UNIVERSIDADE ESTADUAL DA PARAÍBA
CAMPUS I
CENTRO DE CIÊNCIAS E TECNOLOGIA
DEPARTAMENTO DE MATEMÁTICA
CURSO DE LICENCIATURA EM MATEMÁTICA

IGOR BRENDOL PEREIRA MORAIS

UM BREVE ESTUDO SOBRE O TEOREMA DA FUNÇÃO IMPLÍCITA

CAMPINA GRANDE
2021

IGOR BRENDOL PEREIRA MORAIS

UM BREVE ESTUDO SOBRE O TEOREMA DA FUNÇÃO IMPLÍCITA

Trabalho de Conclusão de Curso apresentado ao Departamento de Matemática do Centro de Ciências e Tecnologia da Universidade Estadual da Paraíba como requisito parcial à obtenção do título de Licenciado(a) em Matemática.

Área de concentração: Análise.

Orientador: Prof. Dr. Gustavo da Silva Araújo.

CAMPINA GRANDE
2021

M827b Morais, Igor Brendol Pereira.
Um breve estudo sobre o Teorema da Função Implícita [manuscrito] / Igor Brendol Pereira Morais. - 2021.
63 p. : il. colorido.

Digitado.
Trabalho de Conclusão de Curso (Graduação em Matemática) - Universidade Estadual da Paraíba, Centro de Ciências e Tecnologia , 2021.
"Orientação : Prof. Dr. Gustavo da Silva Araújo , Coordenação do Curso de Matemática - CCT."

1. Teorema da Função Implícita. 2. Diferenciabilidade. 3. Função inversa. 4. Bilhares. I. Título

21. ed. CDD 515.8

Elaborada por Lucas H. A. da Silva - CRB - 15/898 BC/UEPB

IGOR BRENDOL PEREIRA MORAIS

UM BREVE ESTUDO SOBRE O TEOREMA DA FUNÇÃO IMPLÍCITA

Trabalho de Conclusão de Curso apresentado ao Departamento de Matemática do Centro de Ciências e Tecnologia da Universidade Estadual da Paraíba como requisito parcial à obtenção do título de Licenciado(a) em Matemática.

Área de concentração: Análise.

Aprovado em: 14/10/2021.

BANCA EXAMINADORA

Gustavo da Silva Araujo

Prof. Dr. Gustavo da Silva Araújo (Orientador)
Universidade Estadual da Paraíba (UEPB)

Aldo trajano Lourêdo

Prof. Dr. Aldo Trajano Lourêdo
Universidade Estadual da Paraíba (UEPB)

Vilmar Vaz da silva

Prof. Dr. Vilmar Vaz da Silva
Universidade Estadual da Paraíba (UEPB)

AGRADECIMENTOS

Agradeço a Deus, por ter me dado determinação em todos estes anos de estudo, por ter aberto as portas para que eu pudesse atingir mais objetivo na minha vida.

A minha família, meu irmão, minhas irmãs, em especial os meus pais que sempre estiveram do meu lado todo este tempo, me incentivaram nos momentos difíceis e me ensinaram todos os valores que fazem de mim a pessoa que sou.

Agradeço aos meus colegas de graduação, em especial aos que formei amizade, me ajudaram a alcançar mais este objetivo e que me faziam sorrir quando tudo dava errado.

Aos professores do DM-CCT-UEPB que me incentivaram a buscar melhorar a cada dia, em especial o professor Gustavo que se dedicou a me orientar neste trabalho e contribuiu com grandes ensinamentos durante a minha graduação.

Agradeço a todos os que contribuíram de alguma forma, direta ou indiretamente, na minha formação.

RESUMO

Este trabalho tem como principal objetivo construir, com detalhes, o caminho necessário para que seja possível compreender tanto a demonstração quanto a importância do Teorema da Função Implícita. Iremos introduzir os principais conceitos topológicos do Espaço Euclidiano $\mathbb{R}^n$, a noção de distância entre seus pontos, as aplicações que têm como domínio algum subconjunto deste espaço e a noção de continuidade dessas aplicações. Estes conceitos são necessários para o desenvolvimento do trabalho e para generalizar outros conceitos do Cálculo, como, por exemplo, a noção de diferenciabilidade para funções de mais de uma variável e o vetor gradiente, que são ferramentas fundamentais para podermos apresentar o Teorema da Função Implícita e algumas de suas aplicações.

Palavras-chave: Teorema da Função Implícita. Diferenciabilidade. Função Inversa. Bilhares.

ABSTRACT

The main objective of this work is to construct, in detail, the necessary tools to understand both the proof and the importance of the Implicit Function Theorem. We will introduce the main topological concepts of the Euclidean Space $\mathbb{R}^n$, the notion of distance between its points, the applications that have as a domain some subset of this space and the notion of continuity of that applications. These concepts are necessary for the development of the work and to generalize other concepts of Calculus, such as, for example, the notion of differentiability for functions of several variables and the notion of gradient vector, which are fundamental tools for us to present the Implicit Function Theorem and some of its applications.

Keywords: Implicit Function Theorem. Differentiability. Inverse function. Billiards.

SUMÁRIO

1 INTRODUÇÃO

No estudo de equações de mais de uma variável real, como por exemplo o caso de duas variáveis, são estudadas equações do tipo $f(x,y) = c$, com $c \in \mathbb{R}$, como por exemplo $f : \mathbb{R}^2 \to \mathbb{R}$ definida por $f(x,y) = y + yx + x - 1$, ao verificar a existência de soluções para $f(x,y) = 0$ por exemplo, uma vez que pode-se por métodos elementares explicitar a variável y em função de x, ou seja, $y + yx + x - 1 = 0$ implica $y = \frac{1-x}{1+x}$, então pode-se verificar que esta nova expressão é de fato uma função para todo $x \neq -1$, pode-se até mesmo esboçar seu gráfico, é dito então que a equação $f(x,y) = 0$ define implicitamente a variável y em função da variável x, pois as variáveis de f se relacionam de forma que y é de fato uma função de x, ou seja, para cada $x \neq -1$ existe um único y que seja solução para $f(x,y) = 0$. Acontece que, muitas vezes no estudo de equações de várias variáveis existem equações muito complexas de se trabalhar, tais equações são muito difíceis ou até mesmo impossíveis de se verificar, por métodos elementares, se existe alguma função que relacione unicamente as suas variáveis mas, apesar disto, ainda deseja-se obter informações sobre o seu comportamento.

Naturalmente surge a pergunta: sob quais hipóteses uma equação do tipo $f(x,y) = 0$ define implicitamente alguma função, ou melhor, quais características uma equação de mais de uma variável deve possuir para que se possa afirmar que seu conjunto de soluções define, unicamente, uma das suas variáveis em função das demais? Caso exista tal função, quais as informações a respeito do seu comportamento pode-se obter?

Neste trabalho, foi construído todo o caminho necessário para que se possa entender a então apresentação do Teorema da Função Implícita, mais precisamente, a versão do Teorema no Espaço Euclidiano ($\mathbb{R}^n$) que, por ser um espaço vetorial normado, possui propriedades topológicas de fácil compreensão.

Este trabalho teve como principal objetivo estudar as propriedades que uma equação de n variáveis reais deve possuir para que a mesma defina implicitamente alguma função, ou seja, sob quais hipóteses o conjunto de soluções de uma equação define unicamente uma variável em função das demais. Para atingir tal objetivo foi necessário primeiramente apresentar a estrutura e alguns conceitos topológicos do espaço vetorial $\mathbb{R}^n$, logo depois generalizar algumas noções e alguns teoremas clássicos do Cálculo de uma variável real para o Cálculo de várias variáveis reais e, por fim, demonstrar o Teorema da Função Implícita.

