcos, Marcelo, Othon, Alberto e Jonatan.

Ao meu colega de sala, Luiz Edson, pela solidariedade em ajudar nas diversas vezes que surgiram dúvidas sobre a formatação desse trabalho.

A toda a minha turma, que se mostrou unida do início ao fim: Anthony, Antônio Zacarias, Alfredo, Alex Arley, Edson, Jéssyka, Mariana, Marcelo, Fábio, Elias, Pádua, Inácio, Annelise, Paulo, Narcélio, Erineu, Tyara, Nílbio, Neto e Marcus.

"Ganhar sabedoria é melhor do que ganhar  
prata, e obtê-la é melhor do que obter ouro."  
(BÍBLIA, 2015, p. 901)

## RESUMO

O Teorema da Aplicação Inversa (TAI) é um importante resultado de análise, que estabelece, mesmo que localmente, a existência de uma aplicação inversa continuamente diferenciável de uma dada aplicação. Tendo em vista a sua importância, esse trabalho visa sua detalhada demonstração. Primeiramente, faremos uma exposição de todos os pré-requisitos necessários. Em seguida, provaremos alguns resultados que nos serão úteis na posterior demonstração do TAI. Também analisaremos algumas aplicações, a saber, o Teorema da Função Implícita, o método dos Multiplicadores de Lagrange e a Forma Local das Imersões e Submersões. A fim de dar uma ideia mais ampla e mostrar que o TAI tem validade em outros ambientes, daremos uma visão básica, mas suficiente, de Superfícies Euclidianas. Posteriormente, enunciaremos e daremos a devida prova do TAI entre Superfícies Euclidianas. A partir daí, estenderemos mais ainda nosso ambiente para Variedades Diferenciáveis.

**Palavras-chave:** teorema da aplicação inversa; teorema da função implícita; multiplicadores de Lagrange; forma local das imersões; forma local das submersões; superfícies euclidianas; variedades diferenciáveis.

## ABSTRACT

The Inverse Function Theorem (IFT) is an important result of analysis, which establishes, even locally, the existence of a continuously differentiable inverse application of a given map. In view of its importance, this work aims at its detailed demonstration. First, we will make an exhibition of all the necessary prerequisites. Next, we will prove some results that will be useful in the subsequent demonstration of the IFT. We will also analyze some applications, including the Implicit Function Theorem, the Lagrange Multipliers Method, and the Local Form of Immersions and Submersions. To give a broader idea and show that the IFT has validity in other environments, we will give a basic but sufficient view of Euclidean Surfaces. Later, we will enunciate and give due proof of the IFT between Euclidean Surfaces. From there, we will further extend our environment to Smooth Manifolds.

**Keywords:** inverse application theorem; implicit function theorem; Lagrange multipliers; local form of immersions; local form of submersions; euclidean surfaces; smooth manifolds.

## SUMÁRIO

1	<b>INTRODUÇÃO</b> . . . . .	9
2	<b>PRELIMINARES</b> . . . . .	11
2.1	<b>Noções topológicas no espaço <math>\mathbb{R}^n</math></b> . . . . .	11
3	<b>O TEOREMA DA APLICAÇÃO INVERSA NO <math>\mathbb{R}^n</math> E APLICACÕES</b> .	32
3.1	<b>O Teorema da Aplicação Inversa no <math>\mathbb{R}^n</math></b> . . . . .	32
3.2	<b>Aplicações</b> . . . . .	38
3.2.1	<i>Teorema da Função Implícita</i> . . . . .	38
3.2.2	<i>Multiplicadores de Lagrange</i> . . . . .	41
3.2.3	<i>Forma Local das Imersões</i> . . . . .	45
3.2.4	<i>Forma Local das Submersões</i> . . . . .	49
3.3	<b>Superfícies no Espaço Euclidiano <math>\mathbb{R}^n</math></b> . . . . .	51
4	<b>O TEOREMA DA APLICAÇÃO INVERSA EM VARIEDADES DIFE- RENCIÁVEIS</b> . . . . .	56
5	<b>CONCLUSÃO</b> . . . . .	61
	<b>REFERÊNCIAS</b> . . . . .	62

## 1 INTRODUÇÃO

A Análise é uma teoria matemática que surgiu das diversas transformações do Cálculo. Os conceitos iniciais 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 ou outra teoria, por exemplo, a Topologia (ramo da matemática que estuda os espaços topológicos<sup>1</sup>), que surgiu em meados do século XIX. Sendo assim, do ponto de vista da Análise, uma abordagem atual do estudo de funções passa, inevitavelmente, pelo estudo dos aspectos topológicos, ou seja, da topologia dos conjuntos, onde estas são definidas.

Ao longo dos últimos três séculos, matemáticos de várias partes do mundo foram protagonistas para tornar a Análise como conhecemos hoje. Dentre esses, podemos destacar o eminente matemático italiano Joseph Louis Lagrange (1736-1813). Acredita-se que foi devido a ele o que conhecemos hoje como Teorema da Aplicação Inversa. Esse resultado é um dos mais importantes dentre aqueles ligados ao conceito de Diferenciabilidade. O presente trabalho, assim, tem como objetivo geral fornecer ao leitor todo o aparato necessário para a compreensão do enunciado e demonstração do TAI no  $\mathbb{R}^n$ . Adicionalmente, mostraremos algumas aplicações. Além disso, apresentaremos brevemente como este teorema pode ser visto em Variedades Diferenciáveis.

O TAI estabelece que uma aplicação de classe  $C^k$ ,  $f : U \subset \mathbb{R}^m \rightarrow \mathbb{R}^m$ , definida num aberto  $U \subset \mathbb{R}^m$ , cuja derivada num ponto  $a \in U$  é um isomorfismo, quando restrita a um certo aberto  $B \ni a$ , tem como imagem um aberto  $V = f(B) \ni f(a)$ , sendo tal restrição um difeomorfismo de  $B$  sobre  $f(B)$ . A força deste teorema, que é de caráter local, reside no fato de que uma propriedade facilmente verificável, a saber, a bijetividade da derivada de uma função num ponto, implica numa propriedade bastante especial da função numa vizinhança deste ponto, a de ser um difeomorfismo.

Para a leitura desse trabalho, como pré-requisitos, recomendamos conhecimentos elementares de Álgebra Linear e um bom domínio dos conteúdos iniciais de Análise na Reta, a saber: números reais, seqüências de números reais, topologia da reta, limites de funções, funções contínuas e derivada. Sugerimos a leitura, para Álgebra Linear, (LIMA, 2012) e, para Análise na Reta, (ÁVILA, 1999), (FIGUEIREDO, 2015), (LIMA, 2020a), (LIMA, 2007) ou (RUDIN, 1976).

---

<sup>1</sup> Espaços topológicos são estruturas que permitem a formalização de conceitos tais como convergência, conexidade e continuidade.

No capítulo 2 desse trabalho, apoiamo-nos, principalmente, em (LIMA, 2009) e (LIMA, 2006), onde fornecemos as preliminares necessárias para a leitura do capítulo 3. Iniciamos definindo produto interno e norma de um vetor do espaço euclidiano  $\mathbb{R}^n$ ; logo em seguida, ampliamos praticamente todos os principais conceitos da topologia na reta ao espaço euclidiano  $n$ -dimensional.

No capítulo 3, nossa base para o TAI é (LIMA, 2009) e, para as aplicações, (LIMA, 2015). Procuramos detalhar o máximo possível as demonstrações com o objetivo de torná-las compreensíveis ao leitor que acessa o teorema em um primeiro momento. Adiante, damos algumas aplicações do teorema, dentre outras, a Forma Local das Imersões e Submersões, e finalizamos definindo Superfícies no Espaço Euclidiano, abordando alguns conceitos nesses espaços: espaço vetorial tangente e aplicação diferenciável entre superfícies, por exemplo.

No capítulo 4, damos mais algumas definições necessárias para demonstrar o TAI entre Superfícies no Espaço Euclidiano. Logo em seguida, ampliamos o conceito do teorema para Variedades Diferenciáveis.

## 2 PRELIMINARES

O objetivo geral deste capítulo é fornecer ao leitor os conceitos básicos para o entendimento dos diversos aspectos que serão abordados neste trabalho, à partir do capítulo 3. Nos limitaremos em expor, aqui, apenas assuntos diretamente relacionados com os teoremas e definições que virão a ser mencionados adiante. É natural que seja uma tarefa impossível detalhar todos os assuntos precedentes ao Teorema da Aplicação Inversa. No entanto, fizemos o possível para colocar as principais ferramentas que serão usadas no ponto principal dessa dissertação.

Iniciaremos abordando algumas noções topológicas no espaço vetorial  $\mathbb{R}^n$ .<sup>1</sup>

### 2.1 Noções topológicas no espaço $\mathbb{R}^n$

Nas definições a seguir, veremos conceitos fundamentais que caracterizam o  $\mathbb{R}^n$ .

**Definição 2.1** *Um produto interno em um espaço vetorial real qualquer  $E$  é uma função  $\langle \cdot, \cdot \rangle : E \times E \rightarrow \mathbb{R}$  que faz corresponder a cada par de vetores  $x, y \in E$  um número real, indicado por  $\langle x, y \rangle$ , de tal modo que, para quaisquer  $x, y, z \in E$  e  $\alpha \in \mathbb{R}$ , valem:*

1.  $\langle x, y \rangle = \langle y, x \rangle$ ;
2.  $\langle x, y + z \rangle = \langle x, y \rangle + \langle x, z \rangle$ ;
3.  $\langle \alpha x, y \rangle = \alpha \cdot \langle x, y \rangle = \langle x, \alpha y \rangle$ ;
4.  $\langle x, x \rangle > 0$  se  $x \neq 0$ ;
5.  $\langle x, x \rangle = 0 \Leftrightarrow x = 0$ .

O exemplo mais importante é o *produto escalar* do espaço vetorial  $\mathbb{R}^n$ , que é definido por

$$\langle x, y \rangle := x_1y_1 + x_2y_2 + \cdots + x_ny_n,$$

onde  $x = (x_1, x_2, \dots, x_n)$  e  $y = (y_1, y_2, \dots, y_n) \in \mathbb{R}^n$ . Existem vários tipos de produtos internos, mas, neste trabalho, quando fizermos menção do termo produto interno estaremos considerando o produto escalar no espaço euclidiano  $\mathbb{R}^n$ .

Quando  $\langle x, y \rangle = 0$ , com  $x, y \in \mathbb{R}^n$ , dizemos que  $x$  e  $y$  são *ortogonais*, (indicamos por  $x \perp y$ ).

<sup>1</sup> Assumiremos já conhecidas as propriedades que definem espaço vetorial. O leitor pode consultar o capítulo 1 de (LIMA, 2009) ou (LIMA, 2006) para conferir estas informações.

**Definição 2.2** Uma *norma* em um espaço vetorial real qualquer  $E$  é uma função  $|\cdot| : E \rightarrow \mathbb{R}$  que faz corresponder a cada vetor  $x \in E$  um número real indicado por  $|x|$ , de tal modo que, para quaisquer  $x, y \in E$  e  $\alpha \in \mathbb{R}$ , valem:

1.  $x \neq 0 \Rightarrow |x| > 0$ ;
2.  $|x| = 0 \Leftrightarrow x = 0$ ;
3.  $|\alpha \cdot x| = |\alpha||x|$  ( $|\alpha|$  significa o valor absoluto do número real  $\alpha$ );
4.  $|x + y| \leq |x| + |y|$ .

Por definição, temos que  $\langle x, x \rangle = |x|^2$ . Quando  $|x| = 1$ , dizemos que  $x$  é um vetor unitário.

Denominamos a norma  $|x| = \sqrt{x_1^2 + x_2^2 + \dots + x_n^2}$  de *norma euclidiana*, ela é motivada pela fórmula do comprimento de um vetor no plano em coordenadas cartesianas, que se prova com o Teorema de Pitágoras. Para noções geométricas, ela é a mais natural.

Vale ressaltar que no espaço vetorial  $\mathbb{R}^n$  existem diversas normas além da euclidiana. Por exemplo, dado  $x \in \mathbb{R}^n$  temos a *norma do máximo* definida por:  $|x|_M := \max\{|x_1|, |x_2|, \dots, |x_n|\}$ , e a *norma da soma* definida por:  $|x|_S := |x_1| + |x_2| + \dots + |x_n|$ , evidentemente, estas duas normas cumprem as quatro propriedades mencionadas acima. No capítulo 2 de (LIMA, 2020b), o autor mostra que quaisquer normas em  $\mathbb{R}^n$  são *equivalentes*.

Durante todo esse trabalho, dado  $x \in \mathbb{R}^n$ , sempre que usarmos  $|x|$ , estaremos nos referindo à norma euclidiana  $|x| = \sqrt{x_1^2 + x_2^2 + \dots + x_n^2}$ .

Para concluir esses pontos acerca da norma de um vetor  $x \in \mathbb{R}^n$ , vamos mostrar que, para toda norma, vale a desigualdade:  $||x| - |y|| \leq |x - y|$ . De fato, de  $x = (x - y) + y$  resulta que  $|x| \leq |x - y| + |y|$ , daí,  $|x| - |y| \leq |x - y|$ . Invertendo,  $x$  e  $y$  vem  $|y| - |x| \leq |y - x|$ . Sabendo que  $|y - x| = |x - y|$ , teremos  $|y| - |x| \leq |x - y|$ , portanto,  $||x| - |y|| \leq |x - y|$ .

No conjunto  $\mathbb{R}$ , existe a importante noção de intervalo, por exemplo, dados  $a, b \in \mathbb{R}$ , com  $a < b$ , representamos por  $(a, b)$  e  $[a, b]$ , respectivamente, os *intervalos abertos e fechados* de extremos  $a$  e  $b$ . Assim como na reta real, no  $\mathbb{R}^n$ , também, existe a definição de "intervalo", mas neste caso, eles são generalizados com o conceito de *bola*. Dados  $a \in \mathbb{R}^n$  e um número real  $r > 0$ , a *bola aberta* de centro  $a$  e raio  $r$ , no espaço  $\mathbb{R}^n$ , é definida por

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

Analogamente, a *bola fechada* de centro  $a$  e raio  $r$ , será

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

Finalmente, a *esfera*<sup>2</sup> de centro  $a$  e raio  $r$  é o conjunto

$$S[a; r] = \{x \in \mathbb{R}^n; |x - a| = r\}.$$

Em  $\mathbb{R}$ , a noção de “esfera” se reduz simplesmente ao conjunto de dois pontos  $\{a, b\}$ . Adotaremos uma notação especial para a *esfera unitária* de dimensão  $n - 1$ :

$$\mathbb{S}^{n-1} = \{x \in \mathbb{R}^n; |x| = 1\}.$$

**Definição 2.3** Um ponto  $a \in X \subseteq \mathbb{R}^n$  é **interior** ao conjunto  $X$  quando existe uma bola aberta de  $\mathbb{R}^n$  com centro  $a$  e contida em  $X$ . Indicamos por  $\text{int } X$  o conjunto formado por todos os pontos que são interiores a  $X$ .

**Definição 2.4** Um conjunto  $A \subseteq \mathbb{R}^n$  é **aberto** quando todos os seus pontos são interiores, ou seja,  $\text{int } A = A$ .

Dizemos que  $V$  é uma *vizinhança* de um ponto  $a \in \mathbb{R}^n$  quando  $V$  contém um aberto que contém  $a$ .

**Exemplo 2.5** Toda bola aberta  $B(a; r)$  é um conjunto aberto.

De fato, se  $x \in B(a; r)$ , então  $|x - a| < r$ . Tome  $s = r - |x - a| > 0$ . Afirmamos que  $B(x; s) \subset B(a; r)$ . Para ver isto, vamos mostrar que dado  $y \in B(x; s)$ , teremos  $y \in B(a; r)$ . Seja  $y \in B(x; s)$ , então

$$|y - x| < s \Rightarrow |y - x| < r - |x - a| \Rightarrow |y - x| + |x - a| < r.$$

Como

$$|y - a| = |(y - x) + (x - a)| \leq |y - x| + |x - a|,$$

então

$$|y - a| \leq |y - x| + |x - a| < r \Rightarrow |y - a| < r \Rightarrow y \in B(a; r).$$

**Exemplo 2.6** Nenhuma bola fechada, ou esfera de  $\mathbb{R}^n$ , é um conjunto aberto.

<sup>2</sup> Por **esfera** estamos nos referindo à **superfície esférica** de uma “bola sólida”, em que a bola sólida seria a união do conjunto dos pontos interiores à bola com o conjunto de pontos da superfície da bola, ou seja, é a bola fechada  $B[a; r] = B(a; r) \cup S[a; r]$ .

De fato, sejam  $B[a; r] \subset \mathbb{R}^n$  uma bola fechada e  $b \in B[a; r]$ , tal que  $|b - a| = r$ , isto é,  $b \in S[a; r]$ . Dada, então, uma bola  $B(b; \varepsilon)$  com  $\varepsilon < r$ , tomemos

$$x = \lambda(b - a) + a = (\lambda - 1)(b - a) + b \in \mathbb{R}^n,$$

onde  $1 < \lambda < \frac{\varepsilon}{r} + 1$ . Temos que

- $|x - a| = \lambda|b - a| = \lambda r > r$ ;
- $|x - b| = |(\lambda - 1)(b - a)| = (\lambda - 1)r < \varepsilon$ ,

donde,  $x \in B(b; \varepsilon)$  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 anterior, podemos concluir que  $\text{int}(S[a; r]) = \emptyset$  e  $\text{int}(B[a; r]) = B(a; r)$ . Em particular,  $S[a; r]$  e  $B[a; r]$  não são abertos.

Fixado um conjunto  $X \subset \mathbb{R}^n$ , um subconjunto  $A \subset X$  é *aberto em X* quando, para cada  $a \in A$ , existe  $\delta > 0$  tal que  $B(a; \delta) \cap X \subset A$ . Em outras palavras, para cada  $a \in A$  existe  $\delta > 0$  tal que os pontos  $x$ , pertencentes a  $X$ , que cumprem a condição  $|x - a| < \delta$ , estão em  $A$ . Por exemplo,  $A = (0, 1]$  é aberto em  $X = [0, 1]$ .

**Definição 2.7** Um conjunto  $X \subset \mathbb{R}^n$  é **limitado** quando está contido em alguma bola  $B[a; r]$ . Dizer que  $X$  é limitado equivale a dizer que existe  $k > 0$  tal que  $|x| \leq k$  para todo  $x \in X$ .

Uma aplicação  $f : X \rightarrow \mathbb{R}^n$  é *limitada* no conjunto  $X \subseteq \mathbb{R}^n$  quando sua imagem  $f(X) \subset \mathbb{R}^n$  é um conjunto limitado, isto é, quando existe  $c > 0$  tal que  $|f(x)| \leq c$  para todo  $x \in X$ .

Uma sequência  $(x_k)_{k \in \mathbb{N}}$  é *limitada* quando existe uma bola em  $\mathbb{R}^n$  que contém todos os termos  $x_k$ ,  $k = 1, 2, \dots$ . Isto significa que existe  $c > 0$  tal que  $|x_k| \leq c$  para todo  $k \in \mathbb{N}$ .

**Definição 2.8** Uma sequência de pontos  $(x_k)_{k \in \mathbb{N}} \subset \mathbb{R}^n$  é uma **sequência de Cauchy**<sup>3</sup> quando, para todo  $\varepsilon > 0$  dado, existe  $k_0 \in \mathbb{N}$  tal que  $k, r > k_0 \Rightarrow |x_k - x_r| < \varepsilon$ .

Um fato interessante sobre as sequências de Cauchy é que elas são limitadas. De fato, na definição acima, tomando  $\varepsilon = 1$ , vemos que existe um índice  $k_0$  tal que, salvo, possivelmente, os pontos  $x_1, x_2, \dots, x_{k_0}$ , todos os demais termos  $x_k$  pertencem à bola  $B(x_{k_0-1}; 1)$ . Portanto, o conjunto dos termos da sequência é limitado.

A condição para que a sequência  $(x_k)_{k \in \mathbb{N}}$  seja de Cauchy pode ser vista também assim:  $\lim_{k, r \rightarrow \infty} |x_k - x_r| = 0$ , ou seja,  $\lim_{k, r \in \mathbb{N}} |x_k - x_r| = 0$ . Daí, resulta que se  $\mathbb{N}' \subset \mathbb{N}$  é um subconjunto infinito, isto é, se  $(x_k)_{k \in \mathbb{N}'}$  é uma subsequência de  $(x_k)_{k \in \mathbb{N}}$  então  $\lim_{k, r \in \mathbb{N}'} |x_k - x_r| = 0$ .

<sup>3</sup> Augustin-Louis Cauchy (1789-1857) foi um matemático francês.

Uma sequência  $(x_k)_{k \in \mathbb{N}} \subset \mathbb{R}^n$  é dita *convergente* para um ponto  $a \in \mathbb{R}^n$  quando existe  $\lim_{k \rightarrow \infty} x_k = a$ .

As sequências de Cauchy nos fornecem um importante critério de convergência que enunciamos e mostramos abaixo.

**Teorema 2.9 (Critério de Cauchy)** *Um sequência em  $\mathbb{R}^n$  converge se, e somente se, é uma sequência de Cauchy.*

**Demonstração.** Seja  $(x_k)_{k \in \mathbb{N}}$  uma sequência de Cauchy em  $\mathbb{R}^n$ . Sendo limitada, pelo Teorema de Bolzano-Weierstrass,<sup>4</sup> ela possui uma subsequência convergente  $(x_r)_{r \in \mathbb{N}'}$ . Seja  $a = \lim_{r \in \mathbb{N}'} x_r$ . Temos  $\lim_{r \in \mathbb{N}'} |x_r - a| = 0$  e  $\lim_{k \in \mathbb{N}, r \in \mathbb{N}'} |x_k - x_r| = 0$ , conforme mencionado acima. Então, de

$$|x_k - a| \leq |x_k - x_r| + |x_r - a|,$$

resulta que  $\lim_{k \in \mathbb{N}} |x_k - a| = 0$ , ou seja,  $\lim_{k \rightarrow \infty} x_k = a$ . Reciprocamente, se  $(x_k)_{k \in \mathbb{N}}$  é convergente, com  $\lim_{k \rightarrow \infty} x_k = a$ , então, como

$$|x_k - x_r| \leq |x_k - a| + |x_r - a|,$$

concluimos que  $\lim_{k, r \rightarrow \infty} |x_k - x_r| = 0$ , ou seja,  $(x_k)_{k \in \mathbb{N}}$  é uma sequência de Cauchy. ■

**Definição 2.10** *Um conjunto  $X \subseteq \mathbb{R}^n$  chama-se **convexo** quando o segmento de reta que une dois quaisquer de seus pontos está inteiramente contido em  $X$ , isto é equivalente a  $x, y \in X, 0 \leq t \leq 1 \Rightarrow (1-t)x + ty \in X$ .*

Naturalmente, o espaço  $\mathbb{R}^n$  é convexo.

**Exemplo 2.11** *Toda bola  $B(a; r) \subset \mathbb{R}^n$  é convexa.*

Seja  $B = B(a; r)$  a bola aberta de centro  $a$  e raio  $r > 0$ . Se  $x, y \in B$ , então  $|x - a| < r$  e  $|y - a| < r$ . Para qualquer  $t \in [0, 1]$ , temos

$$|(1-t)x + ty - a| = |(1-t)(x - a) + t(y - a)| \leq (1-t)|x - a| + t|y - a| < r.$$

A demonstração é análoga no caso da bola fechada  $B[a; r]$ .

<sup>4</sup> Bernard Bolzano (1781-1848), dentre muitas coisas, foi um matemático da antiga Boêmia. Karl Weierstrass (1815-1897) foi um matemático alemão. A demonstração do Teorema de Bolzano-Weierstrass pode ser encontrada em (LIMA, 2009) e (LIMA, 2006). Também, recomendamos ao leitor ver a mesma versão dele na reta em (LIMA, 2020a), (LIMA, 2007) ou (ÁVILA, 1999).

Falaremos, agora, de algumas noções necessárias para a definição de conjunto fechado. Dado um ponto  $a \in \mathbb{R}^n$ , dizemos que ele é *aderente* a um conjunto  $X \subseteq \mathbb{R}^n$  quando ele é limite de uma sequência de pontos  $(x_k)_{k \in \mathbb{N}}$  desse conjunto. Por exemplo, todo ponto  $a \in X$  é aderente a  $X$  pois podemos escrever  $\lim_{k \rightarrow \infty} x_k = a$ , com  $x_k = a$  para todo  $k \in \mathbb{N}$ .

O conjunto dos pontos aderentes a  $X$  chama-se o *fecho* de  $X$  e é indicado com a notação  $\overline{X}$ .

**Definição 2.12** *Um conjunto  $X$  é dito **fechado** quando contém todos os seus pontos aderentes, ou seja, quando  $X = \overline{X}$ .*

Além da definição acima, existe um outro fato sobre conjunto fechado que é importante mencionar: se  $X \subset \mathbb{R}^n$  é fechado, então um subconjunto  $F \subset X$  é fechado em  $X$  se, e somente se, é fechado em  $\mathbb{R}^n$ . A fim de que o subconjunto  $F \subset X$  seja fechado em  $X$ , é necessário, e suficiente, que  $F$  contenha todos os seus pontos aderentes que pertençam a  $X$ .

**Exemplo 2.13** *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  $x_k = \left(1 - \frac{1}{k}\right)(x - a) + a$ ,  $k \in \mathbb{N}$ , 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,  $x \in \overline{B(a; r)}$ , onde  $S[a; r] \subset \overline{B(a; r)}$ . Uma vez que  $B(a; r) \in \overline{B(a; r)}$ , 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  $\varepsilon = \frac{|y - a| - r}{2}$  é disjunta de  $B(a; r)$ . Sendo assim, nenhuma sequência de pontos de  $B(a; r)$  pode convergir para  $y$ . Donde, concluímos que  $y \notin \overline{B(a; r)}$  e, portanto, que  $\overline{B(a; r)} = B[a; r]$ .

Uma outra definição, equivalente a anterior, dada a um conjunto fechado é: *um conjunto  $X \subset \mathbb{R}^n$  é fechado quando seu complementar em  $\mathbb{R}^n$ ,  $\mathbb{R}^n - X$ , é aberto.*

**Definição 2.14** *Seja  $X$  um subconjunto de  $\mathbb{R}^n$ . O ponto  $a \in \mathbb{R}^n$  é **ponto de acumulação** de  $X$  quando toda bola aberta de centro  $a$  contém algum ponto de  $X$ , diferente do ponto  $a$ . Em outros*

termos:

$$\forall \varepsilon > 0, \exists x \in X; 0 < |x - a| < \varepsilon.$$

Uma maneira equivalente de definir seria: dado  $X \subset \mathbb{R}^n$ ,  $a \in \mathbb{R}^n$  é ponto de acumulação de  $X$  se  $a \in \overline{X - \{a\}}$ , ou seja, se existe uma sequência  $(x_k)_{k \in \mathbb{N}}$  em  $\mathbb{R}^n$  tal que  $x_k \in X - \{a\} \forall k \in \mathbb{N}$ ;

$$\lim_{k \rightarrow \infty} x_k = a.$$

Um ponto de acumulação de um conjunto  $X$  pode não pertencer a  $X$  conforme o exemplo a seguir.

**Exemplo 2.15** *Seja  $X \subset \mathbb{R}^n$  a bola aberta de centro na origem e raio  $r > 0$ . Todo ponto  $a \in \mathbb{R}^n$  com  $|a| = r$  (veja que  $a \notin X$ ) é ponto de acumulação de  $X$ .*

Com efeito, dado  $\varepsilon > 0$ , podemos, sem perda de generalidade, supor  $\varepsilon < 2r$ . Então, o ponto  $x = \left(1 - \frac{\varepsilon}{2r}\right)a$ , que pertence à bola  $X$ , é diferente de  $a$ , e tem-se  $|a - x| = \frac{\varepsilon}{2} < \varepsilon$ .

Denotamos o conjunto dos pontos de acumulação de um conjunto  $X$  por  $X'$ . Quando  $a \in X - X'$ , dizemos que  $a$  é um *ponto isolado* de  $X$ . Se todos os pontos de  $X$  são isolados, chamamos  $X$  de um conjunto *discreto*. Exemplos de conjuntos discretos são  $\mathbb{Z}^n$  e os subconjuntos finitos de  $\mathbb{R}^n$ . Existe um importante teorema que nos ajuda a identificar se um determinado ponto é ponto de acumulação de um subconjunto do  $\mathbb{R}^n$ . Enunciamos e provamos ele logo abaixo.

**Teorema 2.16** *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 em  $\mathbb{R}^n$  contém uma infinidade de elementos de  $X$ .*

**Demonstração.** Sejam  $X \subset \mathbb{R}^n$  e  $a \in X'$ . Então, existe uma sequência  $(x_k)_{k \in \mathbb{N}}$  em  $X - \{a\}$ , tal que  $\lim_{k \rightarrow \infty} x_k = a$ . Dada, então, uma vizinhança  $V$  de  $a$  em  $\mathbb{R}^n$ , à partir de um certo  $k_0$ , todos os termos de  $(x_k)_{k \in \mathbb{N}}$  estão em  $V$ . Além disso, o conjunto  $\{x \in X; x = x_k \text{ para algum } k \geq k_0\}$  é infinito, pois nenhum termo da sequência  $(x_k)_{k \in \mathbb{N}}$  é igual a  $a$ . Reciprocamente, se toda vizinhança de  $a$  contém uma infinidade de pontos de  $X$ , temos que  $a \in X'$ , pois, para cada  $k \in \mathbb{N}$ , podemos escolher um  $x_k \in X$ , satisfazendo  $x_k \in B\left(a; \frac{1}{k}\right)$  e  $x_k \neq a$ , onde teremos  $a \in \overline{X - \{a\}}$ . ■

**Definição 2.17** *A fronteira de um conjunto  $X \subset \mathbb{R}^n$  é o conjunto  $\partial X$ , formado pelos pontos de  $X$  que não são interiores a  $X$ , juntamente com os pontos de  $\mathbb{R}^n - X$  que não são interiores a  $\mathbb{R}^n - X$ . Em outras palavras, temos que,  $x \in \partial X$  quando toda bola de centro  $x$  contém pontos de  $X$  e pontos de  $\mathbb{R}^n - X$ .*

**Exemplo 2.18** Um conjunto  $X \subset \mathbb{R}^n$  é fechado se, e somente se,  $\partial X \subset X$ . Por outro lado,  $A \subset \mathbb{R}^n$  é aberto se, e somente se,  $A \cap \partial A = \emptyset$ .

De fato, seja  $X \subset \mathbb{R}^n$  fechado e  $a \in \partial X$ , então, para todo  $r > 0$ , temos que  $B(a; r) \cap X \neq \emptyset$ , isto nos dá  $a \in \overline{X}$ , como  $X$  é fechado, segue que  $a \in X$ . Logo,  $\partial X \subset X$ . Reciprocamente, considerando  $\partial X \subset X$ , seja  $a \in \overline{X}$ , então para todo  $r > 0$ , temos que  $B(a; r) \cap X \neq \emptyset$ . Se for  $B(a; r) \subset X$ , então não há mais o que fazer e o problema encerra. Mas, se  $B(a; r) \not\subset X$ , então  $B(a; r) \cap (\mathbb{R}^n - X) \neq \emptyset$ , isto nos dá  $a \in \partial X$ . Logo,  $\overline{X} \subset X$ . A inclusão  $X \subset \overline{X}$  é imediata. Portanto,  $X$  é fechado. Por outro lado, se  $A \subset \mathbb{R}^n$  é um conjunto aberto, suponha, por absurdo, que  $A \cap \partial A \neq \emptyset$ , então existe  $a \in (A \cap \partial A)$ . Daí, teremos  $a \in A \Rightarrow \exists r > 0$  tal que  $B(a; r) \subset A$ . No entanto,  $a \in \partial A$ , o que nos dá  $B(a; r) \cap (\mathbb{R}^n - A) \neq \emptyset$ , um absurdo. Portanto,  $A \cap \partial A = \emptyset$ . Reciprocamente, considerando  $A \subset \mathbb{R}^n$  um conjunto tal que  $A \cap \partial A = \emptyset$ , suponha, por absurdo, que  $A$  não é aberto, então existe  $a \in A$  tal que  $a$  não é ponto interior de  $A$ , ou seja, para todo  $r > 0$  temos que  $B(a; r) \not\subset A$ , daí, teremos  $B(a; r) \cap (\mathbb{R}^n - A) \neq \emptyset$ . Como  $B(a; r) \cap A \neq \emptyset$ , então  $a \in \partial A$ , um absurdo. Portanto,  $A$  é aberto.

Veremos agora um importante conceito topológico: a compacidade de conjuntos. Um conjunto compacto admite diversas caracterizações, e algumas delas serão abordadas no intuito de tornar mais frutífero o entendimento de certas propriedades especiais, adquiridas pelas aplicações contínuas (ou diferenciáveis) quando seus domínios são conjuntos compactos.

Dado um conjunto  $X \subset \mathbb{R}^n$ , uma família  $\mathcal{A} = \{A_\lambda\}_{\lambda \in \Lambda}$  de subconjuntos de  $\mathbb{R}^n$  é dita uma *cobertura* de  $X$  se  $X \subset \bigcup A_\lambda$ . Uma *subcobertura* de  $\mathcal{A}$  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 subcobertura  $\mathcal{A} = \{A_\lambda\}_{\lambda \in \Lambda}$  é dita *aberta*, se cada  $A_\lambda$  é aberto, e finita se  $\Lambda$  é um conjunto finito.

**Definição 2.19** Um conjunto  $X \subset \mathbb{R}^n$  é **compacto** quando toda cobertura aberta  $\mathcal{A} = \{A_\lambda\}_{\lambda \in \Lambda}$  de  $X$  admite uma subcobertura finita, isto é, quando existem  $\lambda_1, \lambda_2, \dots, \lambda_i \in \Lambda$ , tais que  $X \subset A_{\lambda_1} \cup A_{\lambda_2} \cup \dots \cup A_{\lambda_i}$ .

**Exemplo 2.20** Nenhuma bola aberta do  $\mathbb{R}^n$  é um conjunto compacto.

De fato, dados  $a \in \mathbb{R}^n$  e  $r > 0$ , façamos, para cada  $k \in \mathbb{N}$ ,  $A_k = B(a; r - (r/k + 1))$ . Claramente,  $\mathcal{A} = \{A_k\}_{k \in \mathbb{N}}$  é uma cobertura aberta de  $B(a; r)$ . Segue-se, daí, que esta cobertura não admite subcobertura finita. Com efeito, dada qualquer quantidade finita de abertos de  $\mathcal{A}$ , dentre eles,

existirá um  $A_k$ , que conterà 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. É fácil ver que a família  $\mathcal{B} = \{B(0; k)\}_{k \in \mathbb{N}}$  constitui uma cobertura aberta de  $\mathbb{R}^n$  e, em particular, de  $X$ . Uma vez que, para todo  $k \in \mathbb{N}$ ,  $B(0; k) \subset B(0; k+1)$ , conclui-se, daí, que  $\mathcal{B}$  não admite subcobertura finita, pois nenhuma bola de  $\mathbb{R}^n$  contém  $X$ . Com base nisto, segue o seguinte resultado: *todo conjunto compacto de  $\mathbb{R}^n$  é, necessariamente, limitado.*

Existe uma relação interessante entre conjunto fechado e compacto que será vista no resultado a seguir.

**Teorema 2.21** *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  $F$ , digamos  $\mathcal{A} = \{A_\lambda\}_{\lambda \in \Lambda}$ , temos que os abertos de  $\mathcal{A}$ , juntamente com  $\mathbb{R}^n - F$ , constituem uma cobertura aberta de  $K$ . Como  $K$  é compacto, desta cobertura podemos extrair uma subcobertura finita  $A_{\lambda_1}, A_{\lambda_2}, \dots, A_{\lambda_i}, \mathbb{R}^n - F$ , que, conseqüentemente, cobrirá  $F$ . Porém,  $\mathbb{R}^n - F$  é disjunto de  $F$ . Logo,  $\{A_{\lambda_1}, A_{\lambda_2}, \dots, A_{\lambda_i}\}$  é uma subcobertura finita de  $\mathcal{A}$ . Portanto,  $F$  é compacto. ■

Pelo fato do espaço vetorial  $\mathbb{R}^n$  ter dimensão finita, os conjuntos compactos deste espaço admitem caracterizações, as quais não envolvem coberturas. Uma delas diz que a condição de ser compacto é equivalente à de ser fechado e limitado. As bolas abertas, que não são compactas, são limitadas, porém, não são fechadas. O espaço  $\mathbb{R}^n$ , que não é compacto, é fechado, mas não é limitado. Temos esse importante resultado a respeito de conjuntos compactos. A demonstração pode ser vista no capítulo 3 de (LIMA, 2015).

**Teorema de Heine-Borel.**<sup>5</sup> *Um subconjunto de  $\mathbb{R}^n$  é compacto se, e somente se, é fechado e limitado.*

Segue do Teorema de Heine-Borel, que todas as bolas fechadas e esferas de  $\mathbb{R}^n$ , são conjuntos compactos. No entanto, 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. Pela mesma razão, uma bola fechada menos um número finito de pontos não é compacta.

Será abordada uma outra caracterização para compactos que falaremos agora.

<sup>5</sup> Heinrich Eduard Heine (1821-1881) foi um matemático alemão. Félix Émile Borel (1871-1956) foi um político e matemático francês.

**Teorema 2.22** *Um conjunto  $K \subset \mathbb{R}^n$  é compacto se, e somente se, toda sequência de pontos de  $K$  possui uma subsequência que converge para um ponto de  $K$ .*

**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  $K$ , visto que, pelo Teorema de Heine-Borel,  $K$  é fechado. Segue, daí, que toda sequência de pontos de  $K$  possui uma subsequência que converge para um ponto de  $K$ . Reciprocamente, se toda sequência de pontos de  $K$  possui uma subsequência que converge para um ponto de  $K$ , 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  $(x_k)_{k \in \mathbb{N}}$  não teria, dessa forma, subsequências limitadas e, portanto, nenhuma delas seria convergente. Isso seria uma contradição. Ainda,  $\bar{K} = K$ , pois se existisse  $a \in \bar{K} - K$ , existiria uma sequência de elementos de  $K$  convergindo para um elemento não pertencente a  $K$ , o que também seria uma contradição. Assim,  $K$  é fechado e limitado. Segue, do Teorema de Heine-Borel, que  $K$  é compacto. ■

Tratemos, agora, sobre o comportamento das aplicações contínuas definidas em conjuntos compactos.

**Definição 2.23** *Uma aplicação  $f : X \rightarrow \mathbb{R}^n$ , definida no conjunto  $X \subseteq \mathbb{R}^m$  é **contínua**, no ponto  $a \in X$ , quando, para qualquer  $\varepsilon > 0$  dado, se pode obter  $\delta > 0$ , tal que todo ponto  $x \in X$ , cuja distância ao ponto  $a$  seja menor do que  $\delta$ , é transformado por  $f$  num ponto  $f(x)$  que dista de  $f(a)$  menos que  $\varepsilon$ . Simbolicamente:*

$$\forall \varepsilon > 0, \exists \delta > 0; x \in X, |x - a| < \delta \Rightarrow |f(x) - f(a)| < \varepsilon.$$

Outra forma de dizer que  $f$  é contínua no ponto  $a \in X$ : para cada bola  $B(f(a); \varepsilon)$  dada, existe uma bola  $B(a; \delta)$  tal que  $f(B(a; \delta) \cap X) \subset B(f(a); \varepsilon)$ .

Se  $f : X \rightarrow \mathbb{R}^n$  é contínua em todos os pontos do conjunto  $X$ , dizemos, simplesmente, que  $f$  é uma aplicação *contínua* em  $X$ .

**Teorema 2.24** *A aplicação  $f : X \rightarrow \mathbb{R}^n$  é contínua no ponto  $a \in X$  se, e somente se, para toda sequência de pontos  $(x_k)_{k \in \mathbb{N}} \subset X$  com  $\lim_{k \rightarrow \infty} x_k = a$ , tem-se  $\lim_{k \rightarrow \infty} f(x_k) = f(a)$ .*

**Demonstração.** Seja  $f : X \rightarrow \mathbb{R}^n$  contínua no ponto  $a$ . Dada a sequência de pontos  $(x_k)_{k \in \mathbb{N}} \subset X$  com  $\lim_{k \rightarrow \infty} x_k = a$ , para todo  $\varepsilon > 0$  existe  $\delta > 0$  tal que  $f(B(a; \delta)) \subset B(f(a); \varepsilon)$ . Correspondente,

a  $\delta$ , existe  $k_0 \in \mathbb{N}$ , tal que  $k > k_0 \Rightarrow x_k \in B(a; \delta)$ . Logo,  $k > k_0 \Rightarrow f(x_k) \in B(f(a); \varepsilon)$ . Isso mostra que  $\lim_{k \rightarrow \infty} f(x_k) = f(a)$ . Reciprocamente, suponhamos, por absurdo, que  $\lim_{k \rightarrow \infty} x_k = a \Rightarrow \lim_{k \rightarrow \infty} f(x_k) \neq f(a)$ . Porém,  $f$  seja descontínua no ponto  $a$ . Então, existe  $\varepsilon > 0$  com a seguinte propriedade: para todo  $k \in \mathbb{N}$ , podemos encontrar  $x_k \in X$  com  $|x_k - a| < \frac{1}{k}$  e  $|f(x_k) - f(a)| \geq \varepsilon$ . Assim, temos  $\lim_{k \rightarrow \infty} x_k = a$ , mas não temos  $\lim_{k \rightarrow \infty} f(x_k) = f(a)$ , uma contradição. ■

Uma outra caracterização de aplicações contínuas que citaremos, mas não iremos demonstrar, é a seguinte: *uma aplicação  $f : X \subset \mathbb{R}^n \rightarrow Y \subset \mathbb{R}^m$  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  $X$ . A aplicação  $f$  é dita *aplicação aberta* quando dado qualquer  $U \subset X$  aberto,  $f(U) \subset Y$  é aberto. Usando essa mesma situação, podemos substituir “aberto” por “fechado”, ou seja, podemos afirmar, também, que:  $f : X \subset \mathbb{R}^n \rightarrow Y \subset \mathbb{R}^m$  é contínua em  $a \in X$  se, e somente se, para todo fechado  $F$  de  $Y$ ,  $f^{-1}(F)$  é fechado em  $X$ . A aplicação  $f$  é dita *aplicação fechada* quando dado qualquer  $F \subset Y$  fechado,  $f^{-1}(F) \subset X$  é fechado. A prova desses dois fatos podem ser encontradas no capítulo 4 de (LIMA, 2015).*

**Teorema 2.25** *A imagem de um conjunto compacto por uma aplicação contínua é um conjunto compacto.*

**Demonstração.** Considere  $f : X \rightarrow \mathbb{R}^n$  uma aplicação contínua e  $K \subset X$  um conjunto compacto. Seja  $(y_k)_{k \in \mathbb{N}}$  uma sequência de pontos em  $f(K)$ . Para cada  $k \in \mathbb{N}$ , existe  $x_k \in K$ , tal que  $f(x_k) = y_k$ . Como  $K$  é compacto, uma subsequência  $(x_k)_{k \in \mathbb{N}'}$  converge para um ponto  $a \in K$ . Sendo  $f$  contínua nesse ponto  $a$ , de  $\lim_{k \in \mathbb{N}'} x_k = a$  resulta, pelo Teorema 2.24, que  $\lim_{k \in \mathbb{N}'} f(x_k) = f(a)$ . Logo, toda sequência de pontos  $(y_k)_{k \in \mathbb{N}} = (f(x_k))_{k \in \mathbb{N}} \subset f(K)$  possui uma subsequência  $(y_k)_{k \in \mathbb{N}'}$  convergente para um ponto  $f(a) \in f(K)$ , ou seja,  $f(K)$  é compacto. ■

Para finalizar os pontos abordados sobre compactos, iremos apresentar o importante resultado abaixo. Logo em seguida, definiremos homeomorfismo.

**Teorema 2.26 (Weierstrass)** *Toda função contínua definida num compacto é limitada e atinge seus extremos no compacto, ou seja, se  $K \subset \mathbb{R}^n$  é compacto e  $f : K \rightarrow \mathbb{R}$  é contínua, então existem  $a, b \in K$ , tais que  $f(a) \leq f(x) \leq f(b)$ , para todo  $x \in K$ .*

**Demonstração.** Pelo Teorema 2.25, temos que  $f(K)$  é compacto e, portanto, fechado e limitado. Daí, segue que  $f$  é limitada. Sejam  $\lambda$  e  $\mu$  o ínfimo e o supremo de  $f(K)$ , respectivamente. Então,

existem sequências  $(x_k)_{k \in \mathbb{N}}$ ,  $(y_k)_{k \in \mathbb{N}}$ , em  $K$ , tais que  $\lim_{k \rightarrow \infty} f(x_k) = \lambda$  e  $\lim_{k \rightarrow \infty} f(y_k) = \mu$ . Uma vez que  $K$  é compacto, tomando subsequências, se necessário, podemos supor que  $\lim_{k \rightarrow \infty} x_k = a \in K$  e  $\lim_{k \rightarrow \infty} y_k = b \in K$ , onde  $\lim_{k \rightarrow \infty} f(x_k) = f(a)$  e  $\lim_{k \rightarrow \infty} f(y_k) = f(b)$ , pois  $f$  é contínua. Portanto,  $\lambda = f(a)$  e  $\mu = f(b)$ , como queríamos demonstrar. ■

Decorre do Teorema de Weierstrass que, se  $f$  é contínua no compacto  $K \subset \mathbb{R}^m$  e  $f(x) > 0$  para todo  $x \in K$ , então existe  $\varepsilon > 0$  tal que  $f(x) \geq \varepsilon$  para todo  $x \in K$ .

**Definição 2.27** *Dados os conjuntos  $X \subset \mathbb{R}^m$  e  $Y \subset \mathbb{R}^n$ , um **homeomorfismo** entre  $X$  e  $Y$  é uma bijeção contínua  $f : X \rightarrow Y$ , cuja inversa  $f^{-1} : Y \rightarrow X$  também é contínua. Dizemos, então, que  $X$  e  $Y$  são **homeomorfos**.*

**Exemplo 2.28** *A bola aberta  $B = B(0; 1) \subset \mathbb{R}^n$  é homeomorfa ao espaço  $\mathbb{R}^n$ .*

De fato, as aplicações  $f : \mathbb{R}^n \rightarrow B$  e  $g : B \rightarrow \mathbb{R}^n$ , definidas por  $f(x) = \frac{x}{1+|x|}$  e  $g(y) = \frac{y}{1-|y|}$ , são contínuas. Além disso, verifica-se, facilmente, que  $g(f(x)) = x$  e  $f(g(y)) = y$  para quaisquer  $x \in \mathbb{R}^n$  e  $y \in B$ . Assim,  $g = f^{-1}$ .

**Exemplo 2.29** *Dada uma aplicação  $f : X \subset \mathbb{R}^n \rightarrow Y \subset \mathbb{R}^m$ , sabemos que o seu gráfico é o conjunto*

$$\text{graf}(f) = \{(x, f(x)); x \in X\} \subset X \times Y \subset \mathbb{R}^n \times \mathbb{R}^m.$$

*Vejamos que, quando  $f$  é contínua, o gráfico de  $f$  é homeomorfo a  $X$ .*

De fato, a aplicação  $\varphi : X \rightarrow \text{graf}(f) \subset \mathbb{R}^n \times \mathbb{R}^m$ , dada por  $\varphi(x) = (x, f(x))$ , é bijetiva e contínua, pois suas coordenadas o são. Uma vez que  $\varphi^{-1} = P|_{\text{graf}(f)}$ , em que  $P : \mathbb{R}^n \times \mathbb{R}^m \rightarrow \mathbb{R}^n$  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  $\varphi^{-1}$  é contínua. Logo,  $\varphi$  é um homeomorfismo.

Existe uma interessante relação entre compacidade e homeomorfismo conforme apresentamos abaixo.

**Teorema 2.30** *Toda bijeção contínua definida num compacto é um homeomorfismo sobre sua imagem.*

**Demonstração.** Sejam  $K \subset \mathbb{R}^n$  e  $f : K \rightarrow f(K) \subset \mathbb{R}^m$  uma bijeção contínua. Pela definição,  $f$  é um homeomorfismo se, e somente se,  $f^{-1}$  é contínua. Uma vez que a inversa de  $f^{-1}$  é a própria  $f$ , basta mostrarmos que  $f$  é uma aplicação fechada. Tome  $F \subset K$  fechado em  $K$ . Uma vez que

$K$  é fechado em  $\mathbb{R}^n$ , tem-se que  $F$  também é. Logo,  $F$  por ser um subconjunto fechado de um compacto, é compacto. Segue do Teorema 2.25 que  $f(F)$  é compacto, logo fechado. Portanto,  $f$  é uma aplicação fechada. ■

**Exemplo 2.31** A aplicação  $f : [0, 2\pi) \rightarrow \mathbb{S}^1$ , onde  $f(t) = (\cos t, \sin t)$  é uma bijeção contínua, mas não é um homeomorfismo.