No capítulo dois, foi apresentado o espaço vetorial $\mathbb{R}^n$, a estrutura dos pontos deste espaço e as suas principais propriedades topológicas como, as formas de medir a distância entre dois pontos e o comprimento de um vetor, as principais operações entre dois vetores do mesmo tipo, as aplicações que tem como domínio algum subconjunto deste espaço e contradomínio algum subconjunto da reta e vice-versa, assim como a continuidade destas aplicações.

No capítulo três, foi estabelecido um extensão adequada de derivada para funções de várias variáveis para poder então fazer uma generalização mais adequada para alguns teoremas clássicos do Cálculo, como por exemplo a Regra da Cadeia e a composição de aplicações contínuas. Neste também foi apresentado o Vetor Gradiente.

O capítulo quatro foi dedicado à demonstrar o Teorema da Função Implícita e explanar alguns resultados do mesmo.

Por fim, no capítulo cinco foram apresentadas algumas aplicações do Teorema da Função Implícita.

A partir deste estudo, pode-se estabelecer sob quais hipóteses uma equação de n variáveis define implicitamente alguma função, é concluído ainda que quando tal função existe, ela é única.

2 NOÇÕES PRELIMINARES

2.1 O espaço Euclidiano

O Espaço Euclidiano $\mathbb{R}^n$, em que $n \in \mathbb{N}$, é formado pelo produto cartesiano

$$\mathbb{R}^n = \mathbb{R} \times \mathbb{R} \times \mathbb{R} \times \cdots \times \mathbb{R}$$

em que seus pontos são as listas $x = (x_1, x_2, x_3, \ldots, x_n)$, com cada $x_i \in \mathbb{R}$. Pode-se interpretar cada $x \in \mathbb{R}^n$ como um ponto ou como um vetor, dependendo do contexto, utilizar-se-á a notação $\vec{0}$ para representar o vetor que possui n coordenada iguais a 0, o *vetor nulo*. Este é um espaço vetorial em que estão bem definidas as operações de soma e multiplicação por escalar. Sejam $x, y \in \mathbb{R}^n$ e $\alpha \in \mathbb{R}$, vale

$$x + y = (x_1 + y_1, x_2 + y_2, x_3 + y_3, \ldots, x_n + y_n)$$

e

$$\alpha \cdot x = (\alpha \cdot x_1, \alpha \cdot x_2, \alpha \cdot x_3, \ldots, \alpha \cdot x_n).$$

Definição 2.1. Uma *aplicação* de um conjunto $A \subseteq \mathbb{R}^m$ em $\mathbb{R}^n$, é uma regra $f : A \to \mathbb{R}^n$ que associa cada vetor x de $\mathbb{R}^m$ a sua imagem

$$f(x) = (f_1(x), f_2(x), f_3(x), \ldots, f_n(x)),$$

em que $f_1, f_2, f_3, \ldots, f_n : A \subseteq \mathbb{R}^m \to \mathbb{R}$ são as funções coordenadas de f.

Uma aplicação diz-se linear quando para cada $x, y \in A$ e $\alpha \in \mathbb{R}$, tem-se:

1. $f(x + y) = f(x) + f(y)$;

2. $f(\alpha \cdot x) = \alpha \cdot f(x)$.

Um tipo importantíssimo e particular de aplicação linear, são os *funcionais lineares* $f : \mathbb{R}^m \to \mathbb{R}$.

O *dual algébrico* (ou *espaço dual*) de um espaço vetorial $\mathbb{R}^m$ é o conjunto $(\mathbb{R}^m)^* = \{f : \mathbb{R}^m \to \mathbb{R} \mid f \text{ é um funcional linear}\}$.

Observação 2.1. Note que o Dual algébrico é não-vazio, pois a aplicação nula (constante igual a 0) pertence a este conjunto.

Definição 2.2. Uma aplicação $\varphi : \mathbb{R}^m \times \mathbb{R}^n \to \mathbb{R}^p$ diz-se bilinear quando é linear em ambas as suas variáveis, ou seja, dados $x, y \in \mathbb{R}^m$, $x', y' \in \mathbb{R}^n$ e $\alpha \in \mathbb{R}$, vale:

1. $\varphi(\alpha x + y, x') = \alpha \cdot \varphi(x, x') + \varphi(y, x')$;

2. $\varphi(x, x' + \alpha y') = \varphi(x, x') + \alpha \cdot \varphi(x, y')$.

Será utilizada aqui frequentemente a noção de distância entre pontos de um espaço vetorial. A *distância* é um número real $d(x,y)$ obtido por meio de uma *métrica* que é uma função $d : M \times M \to \mathbb{R}$ em um conjunto M, em que para cada $x, y, z \in M$, tal função satisfaz as seguintes condições:

D1. $d(x,x) = 0$;

D2. Se $x \neq y$ então $d(x,y) > 0$;

D3. $d(x,y) = d(y,x)$;

D4. $d(x,z) \leq d(x,y) + d(y,z)$.

Definição 2.3. Uma *norma* em $\mathbb{R}^n$ é uma função $\| \quad \| : \mathbb{R}^n \to \mathbb{R}$ que associa cada vetor x de $\mathbb{R}^n$ um número real $\|x\|$, que satisfaz as seguintes propriedades, para quaisquer $x, y \in \mathbb{R}^n$ e um escalar λ:

N1. $x \neq \overrightarrow{0} \;\Rightarrow\; \|x\| \neq 0$;

N2. $\|\lambda \cdot x\| = |\lambda| \cdot \|x\|$;

N3. $\|x + y\| \leq \|x\| + \|y\|$.

Para $\lambda \in \mathbb{R}$, o número $|\lambda|$ é o valor absoluto de λ.

Exemplo 2.1. Norma da Soma, em $\mathbb{R}^n$

$$\|x\|_S = \sum_{i=1}^{n} |x_i|.$$

Para todo $x, y \in \mathbb{R}^n$ e $\lambda \in \mathbb{R}$, se $x \neq \overrightarrow{0}$ então para algum $k = 1, 2, 3, \ldots, n$, o valor absoluto da coordenada x_k de x vai ser maior do que 0 e para todas as demais coordenadas vai valer $|x_i| \geq 0$, logo $\|x\|_S = |x_1| + |x_2| + \cdots + |x_k| + \cdots + |x_n| \geq |x_k| > 0$ e isto verifica a propriedade (**N1.**). $\|\lambda \cdot x\|_S = \|(\lambda \cdot x_1, \lambda \cdot x_2, \lambda \cdot x_3, \ldots, \lambda \cdot x_n)\|_S = |\lambda \cdot x_1| + |\lambda \cdot x_2| + |\lambda \cdot x_3| + \cdots + |\lambda \cdot x_n| = |\lambda| \cdot |x_1| + |\lambda| \cdot |x_2| + |\lambda| \cdot |x_3| + \cdots + |\lambda| \cdot |x_n| = |\lambda| \cdot (|x_1| + |x_2| + |x_3| + \cdots + |x_n| = |\lambda| \cdot \|x\|_S$, o que verifica a propriedade (**N2.**). Por fim, pela definição e pela desigualdade triangular para números reais, segue que $\|x + y\|_S = \|(x_1 + y_1, x_2 + y_2, x_3 + y_3, \cdots, x_n + y_n)\|_S = |x_1 + y_1| + |x_2 + y_2| + |x_3 + y_3| + \cdots + |x_n + y_n| \leq |x_1| + |y_1| + |x_2| + |y_2| + |x_3| + |y_3| + \cdots + |x_n| + |y_n| = \|x\|_S + \|y\|_S$ e isto verifica a propriedade (**N3.**). Isto mostra que de fato, a norma da soma é uma norma em $\mathbb{R}^n$.

Exemplo 2.2. Norma do Máximo, em $\mathbb{R}^n$

$$\|x\|_M = \max_{1 \leq i \leq n} |x_i|.$$

A Norma do Máximo também é uma norma em $\mathbb{R}^n$, pois para todo $x, y \in \mathbb{R}^n$ e $\lambda \in \mathbb{R}$:

(N1.) Se $x \neq \overrightarrow{0}$, então para algum $k = 1, 2, 3, \ldots, n$, vai valer $|x_k| > 0$, mas por definição $|x_k| \leq \max\limits_{1 \leq i \leq n} |x_i|$, logo $0 < |x_k| \leq \max\limits_{1 \leq i \leq n} |x_i| = \|x\|_M$;

(N2.) $\|\lambda \cdot x\|_M = \|(\lambda \cdot x_1, \lambda \cdot x_2, \lambda \cdot x_3, \ldots, \lambda \cdot x_n)\|_M = \max\limits_{1 \leq i \leq n} |\lambda \cdot x_i| = \max\limits_{1 \leq i \leq n} |\lambda| \cdot |x_i| = |\lambda| \cdot \left(\max\limits_{1 \leq i \leq n} |x_i| \right) = |\lambda| \cdot \|x\|_M$;

(N3.) Por definição e pela desigualdade triangular para números reais, segue que $\|x + y\|_M = \|(x_1 + y_1, x_2 + y_2, x_3 + y_3, \ldots, x_n + y_n)\|_M = \max\limits_{1 \leq i \leq n} |x_i + y_i| \leq \max\limits_{1 \leq i \leq n} (|x_i| + |y_i|) = \max\limits_{1 \leq i \leq n} |x_i| + \max\limits_{1 \leq i \leq n} |y_i| = \|x\|_M + \|y\|_M$.

Exemplo 2.3. Norma Euclidiana ou Comprimento de um Vetor em $\mathbb{R}^n$

$$\|x\| = \left(\sum_{i=1}^{n} x_i^2 \right)^{\frac{1}{2}}.$$

A maneira mais natural de calcular a distância entre dois pontos $x, y \in \mathbb{R}^n$ é generalizando a distância entre dois pontos da reta. Quando $x, y \in \mathbb{R}$, a distância entre eles é o valor absoluto da sua diferença, enquanto para $x, y \in \mathbb{R}^n$, a distância entre x e y é a norma do vetor $x - y = (x_1 - y_1, x_2 - y_2, x_3 - y_3, \ldots, x_n - y_n)$. Daí, se verifica que cada norma fornece uma métrica, ou uma forma diferente de calcular a distância entre dois pontos do Espaço Euclidiano, e mais, note que a norma de um vetor $x \in \mathbb{R}^n$ é a distância de x até a origem, ou seja, $\|x\| = d(x, 0)$, sendo 0 a origem de $\mathbb{R}^n$.

Estas três normas apresentadas anteriormente são as mais comuns (ou usuais), as duas primeiras, apesar de serem mais simples, tem sua devida importância como será visto no Teorema 2.1.1, já a Norma Euclidiana é a mais natural na Geometria Euclidiana, diferente das outras duas, ela provêm de um produto interno.

No que segue, será mostrado que as três normas que foram apresentadas aqui, cumprem a seguinte relação, para todo $x \in \mathbb{R}^n$

$$\|x\|_M \leq \|x\| \leq \|x\|_S \leq n \cdot \|x\|_M.$$

Com efeito, seja $|x_k| = \max\limits_{1 \leq i \leq n} |x_i| = \|x\|_M$, como $x \in \mathbb{R}^n$ então $x = (x_1, x_2, x_3, \ldots, x_n)$ e cada $x_i \in \mathbb{R}$, então

$$\|x\|_M = \max_{1 \leq i \leq n} |x_i| = \sqrt{x_k^2} \leq \sqrt{x_1^2 + x_2^2 + \cdots + x_k^2 + \cdots + x_n^2} = \|x\|.$$

Note que $\|x\|_S^2 = |x_1|^2 + |x_2|^2 + |x_3|^2 + \cdots + |x_n|^2 + 2|x_1 \cdot x_2| + 2|x_1 \cdot x_3| + \cdots + 2|x_2 \cdot x_3| + 2|x_2 \cdot x_4| + \cdots + 2|x_{n-1} \cdot x_n| = x_1^2 + x_2^2 + x_3^2 + \cdots + x_n^2 + 2|x_1 \cdot x_2| + 2|x_1 \cdot x_3| + \cdots + 2|x_2 \cdot x_3| + 2|x_2 \cdot x_4| + \cdots + 2|x_{n-1} \cdot x_n| \geq x_1^2 + x_2^2 + x_3^2 + \cdots + x_n^2 = \|x\|^2$, logo

$$\|x\| = \sqrt{x_1^2 + x_2^2 + x_3^2 + \cdots + x_n^2} \leq \sqrt{\|x\|_S^2} = \|x\|_S.$$

Por fim, se $|x_k| = \max\limits_{1\leq i\leq n} = \|x\|_M$, então para todo $i = 1, 2, 3, \cdots, n$ vai valer $|x_k| \geq |x_i|$, ou seja, $|x_k| \geq |x_1|$, $|x_k| \geq |x_2|$, $|x_k| \geq |x_3|, \ldots, |x_k| \geq |x_n|$, logo

$$\|x\|_S = |x_1| + |x_2| + |x_3| + \cdots + |x_n| \leq n \cdot |x_k| = n \cdot \left(\max_{1\leq i\leq n} |x_i|\right) = n \cdot \|x\|_M.$$

Estas igualdades são extremamente importante, pois servem para mostrar que as três normas usuais em $\mathbb{R}^n$ são equivalentes. É dito que duas normas $\| \ \|_1$ e $\| \ \|_2$ quaisquer em $\mathbb{R}^n$ são equivalentes quando existem duas constantes $a > 0$ e $b > 0$ de tal forma que, para todo $x \in \mathbb{R}^n$, tenhamos

$$\|x\|_1 \leq a \cdot \|x\|_2 \text{ e } \|x\|_2 \leq b \cdot \|x\|_1.$$

Isto é extremamente importante pois, em alguns casos pode ser muito difícil de se trabalhar com a Norma Euclidiana, mas fácil de se trabalhar com outra norma mais simples, porém equivalente. A equivalência de normas é uma relação *REFLEXIVA*, *SIMÉTRICA* e *TRANSITIVA* que permite escolher a norma que for mais conveniente para demonstrar algum resultado e sua veracidade permanecerá para todas as normas equivalentes a norma adotada.

Teorema 2.1.1. *Duas normas quaisquer em $\mathbb{R}^n$ são equivalentes.*

A demonstração deste Teorema está no livro Curso de análise vol.2 (LIMA, 2010, p. 19), ele permite substituir a norma que está sendo utilizada por qualquer outra equivalente, porém neste trabalho, é suficiente que as norma usuais (Euclidiana, da Soma e do Máximo) sejam equivalentes pois são as únicas que utilizadas aqui.