Sua inversa  $f^{-1} : \mathbb{S}^1 \rightarrow [0, 2\pi)$  aplica o compacto  $\mathbb{S}^1$  sobre o intervalo  $[0, 2\pi)$ , que não é compacto. Assim,  $f^{-1}$  é descontínua. Mais precisamente,  $f^{-1}$  é descontínua no ponto  $a = (1, 0) = f(0) \in \mathbb{S}^1$ . Com efeito, se para cada  $k \in \mathbb{N}$  colocarmos  $t_k = \left(1 - \frac{1}{k}\right) \cdot 2\pi$  e  $z_k = (\cos t_k, \sin t_k)$ , teremos  $\lim_{k \rightarrow \infty} z_k = a$ , mas  $\lim_{k \rightarrow \infty} f^{-1}(z_k) = \lim_{k \rightarrow \infty} t_k = 2\pi$ . Assim, não vale  $\lim_{k \rightarrow \infty} f^{-1}(z_k) = f^{-1}(a) = 0$ .

Falaremos agora de outra propriedade importante de conjuntos do  $\mathbb{R}^n$ , a *conexidade*.

Uma *cisão* de um conjunto  $X \subset \mathbb{R}^n$  é uma decomposição  $X = A \cup B$ , onde  $A \cap B = \emptyset$  e os conjuntos  $A, B$  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*.

**Definição 2.32** Um conjunto  $X \subseteq \mathbb{R}^n$  chama-se **conexo** quando não admite outra cisão além da trivial. Assim, quando  $X$  é conexo,  $X = A \cup B$ , com  $A, B$  disjuntos e abertos em  $X$ , implica  $A = \emptyset$  ou  $B = \emptyset$ .

Quando existir uma cisão não-trivial  $X = A \cup B$ ,  $X$  é *desconexo*. Por exemplo,  $\mathbb{R} - \{0\}$  é desconexo assim como qualquer conjunto discreto do  $\mathbb{R}^n$  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 ele ter apenas dois subconjuntos, o próprio  $X$  e o conjunto vazio.

Observe que, se  $A \subset X$  é um subconjunto próprio e não vazio, que é aberto e não fechado em  $X$ , então a decomposição  $X = A \cup (X - A)$  é uma cisão não trivial de  $X$ . Assim, vale o seguinte resultado de fácil verificação: *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*. Visto que o espaço  $\mathbb{R}^n$  é convexo, um resultado interessante o qual mostra que ele também é conexo é: *todo conjunto convexo de  $\mathbb{R}^n$  é conexo*.<sup>6</sup>

Agora iremos definir aplicação diferenciável, mostrar como a derivada pode ser vista como transformação linear, bem como analisar outros pontos que nos serão de extrema

<sup>6</sup> A demonstração pode ser vista no capítulo 3 de (LIMA, 2015).

importância no capítulo seguinte. Aqui, vamos admitir que o leitor tenha conhecimentos prévios de alguns conceitos a cerca de funções reais de  $n$  variáveis, por exemplo, derivadas parciais, função diferenciável, etc. Como referência, o capítulo 3 de (LIMA, 2009) ou (LIMA, 2006) podem ser consultados.

No caso das funções de uma variável, a derivada admite uma interpretação que vem da ideia de se considerar a reta tangente ao gráfico de uma função (derivável) num determinado ponto, como, por exemplo, o gráfico de uma função linear que a aproxima numa vizinhança deste ponto. Com base nisto, podemos estender o conceito de diferenciabilidade às aplicações entre espaços euclidianos arbitrários, ou seja, uma tal aplicação é diferenciável num ponto de seu domínio quando existe uma determinada transformação linear que a aproxima numa vizinhança deste ponto.

A derivada caracteriza-se pelo fato de permitir, em muitos aspectos, deduzir o comportamento da aplicação numa vizinhança de um ponto à partir das propriedades de uma transformação linear, que são de fácil verificação. Vejamos agora, as definições formais.

Uma aplicação  $f : U \rightarrow \mathbb{R}^n$ , definida no aberto  $U \subset \mathbb{R}^m$ , diz-se *diferenciável* no ponto  $a \in U$  quando cada uma de suas funções-coordenadas  $f_1, f_2, \dots, f_n : U \rightarrow \mathbb{R}$  é diferenciável nesse ponto. Mais precisamente, para todo  $v = (\alpha_1, \alpha_2, \dots, \alpha_m)$  tal que  $a + v \in U$  e, para cada  $i = 1, 2, \dots, n$ , tem-se  $f_i(a + v) - f_i(a) = \sum_{j=1}^m \frac{\partial f_i}{\partial x_j}(a) \cdot \alpha_j + r_i(v)$ , com  $\lim_{v \rightarrow 0} \frac{r_i(v)}{|v|} = 0$ , ou seja,

$$f_i(a + v) - f_i(a) = \frac{\partial f_i}{\partial x_1}(a) \cdot \alpha_1 + \frac{\partial f_i}{\partial x_2}(a) \cdot \alpha_2 + \dots + \frac{\partial f_i}{\partial x_m}(a) \cdot \alpha_m + r_i(v),$$

onde  $\lim_{v \rightarrow 0} \frac{r_i(v)}{|v|} = 0$ .

A matriz  $Jf(a) = \left[ \frac{\partial f_i}{\partial x_j}(a) \right] \in M(n \times m)$  chama-se a *matriz jacobiana*<sup>7</sup> de  $f$  no ponto  $a$ :

$$Jf(a) = \begin{bmatrix} \frac{\partial f_1}{\partial x_1}(a) & \frac{\partial f_1}{\partial x_2}(a) & \dots & \frac{\partial f_1}{\partial x_m}(a) \\ \frac{\partial f_2}{\partial x_1}(a) & \frac{\partial f_2}{\partial x_2}(a) & \dots & \frac{\partial f_2}{\partial x_m}(a) \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f_n}{\partial x_1}(a) & \frac{\partial f_n}{\partial x_2}(a) & \dots & \frac{\partial f_n}{\partial x_m}(a) \end{bmatrix}.$$

<sup>7</sup> A matriz jacobiana recebe este nome em homenagem ao matemático alemão Carl Gustav Jakob Jacobi (1804-1851).

Chamamos de *derivada da aplicação*  $f$  no ponto  $a$  a transformação linear  $f'(a) : \mathbb{R}^m \rightarrow \mathbb{R}^n$ , cuja matriz em relação às bases canônicas de  $\mathbb{R}^m$  e  $\mathbb{R}^n$  é  $Jf(a)$ . Pela definição de matriz de uma transformação linear, para todo vetor  $v = (\alpha_1, \alpha_2, \dots, \alpha_m) \in \mathbb{R}^m$ , temos

$$f'(a) \cdot v = (\beta_1, \beta_2, \dots, \beta_n), \text{ onde } \beta_i = \sum_{j=1}^m \frac{\partial f_i}{\partial x_j}(a) \cdot \alpha_j = \frac{\partial f_i}{\partial v}(a) \text{ para cada } i = 1, 2, \dots, n.$$

Desta forma, definindo a *derivada direcional* da aplicação  $f$  no ponto  $a$  na direção do vetor  $v$  como

$$\frac{\partial f}{\partial v}(a) = \lim_{t \rightarrow 0} \frac{f(a + tv) - f(a)}{t},$$

teremos

$$\frac{\partial f}{\partial v}(a) = \left( \frac{\partial f_1}{\partial v}(a), \frac{\partial f_2}{\partial v}(a), \dots, \frac{\partial f_n}{\partial v}(a) \right) = f'(a) \cdot v.$$

Podemos resumir as  $n$  igualdades numéricas que dão a diferenciabilidade das funções-coordenadas  $f_i$  na igualdade abaixo, entre vetores de  $\mathbb{R}^n$ :

$$f(a + v) - f(a) = f'(a) \cdot v + r(v), \text{ onde } \lim_{v \rightarrow 0} \frac{r(v)}{|v|} = 0.$$

A expressão acima, nos fornece, sem ambiguidades, a diferenciabilidade da aplicação  $f$  da seguinte forma: a aplicação  $f : U \rightarrow \mathbb{R}^n$ , definida no aberto  $U \subset \mathbb{R}^m$ , é diferenciável no ponto  $a \in U$  quando existe uma transformação linear  $T : \mathbb{R}^m \rightarrow \mathbb{R}^n$  tal que

$$f(a + v) - f(a) = T \cdot v + r(v), \text{ onde } \lim_{v \rightarrow 0} \frac{r(v)}{|v|} = 0.$$

Então, podemos concluir que  $T = f'(a)$ . Daí, tomando  $tv$  em vez de  $v$ , teremos:

$$\frac{f(a + tv) - f(a)}{t} = T \cdot v \pm \frac{r(tv)}{|tv|} \cdot |v|.$$

Assim,

$$T \cdot v = \lim_{t \rightarrow 0} \frac{f(a + tv) - f(a)}{t} = \frac{\partial f}{\partial v}(a) = f'(a) \cdot v. \quad (2.1)$$

Quando  $f : U \rightarrow \mathbb{R}^n$  é diferenciável em todos os seus pontos de  $U$ , dizemos que  $f$  é diferenciável em  $U$ . Daí, podemos definir uma aplicação  $f' : U \rightarrow \mathcal{L}(\mathbb{R}^m; \mathbb{R}^n)$ , que faz corresponder a cada  $x \in U$  a transformação linear  $f'(x) : \mathbb{R}^m \rightarrow \mathbb{R}^n$ .

**Exemplo 2.33** Vamos verificar que a aplicação  $f : \mathbb{R}^2 \rightarrow \mathbb{R}^2$  dada por  $f(x, y) = (y, x^2)$  é diferenciável em  $(1, 2)$ . Também, determinar a sua derivada neste ponto.

Inicialmente, iremos encontrar um candidato à derivada, isto é, uma transformação linear  $T : \mathbb{R}^2 \rightarrow \mathbb{R}^2$  que cumpra as condições que foram definidas anteriormente. Fazendo  $(x, y) = (1, 2)$  e  $v = (\alpha_1, \alpha_2)$ , se  $f$  for diferenciável no ponto  $(1, 2)$ , pelo que foi definido na segunda igualdade de 2.1, temos que

$$T \cdot v = T(\alpha_1, \alpha_2) = \lim_{t \rightarrow 0} \frac{f((1, 2) + t(\alpha_1, \alpha_2)) - f(1, 2)}{t} = (\alpha_2, 2\alpha_1).$$

Nesse caso,  $r(v)$  é dado por

$$r(v) = r(\alpha_1, \alpha_2) = f((1, 2) + (\alpha_1, \alpha_2)) - f(1, 2) - T(\alpha_1, \alpha_2) = (0, \alpha_1^2).$$

Como  $\frac{\alpha_1^2}{|v|} = |\alpha_1| \frac{|\alpha_1|}{|v|} \leq |\alpha_1|$ , então  $\lim_{v \rightarrow 0} \frac{r(v)}{|v|} = 0$ . Portanto,  $f$  é diferenciável em  $(1, 2)$  e  $f'(1, 2) : \mathbb{R}^2 \rightarrow \mathbb{R}^2$  é a transformação linear dada por  $f'(1, 2)(\alpha_1, \alpha_2) = (\alpha_2, 2\alpha_1)$ .

**Exemplo 2.34** Toda transformação linear  $T : \mathbb{R}^m \rightarrow \mathbb{R}^n$  é diferenciável em cada ponto  $x \in \mathbb{R}^m$  e  $T'(x) = T$ .

Por linearidade teremos  $T(x + v) = T \cdot x + T \cdot v$ . Segue, então, que  $r(v) = 0$  e  $T'(x) = T$ .

**Exemplo 2.35** Sejam  $f : U \rightarrow \mathbb{R}^n$  diferenciável no aberto  $U \subset \mathbb{R}^m$  e  $M > 0$  uma constante. Se  $|f(x) - f(y)| \leq M \cdot |x - y|$  para quaisquer  $x, y \in U$ , então,  $|f'(x)| \leq M$  para todo  $x \in U$ .

De fato, suponha, por absurdo, que  $|f'(x)| > M$ , ou seja, existe  $\varepsilon > 0$ , tal que  $|f'(x)| = M + \varepsilon$ , para algum  $x \in U$ . Como  $|f'(x)|$  é o máximo de  $|f'(x) \cdot u|$  para  $|u| = 1$  e  $\mathbb{S}^{m-1}$  é compacta, existe  $u \in \mathbb{R}^m$  com norma 1, tal que  $|f'(x) \cdot u| = M + \varepsilon$ . Pela definição de diferenciabilidade, a este  $\varepsilon$  corresponde  $\delta > 0$  tal que

$$0 < t < \delta \Rightarrow |f(x + tu) - f(x)| = |f'(x) \cdot tu + r(tu)| \geq t|f'(x) \cdot u| - |r(tu)|,$$

onde  $|r(tu)| < t\varepsilon$ . Então,

$$0 < t < \delta \Rightarrow |f(x + tu) - f(x)| > t(M + \varepsilon) - t\varepsilon = t \cdot M.$$

Fazendo  $v = tu$ , temos  $|v| = t$ . Logo,  $|f(x + v) - f(x)| > M \cdot |v|$  o que é uma contradição.

Uma característica de grande relevância sobre aplicações diferenciáveis reside no fato de que diferenciabilidade implica continuidade. Considere uma aplicação  $f : U \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$  definida no aberto  $U \subset \mathbb{R}^n$ , que é diferenciável num ponto  $a \in U$ . Sendo assim, temos que  $r(v) = f(a + v) - f(a) - f'(a) \cdot v$  satisfaz  $\lim_{v \rightarrow 0} \frac{r(v)}{|v|} = 0$ . Assim,  $\lim_{v \rightarrow 0} r(v) = \lim_{v \rightarrow 0} \frac{r(v)}{|v|} |v| = 0$ , onde  $\lim_{v \rightarrow 0} f(a + v) = \lim_{v \rightarrow 0} (f(a) + f'(a) \cdot v + r(v)) = f(a)$ . Portanto,  $f$  é contínua em  $a$ .

Dizemos que uma aplicação  $f : U \rightarrow \mathbb{R}^n$  é de classe  $C^1$  no aberto  $U \subset \mathbb{R}^m$  (indicamos por  $f \in C^1$ ) quando  $f$  for diferenciável e a aplicação derivada  $f' : U \rightarrow \mathcal{L}(\mathbb{R}^m; \mathbb{R}^n)$  é contínua. Equivalentemente,  $f : U \rightarrow \mathbb{R}^n$  é de classe  $C^1$  no aberto  $U \subset \mathbb{R}^m$  se suas funções-coordenadas  $f_1, f_2, \dots, f_n : U \rightarrow \mathbb{R}$  possuem derivadas parciais  $\frac{\partial f_i}{\partial x_j} : U \rightarrow \mathbb{R}$  contínuas. Uma aplicação  $f$  será de classe  $C^0$  quando for contínua. Aplicações  $f : U \rightarrow \mathbb{R}^n$  de classe  $C^k$  são definidas por indução: dizemos que  $f$  é de classe  $C^k$ , com  $k \in \mathbb{N}$ , quando  $f$  é diferenciável e sua derivada  $f' : U \rightarrow \mathcal{L}(\mathbb{R}^m; \mathbb{R}^n)$  é de classe  $C^{k-1}$ . Se  $f \in C^k$  para todo  $k \in \mathbb{N}$ , dizemos que  $f$  é de classe  $C^\infty$  ( $f \in C^\infty$ ). Nesse caso, também temos  $f' \in C^\infty$ . Um caso trivial de funções de classe  $C^\infty$  é o das aplicações constantes, elas são de classe  $C^\infty$  e suas derivadas são nulas.

Considerando as propriedades operatórias da derivada, a diferenciabilidade das coordenadas de uma aplicação diferenciável e usando indução, verifica-se que:<sup>8</sup>

- uma aplicação é de classe  $C^k$  se, e somente se, suas coordenadas o são;
- 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  $C^k$ .

**Exemplo 2.36** *Toda transformação linear  $T : \mathbb{R}^m \rightarrow \mathbb{R}^n$  é de classe  $C^\infty$ .*

De fato,  $f' : \mathbb{R}^m \rightarrow \mathcal{L}(\mathbb{R}^m; \mathbb{R}^n)$  é constante, ou seja,  $f'(x) = f$  para todo  $x \in \mathbb{R}^m$ . Portanto,  $f^{(k)} = 0$  para  $k \in \mathbb{N}$  com  $k > 1$ .

**Exemplo 2.37** *Todo polinômio é uma função de classe  $C^\infty$ .*

De fato, todo polinômio é, evidentemente, uma função contínua e possui derivadas parciais (de qualquer ordem) às quais ainda são polinômios e, portanto, são funções contínuas em  $\mathbb{R}^m$ . Logo, todo polinômio  $f : \mathbb{R}^m \rightarrow \mathbb{R}$  é de classe  $C^1$ . As derivadas de  $f$ , sendo polinômios, são de classe  $C^1$ . Assim,  $f \in C^2$ . Repetindo o argumento, concluímos que  $f \in C^k$  para todo  $k \in \mathbb{N}$ . Logo, todo polinômio é uma função de classe  $C^\infty$ .

Definimos como aplicação polinomial  $f : \mathbb{R}^m \rightarrow \mathbb{R}^n$  aquela cujas coordenadas são funções polinomiais. Aplicações polinomiais são, portanto, de classe  $C^\infty$ .

**Exemplo 2.38** *Seja  $\varphi : GL_n(\mathbb{R}^n) \rightarrow GL_n(\mathbb{R}^n)$  onde  $GL_n(\mathbb{R}^n)$  é o conjunto das matrizes quadradas de ordem  $n$  invertíveis e  $\varphi$  associa cada  $M \in GL_n(\mathbb{R}^n)$  à sua inversa  $M^{-1} \in GL_n(\mathbb{R}^n)$ . Afirmamos que  $\varphi$  é de classe  $C^\infty$ .*

<sup>8</sup> A demonstração desses fatos podem ser encontradas no capítulo 5 de (LIMA, 2015).

Para cada  $M = (x_{ij}) \in GL_n(\mathbb{R}^n)$ , fazamos  $M^{-1} = (y_{kl})$ . Como  $M^{-1} = \frac{1}{\det M} \text{adj } M$ , cada entrada  $y_{kl}$  de  $M^{-1}$  será dada por  $y_{kl} = (-1)^{l+k} \cdot \frac{\det M_{lk}}{\det M}$ , em que  $M_{lk}$  é a matriz cujos vetores-coluna são os mesmos de  $M$ , exceto pelo  $k$ -ésimo, que é o  $l$ -ésimo vetor da base canônica de  $\mathbb{R}^n$ ,  $e_l$ . Temos que o determinante, por ser um polinômio, é de classe  $C^\infty$ . Além disso, uma vez que, para quaisquer  $k, l \in \{1, \dots, n\}$ , o  $\det M_{lk}$  é um polinômio de variáveis  $x_{ij}$ , tem-se que a função  $\xi : GL_n(\mathbb{R}^n) \rightarrow \mathbb{R}$ , que associa cada matriz  $M \in GL_n(\mathbb{R}^n)$  ao  $\det M_{lk}$ , é também de classe  $C^\infty$ . Logo, por ser o quociente de funções de classe  $C^\infty$ , cada coordenada de  $\varphi$  é de classe  $C^\infty$ . Portanto,  $\varphi : GL_n(\mathbb{R}^n) \rightarrow GL_n(\mathbb{R}^n) \in C^\infty$ .

Seja  $f : U \rightarrow \mathbb{R}$  uma função diferenciável definida no aberto  $U \subset \mathbb{R}^m$ , o *gradiente* de  $f$  no ponto  $a \in U$  é o vetor

$$\nabla f(a) = \left( \frac{\partial f}{\partial x_1}(a), \dots, \frac{\partial f}{\partial x_m}(a) \right).$$

Vamos destacar aqui três propriedades importantes do gradiente de uma função diferenciável  $f$ . Fixado um ponto  $a$  e supondo que  $\nabla f(a) \neq 0$ , temos o seguinte:<sup>9</sup>

- o gradiente aponta para uma direção segundo a qual  $f$  é crescente;
- dentre todas as direções ao longo das quais a função  $f$  cresce, a direção do gradiente é a de crescimento mais rápido;
- o gradiente de  $f$  no ponto  $a$  é perpendicular à superfície de nível <sup>10</sup> de  $f$  que passa por esse ponto.

No caso de uma aplicação diferenciável  $f : U \rightarrow \mathbb{R}^n$  definida no aberto  $U \subset \mathbb{R}^m$ , para cada  $i = 1, 2, \dots, n$ , teremos o gradiente  $\nabla f_i(a) = \left( \frac{\partial f_i}{\partial x_1}(a), \dots, \frac{\partial f_i}{\partial x_m}(a) \right)$ .

**Definição 2.39** Um conjunto  $M \subset \mathbb{R}^{n+1}$  chama-se uma **hipersuperfície** de classe  $C^k$  quando é localmente o gráfico de uma função real de  $n$  variáveis de classe  $C^k$ .

Isto significa que, para cada  $p \in M$ , deve existir um aberto  $V \subset \mathbb{R}^{n+1}$  e uma função  $\xi : U \rightarrow \mathbb{R}$ , de classe  $C^k$  num aberto  $U \subset \mathbb{R}^n$ , tais que  $p \in V$  e  $V \cap M = \text{gráfico de } \xi$ . Quando falamos “ $V \cap M = \text{gráfico de } \xi$ ”, queremos dizer que, para um certo inteiro  $i \in [1, n]$ , tem-se

$$V \cap M = \{(x_1, x_2, \dots, x_{n+1}) \in \mathbb{R}^{n+1}; x_i = \xi(x_1, x_2, \dots, x_{i-1}, x_{i+1}, \dots, x_{n+1})\}.$$

<sup>9</sup> O leitor pode encontrar as provas das afirmações citadas no capítulo 3 de (LIMA, 2006).

<sup>10</sup> Se  $f$  é de classe  $C^1$ , o conjunto  $f^{-1}(c) = \{x \in U; f(x) = c\}$  é, para todo  $c \in \mathbb{R}$ , chamado o *conjunto de nível*  $c$  da função  $f$ . Quando  $U \subset \mathbb{R}^m$  e  $m = 2$ , esse conjunto geralmente é chamado de *curva* ou *linha de nível*  $c$  de  $f$ , a qual é definida pela equação  $f(x, y) = c$ . Analogamente, quando  $m = 3$ , o conjunto  $f^{-1}(c)$ , definido pela equação  $f(x, y, z) = c$ , costuma ser chamado a *superfície de nível*  $c$  da função  $f$ .

Dada qualquer função  $f : U \rightarrow \mathbb{R}$  de classe  $C^k$  no aberto  $U \subset \mathbb{R}^n$ , seu gráfico é uma hipersuperfície  $M = \{(x, f(x)) \in \mathbb{R}^{n+1}; x \in U\}$ . Quando  $n = 1$ , uma hipersuperfície em  $\mathbb{R}^2$  chama-se uma curva e, quando  $n = 2$ , tem-se uma *superfície* em  $\mathbb{R}^3$ .

**Exemplo 2.40** A esfera  $\mathbb{S}^n = \{x \in \mathbb{R}^{n+1}; \langle x, x \rangle = 1\}$  é uma hipersuperfície de classe  $C^\infty$  em  $\mathbb{R}^{n+1}$ .

De fato, chamando  $U$  a bola aberta de centro 0 e raio 1 em  $\mathbb{R}^n$ , pondo, para cada  $i = 1, 2, \dots, n+1$ ,  $V_i = \{x \in \mathbb{R}^{n+1}; x_i > 0\}$ ,  $W_i = \{x \in \mathbb{R}^{n+1}; x_i < 0\}$  e escrevendo  $x^* = (x_1, x_2, \dots, x_{i-1}, x_{i+1}, \dots, x_{n+1})$ , temos

$$x \in \mathbb{S}^n \cap V_i \Leftrightarrow |x^*| < 1 \text{ e } x_i = \sqrt{1 - \langle x^*, x^* \rangle};$$

$$x \in \mathbb{S}^n \cap W_i \Leftrightarrow |x^*| < 1 \text{ e } x_i = -\sqrt{1 - \langle x^*, x^* \rangle}.$$

Logo, considerando a função  $\xi : U \rightarrow \mathbb{R}$ , de classe  $C^\infty$ , definida por  $\xi(u) = \sqrt{1 - \langle u, u \rangle}$ , vemos que, para cada  $i = 1, 2, \dots, n+1$ ,  $\mathbb{S}^n \cap V_i$  é o gráfico da função  $x_i = \xi(x^*)$ , enquanto que  $\mathbb{S}^n \cap W_i$  é o gráfico de  $x_i = -\xi(x^*)$ . Como todo ponto  $p \in \mathbb{S}^n$  pertence a algum  $V_i$  ou a algum  $W_i$ , concluímos que  $\mathbb{S}^n$  é uma hipersuperfície de classe  $C^\infty$  em  $\mathbb{R}^{n+1}$ .