Definição 2.4. Um *Produto Interno* em $\mathbb{R}^n$ é uma aplicação $\langle \ \ , \ \ \rangle : \mathbb{R}^n \to \mathbb{R}$ que associa cada par de vetores $x, y \in \mathbb{R}^n$, o número real $\langle x, y\rangle$ chamado produto interno de x por y, em que para quaisquer $x, x', y \in \mathbb{R}^n$ e $\lambda \in \mathbb{R}$, são válidas as propriedades:

P1. $\langle x + x', y\rangle = \langle x, y\rangle + \langle x', y\rangle$

P2. $\langle \lambda \cdot x, y\rangle = \lambda \cdot \langle x, y\rangle$

P3. $\langle x, y\rangle = \langle y, x\rangle$

P4. $x \neq \overrightarrow{0} \Rightarrow \langle x, x\rangle > 0$

As propriedades P1, P2 e P3 entregam a bilinearidade do produto interno. Como já havia sido mencionado antes, o produto interno fornece a norma euclidiana

$$\langle x, x\rangle = \sum_{i=1}^{n}(x_i \cdot x_i) = \sum_{i=1}^{n}(x_i^2) = \left[\left(\sum_{i=1}^{n} x_i^2\right)^{\frac{1}{2}}\right]^2 = \|x\|^2$$

para todo $x \in \mathbb{R}^n$.

Perceba que $\langle x, x\rangle = \|x\|^2$ se, e somente se, $\sqrt{\langle x, x\rangle} = \|x\|$, neste sentido a noma euclidiana é de fato uma norma em $\mathbb{R}^n$. Com efeito, para todo $x \neq \overrightarrow{0}$, vale

$$\langle x, x\rangle > 0 \;\Rightarrow\; \sqrt{\langle x, x\rangle} > 0 \;\Rightarrow\; \|x\| > 0,$$

o que verifica a propriedade (**N1.**). Perceba também que para todo $x \in \mathbb{R}^n$ e todo $\lambda \in \mathbb{R}$, vale

$$\|\lambda \cdot x\| = \sqrt{\langle \lambda \cdot x, \lambda \cdot x\rangle}\sqrt{\lambda^2 \cdot \langle x, x\rangle} = \sqrt{\lambda^2} \cdot \sqrt{\langle x, x\rangle} = |\lambda| \cdot \sqrt{\|x\|^2} = |\lambda| \cdot \|x\|$$

e isto verifica a propriedade (**N2.**). Já a propriedade (**N3.**) é consequência da Desigualdade de Cauchy-Schwarz, portanto será admitida a sua veracidade para mostrar a propriedade (**N3.**) e em seguida será demonstrada tal desigualdade. Observe que para todo $x, y \in \mathbb{R}^n$, vale

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

Pela Desigualdade de Cauchy-Schwarz $|\langle x, y\rangle| \leq \|x\| \cdot \|y\|$, mas $\langle x, y\rangle \leq |\langle x, y\rangle|$, pois esta é uma propriedade do valor absoluto, em particular $\langle x, y\rangle \leq \|x\| \cdot \|y\|$, logo

$$\|x + y\|^2 = \|x\|^2 + 2\langle x, y\rangle + \|y\|^2 \leq \|x\|^2 + 2 \cdot \|x\| \cdot \|y\| + \|y\|^2 = (\|x\| + \|y\|)^2.$$

Como $\|x + y\|^2 \geq 0$ e $(\|x\| + \|y\|)^2 \geq 0$, então não haverá problema algum em extrair a raiz quadrada em ambos os lados da desigualdade acima. Logo

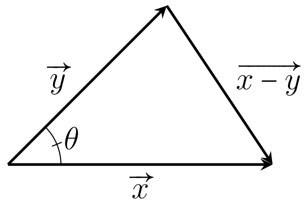
$$\|x + y\|^2 \leq (\|x\| + \|y\|)^2 \;\Rightarrow\; \|x + y\| \leq \|x\| + \|y\|$$

e isto verifica a propriedade (**N3.**).

Observação 2.2. Será considerado aqui apenas o produto interno canônico $\langle x, y\rangle = \sum_{i=1}^{n}(x_i \cdot y_i)$, com $x, y \in \mathbb{R}^n$. Note que, se $x = \overrightarrow{0}$ então, $\langle x, x\rangle = \underbrace{0 \cdot 0 + 0 \cdot 0 + \cdots + 0 \cdot 0 + 0 \cdot 0}_{n \text{ vezes}} = 0$.

Relembrando alguns conceitos do Cálculo Vetorial e generalizando estes para dimensão n, dois vetores $x, y \in \mathbb{R}^n$ são ditos *linearmente dependentes* quando têm a mesma direção, ou seja, quando um é múltiplo escalar do outro. A definição geométrica do produto interno aqui é bem semelhante a que é vista no Cálculo Vetorial, dados $x, y \in \mathbb{R}^n$ não nulos, vale $\langle x, y\rangle = \|x\| \cdot \|y\| \cdot \cos\theta$, em que θ é o ângulo formado pelos vetores x e y. A ideia é a seguinte, sejam x e y vetores do $\mathbb{R}^n$, ambos diferentes de $\overrightarrow{0}$, e seja θ o ângulo formado por estes vetores, então os vetores x, y e $x - y$ fomentam um triângulo de tal forma que, o vetor $x - y$ representa o lado oposto a θ.

Figura 1 – Interpretação geométrica do produto interno.



Fonte: Elaborada pelo autor.

Aplicando agora a lei dos cossenos no triângulo da Figura 1:

$$\begin{aligned} \|x-y\|^2 &= \|x\|^2 + \|y\|^2 - 2\cdot\|x\|\cdot\|y\|\cdot\cos\theta \\ \Rightarrow\ \langle x-y, x-y\rangle &= \|x\|^2 + \|y\|^2 - 2\cdot\|x\|\cdot\|y\|\cdot\cos\theta \\ \Rightarrow\ \langle x,x\rangle - 2\cdot\langle x,y\rangle + \langle y,y\rangle &= \|x\|^2 + \|y\|^2 - 2\cdot\|x\|\cdot\|y\|\cdot\cos\theta \\ \Rightarrow\ \|x\|^2 - 2\cdot\langle x,y\rangle + \|y\|^2 &= \|x\|^2 + \|y\|^2 - 2\cdot\|x\|\cdot\|y\|\cdot\cos\theta \\ \Rightarrow\ -2\cdot\langle x,y\rangle &= -2\cdot\|x\|\cdot\|y\|\cdot\cos\theta \\ \Rightarrow\ \langle x,y\rangle &= \|x\|\cdot\|y\|\cdot\cos\theta. \end{aligned}$$

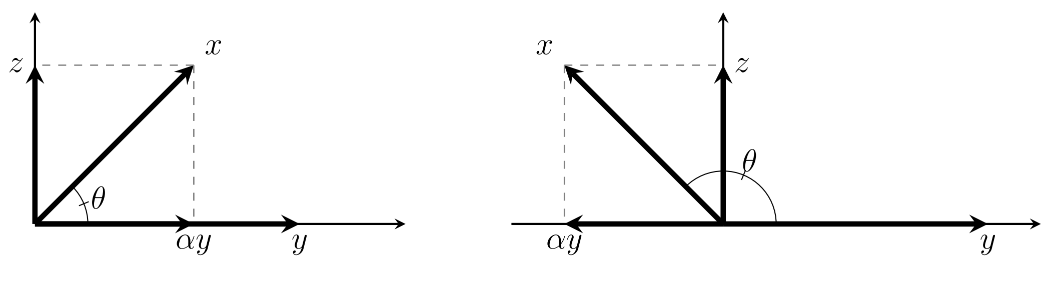
Como os vetores x e y são ambos diferentes de $\overrightarrow{0}$ então $\|x\| > 0$ e $\|y\| > 0$, logo a única forma do produto interno entre eles resultar em 0 é se $\cos\theta = 0$, ou seja $\theta = \frac{\pi}{2}$, isto quer dizer que dois vetores são *ortogonais* quando o seu produto interno é igual a 0.

Teorema 2.1.2. (Desigualdade de Cauchy – Schwarz). $|\langle x,y\rangle| \leq \|x\|\cdot\|y\|$, *para quaisquer* $x, y \in \mathbb{R}^n$. *Valendo a igualdade se, e somente se, os vetores são Linearmente Dependentes.*

Demonstração. Se $y = \overrightarrow{0}$, então seja $x = (x_1, x_2, x_3, \ldots, x_n)$, logo $|\langle x,y\rangle| = |\langle x,0\rangle| = |x_1\cdot 0 + x_2\cdot 0 + x_3\cdot 0 + \cdots + x_n\cdot 0| = |0| = 0$, por outro lado $\|x\|\cdot\|y\| = \|x\|\cdot\|\overrightarrow{0}\| = \|x\|\cdot 0 = \|x\cdot 0\| = \|(x_1\cdot 0, x_2\cdot 0, x_3\cdot 0, \ldots, x_n\cdot 0)\| = \|\overrightarrow{0}\| = 0$ e fica provado para o caso em que um dos vetores é o vetor nulo.

Se $y \neq \overrightarrow{0}$, então considere x e y dois vetores quaisquer de $\mathbb{R}^n$. Ao projetar ortogonalmente o vetor x sobre o vetor y, esta projeção é um múltiplo escalar de y, ou seja, a projeção de x sobre y é um vetor αy.

Figura 2 – Projeção ortogonal.



Fonte: Elaborada pelo autor.

Sendo θ o ângulo formado por x e y, a Figura 2 apresenta tanto o caso em que tal ângulo é agudo quanto o caso quando o mesmo é obtuso. Dos conhecimentos do Cálculo Vetorial, a fim de conseguir um vetor z que seja ortogonal à y, basta tomar $z = x - \alpha y$, desde que αy seja a projeção ortogonal de x sobre y, mas por construção isto ocorre, portanto o vetor $z = x - \alpha y$ é de fato ortogonal à y, isto quer dizer que $\langle z, y\rangle = 0$, logo

$$0 = \langle z, y\rangle = \langle x - \alpha y, y\rangle = \langle x, y\rangle - \alpha \cdot \langle y, y\rangle = \langle x, y\rangle - \alpha \cdot \|y\|^2 \quad \Leftrightarrow \quad \alpha = \frac{\langle x, y\rangle}{\|y\|^2}.$$

Agora note que $z = x - \alpha y$ se, e somente se, $x = z + \alpha y$, perceba também que $\|z\|^2 \geq 0$ para todo $z \in \mathbb{R}^n$, logo

$$\begin{aligned}
\|x\|^2 &= \langle x, x\rangle = \langle z + \alpha y, z + \alpha y\rangle = \langle z, z\rangle + \langle z, \alpha y\rangle + \langle \alpha y, z\rangle + \langle \alpha y, \alpha y\rangle \\
&= \langle z, z\rangle + 2\alpha \cdot \langle z, y\rangle + \alpha^2 \cdot \langle y, y\rangle = \|z\|^2 + \alpha^2\|y\|^2 \\
&\geq \alpha^2\|y\|^2.
\end{aligned}$$

Multiplicando todos os membros da equação acima por $\|y\|^2$, segue que

$$\begin{aligned}
\|x\|^2 \cdot \|y\|^2 \geq \alpha^2\|y\|^2 \cdot \|y\|^2 = \left(\frac{\langle x, y\rangle}{\|y\|^2}\right)^2 (\|y\|^2)^2 = \langle x, x\rangle^2 \quad &\Rightarrow \quad \sqrt{\|x\|^2 \cdot \|y\|^2} \geq \sqrt{\langle x, x\rangle^2} \\
&\Rightarrow \quad \|x\| \cdot \|y\| \geq |\langle x, x\rangle|
\end{aligned}$$

e então fica provada a primeira afirmação.

Agora, se vale a igualdade, então

$$\begin{aligned}
|\langle x, y\rangle| = \|x\| \cdot \|y\| \quad \Rightarrow \quad \langle x, y\rangle^2 &= \|x\|^2 \cdot \|y\|^2 \\
\Rightarrow \quad \frac{\langle x, y\rangle^2}{\|y\|^2} &= \|x\|^2 = \langle x, x\rangle = \langle z + \alpha y, z + \alpha y\rangle = \|z\|^2 + \alpha^2 \cdot \|y\|^2 \\
&= \|z\|^2 + \left(\frac{\langle x, y\rangle}{\|y\|^2}\right)^2 \cdot \|y\|^2 = \|z\|^2 + \frac{\langle x, y\rangle^2}{\|y\|^2},
\end{aligned}$$

mas

$$\frac{\langle x, y\rangle^2}{\|y\|^2} = \|z\|^2 + \frac{\langle x, y\rangle^2}{\|y\|^2} \quad \Leftrightarrow \quad \|z\|^2 = 0 \quad \Leftrightarrow \quad \|z\| = 0 \quad \Leftrightarrow \quad z = \overrightarrow{0}$$

e portanto, se $z = \overrightarrow{0}$, então $x = \alpha y$, logo x e y são linearmente dependentes.

Reciprocamente, se x e y são linearmente dependentes então a projeção de x sobre y

é o próprio x, ou seja, $\alpha y = x$, logo

$$\begin{aligned}
\|x\|^2 &= \langle x, x\rangle = \langle \alpha y, \alpha y\rangle = \alpha^2 \cdot \langle y, y\rangle = \alpha^2 \cdot \|y\|^2 \\
\Rightarrow \|x\|^2 \cdot \|y\|^2 &= \alpha^2 \cdot \|y\|^2 \cdot \|y\|^2 = \alpha\|y\|^2 \cdot \alpha\|y\|^2 \\
&= \langle \alpha y, y\rangle \cdot \langle \alpha y, y\rangle = \langle \alpha y, y\rangle^2 = \langle x, y\rangle^2 \\
\Rightarrow \sqrt{\|x\|^2 \cdot \|y\|^2} &= \sqrt{\langle x, y\rangle^2} \\
\Rightarrow \|x\| \cdot \|y\| &= |\langle x, y\rangle|
\end{aligned}$$

e portanto vale a igualdade. □

À título de curiosidade, uma norma provém de um produto interno quando satisfaz a *identidade do paralelogramo*

$$\|x + y\|^2 + \|x - y\|^2 = 2(\|x\|^2 + \|y\|^2). \tag{2.1}$$

Perceba que $\|x+y\|^2 + \|x-y\|^2 = \langle x+y, x+y\rangle + \langle x-y, x-y\rangle = \langle x, x\rangle + 2\cdot\langle x, y\rangle + \langle y, y\rangle + \langle y, y\rangle - 2\cdot\langle x, y\rangle + \langle x, x\rangle = \langle x, x\rangle + \langle y, y\rangle + \langle y, y\rangle + \langle x, x\rangle = 2(\|x\|^2 + \|y\|^2)$. Portanto a Norma Euclidiana provém de um produto interno, porém isto não acontece com as outras duas que foram apresentadas, ao considerar $x, y \in \mathbb{R}^2$ com $x = (1, 0)$ e $y = (0, 1)$, e trocar na equação (2.1), a Norma Euclidiana pela Norma da Soma ou pela Norma do Máximo, segue que estas não verificam a identidade.

Definição 2.5. Chama-se *bola aberta* de centro em a e raio r, o conjunto de todos os pontos de $\mathbb{R}^n$ que estão a uma distância menor do que r do ponto a. Respectivamente a *bola fechada* é o conjunto de todos os pontos que estão a uma distância menor do que ou igual a r do ponto a e denota-se por

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

e

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

Observação 2.3. Neste segundo caso, se restringir o conjunto dos pontos cuja distância em relação ao ponto a seja exatamente igual a r, obtém-se então a esfera de centro em a e raio r, a qual é denotada por $S[a; r]$.