Seja  $M \subset \mathbb{R}^{n+1}$  uma hipersuperfície de classe  $C^k$  ( $k \geq 1$ ). A cada ponto  $p \in M$  associaremos o conjunto  $T_p M$ , formado por todos os vetores-velocidade  $v = \lambda'(0)$  dos caminhos<sup>11</sup>  $\lambda : (-\delta, \delta) \rightarrow M$  que são diferenciáveis no ponto 0 e cumprem a condição  $\lambda(0) = p$ . O conjunto  $T_p M$  é chamado de *espaço vetorial tangente* de  $M$  no ponto  $p$ .

**Definição 2.41** Um número  $c \in \mathbb{R}$  chama-se **valor regular** de uma função  $f : U \rightarrow \mathbb{R}$  de classe  $C^1$ , definida no aberto  $U \subset \mathbb{R}^{n+1}$ , quando não há pontos críticos de  $f$  no nível  $c$ , isto é, quando  $f(x) = c \Rightarrow \nabla f(x) \neq 0$ . Dizemos também que  $c$  é um nível regular de  $f$ . Quando existe  $x \in U$  tal que  $f(x) = c$  e  $\nabla f(x) = 0$ , chamamos  $c$  de um nível crítico de  $f$ .

**Definição 2.42** Sejam  $U$  e  $V$  abertos do espaço euclidiano  $\mathbb{R}^n$ . Uma bijeção  $f : U \rightarrow V$  chama-se **difeomorfismo** de  $U$  sobre  $V$  quando é diferenciável e sua inversa  $f^{-1} : V \rightarrow U$  também é diferenciável. Dizemos que  $U$  e  $V$  são **difeomorfos**.

Uma aplicação diferenciável  $f : U \rightarrow \mathbb{R}^m$ , definida no aberto  $U \subset \mathbb{R}^m$ , é um *difeomorfismo local* quando, para cada  $x \in U$ , existe uma bola aberta  $B = B(x; \delta) \subset U$  tal que  $f$  aplica

<sup>11</sup> Um caminho é uma aplicação contínua  $f : I \rightarrow \mathbb{R}^n$  cujo domínio é um intervalo da reta.

$B$  difeomorficamente sobre um aberto  $V$  contendo  $f(x)$ . Daí, se  $f : U \rightarrow \mathbb{R}^m$  é um difeomorfismo local, então  $f'(x) : \mathbb{R}^m \rightarrow \mathbb{R}^m$  é um isomorfismo para todo  $x \in U$ . O Teorema da Aplicação Inversa, que provaremos no capítulo a seguir, estabelecerá a recíproca, no caso em que  $f$  é uma aplicação de classe  $C^k, k \geq 1$ : se  $f'(x)$  é um isomorfismo,<sup>12</sup> ou seja, se  $f'(x)$  invertível para todo  $x \in U$ , então  $f$  é um difeomorfismo local.

Além dos fatos que acabamos de abordar, será necessário o uso de alguns resultados mais elaborados na apresentação do tema principal desse trabalho. Dessa forma, enunciaremos e provaremos os teoremas abaixo. O primeiro deles nos diz respeito à desigualdade do valor médio para caminhos.

O Teorema do Valor Médio para uma função diferenciável  $f : [a, b] \rightarrow \mathbb{R}$  diz que existe  $c$ , com  $a < c < b$ , tal que  $f(b) - f(a) = f'(c)(b - a)$ . No entanto, essa igualdade nem sempre vale para caminhos  $f : I \rightarrow \mathbb{R}^n$ . De fato, se considerarmos o caminho  $f : [0, 2\pi] \rightarrow \mathbb{R}^2$ , dado por  $f(t) = (\cos t, \sin t)$ , temos  $f(2\pi) - f(0) = 0$ . Mas, como  $|f'(t)| = 1$  para todo  $t \in [0, 2\pi]$  não pode existir  $c \in [0, 2\pi]$  tal que  $f(2\pi) - f(0) = f'(c) \cdot (2\pi - 0)$ . Para caminhos, o Teorema do Valor Médio assume a forma de uma desigualdade. Assim, temos o seguinte resultado:

**Teorema 2.43 (Desigualdade do valor médio para caminhos)** *Seja  $f : [a, b] \rightarrow \mathbb{R}^n$  um caminho, diferenciável no intervalo aberto  $(a, b)$ , com  $|f'(t)| \leq M$  para todo  $t \in (a, b)$ . Então  $|f(b) - f(a)| \leq M \cdot (b - a)$ .*

**Demonstração.** Defina a função  $\varphi : [a, b] \rightarrow \mathbb{R}$ , fazendo  $\varphi(t) = \langle f(t), f(b) - f(a) \rangle$ . Como  $f$  é diferenciável, então  $\varphi$  é contínua e derivável em  $(a, b)$ . Logo, pelo Teorema do Valor Médio, existe  $c \in (a, b)$  tal que  $\varphi(b) - \varphi(a) = \varphi'(c) \cdot (b - a)$ . Além disso, tem-se que  $\varphi'(t) = \langle f'(t), f(b) - f(a) \rangle$ . Verifica-se facilmente que  $\varphi(b) - \varphi(a) = |f(b) - f(a)|^2$ . Assim, de  $\varphi(b) - \varphi(a) = \varphi'(c) \cdot (b - a)$ , vem

$$\begin{aligned} |f(b) - f(a)|^2 &= \langle f'(c), f(b) - f(a) \rangle \cdot (b - a) \\ &= |\langle f'(c), f(b) - f(a) \rangle| \cdot (b - a) \\ &\leq |f'(c)| \cdot |f(b) - f(a)| \cdot (b - a). \end{aligned}$$

Fazendo  $|f'(c)| = M$  e cancelando o termo  $|f(b) - f(a)|$ , obtemos  $|f(b) - f(a)| \leq M \cdot (b - a)$ , como queríamos demonstrar. ■

O próximo resultado nos dá um caso mais geral do teorema acima, onde o domínio da aplicação é um conjunto aberto do  $\mathbb{R}^m$ .

<sup>12</sup> Dados dois espaços vetoriais  $X$  e  $Y$ , se existe uma bijeção linear  $f : X \rightarrow Y$ , dizemos, então, que  $f : X \rightarrow Y$  é um isomorfismo e que os espaços vetoriais  $X$  e  $Y$  são isomorfos.

**Teorema 2.44 (Desigualdade do valor médio)** *Seja  $f : U \rightarrow \mathbb{R}^n$ , definida num aberto  $U \subset \mathbb{R}^m$ , diferenciável em todos os pontos do segmento de reta  $[a, a+v] \subset U$ . Se, para todo  $t \in [0, 1]$ , tem-se  $|f'(a+tv)| \leq M$ , então,  $|f(a+v) - f(a)| \leq M \cdot |v|$ .*

**Demonstração.** Defina o caminho  $\lambda : [0, 1] \rightarrow \mathbb{R}^n$  dado por  $\lambda(t) = f(a+tv)$ . Note que  $\lambda$  é diferenciável, com  $\lambda'(t) = f'(a+tv) \cdot v$ . Daí,  $|\lambda'(t)| \leq |f'(a+tv)| \cdot |v| \leq M \cdot |v|$  para todo  $t \in [0, 1]$ .<sup>13</sup> Segue do Teorema 2.43 que  $|\lambda(1) - \lambda(0)| \leq M \cdot |v| \cdot (1 - 0)$ , isto é,  $|f(a+v) - f(a)| \leq M \cdot |v|$ . ■

<sup>13</sup> Na primeira desigualdade, usamos o fato de que a norma de uma transformação linear  $T : \mathbb{R}^m \rightarrow \mathbb{R}^n$  é o número  $\|T\| = \sup\{|T \cdot u|; u \in \mathbb{S}^{m-1}\}$ . Dessa definição resulta que, para todo  $v \in \mathbb{R}^m$ , tem-se  $|T \cdot v| \leq \|T\| \cdot |v|$  e que, se  $S : \mathbb{R}^n \rightarrow \mathbb{R}^p$  é outra transformação linear, então,  $\|S \cdot T\| \leq \|S\| \cdot \|T\|$ .

### 3 O TEOREMA DA APLICAÇÃO INVERSA NO $\mathbb{R}^n$ E APLICACÕES

Nesse capítulo, apresentaremos o teorema no espaço  $\mathbb{R}^n$ , bem como sua prova. Logo em seguida, iremos rever o mesmo em outro ambiente.

Considere  $U, V \subset \mathbb{R}^m$  conjuntos abertos. Já vimos que um difeomorfismo  $f : U \rightarrow V$  é uma bijeção diferenciável cuja inversa é diferenciável. Conforme veremos no Teorema 3.1, quando um difeomorfismo é de classe  $C^k$ , seu inverso também será. A regra da cadeia nos garante que sua derivada  $f'(x) : \mathbb{R}^m \rightarrow \mathbb{R}^m$ , em cada ponto  $x \in U$ , é um isomorfismo (invertível).<sup>1</sup> Em termos do determinante jacobiano, isso significa que  $\det Jf(x) \neq 0$  para todo  $x \in U$ , onde  $Jf(x) = \left( \frac{\partial f_i}{\partial x_j} \right)$  é a matriz jacobiana de  $f$  no ponto  $x$ . É natural questionar se a recíproca é verdadeira, é sobre isso que começaremos a tratar agora.

#### 3.1 O Teorema da Aplicação Inversa no $\mathbb{R}^n$

De posse dos resultados preliminares, seguiremos agora para a demonstração. Antes, enunciaremos e provaremos algumas afirmações que nos serão úteis. A primeira delas nos diz respeito à continuidade das derivadas parciais da inversa de um difeomorfismo.

**Teorema 3.1** *Sejam  $U, V \subset \mathbb{R}^m$  abertos. Se o difeomorfismo  $f : U \rightarrow V$  é de classe  $C^k$ , com  $k \geq 1$ , então, seu inverso  $g = f^{-1} : U \rightarrow V$  também é de classe  $C^k$ .*

**Demonstração.** Faremos a demonstração usando indução em  $k$ . Antes, observe que, para todo  $y = f(x) \in V$ , usando derivada da função inversa, temos que  $g'(y) = f'(x)^{-1}$ . Agora, aplicando a inversa  $f^{-1}$  em  $y = f(x)$ , teremos  $f^{-1}(y) = f^{-1}(f(x)) = x$ . Então,  $g'(y) = [f'(f^{-1}(y))]^{-1}$ , ou seja, a aplicação  $g' : V \rightarrow \mathcal{L}(\mathbb{R}^m)$  pode ser escrita como a composta:

$$g' = \varphi \circ f' \circ f^{-1} : U \rightarrow V$$

em que

- $\varphi : GL_n(\mathbb{R}^m) \rightarrow GL_n(\mathbb{R}^m)$  onde  $GL_n(\mathbb{R}^m)$  é o conjunto de todos os operadores invertíveis de  $\mathcal{L}(\mathbb{R}^m)$ ;
- $f' : U \rightarrow \mathcal{L}(\mathbb{R}^m)$  leva todo ponto  $x \in U$  na derivada  $f'(x) : \mathbb{R}^m \rightarrow \mathbb{R}^m$ ;
- $f^{-1} : U \rightarrow V$  é a aplicação inversa de  $f$ .

<sup>1</sup> Este fato é um corolário da regra da cadeia. Mais detalhes podem ser encontrados na seção 3 do capítulo 5 de (LIMA, 2006).

Observe que a aplicação  $\varphi$  é de classe  $C^\infty$ , conforme vimos no Exemplo 2.38. Seguindo com a prova por indução, o teorema é válido para  $k = 1$ . De fato, se o difeomorfismo  $f : U \rightarrow V$  é de classe  $C^1$ , então  $f'$  é de classe  $C^0$ . Sendo  $f$  um difeomorfismo, sua inversa  $f^{-1}$  é diferenciável, logo contínua, conseqüentemente, de classe  $C^0$ . Sendo  $\varphi$  de classe  $C^0$ , temos que  $g'$  é composta de três aplicações de classe  $C^0$ . Logo,  $g'$  também será. O que nos dá  $g = f^{-1} \in C^1$ . Suponha, agora, que o teorema seja válido para um certo  $k - 1$ . Provemos para  $k$ . Se  $f$  é de classe  $C^k$ , temos que  $f'$  é de classe  $C^{k-1}$ . Por hipótese de indução,  $f^{-1}$  é de classe  $C^{k-1}$ . Por sua vez,  $\varphi$  também é de classe  $C^{k-1}$ . Assim,  $g' = \varphi \circ f' \circ f^{-1}$  é composta por três funções de classe  $C^{k-1}$ , o que nos dá  $g' \in C^{k-1}$ . Portanto,  $g = f^{-1}$  é de classe  $C^k$ . ■

O próximo teorema nos fornece uma estimativa para  $f : U \rightarrow \mathbb{R}^n$  na vizinhança de algum ponto do aberto  $U$  de  $\mathbb{R}^m$ .

**Teorema 3.2** *Seja  $f : U \rightarrow \mathbb{R}^n$  de classe  $C^1$  no aberto  $U \subset \mathbb{R}^m$ . Se, para algum  $a \in U$ , a derivada  $f'(a) : \mathbb{R}^m \rightarrow \mathbb{R}^n$  é injetiva, então existem  $\delta > 0$  e  $c > 0$  tais que  $B = B(a; \delta) \subset U$  e, para quaisquer  $x, y \in B$  tem-se  $|f(x) - f(y)| \geq c \cdot |x - y|$ . Em particular, a restrição  $f|_B$  é injetiva.*

**Demonstração.** Considere a função real  $g : \mathbb{S}^{m-1} \rightarrow \mathbb{R}$  que leva cada ponto  $u$  da esfera unitária  $\mathbb{S}^{m-1}$  no número real  $|f'(a) \cdot u|$ . Pelo Teorema de Weierstrass e pela injetividade de  $f$ , existe  $c > 0$  tal que  $|f'(a) \cdot u| \geq 2c$ , para todo  $u \in \mathbb{S}^{m-1}$ . Tomando o vetor unitário  $\frac{v}{|v|} \in \mathbb{S}^{m-1}$ , segue que  $|f'(a) \cdot v| \geq 2c|v|$  para todo vetor não nulo  $v \in \mathbb{R}^m$ . Defina  $r : U \rightarrow \mathbb{R}^n$  pondo

$$r(x) := f(x) - f(a) - f'(a)(x - a), \quad x \in U.$$

Assim, para quaisquer  $x, y \in U$  temos que

$$r(x) - r(y) = f(x) - f(y) - f'(a)(x - y),$$

o que nos dá

$$f(x) - f(y) = f'(a)(x - y) + r(x) - r(y).$$

Lembrando que  $|w - s| \geq |w| - |s|$  com  $w, s \in \mathbb{R}^m$ , vem

$$\begin{aligned} |f(x) - f(y)| &= |f'(a)(x - y) + r(x) - r(y)| \\ &= |f'(a)(x - y) - (r(y) - r(x))| \\ &\geq |f'(a)(x - y)| - |r(x) - r(y)| \\ &\geq 2c|x - y| - |r(x) - r(y)|. \end{aligned}$$

Como  $f$  é de classe  $C^1$ , a aplicação  $r$  definida também será de classe  $C^1$ , ou seja,  $r'$  é contínua. Além disso,  $r(a) = 0$ , o que implica em  $r'(a) = 0$ . Pela continuidade de  $r'$ , dado  $c > 0$ , existe  $\delta > 0$  tal que

$$|x - a| < \delta \Rightarrow |r'(x) - r'(a)| < c \Rightarrow |r'(x)| < c.$$

Sendo  $r$  diferenciável em  $U$ , em particular, no conjunto convexo  $B = B(a; \delta)$ , pela desigualdade do valor médio, temos que

$$x, y \in B(a; \delta) \Rightarrow |r(x) - r(y)| \leq c|x - y|,$$

o que nos dá

$$-|r(x) - r(y)| \geq -c|x - y|.$$

Como  $|f(x) - f(y)| \geq 2c|x - y| - |r(x) - r(y)|$ , então,

$$x, y \in B \Rightarrow |f(x) - f(y)| \geq 2c|x - y| - c|x - y| \Rightarrow |f(x) - f(y)| \geq c|x - y|,$$

como queríamos demonstrar. Em particular,  $f|_B$  é injetiva. De fato, sejam  $x, y \in B$  tais que  $f(x) = f(y)$ , então teremos

$$|f(y) - f(y)| \geq c|x - y| \Rightarrow 0 \geq c|x - y| \Rightarrow |x - y| \leq 0.$$

Se fosse  $|x - y| < 0$ , teríamos, claramente, um absurdo. Logo, só podemos ter  $|x - y| = 0$ , o que nos leva a concluir que  $x = y$ . ■

Mostraremos, agora, um fato acerca da diferenciabilidade do homeomorfismo inverso.

**Teorema 3.3** *Seja  $f : U \rightarrow V$  um homeomorfismo de classe  $C^1$  entre os abertos  $U, V \subset \mathbb{R}^m$ . Se, para algum  $x \in U$ , a derivada  $f'(x) : \mathbb{R}^m \rightarrow \mathbb{R}^m$  é um operador invertível, então o homeomorfismo inverso  $g = f^{-1} : V \rightarrow U$  é diferenciável no ponto  $f(x)$ , com  $g'(f(x)) = f'(x)^{-1}$ .*

**Demonstração.** Sejam  $x$  e  $v \neq 0$ , pontos de  $U$ , tais que  $x + v \in U$ . Escolhamos  $y$  e  $w \neq 0 \in V$  de tal forma que  $f(x) = y$  e  $f(x + v) = y + w$ . Então,

$$f(x + v) = f(x) + w \Rightarrow w = f(x + v) - f(x).$$

Sendo  $f$  de classe  $C^1$ , logo diferenciável, temos que

$$w = f(x + v) - f(x) = f'(x) \cdot v + r(v), \text{ onde } \lim_{v \rightarrow 0} \frac{r(v)}{|v|} = 0.$$

Como  $g = f^{-1}$ , de  $f(x) = y$  vem  $g(f(x)) = g(y) \Rightarrow x = g(y)$  e, de  $f(x+v) = y+w$  teremos  $g(f(x+v)) = g(y+w) \Rightarrow x+v = g(y+w) \Rightarrow v = g(y+w) - x \Rightarrow$

$$g(y+w) - g(y) = v.$$

Como  $f'(x)$  é um operador invertível, definimos

$$g(y+w) - g(y) = f'(x)^{-1} \cdot w + s(w). \quad (3.1)$$

Para mostrar que  $g$  é diferenciável e que  $f'(x)^{-1}$  é a derivada de  $g$  no ponto  $y$ , devemos provar que  $\lim_{w \rightarrow 0} \frac{s(w)}{|w|} = 0$ . Substituindo  $w = f'(x) \cdot v + r(v)$  e  $g(y+w) - g(y) = v$  na igualdade (3.1), obtemos

$$\begin{aligned} v &= f'(x)^{-1} [f'(x) \cdot v + r(v)] + s(w) \Rightarrow v = f'(x)^{-1} \cdot f'(x) \cdot v + f'(x)^{-1} \cdot r(v) + s(w) \Rightarrow \\ &v = v + f'(x)^{-1} \cdot r(v) + s(w) \Rightarrow s(w) = -f'(x)^{-1} \cdot r(v). \end{aligned}$$

Daí, podemos escrever  $\frac{s(w)}{|w|} = -f'(x)^{-1} \cdot \frac{r(v)}{|v|} \cdot \frac{|v|}{|w|}$ . Como  $w$  foi definido inicialmente como  $f(x+v) - f(x)$ , segue que

$$\frac{s(w)}{|w|} = -f'(x)^{-1} \cdot \frac{r(v)}{|v|} \cdot \frac{|v|}{|f(x+v) - f(x)|}.$$

Pelo Teorema 2.2, existem  $\delta > 0$  e  $c > 0$  tais que  $|f(x+v) - f(x)| \geq c \cdot |v|$ , isto é,  $\frac{|v|}{|f(x+v) - f(x)|} \leq \frac{1}{c}$ . Dessa última desigualdade, podemos concluir que  $\frac{|v|}{|f(x+v) - f(x)|}$  é limitada. Além disso, como  $f$  é um homeomorfismo, então  $g = f^{-1}$  é contínua. Logo, se  $w \rightarrow 0$ , então  $v \rightarrow 0$ , o que nos leva a  $\lim_{v \rightarrow 0} \frac{r(v)}{|v|} = 0$ . Assim,

$$\lim_{w \rightarrow 0} \frac{s(w)}{|w|} = -f'(x)^{-1} \cdot \lim_{v \rightarrow 0} \frac{r(v)}{|v|} \cdot \frac{|v|}{|f(x+v) - f(x)|} = -f'(x)^{-1} \cdot 0 = 0.$$

■

A fim de concluir com as ferramentas necessárias para a demonstração principal desse capítulo, iremos demonstrar um lema que mostrará um resultado importante sobre mínimo local de uma função real. Antes, porém, vamos recordar um importante conceito, máximos e mínimos locais.

Dada uma função  $f : U \subset \mathbb{R}^n \rightarrow \mathbb{R}$ , um ponto  $a \in \mathbb{R}^n$  é dito *máximo local* de  $f$  se existe uma vizinhança  $V$  de  $a$ , tal que  $f(x) \leq f(a)$  para todo  $x \in U \cap V$ . Analogamente, um

ponto  $b \in \mathbb{R}^n$  é dito *mínimo local* de  $f$  se existe uma vizinhança  $V$  de  $b$ , tal que  $f(x) \geq f(b)$  para todo  $x \in U \cap V$ . Todo ponto que é um máximo ou mínimo local de uma função é dito *extremo local* da mesma. Se  $U$  é aberto,  $f$  é diferenciável em  $a \in U$  e  $f'(a) = 0$ , dizemos que  $a$  é um *ponto crítico* de  $f$ .<sup>2</sup> Dados esses lembretes, vamos ao lema.

**Lema 3.4** *Sejam  $U \subset \mathbb{R}^m$  aberto e  $g : U \rightarrow \mathbb{R}^n$  diferenciável no ponto  $a \in U$ , com  $g'(a) : \mathbb{R}^m \rightarrow \mathbb{R}^n$  sobrejetiva. Se  $a$  é um ponto de mínimo local de  $|g(x)|$ , com  $x \in U$ , então  $g(a) = 0$ .*

**Demonstração.** Sendo  $a$  um ponto de mínimo local para  $|g(x)|$ , então, ele também será um ponto de mínimo local para a função  $\varphi : U \rightarrow \mathbb{R}$  definida por  $\varphi(x) = |g(x)|^2 = \langle g(x), g(x) \rangle$ . Assim,  $\varphi'(a) = 0$ . No entanto, como  $\varphi'(a) \cdot v = 2\langle g'(a) \cdot v, g(a) \rangle$ , temos que  $\langle g'(a) \cdot v, g(a) \rangle = 0$ . Isso significa que  $g(a)$  é ortogonal à imagem de  $g'(a)$ , que, por ser sobrejetiva, é o  $\mathbb{R}^n$ . Portanto, só podemos concluir que  $g(a) = 0$ . ■

Dados os resultados principais, vamos ao enunciado e demonstração do Teorema da Aplicação Inversa no  $\mathbb{R}^n$ .