Definição 2.6. Um conjunto $X \subset \mathbb{R}^n$ é dito *limitado* se X está contido em alguma bola.

Em outras palavras, é dito que um conjunto X é limitado se existir um número real $r > 0$ e um ponto $a \in \mathbb{R}^n$, de tal forma que todos os pontos de X estejam contidos na bola de centro em a e raio r. Neste sentido pode-se dizer que toda bola em $\mathbb{R}^n$ é um conjunto limitado, pois satisfaz a definição de imediato.

Dado um ponto $a \in X$, é dito que a é um *ponto interior* de X quando existe algum δ positivo tal que $B(a;\delta) \subset X$ e denota-se por $\text{int}X$ o conjunto dos pontos interiores a X. Para todo $x \in \text{int}X$, é dito que X é uma *vizinhança* do ponto x.

Definição 2.7. Um conjunto X é dito aberto se todos os seus pontos são interiores, ou seja:

$$\forall x \in X,\ \exists\ \delta > 0 \ \text{ tal que } \ B(x;\delta) \subset X.$$

Uma consequência imediata da definição é que toda bola aberta é um conjunto aberto, pois para todo real $r > 0$, se $x \in B(a;r)$, então $\|x-a\| < r$ implica $0 < r - \|x-a\|$. Como o número real $r - \|x-a\|$ é estritamente positivo, ao definir uma bola $B(x; r - \|x-a\|)$, segue que, se $y \in B(x; r - \|x-a\|)$, então $\|y-x\| < r - \|x-a\|$, portanto $\|y-a\| = \|y-x+x-a\| \leq \|y-x\| + \|x-a\| < r - \|x-a\| + \|x-a\| = r$ implica $y \in B(a;r)$. Como y foi tomado arbitrariamente, então $B(x; r - \|x-a\|) \subset B(a;r)$.

Um tipo de aplicação muito importante na matemática em geral é a sequência de pontos.

Uma *sequência* em $\mathbb{R}^n$ é uma aplicação

$$\begin{aligned} x : \mathbb{N} &\longrightarrow \mathbb{R}^n \\ k &\longmapsto x_k \in \mathbb{R}^n \end{aligned}$$

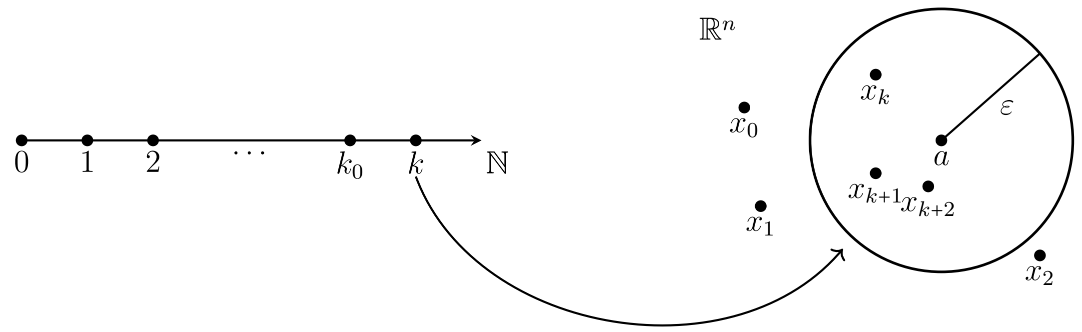
em que x_k é o k-ésimo termo da sequência ou termo de índice k. Note que cada termo x_k de uma sequência de pontos de $\mathbb{R}^n$ possui n coordenadas, isto quer dizer que, $x_k = (x_{k1}, x_{k2}, x_{k3}, \ldots, x_{kn})$. Neste sentido, cada coordenada x_{ki} de x_k representa um termo de uma sequência de pontos da reta, ou seja, cada coordenada de uma sequência de pontos em $\mathbb{R}^n$ é uma sequência da reta. Para representar uma sequência em $\mathbb{R}^n$, utiliza-se a notação $(x_k)_{k\in\mathbb{N}}$ ou simplesmente (x_k).

Uma subsequência de (x_k) é uma restrição infinita dos seus termos, de tal forma que, dado um subconjunto $\mathbb{N}' = \{k_1, k_2, k_3, \ldots, k_n, \ldots\} \subset \mathbb{N}$ com $k_n < k_{n+1}$ para todo $n \in \mathbb{N}$ se tenha $(x_{k_1}, x_{k_2}, x_{k_3}, \ldots, x_{k_n}, \ldots)$. Utiliza-se as notações $(x_k)_{k\in\mathbb{N}'}$, $(x_{k_i})_{i\in\mathbb{N}}$, ou simplesmente (x'_k) para representar uma subsequência de (x_k).

É dito que uma sequência é *limitada* quando existe um majorante real estritamente positivo para os seus termos em norma, ou seja, se existe $c > 0$, com $c \in \mathbb{R}$ tal que $\|x_k\| \leq c$ para todo $k \in \mathbb{N}$. É dito também que uma sequência $(x_k) \in \mathbb{R}^n$ *converge* para um ponto $a \in \mathbb{R}^n$ ou que a é o limite de (x_k), $\lim\limits_{k\to\infty} x_k = a$, ou simplesmente quando para todo $\varepsilon > 0$, puder ser exibido um número $k_0 \in \mathbb{N}$, tal que todos os termos de (x_k) que são sucessores a x_{k_0} (em relação a ordem dos índices) estejam contidos em $B(a;\varepsilon)$. Isto equivale a dizer que para todo $\varepsilon > 0$, existe um $k_0 \in \mathbb{N}$ tal que $k > k_0$ implica $\|x_k - a\| < \varepsilon$.

Observação 2.4. Perceba que, nestas condições, para toda sequência de pontos $(x_k)_{k\in\mathbb{N}} \in \mathbb{R}^n$ e $c \in \mathbb{R}^n$, vale $\lim\limits_{k\to\infty} \|x_k - c\|_1 = 0$ se, e somente se, $\lim\limits_{k\to\infty} \|x_k - c\|_2 = 0$, e isto quer dizer que duas normas equivalentes dão origem a mesma noção de limite em $\mathbb{R}^n$ (LIMA, 2010, p, 19).

Figura 3 – Sequência convergente em $\mathbb{R}^n$.



Fonte: Elaborada pelo autor.

Definição 2.8. Um ponto $a \in \mathbb{R}^n$ diz-se *valor de aderência* de uma sequência de pontos $(x_k) \in \mathbb{R}^n$, quando é o limite se alguma subsequência de (x_k).

Teorema 2.1.3 (Bolzano-Weierstrass)**.** *Toda sequência limitada em $\mathbb{R}^n$ possui uma subsequência que converge.*

Teorema 2.1.4. *Uma sequência limitada em $\mathbb{R}^n$ é convergente se, e somente se, possui um único valor de aderência.*

Demonstração. Seja (x_k) uma sequência limitada em $\mathbb{R}^n$ com $\lim\limits_{k\to\infty} x_k = a$ e seja $(x_{k'})$ com $k' \in \mathbb{N}' \subset \mathbb{N}$ uma subsequência qualquer de (x_k) que converge. Da definição de convergência de sequencias, como (x_k) converge então para todo $\varepsilon > 0$ existe $k_0 \in \mathbb{N}$ tal que

$$k > k_0 \Rightarrow \|x_k - a\| < \varepsilon, \tag{2.2}$$

em outras palavras $x_k \in B(a;\varepsilon)$. Como $\mathbb{N}'$ é um conjunto infinito, para $k' \in \mathbb{N}'$ com $k' > k_0$, por (2.2) segue que $\|x_{k'} - a\| < \varepsilon$, ou seja, $x_{k'} \in B(a;\varepsilon)$ para todo $k' > k_0$ e portanto $(x_{k'})$ converge para a, logo (x_k) possui um único valor de aderência.

Reciprocamente, seja a o único valor de aderência de (x_k). Suponha que $\lim\limits_{k\to\infty} x_k \neq a$, isto quer dizer que para algum $\varepsilon > 0$ o conjunto $\mathbb{N}' = \{k \in \mathbb{N}; x_k \notin B(a;\varepsilon)\} \subset \mathbb{N}$ é infinito, ou seja, $\|x_k - a\| \geq \varepsilon > 0$ para todo k pertencente ao conjunto infinito $\mathbb{N}' \subset \mathbb{N}$. Como $\mathbb{N}'$ é infinito, tome uma subsequência $(x_{k'})$ de (x_k) com $k' \in \mathbb{N}'$. Como (x_k) é limitada e todos os termos de $(x_{k'})$ são termos de (x_k), então a subsequência $(x_{k'})$ também é limitada e todos os seus termos satisfazem a condição $|x_{k'} - a| \geq \varepsilon$, pois $k' \in \mathbb{N}'$. Como $(x_{k'})$ é limitada, o Teorema de Bolzano-Weierstrass garante que tal sequência possui uma subsequência $(x_{k''})$ com $k'' \in \mathbb{N}'' \subset \mathbb{N}'$, que converge para um ponto $b \in \mathbb{R}^n$ e todos os termos de $(x_{k''})$ ainda satisfazem a condição $\|x_{k''} - a\| \geq \varepsilon$, pois $k'' \in \mathbb{N}''$ e $\mathbb{N}'' \subset \mathbb{N}'$. Logo

$\lim\limits_{k''\to\infty} \|x_{k''} - a\| = \|b - a\| \geq \varepsilon > 0$, isto quer dizer que $b \neq a$ e (x_k) possui no mínimo dois valores de aderência, o que é um absurdo pois a é o único valor de aderência de (x_k). Esta contradição é consequência de supor que $\lim\limits_{k\to\infty} x_k \neq a$, portanto $\lim\limits_{k\to\infty} x_k = a$. □

Definição 2.9. Diz-se que um ponto $a \in X \subset \mathbb{R}^n$ é um ponto *aderente* a X quando a é limite de alguma sequência de pontos de X. Chama-se *fecho* de X e denota-se por $\overline{X}$ o conjunto de todos os pontos aderentes a X.

Definição 2.10. Um conjunto $X \subset \mathbb{R}^n$ é dito *fechado* quando contém todos os seus pontos aderentes, ou seja, $X = \overline{X}$.

Mostra-se que toda bola fechada é um conjunto fechado observando que, se b é um ponto aderente à $B[a;r]$, então existe $(x_k)_{k\in\mathbb{N}}$ com $\lim\limits_{k\to\infty} x_k = b$ tal que, para todo $k \in \mathbb{N}$, $x_k \in B[a;r]$, isto quer dizer que para todo $k \in \mathbb{N}$ vale $\|x_k - a\| \leq r$, mas $\lim\limits_{k\to\infty} \|x_k - a\| = \|b - a\| \leq r$ implica $b \in B[a;r]$.

Definição 2.11. Um conjunto $K \in \mathbb{R}^n$ diz-se compacto quando é limitado e fechado.

Definição 2.12. Dado um conjunto $X \subseteq \mathbb{R}^n$ e um ponto $a \in X$, quando para todo $\varepsilon > 0$ existir algum $x \in B(a;\varepsilon)$ diferente do próprio a, será dito que a é um *ponto de acumulação de* X. Será dito que a é um ponto *isolado* quando não for ponto de acumulação, ou seja, quando existir $r > 0$ tal que a bola aberta de centro em a e raio r contenha apenas o ponto a, $B(a;r) = \{a\}$.

O espaço $\mathbb{R}^n$ não possui nenhum ponto isolado pois, ao fixar arbitrariamente um ponto $a \in \mathbb{R}^n$ e considerar um ponto $b = a + \dfrac{r}{2\|x\|} \cdot x$, em que x é um ponto qualquer de $\mathbb{R}^n$ e $r > 0$ um real arbitrário, então nestas condições b é um ponto de $\mathbb{R}^n$, e mais, $\|b - a\| = \left\| a + \dfrac{r}{2\|x\|} \cdot x - a \right\| = \dfrac{|r| \cdot \|x\|}{|2| \cdot \|x\|} = \dfrac{r}{2} < r$, ou seja, $b \in B(a;r)$, isto quer dizer que para qualquer $r > 0$, a bola $B(a;r)$ não contem apenas o a. Como o ponto a foi fixado arbitrariamente então nenhum ponto de $\mathbb{R}^n$ é isolado.

2.2 Continuidade

Definição 2.13. É dito que uma aplicação $f : X \subseteq \mathbb{R}^m \to \mathbb{R}^n$ é contínua em um ponto $a \in X$ se para todo $\varepsilon > 0$ existe $\delta > 0$ de tal forma que

$$x \in X, x \in B(a;\delta) \Rightarrow f(x) \in B(f(a);\varepsilon).$$

Quando f é contínua em todos os pontos do seu domínio, é dito simplesmente que f é contínua.

Da definição de bolas que foi vista anteriormente, dizer que $x \in B(a;\delta)$ equivale a dizer que a distância de x até a é menor do que δ, ou seja, $\|x - a\| < \delta$, analogamente $f(x) \in B(f(a);\varepsilon)$ equivale a dizer $\|f(x) - f(a)\| < \varepsilon$.

Teorema 2.2.1. *Sejam $U \subseteq \mathbb{R}^m$, $V \subseteq \mathbb{R}^n$ e $f : U \to \mathbb{R}^n$ uma aplicação contínua no ponto $a \in U$, com $f(U) \subset V$. Seja $g : V \to \mathbb{R}^p$ uma aplicação contínua no ponto $b = f(a) \in V$, então a composição $g \circ f : U \to \mathbb{R}^p$ é contínua no ponto a.*

Demonstração. A continuidade de g fornece para cada $\varepsilon > 0$ um $\delta' > 0$ tal que

$$y \in V,\ \|y - b\| < \delta' \Rightarrow \|g(y) - g(b)\| = \|g(y) - g(f(a))\| < \varepsilon.$$

Por outro lado, a continuidade de f fornece para cada $\delta' > 0$, um $\delta > 0$ tal que

$$x \in U,\ \|x - a\| < \delta \Rightarrow \|f(x) - f(a)\| < \delta'.$$

Mas se $x \in U$, então $f(x) \in V$, logo

$$\begin{aligned} x \in U,\ \|x - a\| < \delta \Rightarrow \|f(x) - f(a)\| < \delta' &\Rightarrow \|g(f(x)) - g(f(a))\| < \varepsilon \\ &\Rightarrow \|(g \circ f)(x) - (g \circ f)(a)\| < \varepsilon. \end{aligned}$$

□

Teorema 2.2.2. *Seja $f : U \subseteq \mathbb{R}^m \to \mathbb{R}^n$ e $a \in U$, com $f(a) = (f_1(a), f_2(a), f_3(a), \ldots, f_n(a))$. f é contínua no ponto a se, e somente se, cada uma das suas funções coordenadas $f_i : U \to \mathbb{R}$ for contínua em a.*

Demonstração. Primeiramente perceba que para todo $v = (\alpha_1, \alpha_2, \alpha_3, \ldots, \alpha_n) \in \mathbb{R}^n$, vale $\|v\| = \sqrt{\alpha_1^2 + \alpha_2^2 + \alpha_3^2 + \cdots + \alpha_n^2} \geq \sqrt{\alpha_i^2} = \|\alpha_i\|$, $\forall i = 1, 2, 3, \ldots, n$. Da continuidade de f segue que para todo $\varepsilon > 0$ existe $\delta > 0$ tal que

$$x \in U, \|x - a\| < \delta \Rightarrow \|f_i(x) - f_i(a)\| \leq \|f(x) - f(a)\| < \varepsilon$$

e portanto cada $f_i : U \to \mathbb{R}$ é contínua em a.