**Teorema 3.5 (Teorema da Aplicação Inversa)** *Seja  $f : U \rightarrow \mathbb{R}^m$  de classe  $C^k$ , com  $k \geq 1$ , no aberto  $U \subset \mathbb{R}^m$ . Se  $a \in U$  é tal que  $f'(a) : \mathbb{R}^m \rightarrow \mathbb{R}^m$  é invertível (equivalentemente:  $\det Jf(a) \neq 0$ ), então existe uma bola aberta  $B = B(a; \delta) \subset U$  tal que a restrição  $f|_B$  é um difeomorfismo sobre um aberto  $V \ni f(a)$ .*

**Demonstração.** Pelo Teorema 3.2, sendo  $f'(a) : \mathbb{R}^m \rightarrow \mathbb{R}^m$  invertível, com  $a \in U$ , existe  $\delta > 0$  tal que  $\bar{B} = B[a; \delta] \subset U$  e  $f$  é injetiva no conjunto compacto  $\bar{B}$ . Então,  $f|_{\bar{B}} : \bar{B} \rightarrow f(\bar{B})$  é uma bijeção contínua e, pelo Teorema 2.30,  $\bar{B}$  é homeomorfo a  $f(\bar{B})$ . Em particular,  $f : B = B(a; \delta) \rightarrow f(B) = V$  é um homeomorfismo. Como  $f'(x)$  depende continuamente de  $x$ , e todo operador linear suficientemente próximo de um invertível é também invertível, então  $f'(x) : \mathbb{R}^m \rightarrow \mathbb{R}^m$  é um isomorfismo para todo  $x \in B = B(a; \delta)$ . Pelo Teorema 3.3, se provarmos que  $V = f(B)$  é aberto, então  $f|_B$  será um difeomorfismo sobre  $V = f(B)$ . Seja  $q = f(p)$  com  $p \in B = B(a; \delta)$ . Chamando de  $S = S[a; \delta]$ , a esfera que é fronteira de  $\bar{B}$ , temos que  $q \notin f(S)$ , pois  $f|_{\bar{B}}$  é injetiva. Além disso, sendo  $f$  contínua no compacto  $S$  e, como  $|f(x) - q| > 0$ , pelo Teorema de Weierstrass, existe  $\varepsilon > 0$  tal que  $|f(x) - q| \geq 2\varepsilon$  para todo  $x \in S$ . Afirmamos que  $B(q; \varepsilon) \subset f(B) = V$ . De fato, dado  $y \in B(q; \varepsilon)$ , defina  $g(x) = f(x) - y$ . Temos que, o mínimo de  $|g(x)|$ , quando  $x$  varia no compacto  $\bar{B}$ , não é atingido no ponto  $x \in S$ , pois  $x \in S \Rightarrow |f(x) - y| \geq \varepsilon$ ,

<sup>2</sup> Também, dizemos que  $a$  é um ponto crítico de  $f$  quando  $\nabla f(a) = 0$ .

enquanto  $|f(p) - y| = |q - y| < \varepsilon$ , onde  $p \in B = B(a; \delta)$ . Assim, o mínimo de  $|f(x) - y|$ , com  $x \in \bar{B}$ , é atingido num ponto  $x_0 \in B = B(a; \delta)$ . De acordo com o Lema 3.4, esse mínimo vale 0, então  $|g(x_0)| = 0 \Rightarrow |f(x_0) - y| = 0 \Rightarrow y = f(x_0) \in f(B) = V$ , ou seja,  $B(q; \varepsilon) \subset f(B) = V$ . Portanto,  $f(B)$  é aberto, como queríamos demonstrar. ■

Iremos agora ver alguns problemas básicos que podem ser resolvidos facilmente com o Teorema da Aplicação Inversa.

**Exemplo 3.6** *Toda matriz próxima da identidade  $I_n$  tem raiz quadrada.*

Considere as matrizes  $X$  e  $Y \in M(n \times n)$ . Dizemos que  $X$  é raiz quadrada de  $Y$  quando  $X^2 = Y$ . Visto que  $\det(X^2) = (\det X)^2$ , uma condição necessária para que  $Y$  tenha uma raiz quadrada é que  $\det Y \geq 0$ . No entanto, essa condição não é suficiente, pois embora a matriz  $Y = \begin{bmatrix} -1 & 0 \\ 1 & -1 \end{bmatrix}$  tenha determinante  $> 0$ , não existe  $X \in M(2 \times 2)$  tal que  $X^2 = Y$  (fato facilmente verificável). Vamos usar o Teorema da Aplicação Inversa para mostrar que toda matriz próxima da identidade possui raiz quadrada. Considere a aplicação  $f : M(n \times n) \rightarrow M(n \times n)$  dada por  $f(X) = X^2$ . Note que  $f$  é uma aplicação de classe  $C^\infty$ . Sua derivada num ponto  $X \in M(n \times n)$  é a transformação linear  $f'(X) : \mathbb{R}^{n^2} \rightarrow \mathbb{R}^{n^2}$ , dada por  $f'(X) \cdot Y = Y \cdot X + X \cdot Y$ . Em particular, se  $X = I_n$ , teremos  $f'(I_n) \cdot Y = 2Y$ . Logo,  $f'(I_n) : \mathbb{R}^{n^2} \rightarrow \mathbb{R}^{n^2}$  é invertível. Pelo Teorema da Aplicação Inversa, existe um aberto  $B$  em  $M(n \times n)$ , com  $I_n \in B$ , tal que  $f|_B$  é um difeomorfismo sobre o aberto  $V = f(B)$  que contem  $f(I_n)$ . Isto significa que, para toda matriz  $Y \in V$ , existe uma única matriz  $X = \sqrt{Y} \in B$  tal que  $X^2 = Y$ . Além disso, pelo Teorema 3.1, a aplicação inversa  $f^{-1} : V \rightarrow U$ , dada por  $Y = \sqrt{Y}$ , é de classe  $C^\infty$ .

**Exemplo 3.7** *Para todo  $k \in \mathbb{N}$ , existem abertos  $U, V \subset M(n \times n) = \mathbb{R}^{n^2}$  tais que toda matriz  $Y \in V$  possui uma única raiz  $k$ -ésima  $X \in U$ , isto é, tal que  $X^k = Y$ .*

A prova é análoga ao exemplo anterior. A aplicação  $f : \mathbb{R}^{n^2} \rightarrow \mathbb{R}^{n^2}$ , dada por  $f(X) = X^k$ , é de classe  $C^\infty$ . Sua derivada no ponto  $X \in \mathbb{R}^{n^2}$  é a transformação linear  $f'(X) \cdot Y = \sum_{i=1}^k X^{i-1} \cdot Y \cdot X^{k-i}$ . No ponto  $X = I_n$ , temos que  $f'(I_n) \cdot Y = k \cdot Y$ . Logo,  $f'(I_n) : \mathbb{R}^{n^2} \rightarrow \mathbb{R}^{n^2}$  é um isomorfismo. Pelo Teorema da Aplicação Inversa, existe um aberto  $B \subset \mathbb{R}^{n^2}$ , com  $I_n \in B$ , tal que  $f|_B$  é um difeomorfismo sobre o aberto  $V = f(B)$  que contém  $f(I_n)$ , ou seja, toda matriz  $Y \in V$  possui uma única raiz  $k$ -ésima  $X \in U$ .

**Exemplo 3.8** *Sejam  $\varphi : U \rightarrow \mathbb{R}^m$  de classe  $C^1$  no aberto  $U \subset \mathbb{R}^m$  e  $c \in [0, 1)$  tais que  $|\varphi(x) - \varphi(y)| \leq c|x - y|$  para quaisquer  $x, y \in U$ . Iremos mostrar que  $f : U \rightarrow \mathbb{R}^m$ , dada por  $f(x) = x + \varphi(x)$ , é um difeomorfismo de  $U$  sobre o aberto  $V = f(U) \subset \mathbb{R}^m$ . Além disso, se  $U = \mathbb{R}^m$ , então  $f(U) = \mathbb{R}^m$ .*

Do Exemplo 2.35, temos que  $|\varphi'(x) \cdot v| < c|v| < |v|$  para todo  $x \in U$  e todo  $v \in \mathbb{R}^m - \{0\}$ . Logo,  $|f'(x) \cdot v| = |v + \varphi'(x) \cdot v| \geq |v| - |\varphi'(x) \cdot v| > 0$ , se  $v \neq 0$ . Isso mostra que  $f'(x) : \mathbb{R}^m \rightarrow \mathbb{R}^m$  é um isomorfismo, para todo  $x \in U$ . Pelo Teorema da Aplicação Inversa,  $f$  é um difeomorfismo local, que transforma cada aberto  $A \subset U$  num aberto  $f(A) \subset \mathbb{R}^m$ . Além disso, dados  $x, y \in U$ , com  $x \neq y$  temos que

$$|f(x) - f(y)| = |x - y + \varphi(x) - \varphi(y)| \geq |x - y| - c|x - y| = (1 - c)|x - y| > 0,$$

o que nos dá  $f$  injetiva. Portanto,  $f : U \rightarrow \mathbb{R}^m$  é um difeomorfismo de  $U$  sobre o aberto  $V = f(U) \subset \mathbb{R}^m$ . Agora, suponha que  $U = \mathbb{R}^m$ . Como  $\mathbb{R}^m$  é conexo, para que tenhamos  $f(\mathbb{R}^m) = \mathbb{R}^m$ , devemos ter  $f(\mathbb{R}^m)$  aberto (fato já mostrado) e fechado. Para provar que  $f(\mathbb{R}^m)$  é fechado, seja  $(x_k)_{k \in \mathbb{N}}$  uma sequência tal que  $\lim_{k \rightarrow \infty} f(x_k) = y \in \mathbb{R}^m$ . Como  $|x_k - x_r| \leq \frac{1}{1-c}|f(x_k) - f(x_r)|$ , então  $(x_k)_{k \in \mathbb{N}}$  é uma sequência de Cauchy e, portanto, converge para um certo  $x$ . Daí,  $f(x) = \lim_{k \rightarrow \infty} f(x_k) = y \in f(\mathbb{R}^m)$ . Logo,  $f(\mathbb{R}^m)$  é fechado.

## 3.2 Aplicações

Nessa seção, iremos apresentar alguns resultados de maior relevância usando o Teorema da Aplicação Inversa. O primeiro deles, o notável Teorema da Função Implícita.

### 3.2.1 Teorema da Função Implícita

Com o intuito de nos servir de motivação, considere o caso particular de uma função  $f$  definida num aberto  $U \subset \mathbb{R}^2$ . O Teorema da Função Implícita estabelece condições para que o conjunto de pontos  $(x, y) \in U$  que cumprem a igualdade  $f(x, y) = c$ ,  $c \in \mathbb{R}$ , 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  $C^k$ , com  $k \geq 1$ , 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  $f(x, y) = c$ ,  $c \in \mathbb{R}$ , define, em muitos casos, o traço de uma curva, a qual denominamos *curva de nível* de  $f$ . Assim, decorre do

Teorema da Função Implícita que uma curva de nível de uma função de classe  $C^k$ , com  $k \geq 1$ , seja, localmente, o gráfico de uma função de uma única variável. Nesse caso, pelo menos uma das derivadas parciais de seus pontos deve ser diferente de zero.

Por exemplo, considerando a função de classe  $C^1$ , dada por

$$f(x, y) = x^2 + y^2, (x, y) \in \mathbb{R}^2,$$

temos que o conjunto de pontos  $(x, y) \in \mathbb{R}^2$  que cumprem a igualdade  $f(x, y) = 1$  é a esfera  $\mathbb{S}^1$ . Geometricamente, para um dado ponto  $(a, b) \in \mathbb{S}^1$ , distinto de  $(1, 0)$  e  $(-1, 0)$ , existe uma vizinhança do mesmo em  $\mathbb{S}^1$  que é o gráfico de uma função  $y = y(x)$ . Observe que  $\frac{\partial f}{\partial y}(a, b) = 2b \neq 0$ , enquanto que  $\frac{\partial f}{\partial y}(1, 0) = \frac{\partial f}{\partial y}(-1, 0) = 0$ .

No enunciado do teorema a seguir,  $f'(a, b)|_{\{0\} \times \mathbb{R}^m}$  é um isomorfismo no sentido de ser invertível a matriz  $m \times n$  cujos vetores-coluna são os  $m$  últimos vetores-coluna da matriz jacobiana de  $f$  em  $(a, b)$ . Dito isto, e sendo  $f$  de classe  $C^k$ , 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  $\xi$ , de classe  $C^k$ , definida num aberto de  $\mathbb{R}^n$  e que toma valores em  $\mathbb{R}^m$ . Vamos ao enunciado.

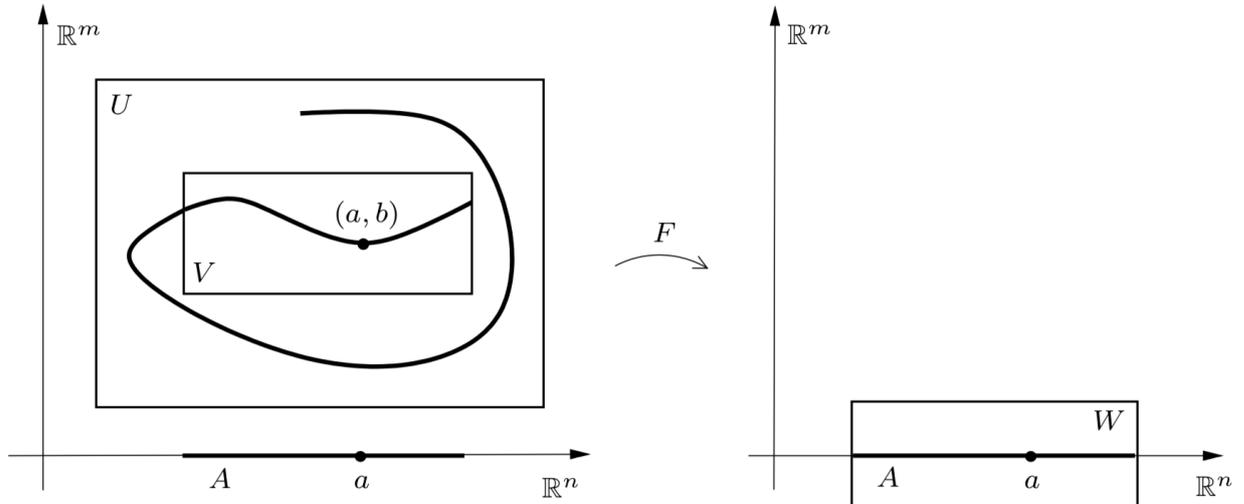
**Teorema 3.9 (Teorema da Função Implícita)** *Sejam  $U \subset \mathbb{R}^{n+m} = \mathbb{R}^n \times \mathbb{R}^m$  um conjunto aberto,  $f : U \rightarrow \mathbb{R}^m$ , uma aplicação de classe  $C^k$ , com  $k \geq 1$ , e  $(a, b) \in U$ , tais que  $f(a, b) = c \in \mathbb{R}^m$ . Nessas condições, 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:*

1. *Para todo  $x \in A$ , existe um único  $y = \xi(x) \in \mathbb{R}^m$ , tal que  $(x, \xi(x)) \in V$  e  $f(x, \xi(x)) = c$ ;*
2. *A aplicação  $\xi(x) \in \mathbb{R}^m$  é de classe  $C^k$  e  $\xi'(a) = -T_2^{-1}T_1$ , em que  $T_1 \in \mathcal{L}(\mathbb{R}^n; \mathbb{R}^m)$  e  $T_2 \in \mathcal{L}(\mathbb{R}^m)$  são definidas por  $T_1 \cdot h = f'(a, b)(h, 0)$  e  $T_2 \cdot k = f'(a, b)(0, k)$ .*

**Demonstração.** Sem perda de generalidade, suponha que  $c = 0$  e defina  $F : U \rightarrow \mathbb{R}^n \times \mathbb{R}^m$ , onde  $F(x, y) = (x, f(x, y))$ , ou seja, as coordenadas de  $F$  são a projeção  $P_1(x, y) = x$  e a aplicação  $f$ , onde  $F$  é de classe  $C^k$  e

$$F'(a, b)(h, k) = (h, f'(a, b)(h, k)),$$

onde  $(h, k) \in \mathbb{R}^n \times \mathbb{R}^m$ . Em particular, se  $F'(a, b)(h, k) = (0, 0)$ , teremos  $h = 0$  e  $0 = f'(a, b)(h, k) = f'(a, b)(0, k)$ , onde  $k = 0$ , visto que, por hipótese,  $f'(a, b)|_{\{0\} \times \mathbb{R}^m}$  é um isomorfismo. Então,  $F'(a, b)$  é um isomorfismo de  $\mathbb{R}^n \times \mathbb{R}^m$  sobre si mesmo. Pelo Teorema da Aplicação Inversa, 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  $F|_V : V \rightarrow W$  é um difeomorfismo de classe  $C^k$ . Observe a figura a seguir.

Figura 1 – Aplicação  $F$ 

Fonte: (LIMA, 2015)

Sem perda de generalidade, vamos supor que  $W = A \times B$ , em que  $A$  e  $B$  são abertos de  $\mathbb{R}^n$  e  $\mathbb{R}^m$ , respectivamente, e  $a \in A$ . Então, dado  $x \in A$ , existe um único  $y = \xi(x) \in V$  e  $F(x, y) = (x, f(x, y)) = (x, 0) \in W$ , pois  $F|_V \rightarrow W$  é bijetiva. Portanto, a aplicação  $\xi$ , assim definida, satisfaz  $f(x, \xi(x)) = 0$  para todo  $x \in A$ , o que conclui a prova de 1. Para a prova de 2, considere  $G = (F|_V)^{-1}$ . Pelo Teorema da Aplicação Inversa,  $G$  é de classe  $C^k$ . Além disso,  $G(x, 0) = (x, \xi(x))$ . Definindo a projeção  $P_2(x, y) = y$  e  $\phi(x) = (x, 0)$ , com  $x \in \mathbb{R}^n$ , teremos  $\xi = P_2 \circ G \circ \phi$ , onde  $\xi$  é de classe  $C^k$ . Note que, para todo  $(h, k) \in \mathbb{R}^n \times \mathbb{R}^m$ , tem-se

$$f'(a, b)(h, k) = f'(a, b)((h, 0) + (0, k)) = f'(a, b)(h, 0) + f'(a, b)(0, k) = T_1 \cdot h + T_2 \cdot k.$$

Derivando ambos os membros de  $f(x, \xi(x)) = 0$  e fazendo  $x = a$ , teremos

$$0 = f'(a, \xi(a))(h, \xi'(a)h) = f'(a, b)(h, \xi'(a)h) = T_1 \cdot h + T_2 \cdot \xi'(a)h, \forall h \in \mathbb{R}^n,$$

donde, concluímos que  $\xi'(a) = -T_2^{-1}T_1$ . Isso encerra a prova de 2 e, assim, concluímos a demonstração do teorema. ■

Vale destacar que dada uma aplicação  $f : U \subset \mathbb{R}^n \times \mathbb{R}^m \rightarrow \mathbb{R}^m$ , pelo Teorema da Função Implícita, se  $f$  é uma submersão<sup>3</sup> 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  $\mathbb{R}^n$  e tomando valores em  $\mathbb{R}^m$ .

<sup>3</sup> Dado um aberto  $U \subset \mathbb{R}^n$ , uma aplicação diferenciável  $f : U \rightarrow \mathbb{R}^m$ , cuja derivada  $f'(x)$  é sobrejetiva para todo  $x \in U$ , é dita uma *submersão*.

### 3.2.2 Multiplicadores de Lagrange

Iniciaremos nossas considerações com as funções reais  $f : U \rightarrow \mathbb{R}$ , de classe  $C^1$  definida no aberto  $U \subset \mathbb{R}^{n+1}$  com somente uma restrição. O método de multiplicador de Lagrange se aplica na seguinte situação: considere uma função  $f : U \rightarrow \mathbb{R}$ , de classe  $C^1$ , uma hipersuperfície  $M = \varphi^{-1}(c)$ , que é imagem inversa de um valor regular  $c$  da função  $\varphi : U \rightarrow \mathbb{R}$ , de classe  $C^1$ , e procura-se determinar quais são os pontos críticos da restrição  $f|_M$ , ou seja, os pontos críticos  $x$  de  $f$  sujeitos à condição  $\varphi(x) = c$ . Não se trata de determinar os pontos críticos de  $f : U \rightarrow \mathbb{R}$  que estão localizados sobre a hipersuperfície  $M$ , mas sim os pontos críticos da função  $f|_M : M \rightarrow \mathbb{R}$ .

Um ponto  $x \in M$  chama-se um *ponto crítico da restrição*  $f|_M$  quando, para todo caminho diferenciável  $\lambda : (-\delta, \delta) \rightarrow M$  com  $\lambda(0) = x$  tem-se  $(f \circ \lambda)'(0) = 0$ . Pondo  $v = \lambda'(0)$ , esta condição significa  $\langle \nabla f(x), v \rangle = 0$ . Como  $v$  é um vetor arbitrário pertencente ao espaço vetorial tangente  $T_x M$ , vemos que  $x \in M$  é um ponto crítico de  $f|_M$  se, e somente se,  $\nabla f(x)$  é ortogonal ao espaço vetorial tangente  $T_x M$ .

Temos que  $\nabla \varphi(x)$  é um vetor não-nulo ortogonal a  $T_x M$ . Como o complemento ortogonal<sup>4</sup> de  $T_x M$  em  $\mathbb{R}^{n+1}$  tem dimensão 1, segue-se que  $\nabla f(x) \perp T_x M$  se, e somente se,  $\nabla f(x)$  é um múltiplo de  $\nabla \varphi(x)$ . Nessas condições, temos o **Teorema do Multiplicador de Lagrange**: *o ponto  $x \in U$  é um ponto crítico da restrição  $f|_M$  de  $f$  à hipersuperfície  $M = \varphi^{-1}(c)$  se, e somente se:*

1.  $\varphi(x) = c$ ;
2.  $\nabla f(x) = \lambda \cdot \nabla \varphi(x)$  para algum  $\lambda \in \mathbb{R}$ .

Faremos a demonstração geral desse teorema logo adiante. As condições do Teorema do Multiplicador de Lagrange representam um sistema de  $n + 2$  equações nas  $n + 2$  incógnitas  $x_1, x_2, \dots, x_{n+1}, \lambda$ . O fator  $\lambda$  é chamado o *multiplicador de Lagrange*. Sua presença torna o número de incógnitas igual ao número de equações, o que viabiliza a solução na prática.

Note que se  $x \in M$  é um ponto de mínimo ou de máximo local de  $f|_M$ , então para todo caminho diferenciável  $\lambda : (-\delta, \delta) \rightarrow M$  com  $\lambda(0) = x$ , a função  $f \circ \lambda : (-\delta, \delta) \rightarrow \mathbb{R}$  tem um mínimo ou um máximo local no ponto 0. Logo,  $(f \circ \lambda)'(0) = 0$ . Portanto, os mínimos e máximos locais de  $f|_M$  estão incluídos na definição de ponto crítico dada pelo Teorema do Multiplicador de Lagrange.

Vale ressaltar, também, que todo ponto crítico  $x \in M$  da função  $f : U \rightarrow \mathbb{R}$  é, com

<sup>4</sup> O complemento ortogonal de  $T_x M$  é o conjunto  $T_x M^\perp = \{w \in \mathbb{R}^n; \langle u, w \rangle = 0 \forall u \in T_x M\}$ .

maior razão, ponto crítico da restrição  $f|_M$  pois, sendo  $\nabla f(x) = 0$ , tem-se  $\langle \nabla f(x), v \rangle = 0$  para todo  $v \in \mathbb{R}^{n+1}$ .

Muitas vezes, a condição adicional  $\varphi(x) = c$  é escrita como  $\varphi(x) = 0$ . Isso não representa perda de generalidade. Basta usar, em vez de  $\varphi$ , a função  $\psi(x) = \varphi(x) - c$ . Então,  $\psi(x) = 0 \Leftrightarrow \varphi(x) = c$  e  $c$  é o valor regular de  $\psi$ .

Falando agora de maneira mais geral, o método dos multiplicadores de Lagrange 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  $C^1$ ,  $g : U \rightarrow \mathbb{R}^m$ . Nessas condições, (falaremos detalhadamente sobre isso no próximo teorema) 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  $g$ :  $g_1, g_2, \dots, g_m$ , em  $p$ , isto é, devem existir  $\lambda_1, \lambda_2, \dots, \lambda_m \in \mathbb{R}$  tais que

$$\nabla f(p) = \lambda_1 \nabla g_1(p) + \lambda_2 \nabla g_2(p) + \dots + \lambda_m \nabla g_m(p).$$

Nesse contexto, os reais  $\lambda_1, \lambda_2, \dots, \lambda_m$  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  $\lambda_1, \lambda_2, \dots, \lambda_m$ , são as variáveis do sistema de equações vetoriais:

$$\begin{cases} \nabla f(p) &= \lambda_1 \nabla g_1(p) + \lambda_2 \nabla g_2(p) + \dots + \lambda_m \nabla g_m(p) \\ g(p) &= c \end{cases}.$$

Dadas essas considerações iniciais, vamos ao enunciado e demonstração do Teorema dos Multiplicadores de Lagrange.

**Teorema 3.10 (Multiplicadores de Lagrange)** *Sejam  $U \subset \mathbb{R}^n \times \mathbb{R}^m$  um conjunto aberto,  $f : U \rightarrow \mathbb{R}$  uma função diferenciável e  $M \subset U$  um conjunto de nível de uma submersão de classe  $C^1$ ,  $g = (g_1, g_2, \dots, g_m) : U \rightarrow \mathbb{R}^m$ . Então, se  $p \in M$  é um extremo local de  $f|_M$ , existem  $\lambda_1, \lambda_2, \dots, \lambda_m \in \mathbb{R}$ , tais que*

$$\nabla f(p) = \lambda_1 \nabla g_1(p) + \lambda_2 \nabla g_2(p) + \dots + \lambda_m \nabla g_m(p).$$

**Demonstração.** Seja  $p = (x_0, y_0) \in M$  um extremo local de  $f|_M$ . Uma vez que  $g$  é uma submersão, temos que  $g'(p) \in \mathcal{L}(\mathbb{R}^{n+m}, \mathbb{R}^m)$  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  $C^1$ ,  $\xi : A \rightarrow \mathbb{R}^m$ , tais que o gráfico de  $\xi$  está contido em  $M$ , ou seja,  $(x, \xi(x)) \in M$  para todo  $x \in A$ . Considere o gráfico de  $\xi'(x_0)$ ,  $\mathbb{G} := \{(h, \xi'(x_0)h); h \in \mathbb{R}^n\} \subset \mathbb{R}^n \times \mathbb{R}^m$ , e observe que este é um subespaço vetorial de  $\mathbb{R}^n \times \mathbb{R}^m$  de dimensão  $n$  (note que  $\{(e_1, \xi'(x_0)e_1), (e_2, \xi'(x_0)e_2), \dots, (e_n, \xi'(x_0)e_n)\}$  é uma base de  $\mathbb{G}$ ). Sendo  $M$  um conjunto de nível de  $g$ , temos que  $g(x, \xi(x))$  é constante. Em particular, dado um vetor  $h \in \mathbb{R}^n$ , tem-se  $g'(x_0, \xi(x_0))(h, \xi'(x_0)h) = 0$ , isto é,  $g'(p)(\mathbb{G}) = 0$ . Logo, para todo  $v = (h, \xi'(x_0)h) \in \mathbb{G}$ , teremos

$$0 = g'(p)v = (g'_1(p)v, g'_2(p)v, \dots, g'_m(p)v) = (\langle \nabla g'_1(p), v \rangle, \langle \nabla g'_2(p), v \rangle, \dots, \langle \nabla g'_m(p), v \rangle),$$

donde concluímos que, para cada  $j \in \{1, 2, \dots, m\}$ ,  $\nabla g'_j(p)$  é ortogonal a  $\mathbb{G}$ . Porém, os vetores-linha da matriz jacobiana de  $g$  em  $p$  são, justamente,  $\nabla g'_1(p), \nabla g'_2(p), \dots, \nabla g'_m(p)$  e, portanto, são linearmente independentes, já que  $g'(p)$  tem posto<sup>5</sup>  $m$ . Segue daí que, o complemento ortogonal de  $\mathbb{G}$  em  $\mathbb{R}^n \times \mathbb{R}^m$  é o subespaço gerado pelos vetores  $\nabla g'_1(p), \nabla g'_2(p), \dots, \nabla g'_m(p)$ . Um 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,  $v = (h, \xi'(x_0)h) \in \mathbb{G}$ , teremos

$$0 = f'(x_0, \xi(x_0))(h, \xi'(x_0)h) = \langle \nabla f(p), v \rangle,$$

onde  $f(p) \in \mathbb{G}^\perp$ . Logo, existem reais  $\lambda_1, \lambda_2, \dots, \lambda_m$ , tais que

$$\nabla f(p) = \lambda_1 \nabla g_1(p) + \lambda_2 \nabla g_2(p) + \dots + \lambda_m \nabla g_m(p),$$

como queríamos demonstrar. ■

**Exemplo 3.11** *Seja  $f : \mathbb{R}^2 \rightarrow \mathbb{R}$  definida por  $f(x, y) = ax + by$ , com  $a^2 + b^2 \neq 0$ . Quais os pontos críticos da restrição de  $f$  ao círculo unitário  $\mathbb{S}^1$ ?*

Temos que  $\nabla f(x, y) = (a, b)$ ,  $\varphi(x, y) = x^2 + y^2$ ,  $\mathbb{S}^1 = \varphi^{-1}(1)$  e  $\nabla \varphi(x, y) = (2x, 2y)$ .

Os pontos críticos de  $f|_{\mathbb{S}^1}$  são aqueles onde os vetores  $(a, b)$  e  $(2x, 2y)$  são colineares. Além disso, deve-se ter  $x^2 + y^2 = 1$ , isso nos dá  $x = \frac{a}{\sqrt{a^2 + b^2}}$ ,  $y = \frac{b}{\sqrt{a^2 + b^2}}$  ou  $x = \frac{-a}{\sqrt{a^2 + b^2}}$ ,  $y = \frac{-b}{\sqrt{a^2 + b^2}}$ . Nesses pontos,  $f|_{\mathbb{S}^1}$  assume, respectivamente, seu valor máximo igual a  $\sqrt{a^2 + b^2}$ , e seu valor mínimo igual a  $-\sqrt{a^2 + b^2}$ .

**Exemplo 3.12** *Dados  $a \in \mathbb{R}^{n+1}$  e uma hipersuperfície  $M \subset \mathbb{R}^{n+1}$ , que não contenha o ponto  $a$ , determinar o ponto  $p \in M$  mais próximo de  $a$ .*

<sup>5</sup> O posto de uma transformação linear é a dimensão de sua imagem.

Considere a função  $f(x) = |x - a|$ , que é de classe  $C^\infty$  em  $\mathbb{R}^{n+1} - \{a\}$ . Procuramos os pontos onde  $f|_M$  assume seu valor mínimo. Eles estão entre os pontos críticos de  $f|_M$ , isto é, entre os pontos nos quais  $\nabla f(x)$  é normal a  $M$ . Ora, como  $f(x) = \sqrt{\sum(x_i - a_i)^2}$ , temos  $\frac{\partial f}{\partial x_i} = \frac{x_i - a_i}{|x - a|}$ , donde  $\nabla f(x) = \frac{x - a}{|x - a|}$ . Assim, os pontos críticos de  $f|_M$  são os pontos  $x \in M$  tais que  $x - a$  é normal a  $M$ .

**Exemplo 3.13** Considere a função  $f : \mathbb{R}^3 \rightarrow \mathbb{R}$ ,  $f(x, y, z) = x \cdot y \cdot z$ , e, para cada  $c > 0$ , procuremos seu valor máximo na superfície  $M = \{(x, y, z) \in \mathbb{R}^3; x + y + z = c, y > 0, z > 0\}$ . Na prática, iremos mostrar que, dados três números reais  $x, y$  e  $z$ , sua média geométrica é menor do que ou igual à média aritmética destes números.

O valor máximo mencionado acima existe em  $\bar{M}$  visto que ele é um conjunto compacto mas, na realidade, está em  $M$  pois  $f$  é positiva em  $M$  e nula em  $\bar{M} - M$ . Temos que  $\nabla f(x, y, z) = (yz, xz, xy)$  e  $M$  é uma parte de  $\varphi^{-1}(c)$ , com  $\varphi(x, y, z) = x + y + z$ , donde  $\nabla \varphi(x, y, z) = (1, 1, 1)$ . Logo, num ponto de  $M$  onde  $f|_M$  seja máxima, devemos ter  $yz = xz = xy = \lambda$  e  $x + y + z = c$ . Assim,  $x = y = z = \frac{c}{3}$  e  $f(x, y, z) = \frac{c^3}{27}$ . Sendo esse valor máximo de  $f$  em  $M$ , devemos ter  $xyz \leq \frac{c^3}{27} = \left(\frac{x + y + z}{3}\right)^3$ , sempre que  $x > 0, y > 0, z > 0$  e  $x + y + z = c$ . Mas  $c$  é arbitrário e três números positivos  $x, y, z$  têm sempre uma soma  $c$ . Podemos, então, afirmar que, dados 3 números positivos  $x, y, z$ , tem-se  $xyz \leq \left(\frac{x + y + z}{3}\right)^3$ , ou seja,  $\sqrt[3]{xyz} \leq \left(\frac{x + y + z}{3}\right)$ , como queríamos demonstrar.

**Exemplo 3.14 (Desigualdade de Hadamard)** <sup>6</sup> Se  $X$  é uma matriz  $n \times n$  cujas linhas são os vetores  $X_i = (x_{i1}, x_{i2}, \dots, x_{in})$ ,  $i = 1, 2, \dots, n$ , então  $\det X \leq |X_1| \cdot |X_2| \cdots |X_n|$ , onde  $|X_i|$  é a norma euclidiana de  $X_i$ .

O resultado é evidente se  $\det X = 0$ . Caso  $\det X \neq 0$ , então todos os vetores-linha são diferentes de zero. Logo,  $X_i = |X_i| \cdot W_i$ , com  $|W_i| = 1$ , para todo  $i = 1, 2, \dots, n$ . Então,  $\det X = |X_1| \cdot |X_2| \cdots |X_n| \cdot \det W$ , onde  $W$  é a matriz cujas linhas são os vetores unitários  $W_1, W_2, \dots, W_n$ . A desigualdade de Hadamard ficará provada se mostrarmos que  $\det W \leq 1$ . De um modo mais geral, mostraremos que se  $W = (w_{ij})$  é uma matriz  $n \times n$ , tal que  $\sum_{i,j} w_{ij}^2 = n$ , então  $\det W \leq 1$ . Defina  $f, \varphi : \mathbb{R}^n \rightarrow \mathbb{R}$ , pondo  $f(X) = \det X$  e  $\varphi(X) = \sum_{i,j} x_{ij}^2$ . Temos  $\frac{\partial \varphi}{\partial x_{ij}}(X) = 2x_{ij}$  e  $\frac{\partial f}{\partial x_{ij}}(X) = (-1)^{i+j} X_{[ij]}$ , onde  $X_{[ij]}$  é o determinante da matriz  $(n-1) \times (n-1)$ , obtida de  $\varphi$  pela omissão

<sup>6</sup> Jacques Salomon Hadamard (1865-1963) foi um matemático francês.

da  $i$ -ésima linha e da  $j$ -ésima coluna. Todo número real diferente de zero é um valor regular de  $\varphi$ . Logo,  $M = \varphi^{-1}(n)$  é uma hipersuperfície compacta de classe  $C^\infty$  em  $\mathbb{R}^{n^2}$ , que é a esfera de centro 0 e raio  $\sqrt{n}$ . Pelo Teorema do Multiplicador de Lagrange, uma matriz  $W = (w_{ij})$  é ponto crítico de  $f|_M$  se, e somente se,  $\sum_{i,j} w_{ij}^2 = n$  e  $\nabla f(W) = \lambda \cdot \nabla \varphi(W)$  para algum  $\lambda$  real. Daí, teremos

$$(-1)^{i+j} W_{[ij]} = 2\lambda \cdot w_{ij} \quad (3.2)$$

para quaisquer  $i, j \in [1, n]$ . Multiplicando por  $w_{ij}$ , somando e levando em conta a regra de expansão de um determinante em relação aos elementos de uma linha, temos

$$n \cdot \det W = \sum_{i,j} (-1)^{i+j} w_{ij} W_{[ij]} = 2\lambda \cdot \sum w_{ij}^2 = 2\lambda \cdot n,$$

donde  $\det W = 2\lambda$ . Agora, multipliquemos (3.2) por  $w_{ij}$ , fixemos  $i$  e somemos em relação a  $j$ .

Resulta em

$$\det W = \sum_j (-1)^{i+j} w_{ij} W_{[ij]} = 2\lambda \cdot \sum_j w_{ij}^2 = \det W \cdot \sum_j w_{ij}^2.$$

Suponhamos que  $W$  seja um ponto onde  $f|_M$  assume seu valor máximo. Então,  $\det W = f(W) \neq 0$  e da igualdade acima vem  $|W_i|^2 = \sum_j w_{ij}^2 = 1$  para todo  $i$ . Em seguida, multipliquemos (3.2) por  $w_{kj}$ , com  $k \neq i$ , e somemos em relação a  $j$ . Teremos

$$\sum_j (-1)^{i+j} w_{kj} W_{[ij]} = 2\lambda \cdot \sum_j w_{kj} w_{ij} = 2\lambda \langle W_k, W_i \rangle.$$

Note que o primeiro somatório acima é zero, por ser o desenvolvimento em relação aos elementos da  $i$ -ésima linha do determinante de uma matriz com duas linhas, a  $i$ -ésima e  $k$ -ésima linhas, iguais a  $W_k$ . Logo,  $\langle W_k, W_i \rangle = 0$  para  $k \neq i$ . Assim, todo ponto  $W \in M$ , onde  $f|_M$  atinja o seu valor máximo, é uma matriz cujas linhas são vetores unitários dois a dois ortogonais, isto é,  $W$  é uma matriz ortogonal. Em particular,  $\det W = \pm 1$ . Por ser  $\det W$  máximo, seu valor é evidentemente 1. Concluimos que  $\det W \leq 1$  para todo  $W \in M$ , o que demonstra a desigualdade de Hadamard.

### 3.2.3 Forma Local das Imersões

Uma *imersão* do aberto  $U \subset \mathbb{R}^m$  no espaço euclidiano  $\mathbb{R}^n$  é uma aplicação diferenciável  $f : U \rightarrow \mathbb{R}^n$  tal que, para cada  $x \in U$ , a derivada  $f'(x) : \mathbb{R}^m \rightarrow \mathbb{R}^n$  é uma transformação linear injetiva, o que ocorre somente quando  $m \leq n$ . A aplicação composta de duas imersões

é ainda uma imersão. Se a derivada  $f' : U \rightarrow \mathcal{L}(\mathbb{R}^m; \mathbb{R}^n)$  for contínua no ponto  $a \in U$ , então, pelo Teorema 3.2, a condição de  $f'(a) : \mathbb{R}^m \rightarrow \mathbb{R}^n$  ser injetiva implica que  $f$  é injetiva numa vizinhança de  $a$  e, mais precisamente,  $f$  transforma uma bola de centro  $a$  homeomorficamente sobre sua imagem. Vejamos algumas situações.

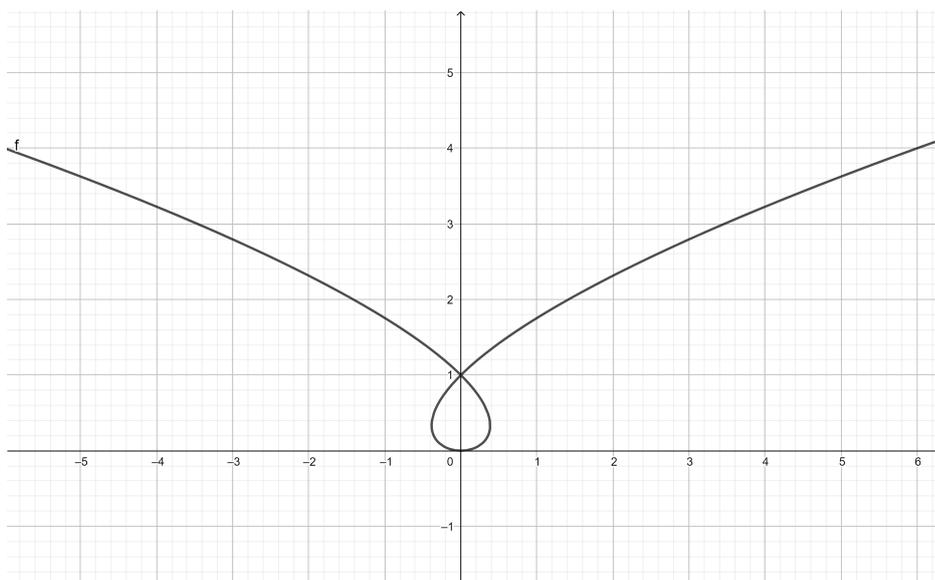
**Exemplo 3.15** Seja  $f : \mathbb{R}^m \rightarrow \mathbb{R}^m \times \mathbb{R}^n$  a aplicação de inclusão, dada por  $f(x) = (x, 0)$ .

Como  $f$  é linear, temos  $f'(x) = f(x)$  para todo  $x \in \mathbb{R}^m$ , logo  $f$  é uma imersão de classe  $C^\infty$ . O Teorema da Forma Local das Imersões mostra que toda imersão de classe  $C^k$  ( $k \geq 1$ ) se comporta localmente como esta.

**Exemplo 3.16** Seja  $J \subset \mathbb{R}$  um intervalo aberto. Um caminho diferenciável  $f : J \rightarrow \mathbb{R}^n$  é uma imersão se, e somente se, seu vetor-velocidade é diferente de zero em cada ponto  $t \in J$ .

Isto significa que a imagem  $f(J)$  possui, em cada ponto  $f(t)$ , uma reta tangente, a saber, a reta  $L = \{f(t) + s \cdot f'(t); s \in \mathbb{R}\}$ . Observe que uma imersão pode não ser injetiva. Então, quando  $f(t_1) = f(t_2)$ , as duas retas tangentes  $L_1 = f(t_1) + \mathbb{R} \cdot f'(t_1)$  e  $L_2 = f(t_2) + \mathbb{R} \cdot f'(t_2)$  podem, ou não, ser distintas. Se considerarmos, entretanto, uma pequena vizinhança  $V$  de  $t_1$ , nenhum outro ponto de  $V$  terá a mesma imagem que  $t_1$  e assim  $L_1$  será a única reta tangente no ponto  $f(t_1)$  para o caminho restrito  $f|_V$ . Um exemplo simples de imersão da reta no plano é dado por  $f : \mathbb{R} \rightarrow \mathbb{R}^2$ ,  $f(t) = (t^3 - t, t^2)$ . A imagem de  $f$  é esboçada na figura abaixo.

Figura 2 – Gráfico de  $f(t) = (t^3 - t, t^2)$

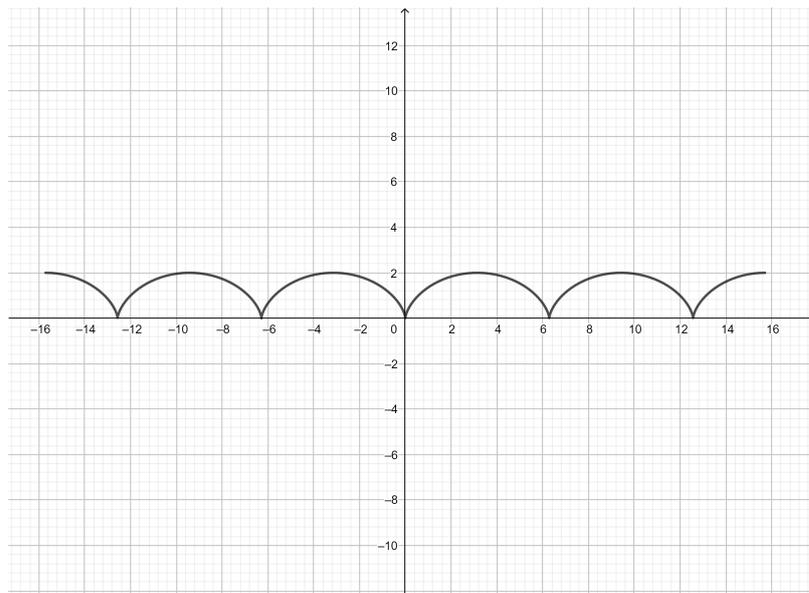


Fonte: Elaborada pelo autor

**Exemplo 3.17** O caminho  $g : \mathbb{R} \rightarrow \mathbb{R}^2$ , definido por  $g(t) = (t - \operatorname{sen} t, 1 - \operatorname{cos} t)$ , é injetivo, de classe  $C^\infty$ , mas não é uma imersão de  $\mathbb{R}$  em  $\mathbb{R}^2$ .

A imagem de  $g$ , uma curva chamada *ciclóide* (figura a seguir), exhibe uma infinidade de pontos angulares chamados de cúspides, nos quais o vetor-velocidade  $g'(t)$  deve obrigatoriamente ser igual a zero.

Figura 3 – Ciclóide



Fonte: Elaborada pelo autor

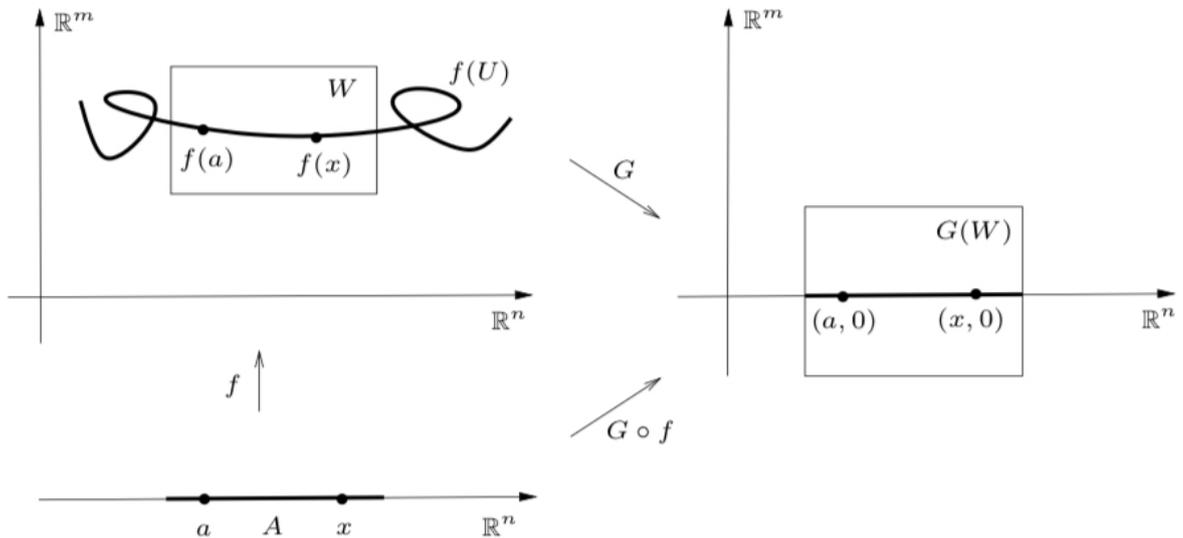
No entanto, note que os pontos onde a derivada de uma aplicação deixa de ser injetiva nem sempre se denunciam pela forma geométrica da imagem. Por exemplo, a imagem do caminho  $f : \mathbb{R} \rightarrow \mathbb{R}^2$ ,  $f(t) = (t^3, t^3)$ , é uma reta. Para  $t = 0$ , o vetor-velocidade  $f'(t)$  se anula, o que não se deve ao aspecto de  $f(\mathbb{R})$ , mas à maneira como tal conjunto está parametrizado por  $f$ .

O teorema abaixo mostrará que em toda imersão de classe  $C^k$  é possível introduzir novas coordenadas na vizinhança de cada ponto da imagem, de modo que  $f$  assumia localmente a forma do Exemplo 3.15.

**Teorema 3.18 (Forma Local das Imersões)** *Sejam  $U \subset \mathbb{R}^n$  um conjunto aberto e  $f : U \rightarrow \mathbb{R}^n \times \mathbb{R}^m = \mathbb{R}^{n+m}$  uma aplicação de classe  $C^k$  ( $k \geq 1$ ), tal que  $f$  é uma imersão em  $a \in U$ , isto é,  $f'(a) \in \mathcal{L}(\mathbb{R}^n; \mathbb{R}^{n+m})$  é injetiva. Então, existem abertos  $A \ni a$ , de  $U$ , e  $W \ni f(a)$ , de  $\mathbb{R}^{n+m}$ , bem como um difeomorfismo  $G : W \rightarrow G(W) \subset \mathbb{R}^{n+m}$ , tais que a composta  $G \circ f : A \rightarrow \mathbb{R}^{n+m}$  está bem definida e, para todo  $x \in A$ , vale a igualdade  $(G \circ f)(x) = (x, 0)$ .*

A demonstração faz uso do Teorema da Aplicação Inversa. A figura seguinte contém o significado do teorema.

Figura 4 – Forma Local das Imersões



Fonte: (LIMA, 2015)

**Demonstração.** Uma vez que  $f'(a)$  é injetiva, podemos supor, sem perda de generalidade, que  $f'(a)(\mathbb{R}^n) = \mathbb{R}^n \times \{0\} \subset \mathbb{R}^{n+m}$ . Defina  $F : U \times \mathbb{R}^m \rightarrow \mathbb{R}^{n+m}$  por

$$F(x, y) = f(x) + (0, y) = f(P_1(x, y)) + P_2(x, y),$$

em que  $P_1$  e  $P_2$  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 por  $U$  e  $U \times \{0\}$ . Dado, então  $(h, k) \in \mathbb{R}^n \times \mathbb{R}^m$ , tem-se

$$F'(a, 0)(h, k) = f'(P_1(a, 0))P_1'(a, 0)(h, k) + P_2'(a, 0)(h, k) = f'(a)h + (0, k),$$

onde concluímos que o conjunto imagem de  $F'(a, 0)$  é  $\mathbb{R}^n \times \{0\} \oplus \{0\} \times \mathbb{R}^m = \mathbb{R}^{n+m}$ <sup>7</sup>. Segue-se que  $F'(a)$  é sobrejetiva e, portanto, um isomorfismo. Logo, pelo Teorema da Aplicação Inversa, existem abertos de  $\mathbb{R}^{n+m}$ ,  $V \ni (a, 0)$  e  $W \ni F(a, 0) = f(a)$ , tais que  $F|_V : V \rightarrow W$  é um difeomorfismo. Desta forma, fazendo-se  $G = (F|_V)^{-1}$  e  $A = P_1(V) \subset U$ , tem-se que  $A$  é um aberto e, para todo  $x \in A$ , teremos

$$(G \circ f)(x) = G(f(x)) = G(f(x) + (0, 0)) = G(F(x, 0)) = (x, 0),$$

como queríamos demonstrar. ■

<sup>7</sup> Recordemos, da Álgebra Linear, que  $\mathbb{R}^n \times \{0\} \oplus \{0\} \times \mathbb{R}^m$  indica a soma direta de  $\mathbb{R}^n \times \{0\}$  com  $\{0\} \times \mathbb{R}^m$ , isto significa que  $\mathbb{R}^n \times \{0\}$  e  $\{0\} \times \mathbb{R}^m$  têm em comum apenas o elemento  $\{0\}$ .

### 3.2.4 Forma Local das Submersões

Seja  $U \subset \mathbb{R}^m$  um aberto. Lembremos que a aplicação  $f : U \rightarrow \mathbb{R}^n$  é uma submersão se, para todo  $x \in U$ , a derivada  $f'(x) : \mathbb{R}^m \rightarrow \mathbb{R}^n$  é sobrejetiva, isto só pode ocorrer quando  $m \geq n$ . Pela regra da cadeia, a composta de duas submersões é ainda uma submersão.

**Exemplo 3.19** Seja  $P : \mathbb{R}^m \times \mathbb{R}^n \rightarrow \mathbb{R}^n$  a projeção definida por  $P(x, y) = y$ .

Então,  $P'(x, y) = P(x, y)$  para todo  $(x, y) \in \mathbb{R}^m \times \mathbb{R}^n$ . Logo,  $P$  é uma submersão.

O Teorema da Forma Local das Submersões mostrará que toda submersão se comporta localmente como a do Exemplo 3.19.

**Exemplo 3.20** Seja  $\mathbb{R}^m = \mathbb{R}^{m-i} \oplus \mathbb{R}^i$  uma decomposição em soma direta. Defina  $f : \mathbb{R}^m \rightarrow \mathbb{R}$  por  $f(x, y) = |x|^2 - |y|^2$ , onde  $|x|^2 = \langle x, x \rangle$  e  $|y|^2 = \langle y, y \rangle$  para  $x \in \mathbb{R}^{m-i}$  e  $y \in \mathbb{R}^i$ .

Então,  $f'(x, y) \cdot (h, k) = 2\langle x, h \rangle - 2\langle y, k \rangle$ . Portanto,  $f'(x, y) \neq 0$ , exceto na origem  $(x, y) = (0, 0)$ . Assim,  $f|_{\mathbb{R}^m - \{0\}}$  é uma submersão pois, como funcional linear,  $f'(x, y)$  pode ser apenas nula ou sobrejetiva.

No teorema a seguir, escreveremos o espaço vetorial  $\mathbb{R}^{m+n}$  como uma soma direta  $\mathbb{R}^{m+n} = E \oplus F$ , onde os elementos do espaço original serão representados por pares  $(x, y)$  de modo que  $x \in E$  e  $y \in F$ .

Se é dada uma decomposição em soma direta do tipo  $\mathbb{R}^{m+n} = E \oplus F$  e  $f : U \rightarrow \mathbb{R}^n$  é definida num aberto  $U \subset \mathbb{R}^{m+n}$ , a derivada de  $f$  no ponto  $z_0 = (x_0, y_0) \in U$  ao longo de  $E$  é indicada por  $\partial_1 f(z_0)$  e a derivada ao longo de  $F$  é representada por  $\partial_2 f(z_0)$ . Essas são as derivadas parciais de  $f$  no ponto  $z_0$  em relação à decomposição  $\mathbb{R}^{m+n} = E \oplus F$ . Qualquer uma dessas derivadas parciais podem existir ou não. Quando existem, e  $f : U \rightarrow \mathbb{R}^n$  é diferenciável no ponto  $z_0$ , elas são transformações lineares  $\partial_1 f(z_0) := f'(z_0)|_E : E \rightarrow \mathbb{R}^n$  e  $\partial_2 f(z_0) := f'(z_0)|_F : F \rightarrow \mathbb{R}^n$ .

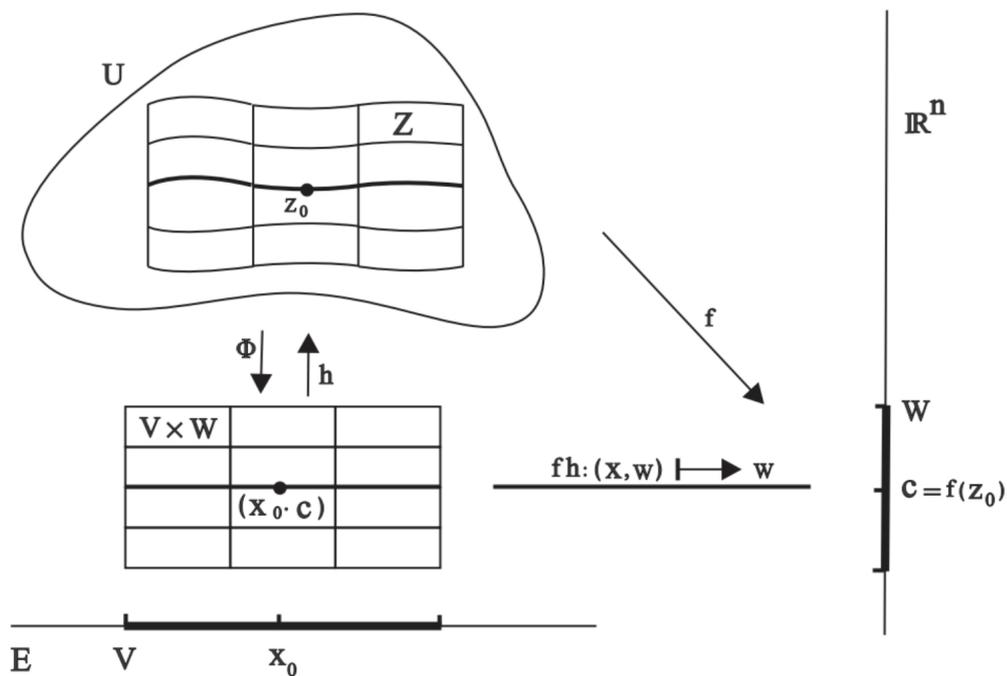
O Teorema da Forma Local das Submersões diz que, dada uma submersão  $f$  de classe  $C^k$  ( $k \geq 1$ ), é possível tomar novas coordenadas em torno de cada ponto de seu domínio de modo que  $f$  seja a projeção sobre as  $n$  últimas coordenadas.

**Teorema 3.21 (Forma Local das Submersões)** Sejam  $U \in \mathbb{R}^{m+n}$  um aberto e  $f : U \rightarrow \mathbb{R}^n$  uma aplicação de classe  $C^k$  ( $k \geq 1$ ), tal que  $f$  é uma submersão em  $z_0 \in U$ , isto é,  $f'(z_0) \in \mathcal{L}(\mathbb{R}^{m+n}; \mathbb{R}^n)$  é sobrejetiva. Dada qualquer decomposição em soma direta  $\mathbb{R}^{m+n} = E \oplus F$ , (com

$z_0 = (x_0, y_0)$ , onde  $x_0 \in E$  e  $y_0 \in F$  tal que  $\partial_2 f(z_0) = f'(z_0)|_F : F \rightarrow \mathbb{R}^n$  é um isomorfismo, existe um difeomorfismo  $h : V \times W \rightarrow Z$ , de classe  $C^k$ , tal que  $f \circ h(x, w) = w$  para todo  $(x, w) \in V \times W$ , onde  $V \ni x_0$  é aberto em  $E$ ,  $W \ni f(z_0)$  é aberto em  $\mathbb{R}^n$  e  $Z \ni z_0$  é aberto em  $\mathbb{R}^{m+n}$  com  $Z \subset U$ .

Como no caso da Forma Local das Imersões, o teorema acima faz uso do Teorema da Aplicação Inversa. Também, é importante relacionar o enunciado do Teorema da Forma Local das Submersões com a figura que o acompanha abaixo.

Figura 5 – Forma Local das Submersões



Fonte: (LIMA, 2014)

**Demonstração.** Defina a aplicação  $\Phi : U \rightarrow E \times \mathbb{R}^n$ , de classe  $C^k$ , fazendo  $\Phi(x, y) = (x, f(x, y))$ .

A derivada  $\Phi'(z_0) : \mathbb{R}^{m+n} \rightarrow E \times \mathbb{R}^n$  é dada por

$$\Phi'(z_0)(h, k) = (h, \partial_1 f(z_0) \cdot h + \partial_2 f(z_0) \cdot k).$$

Para usar o Teorema da Aplicação Inversa, devemos verificar se  $\Phi'(z_0) : \mathbb{R}^{m+n} \rightarrow E \times \mathbb{R}^n$  dada acima é um isomorfismo. Para tanto devemos mostrar que  $\Phi'(z_0)(h, k)$  possui inversa. De fato, fazendo

$$(h, \partial_1 f(z_0) \cdot h + \partial_2 f(z_0) \cdot k) = (u, v),$$

temos que

$$h = u \text{ e } v = \partial_1 f(z_0) \cdot h + \partial_2 f(z_0) \cdot k,$$

o que nos dá

$$k = [\partial_2 f(x_0)]^{-1} \cdot (v - \partial_1 f(x_0) \cdot u).$$

Como, por hipótese,  $\partial_2 f(z_0)$  é um isomorfismo, segue que existe a inversa  $[\partial_2 f(x_0)]^{-1}$ . Assim,

$$[\Phi'(z_0)]^{-1}(u, v) = (u, [\partial_2 f(x_0)]^{-1} \cdot (v - \partial_1 f(x_0) \cdot u))$$

é uma inversa para  $\Phi'(z_0)$ , que é então um isomorfismo. Pelo Teorema da Aplicação Inversa, se escrevermos  $f(z_0) = c$ ,  $\Phi$  é um difeomorfismo de classe  $C^k$  de uma vizinhança de  $z_0$  sobre uma vizinhança de  $(x_0, c)$ . Esta última pode ser escolhida na forma  $V \times W$  onde  $V$  é aberto em  $E$  e  $W$  é aberto em  $\mathbb{R}^n$ , onde  $x_0 \in V$  e  $c \in W$ . Tome  $Z = \Phi^{-1}(V \times W)$  aberto,  $z_0 \in Z \subset U$  e  $h = \Phi^{-1} : V \times W \rightarrow Z$ . Então, para qualquer  $(x, w)$  em  $V \times W$ , afirmamos que  $f \circ h(x, w) = w$ . De fato, como  $\Phi(x, y) = (x, f(x, y))$ , segue que  $h = \Phi^{-1}$  é da forma  $h(x, w) = (x, h_2(x, w))$ . Daí, para todo  $(x, w)$  em  $V \times W$ , teremos

$$\begin{aligned} (x, w) &= \Phi \circ h(x, w) \\ &= \Phi(x, h_2(x, w)) \\ &= (x, f(x, h_2(x, w))) \\ &= (x, f \circ h(x, w)), \end{aligned}$$

donde resulta  $f \circ h(x, w) = w$ , o que finaliza a demonstração. ■

### 3.3 Superfícies no Espaço Euclidiano $\mathbb{R}^n$

Nessa seção, abordaremos alguns pontos acerca de *superfícies*, um ambiente onde naturalmente podemos falar de aplicações do Teorema da Aplicação Inversa. Mais especificamente, no fim dessa seção, abordaremos um teorema que usa diretamente o Teorema da Função Implícita na demonstração. A teoria das superfícies é um tema vasto e não é nosso intuito se delongar neste assunto, visto que, de certa forma, fugiria do objetivo principal desse trabalho. A ideia em apresentar esse tópico é fornecer ao leitor um caminho que faça uma ligação natural com o próximo capítulo.

**Definição 3.22** *Uma parametrização de classe  $C^k$  de um conjunto  $V \subset \mathbb{R}^n$  é um homeomorfismo  $\varphi : V_0 \rightarrow V$ , que é também uma imersão de classe  $C^k$ , definido no aberto  $V_0 \in \mathbb{R}^m$ .<sup>8</sup>*

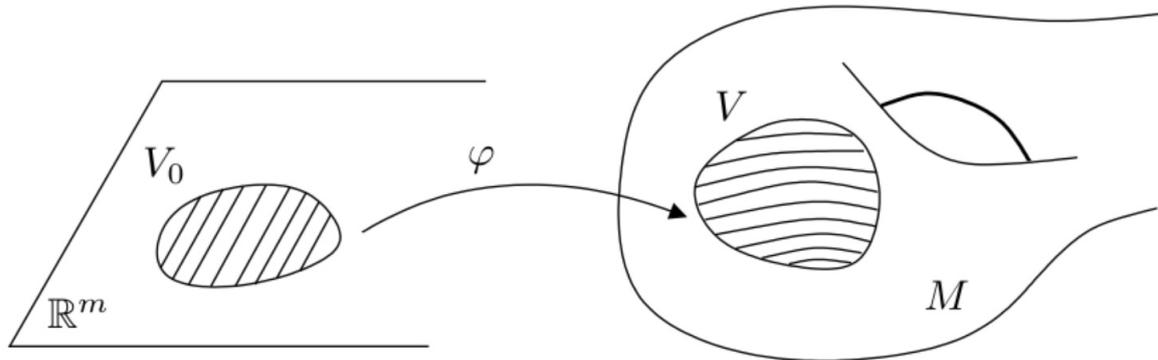
<sup>8</sup> Note que, pela definição dada, toda parametrização é um difeomorfismo.

Se  $\varphi : V_0 \rightarrow V$  é uma parametrização, cada ponto  $y = \varphi(x) \in V$ , embora pertença a  $\mathbb{R}^n$ , necessita apenas de  $m$  números para ter sua posição determinada, a saber, as  $m$  coordenadas de  $x \in V_0$ , chamadas os *parâmetros* que servem para localizar os pontos de  $V$ .

**Definição 3.23** Uma *superfície* de dimensão  $m$  e classe  $C^k$  em  $\mathbb{R}^n$  é um conjunto  $M \subset \mathbb{R}^n$  que pode ser coberto por uma coleção de abertos  $U \in \mathbb{R}^n$ , tais que cada  $V = U \cap M$  admite uma parametrização  $\varphi : V_0 \rightarrow V$ , de classe  $C^k$ , definida num aberto  $V_0 \in \mathbb{R}^m$ .

A figura abaixo ilustra este fato. Cada um desses conjuntos  $V = U \cap M$  é um aberto em  $M$ . Para cada  $p \in M$ , diz-se que  $V$  é uma *vizinhança parametrizada* de  $p$ .

Figura 6 – Definição de superfície



Fonte: (LIMA, 2006)

Por tudo que consideramos acima, podemos concluir que uma superfície de classe  $C^k$  e dimensão  $m$  em  $\mathbb{R}^n$  é um subconjunto tal que cada um dos seus pontos possui uma vizinhança parametrizada, por uma parametrização de classe  $C^k$  e dimensão  $m$ .

Quando  $M \subset \mathbb{R}^n$  é uma superfície de dimensão  $m$ , escrevemos  $M^m$ . Dizemos também que  $M$  tem *codimensão*  $n - m$ . Se  $\dim M = 1$ ,  $M$  chama-se uma *curva* de classe  $C^k$ .

**Exemplo 3.24** Todo aberto  $U \in \mathbb{R}^n$  é uma superfície de dimensão  $n$  e classe  $C^\infty$  em  $\mathbb{R}^n$ , coberta por uma única vizinhança parametrizada  $U$ , sendo  $\varphi : U \rightarrow U$  a aplicação identidade. Reciprocamente, se  $M \subset \mathbb{R}^n$  é uma superfície de dimensão  $n$  e classe  $C^1$ , então cada vizinhança parametrizada  $V \subset M$ , sendo imagem do aberto  $V_0 \subset \mathbb{R}^n$  pelo difeomorfismo  $\varphi : V_0 \rightarrow V$ , é um subconjunto aberto de  $\mathbb{R}^n$ , pelo Teorema da Aplicação Inversa. O conjunto  $M$ , reunião das vizinhanças parametrizadas  $V$ , é portanto aberto em  $\mathbb{R}^n$ . Assim, as superfícies de classe  $C^1$  e dimensão  $n$  em  $\mathbb{R}^n$  são os subconjuntos abertos de  $\mathbb{R}^n$ .

**Exemplo 3.25** No outro extremo,  $M \subset \mathbb{R}^n$  é uma superfície de dimensão 0 se, e somente se, para cada ponto  $p \in M$  existe um aberto  $U \subset \mathbb{R}^n$  tal que  $U \cap M$  admite uma parametrização  $\varphi : V_0 \rightarrow V$ , definida num aberto  $V_0 \subset \mathbb{R}^0 = \{0\}$ . Assim,  $V_0$  é um ponto e  $V = \{p\}$ . Portanto,  $M \subset \mathbb{R}^n$  é uma superfície de dimensão 0 se, e somente se, é um conjunto discreto, ou seja, seus pontos são todos isolados.

**Exemplo 3.26** Ser uma superfície é uma propriedade local, ou seja, se cada ponto  $p$  do conjunto  $M$  pertence a um aberto  $Z \subset \mathbb{R}^n$  tal que  $Z \cap M$  é uma superfície de dimensão  $m$  e classe  $C^k$ , então  $M$  é uma superfície de dimensão  $m$  e classe  $C^k$ . Assim, por exemplo, se para todo ponto  $p \in M$  existe um aberto  $Z \subset \mathbb{R}^n$ , tal que  $p \in Z$  e  $Z \cap M$  é o gráfico de uma aplicação  $f : V_0 \rightarrow \mathbb{R}^n$ , de classe  $C^k$  num aberto  $V_0 \subset \mathbb{R}^m$ , então  $M$  é uma superfície de dimensão  $m$  e classe  $C^k$ . Em particular, toda hipersuperfície de classe  $C^k$  em  $\mathbb{R}^{n+1}$  é uma superfície de dimensão  $n$  e classe  $C^k$ .

Seja  $M \subset \mathbb{R}^n$  uma superfície de dimensão  $m$  e classe  $C^k$  ( $k \geq 1$ ). Dado um ponto  $p \in M$ , o espaço vetorial tangente a  $M$  no ponto  $p$  é o subespaço vetorial  $T_p M \subset \mathbb{R}^n$  que pode ser descrito de duas maneiras:

1. Tomamos uma parametrização  $\varphi : V_0 \rightarrow V$ , de classe  $C^k$ , de uma vizinhança  $V \ni p$ ; seja  $\varphi(a) = p$ . Pomos  $T_p M = \varphi'(a) \cdot \mathbb{R}^m$  como a imagem da transformação linear injetiva  $\varphi'(a) : \mathbb{R}^m \rightarrow \mathbb{R}^n$ ;
2. Consideramos todos os caminhos  $\lambda : (-\varepsilon, \varepsilon) \rightarrow M$ , com  $\lambda(0) = p$ , diferenciáveis no ponto 0. Definimos  $T_p M$  como o conjunto dos vetores-velocidade  $\lambda'(0)$  desses caminhos.

A definição 1 apresenta  $T_p M$  como um subespaço vetorial de dimensão  $m$  do espaço  $\mathbb{R}^n$ , mas tem o inconveniente de depender de uma parametrização  $\varphi$ . Por outro lado, a definição 2 não depende da escolha de parametrização, mas não é evidente que  $T_p M$ , assim definido, seja um espaço vetorial. No entanto, 1 e 2 definem o mesmo conjunto. Não faremos a prova deste fato, mas o leitor pode encontrar a demonstração no capítulo 5 de (LIMA, 2006).

**Exemplo 3.27** Se  $M \subset \mathbb{R}^{m+n}$  é o gráfico de uma aplicação  $f : U \rightarrow \mathbb{R}^n$ , seu espaço vetorial tangente  $T_p M$ , no ponto  $p = (a, f(a))$ , é o gráfico da derivada  $f'(a) : \mathbb{R}^m \rightarrow \mathbb{R}^n$ , ou seja,  $T_p M = \{(v, f'(a) \cdot v); v \in \mathbb{R}^m\}$ . Basta tomar a parametrização natural  $\varphi : U \rightarrow M$ ,  $\varphi(x) = (x, f(x))$ , cuja derivada no ponto  $a$  é dada por  $\varphi'(a) \cdot v = (v, f'(a) \cdot v)$ , para ver que  $T_p M = \varphi'(a) \cdot \mathbb{R}^m$  é o gráfico de  $f'(a)$ . A mesma observação se aplica a uma superfície que é localmente o gráfico de uma aplicação.

Seja  $M \subset \mathbb{R}^n$  uma superfície de classe  $C^k$  ( $k \geq 1$ ). Uma aplicação  $f : M \rightarrow \mathbb{R}^r$  diz-se *diferenciável* no ponto  $p \in M$  quando existe uma parametrização  $C^k$ ,  $\varphi : V_0 \rightarrow V$ , de uma vizinhança  $V$  de  $p$  em  $M$ , tal que  $f \circ \varphi : V_0 \rightarrow \mathbb{R}^r$  (onde  $V_0$  é um aberto do  $\mathbb{R}^m$ ) é diferenciável no ponto  $a = \varphi^{-1}(p)$ . Se tomarmos qualquer outra parametrização  $C^k$ ,  $\psi : W_0 \rightarrow W$ , de uma vizinhança de  $p$  em  $M$ , teremos  $f \circ \psi$  diferenciável no ponto  $\psi^{-1}(p)$  se, e somente se,  $f \circ \varphi$  for diferenciável no ponto  $\varphi^{-1}(p)$ , pois  $f \circ \psi = (f \circ \varphi) \circ (\varphi^{-1} \circ \psi)$  e assumindo que  $\varphi^{-1} \circ \psi : \psi^{-1}(V \cap W) \rightarrow \varphi^{-1}(V \cap W)$  é um difeomorfismo de classe  $C^k$ .<sup>9</sup>

Analogamente, dada a aplicação  $f : M \rightarrow \mathbb{R}^r$ , definida numa superfície de classe  $C^k$ , dizemos que  $f$  é de classe  $C^i$  ( $0 \leq i \leq k$ ) quando, para cada ponto  $p \in M$ , existe uma parametrização  $\varphi : V_0 \rightarrow V \subset M$ , com  $p \in V$ ,  $\varphi \in C^k$ , tal que  $f \circ \varphi : V_0 \rightarrow \mathbb{R}^r$  é de classe  $C^i$  no aberto  $V_0 \subset \mathbb{R}^m$ . Neste caso, para qualquer outra parametrização  $\psi : W_0 \rightarrow W \subset M$ , com  $p \in W$  e  $\psi \in C^k$  tem-se ainda  $f \circ \psi : W_0 \rightarrow \mathbb{R}^r$  de classe  $C^i$ . Mas se for  $i \geq k + 1$ , não tem sentido definir a noção de aplicação de classe  $C^i$  em  $M$  se esta superfície é apenas de classe  $C^k$  e nada mais.

Se  $f : M \rightarrow \mathbb{R}^r$  é diferenciável no ponto  $p \in M$ , sua derivada nesse ponto é a transformação linear  $f'(p) : T_p M \rightarrow \mathbb{R}^r$ , definida do seguinte modo: tomamos uma parametrização  $\varphi : V_0 \rightarrow V$ , com  $p = \varphi(a)$ . Dado um vetor  $v \in T_p M$ , temos  $v = \varphi'(a) \cdot v_0$  para todo vetor  $v_0 \in \mathbb{R}^m$ . Definimos  $f'(p) \cdot v = (f \circ \varphi)'(a) \cdot v_0$ . A transformação linear  $f'(p) : T_p M \rightarrow \mathbb{R}^r$  está bem definida, ou seja, se  $\psi : W_0 \rightarrow W$  for outra parametrização, com  $p = \psi(b)$ , e tivermos  $v = \psi'(b) \cdot w_0$  para algum  $w_0 \in \mathbb{R}^m$ , então  $(f \circ \varphi)'(a) \cdot v_0 = (f \circ \psi)'(b) \cdot w_0$ .<sup>10</sup>

Qualquer vetor  $v \in T_p M$  é o vetor velocidade,  $v = \lambda'(0)$ , de um caminho  $\lambda : (-\varepsilon, \varepsilon) \rightarrow M$ , com  $\lambda(0) = p$ . Então, a imagem  $f'(p) \cdot v = (f \circ \lambda)'(0)$  é o vetor velocidade, no ponto 0, do caminho  $f \circ \lambda : (-\varepsilon, \varepsilon) \rightarrow \mathbb{R}^r$ .

Se  $N \subset \mathbb{R}^r$  é outra superfície de classe  $C^k$  e a aplicação  $f : M \rightarrow \mathbb{R}^r$ , diferenciável no ponto  $p \in M$ , é tal que  $f(M) \subset N$ , diremos que  $f : M \rightarrow N$  é uma *aplicação diferenciável* no ponto  $p$ . O que falamos anteriormente sobre  $f'(p) \cdot v$  ser o vetor velocidade de um caminho mostra que se  $f : M \rightarrow N$  é diferenciável no ponto  $p \in M$ , então a derivada  $f'(p)$  é uma transformação linear de  $T_p M$  em  $T_q N$ , onde  $q = f(p)$ .

Vale a Regra da Cadeia: se  $f : M \rightarrow N$  é diferenciável no ponto  $p \in M$  e  $g : N \rightarrow \mathbb{R}^s$  é diferenciável no ponto  $q = f(p)$ , então  $g \circ f : M \rightarrow \mathbb{R}^s$  é diferenciável no ponto  $p$ , com  $(g \circ f)'(p) = g'(q) \cdot f'(p)$ .

Generalizando a Definição 2.41, diremos que um ponto  $c \in \mathbb{R}^n$  é *valor regular* de

<sup>9</sup> Observe o capítulo 4, em particular, o corolário 4.4.

<sup>10</sup> A demonstração deste fato pode ser vista no capítulo 5 de (LIMA, 2006).

uma aplicação diferenciável  $f : U \rightarrow \mathbb{R}^n$ , definida num aberto  $U \subset \mathbb{R}^m$ , quando, para todo ponto  $x \in U$  tal que  $f(x) = c$ , a derivada  $f'(x) : \mathbb{R}^m \rightarrow \mathbb{R}^n$  for uma transformação linear sobrejetiva.

Considerando as funções-coordenadas  $f_1, f_2, \dots, f_n : U \rightarrow \mathbb{R}$  da aplicação  $f$  do parágrafo anterior, vemos que  $c \in \mathbb{R}^n$  é um valor regular de  $f$  se, e somente se, em todo ponto  $x \in f^{-1}(c)$  os vetores  $\nabla f_1(x), \nabla f_2(x), \dots, \nabla f_n(x)$  são linearmente independentes. De fato, estes vetores são as linhas da matriz jacobiana  $Jf(x)$  e sua independência linear significa que esta matriz  $n \times m$  tem posto  $n$ , ou seja, que a transformação linear  $f'(x)$  é sobrejetiva. Quando  $n = 1$ , essa condição significa que  $f(x) = c \Rightarrow \nabla f(x) \neq 0$  e recaímos na definição de valor regular dada em 2.41.

Finalizando nossas considerações, iremos enunciar e provar o resultado abaixo usando uma das aplicações do Teorema da Aplicação Inversa.

**Teorema 3.28** *Seja  $f : U \rightarrow \mathbb{R}^{n-m}$  de classe  $C^k$  no aberto  $U \subset \mathbb{R}^n$ . Se  $c \in \mathbb{R}^{n-m}$  é valor regular de  $f$ , então  $f^{-1}(c)$  é uma superfície de classe  $C^k$  e dimensão  $m$  em  $\mathbb{R}^n$ . Em cada ponto  $p \in f^{-1}(c)$ , o espaço tangente  $T_p[f^{-1}(c)]$  é o núcleo da derivada  $f'(p) : \mathbb{R}^n \rightarrow \mathbb{R}^{n-m}$ .*

**Demonstração.** Pelo Teorema da Aplicação Implícita, para cada ponto  $p \in f^{-1}(c)$  existe um aberto  $Z \subset \mathbb{R}^n$ , contendo  $p$ , tal que  $Z \cap f^{-1}(c)$  é o gráfico de uma aplicação de classe  $C^k$ , definida num aberto de  $\mathbb{R}^m$ . Logo, cada  $Z \cap f^{-1}(c)$  é também uma superfície de dimensão  $m$  e classe  $C^k$ . Segue-se que  $f^{-1}(c)$  também o é. (Observe o Exemplo 3.26.) Dado qualquer  $p \in f^{-1}(c)$ , como todo vetor  $v$ , tangente a  $f^{-1}(c)$  no ponto  $p$ , é o vetor-velocidade de um caminho  $\lambda$  contido em  $f^{-1}(c)$  (logo  $f \circ \lambda$  é constante), temos  $f'(p) \cdot v = (f \circ \lambda)'(0) = 0$ . Assim, o espaço vetorial tangente  $T_p[f^{-1}(c)]$  está contido no núcleo de  $f'(p)$ . Como ambos são espaços vetoriais de dimensão  $m$ , segue-se que eles são iguais. ■

#### 4 O TEOREMA DA APLICAÇÃO INVERSA EM VARIEDADES DIFERENCIÁVEIS

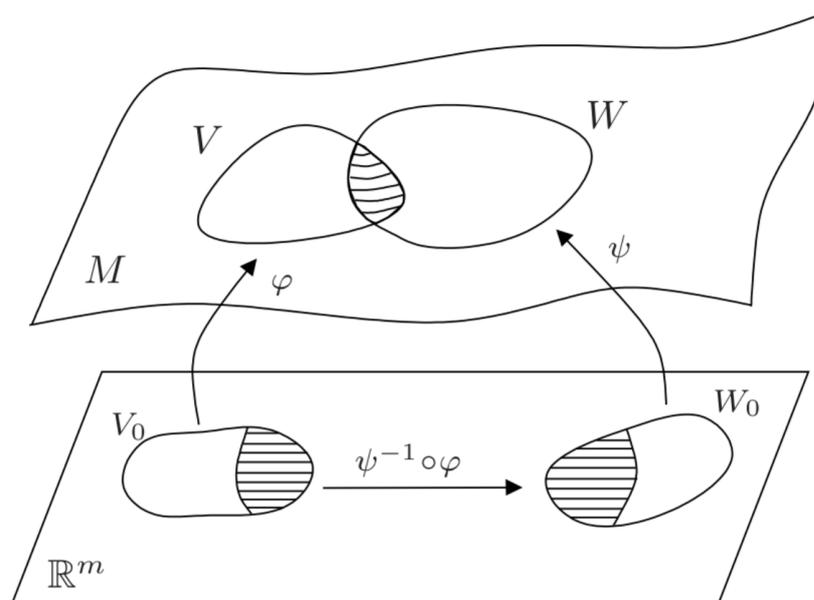
Nesse capítulo, não iremos nos prolongar em demasia dando todos os detalhes técnicos e definições a fim de apresentar o Teorema da Aplicação Inversa em Variedades Diferenciáveis. Nosso intuito aqui é mostrar ao leitor que ele pode ser estendido de forma natural a outros ambientes. Para tanto, retomemos o assunto de superfícies euclidianas visto na última seção do capítulo anterior. Aqui usaremos o Teorema da Aplicação Inversa para mostrar que mudanças de parametrizações são sempre difeomorfismos, mostraremos o mesmo entre superfícies de mesma dimensão e logo em seguida finalizaremos nosso trabalho generalizando esta última ideia para Variedades Diferenciáveis.

Sejam  $\varphi : V_0 \rightarrow V$  e  $\psi : W_0 \rightarrow W$  parametrizações numa superfície  $M$ , de classe  $C^k$  e dimensão  $m$ . Suponha que  $V \cap W \neq \emptyset$ . Então, todo ponto  $p \in V \cap W$  pode ser escrito como  $p = \varphi(x)$ ,  $x \in V_0$  e, como  $p = \psi(y)$ ,  $y \in W_0$ , pode ser representado pelos  $m$  parâmetros que são as coordenadas de  $x$  e pelas  $m$  coordenadas de  $y$ . A correspondência  $x \mapsto y$ , definida pela relação  $\varphi(x) = \psi(y)$ , é a aplicação

$$\psi^{-1} \circ \varphi : \varphi^{-1}(V \cap W) \rightarrow \psi^{-1}(V \cap W),$$

chamada *mudança de parametrização*. Observe a figura a seguir:

Figura 7 – Mudança de parametrização

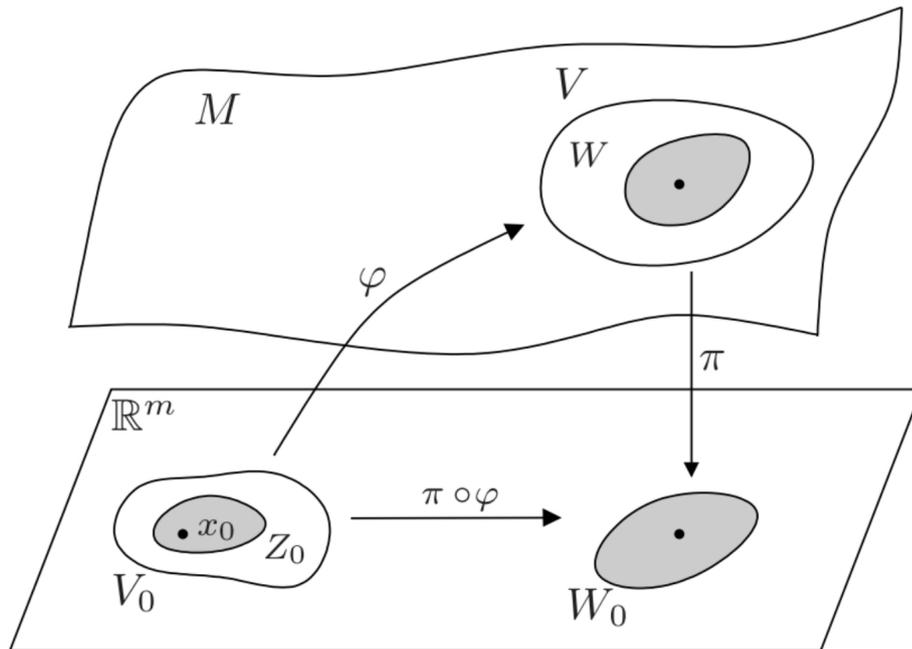


No teorema a seguir,  $M$  é uma superfície de dimensão  $m$  e classe  $C^k$  em  $\mathbb{R}^n$ . Quando mencionarmos “projeção  $\pi : \mathbb{R}^n \rightarrow \mathbb{R}^m$ ”, estaremos nos referindo a uma aplicação dada por  $\pi(x_1, x_2, \dots, x_n) = (x_{i_1}, x_{i_2}, \dots, x_{i_m})$ , definida a partir da escolha de  $m$  índices  $i_1 < i_2 < \dots < i_m$ , compreendidos entre 1 e  $n$ .

**Teorema 4.1** *Seja  $\varphi : V_0 \rightarrow V$  uma parametrização em  $M$ . Para cada  $p = \varphi(x_0) \in V$  existe uma projeção  $\pi : \mathbb{R}^n \rightarrow \mathbb{R}^m$  tal que  $\pi \circ \varphi$  aplica um aberto  $Z_0$ , com  $x_0 \in Z_0 \subset V_0$ , difeomorficamente sobre um aberto  $W_0 \subset \mathbb{R}^m$ .*

A figura abaixo ajuda a interpretar melhor o teorema:

Figura 8 – Teorema 4.1



Fonte: (LIMA, 2009)

**Demonstração.** A matriz jacobiana  $\left[ \frac{\partial \varphi_i}{\partial x_i}(x_0) \right] \in M(n \times m)$  tem  $m$  linhas linearmente independentes, de índices  $i_1 < i_2 < \dots < i_m$ . Essas linhas formam a matriz  $m \times m$ , invertível,  $J = \left[ \frac{\partial \varphi_{i_k}}{\partial x_i}(x_0) \right]$  e os índices  $i_k$  definem uma projeção  $\pi : \mathbb{R}^n \rightarrow \mathbb{R}^m$ . Note que  $J$  é a matriz jacobiana da aplicação  $\pi \circ \varphi : V_0 \rightarrow \mathbb{R}^m$ . Logo, o resultado segue imediatamente do Teorema da Aplicação Inversa. ■

O teorema acima nos dá os seguintes corolários:

**Corolário 4.2** *Toda superfície de classe  $C^k$  é localmente o gráfico de uma aplicação de classe  $C^k$ .*

**Demonstração.** Usando a notação do teorema acima, escreva os elementos de  $\mathbb{R}^n$  sob a forma  $z = (y, y')$ , onde  $y = \pi(z)$ . Defina  $W = \varphi(Z_0)$ . Então, aplicação  $\psi = \varphi \circ (\pi \circ \varphi)^{-1} : W_0 \rightarrow W$  é uma parametrização. Além disso, para todo  $y \in W_0$ , tem-se

$$\pi(\psi(y)) = (\pi \circ \varphi) \circ (\pi \circ \varphi)^{-1}(y) = y.$$

Daí, teremos  $\psi(y) = (y, y')$  e, portanto,  $W$  é o gráfico da aplicação de classe  $C^k$ ,  $f : W_0 \rightarrow \mathbb{R}^{n-m}$ , dada por  $f(y) = y'$ . ■

**Corolário 4.3** *Seja  $M \subset \mathbb{R}^n$  uma superfície de classe  $C^k$  e dimensão  $m$ . Se uma aplicação  $f : V_0 \rightarrow \mathbb{R}^n$ , de classe  $C^k$  no aberto  $V_0 \subset \mathbb{R}^p$ , tiver sua imagem  $f(V_0)$  contida na vizinhança  $W \subset M$ , parametrizada por  $\psi : W_0 \rightarrow W$ , então  $\psi \circ f : V_0 \rightarrow \mathbb{R}^m$  é uma aplicação de classe  $C^k$ .*

**Demonstração.** De fato, para cada ponto  $x_0 \in V_0$ , com  $f(x_0) = \psi(y_0)$ , existe, pelo Teorema 4.1, uma projeção  $\pi : \mathbb{R}^n \rightarrow \mathbb{R}^m$  tal que  $\pi \circ \psi$  é um difeomorfismo de uma vizinhança de  $y_0$  sobre um aberto de  $\mathbb{R}^m$ . Então, numa vizinhança de  $x_0$ , podemos escrever

$$\psi^{-1} \circ f = (\pi \circ \psi)^{-1} \circ \pi \circ f.$$

Portanto,  $\psi^{-1} \circ f$  é de classe  $C^k$ . ■

De acordo com o que definimos no início desse capítulo sobre mudança de parametrização, temos o seguinte resultado.

**Corolário 4.4** *Numa superfície de classe  $C^k$ , toda mudança de parametrização  $\psi^{-1} \circ \varphi$  é um difeomorfismo de classe  $C^k$ .*

**Demonstração.** Pelo corolário 4.3,  $\psi^{-1} \circ \varphi$  é uma aplicação de classe  $C^k$ . Pela mesma razão, sua inversa  $\varphi^{-1} \circ \psi$  também é de classe  $C^k$ . Portanto  $\psi^{-1} \circ \varphi$  é um difeomorfismo. ■

Tendo em vista tudo o que já foi definido da seção 3.3 até aqui, estamos em condições de enunciar e demonstrar o teorema abaixo.

**Teorema 4.5 (O Teorema da Aplicação Inversa entre Superfícies)** *Sejam  $M \subset \mathbb{R}^r$ , e  $N \subset \mathbb{R}^s$  superfícies de classe  $C^k$  ( $k \geq 1$ ) e a aplicação diferenciável  $f : M \rightarrow N$ . Se  $p \in M$  é um ponto tal que  $f'(p) : T_p M \rightarrow T_{f(p)} N$  é um isomorfismo, então existem vizinhanças  $U$  de  $p$  e  $V$  de  $f(p)$  tal que  $f|_U : U \rightarrow V$  é um difeomorfismo.*

**Demonstração.** Como  $f'(p)$  é bijetiva, então  $M$  e  $N$  têm a mesma dimensão, digamos  $m$ . Considere  $B_p$  e  $B_{f(p)}$  bases de  $T_pM$  e  $T_{f(p)}N$ , respectivamente. Sejam  $\varphi : Z_0 \rightarrow Z$  uma parametrização em  $M$  onde  $Z$  é uma vizinhança parametrizada de  $p \in M$  e  $\psi : W_0 \rightarrow W$  uma parametrização em  $N$ , onde  $W$  é uma vizinhança parametrizada de  $f(p) \in N$ . Pela definição de parametrização,  $\varphi$  e  $\psi$  são difeomorfismos. Sabemos que  $\varphi : Z_0 \rightarrow Z$  e  $\psi : W_0 \rightarrow W$  levam à base canônica (do mesmo espaço euclidiano) nas bases  $B_p$  e  $B_{f(p)}$  de  $T_pM$  e  $T_{f(p)}N$ , respectivamente. Temos que a matriz jacobiana de  $f$  no ponto  $p$ ,  $Jf(p)$ , é a matriz jacobiana de  $\psi^{-1} \circ f \circ \varphi$  no ponto  $\varphi^{-1}(p)$ ,  $J\psi^{-1} \circ f \circ \varphi(\varphi^{-1}(p))$ . Pelo Teorema da Aplicação Inversa no  $\mathbb{R}^n$ , teremos  $0 \neq \det Jf(p) = \det J\psi^{-1} \circ f \circ \varphi(\varphi^{-1}(p))$ . Segue, daí, que  $\psi^{-1} \circ f \circ \varphi$  é um difeomorfismo local. Sendo assim,  $f$  também o é, ou seja, existem vizinhanças  $U$  de  $p$  e  $V$  de  $f(p)$  tal que  $f|_U : U \rightarrow V$  é um difeomorfismo. ■

Tudo que acabamos de abordar aqui, a respeito de superfícies, incluindo a seção 3.3, pode ser estendido ao conceito de Variedades Diferenciáveis. As superfícies no espaço euclidiano  $\mathbb{R}^n$  apresentadas em 3.3 constituem variedades  $n$ -dimensionais e, de certa forma, variedade generaliza o conceito de superfície. É lógico que não podemos simplesmente dar um “salto” e admitir como válidos todos os conceitos vistos anteriormente para variedades (algo que nem definimos de maneira formal).

O estudo das Variedades Diferenciáveis requer um aparato mais rebuscado de conceitos que não apresentaremos aqui. Este estudo inicia-se, evidentemente, com a definição precisa de *variedade topológica*. À partir daí, temos um ambiente onde são construídos os conceitos de estruturas diferenciáveis (neste, o Corolário 4.4 é admitido como um axioma), aplicações diferenciáveis entre variedades e vetores tangentes que nos dão uma noção precisa para definir a diferencial de aplicações entre variedades. Com tudo isso bem fundamentado de maneira sólida, tudo que vimos sobre imersões e submersões no espaço euclidiano  $\mathbb{R}^n$  pode ser ampliado para Variedades Diferenciáveis e, claro, não podemos nos esquecer que o mesmo vale para o Teorema da Aplicação Inversa.<sup>1</sup> Abaixo segue seu enunciado.

**Teorema 4.6** *Sejam  $M$  e  $N$  variedades diferenciáveis e  $F : M \rightarrow N$  uma aplicação diferenciável. Se  $p \in M$  é um ponto tal que  $F'(p)$  é isomorfismo, então  $F$  é um difeomorfismo local, isto é, existem vizinhanças conexas<sup>2</sup>  $U_0$  de  $p$  e  $V_0$  de  $F(p)$  tal que  $F|_{U_0} : U_0 \rightarrow V_0$  é um difeomorfismo.*

<sup>1</sup> Veja os capítulos 1, 2, 3 e 4 de (LEE, 2011a).

<sup>2</sup> Aqui, tratamos os abertos como conexos. Observe a Definição 2.32 e os comentários feitos em seguida.

A prova do teorema acima segue a mesma ideia do que fizemos na sua versão entre superfícies euclidianas, logicamente, com os conceitos bem definidos e construídos dentro do ambiente de Variedades Diferenciáveis. Não faremos a demonstração, mas o leitor interessado em prosseguir com o assunto pode consultar as referências (LEE, 2011b) e (LEE, 2011a), que tratam do tema variedades topológicas e diferenciáveis de maneira excepcional.

## 5 CONCLUSÃO

Vimos no capítulo 3 em 3.5, que o Teorema da Aplicação Inversa tem o seguinte enunciado: *Seja  $f : U \rightarrow \mathbb{R}^m$  de classe  $C^k$ , com  $k \geq 1$ , no aberto  $U \subset \mathbb{R}^m$ . Se  $a \in U$  é tal que  $f'(a) : \mathbb{R}^m \rightarrow \mathbb{R}^m$  é invertível (equivalentemente:  $\det Jf(a) \neq 0$ ), então existe uma bola aberta  $B = B(a; \delta) \subset U$  tal que a restrição  $f|_B$  é um difeomorfismo sobre um aberto  $V \ni f(a)$ .* No capítulo 2, apresentamos todos os pré-requisitos que julgamos suficientes para a compreensão do TAI como, por exemplo, as definições de conjunto aberto, aplicação diferenciável, aplicação de classe  $C^k$ , difeomorfismo, dentre outros.

No capítulo 3, também analisamos alguns exemplos onde foram feitas aplicações básicas do TAI. Além disso, apresentamos resultados mais fortes como o Teorema da Função Implícita, o método dos multiplicadores de Lagrange e a Forma Local das Imersões e Submersões.

Finalmente, no capítulo 4, mostramos como o TAI pode ser visto em outros ambientes. Apresentamos o conceito de superfície euclidiana na última seção do capítulo 3 e, a partir dele, expandimos para a ideia de variedade diferenciável. Concluímos o trabalho apresentando, mas não demonstrando, o TAI em variedades diferenciáveis.

## REFERÊNCIAS

- ÁVILA, G. **Introdução à análise matemática**. São Paulo: Editora Blucher, 1999.
- BÍBLIA. A. T. Provérbios. *In*: BÍBLIA. **Tradução do novo mundo da bíblia sagrada**. Brazilian portuguese edition. United States: Watchtower Bible and Tract Society of New York, 2015. p. 901. cap. 3, vers. 14.
- FIGUEIREDO, D. G. **Análise I**. Rio de Janeiro: LTC, 2015.
- LEE, J. M. **Introduction to smooth manifolds**. United States: Springer, 2011. v. 218.
- LEE, J. M. **Introduction to topological manifolds**. United States: Springer, 2011. v. 202.
- LIMA, E. L. **Curso de análise**. Rio de Janeiro: IMPA, 2006. v. 2.
- LIMA, E. L. **Curso de análise**. Rio de Janeiro: IMPA, 2007. v. 1.
- LIMA, E. L. **Análise real, funções de  $n$  variáveis**. Rio de Janeiro: IMPA, 2009. v. 2.
- LIMA, E. L. **Álgebra linear**. Rio de Janeiro: IMPA, 2012.
- LIMA, E. L. **Análise no espaço  $\mathbb{R}^n$** . Rio de Janeiro: IMPA, 2014.
- LIMA, E. L. **Análise real, funções de uma variável**. Rio de Janeiro: IMPA, 2020. v. 1.
- LIMA, E. L. **Espaços métricos**. Rio de Janeiro: IMPA, 2020.
- LIMA, R. F. **Topologia e análise no espaço  $\mathbb{R}^n$** . Rio de Janeiro: SBM, 2015.
- RUDIN, W. **Principles of mathematical analysis**. Singapore: McGraw-hill New York, 1976. v. 3.