Reciprocamente, se cada $f_i : U \to \mathbb{R}$ é contínua em a, então para todo $\varepsilon > 0$ existem $\delta_1 > 0, \delta_2 > 0, \delta_3 > 0, \ldots, \delta_n > 0$ tais que

$$x \in U, \|x - a\| < \delta_i \Rightarrow \|f_i(x) - f_i(a)\| < \varepsilon.$$

Por conveniência, será tomada agora a Norma do Máximo e considerado δ o menor dos δ_i, logo

$$x \in U, \|x - a\|_M < \delta \Rightarrow \|f(x) - f(a)\|_M \leq |f_i(x) - f_i(a)| < \varepsilon$$

e portanto f é contínua em a. □

Exemplo 2.4. A soma de vetores em $\mathbb{R}^n$ é contínua. Considere a aplicação $s : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n$ definida por $s(x, y) = x + y$ considere também a norma da soma. Para todo $\varepsilon > 0$ existe um $\delta = \varepsilon$ tal que, se $(x,y),(v,w) \in \mathbb{R}^n$ e $\|(x,y)-(v,w)\|_S < \delta$, então

$$\begin{aligned}\|s(x,y)-s(v,w)\|_S &= \|x+y-(v+w)\|_S = \|x-v+y-w\|_S \\ &\leq \|x-v\|_S + \|y-w\|_S = \|(x,y)-(v,w)\|_S < \varepsilon.\end{aligned}$$

Exemplo 2.5. A multiplicação por escalar em $\mathbb{R}^n$ é contínua. Considere $\varphi : \mathbb{R}\times\mathbb{R}^n \to \mathbb{R}^n$ definida por $\varphi(t,x) = t\cdot x$. Fixando $t_0 \in \mathbb{R}$ e $x_0 \in \mathbb{R}^n$ arbitrariamente. Dado $\varepsilon > 0$, escolhendo $\delta_1 = \dfrac{\varepsilon}{3(|t_0|+1)}$ e $\delta_2 = \dfrac{\varepsilon}{3(\|x_0\|+1)}$, não haverá perda de generalidade em supor que $\varepsilon < 1$, logo $\varepsilon^2 < \varepsilon$. Portanto, se $|t-t_0| < \delta_2$ e $\|x-x_0\| < \delta_1$, segue que

$$|t_0|\cdot\|x-x_0\| < \delta_1\cdot|t_0| = \frac{\varepsilon\cdot|t_0|}{3(|t_0|+1)} = \frac{\varepsilon}{3}\cdot\frac{|t_0|}{(|t_0|+1)} < \frac{\varepsilon}{3},$$

$$|t-t_0|\cdot\|x_0\| < \delta_2\cdot\|x_0\| = \frac{\varepsilon\cdot\|x_0\|}{3(\|x_0\|+1)} = \frac{\varepsilon}{3}\cdot\frac{\|x_0\|}{(\|x_0\|+1)} < \frac{\varepsilon}{3}$$

e

$$\begin{aligned}|t-t_0|\cdot\|x-x_0\| < \delta_2\cdot\delta_1 &= \frac{\varepsilon}{3(\|x_0\|+1)}\cdot\frac{\varepsilon}{3(|t_0|+1)} = \frac{\varepsilon}{3}\cdot\frac{1}{(\|x_0\|+1)}\cdot\frac{\varepsilon}{3}\cdot\frac{1}{(|t_0|+1)} \\ &< \frac{\varepsilon}{3}\cdot\frac{\varepsilon}{3} = \frac{\varepsilon^2}{9} < \frac{\varepsilon^2}{3} < \frac{\varepsilon}{3}.\end{aligned}$$

Então,

$$\begin{aligned}\|\varphi(t,x)-\varphi(t_0,x_0)\| &= \|t\cdot x - t_0\cdot x_0\| = \|(t-t_0)\cdot(x-x_0) + t_0\cdot(x-x_0) + (t-t_0)\cdot x_0\| \\ &\leq \|(t-t_0)\cdot(x-x_0)\| + \|t_0\cdot(x-x_0)\| + \|(t-t_0)\cdot x_0\| \\ &= |(t-t_0)|\cdot\|x-x_0\| + |t_0|\cdot\|x-x_0\| + |t-t_0|\cdot\|x_0\| \\ &< \frac{\varepsilon}{3}+\frac{\varepsilon}{3}+\frac{\varepsilon}{3} = \frac{3\cdot\varepsilon}{3} = \varepsilon.\end{aligned}$$

Em particular, o produto de números reais é contínuo.

Exemplo 2.6. O produto interno em $\mathbb{R}^n$ é uma aplicação contínua. Ora, seja $\xi : \mathbb{R}^n\times\mathbb{R}^n \to \mathbb{R}$ com $\xi(x,y) = \langle x,y\rangle$, em que $x = (x_1,x_2,x_3,\dots,x_n)$ e $y = (y_1,y_2,y_3,\dots,y_n)$, então $\xi(x,y) = x_1\cdot y_1 + x_2\cdot y_2 + x_3\cdot y_3 + \cdots + x_n\cdot y_n$. Como cada $x_i, y_i \in \mathbb{R}$, o produto de números reais é contínuo e ainda resulta em um vetor da reta, portanto o produto interno é contínuo por se resumir a uma soma de vetores da reta.

Exemplo 2.7. A aplicação $\rho : \mathbb{R}-\{0\} \to \mathbb{R}$ com $\rho(t) = \dfrac{1}{t}$, é contínua em seu domínio. Fixando arbitrariamente $t_0 \in \mathbb{R}-\{0\}$. Dado $\varepsilon > 0$, considere δ o menor dos valores $\dfrac{|t_0|}{2}$ e $\dfrac{\varepsilon\cdot|t_0|^2}{2}$, ou seja, $\delta \leq \dfrac{|t_0|}{2}$ e $\delta \leq \dfrac{\varepsilon\cdot|t_0|^2}{2}$, perceba que escolhendo δ desta forma, segue que

$\delta > 0$. Como $\delta \leq \dfrac{|t_0|}{2}$, então $\left(-\dfrac{t_0}{2}, \dfrac{t_0}{2}\right) \cap (t_0 - \delta, t_0 + \delta) = \varnothing$, logo $t \in (t_0 - \delta, t_0 + \delta)$ implica $t \notin \left(-\dfrac{t_0}{2}, \dfrac{t_0}{2}\right)$ implica $|t| > \dfrac{|t_0|}{2}$, portanto

$$
\begin{aligned}
|t - t_0| < \delta \Rightarrow |f(t) - f(t_0)| &= \left|\frac{1}{t} - \frac{1}{t_0}\right| = \left|\frac{1}{t_0} - \frac{1}{t}\right| = \left|\frac{t - t_0}{t_0 \cdot t}\right| = \frac{|t - t_0|}{|t_0| \cdot |t|} \\
&< \frac{\delta}{|t_0| \cdot |t|} < \frac{\delta}{|t_0| \cdot \dfrac{|t_0|}{2}} = \frac{2\delta}{|t_0|^2} \leq \frac{2 \cdot \dfrac{\varepsilon \cdot |t_0|^2}{2}}{|t_0|^2} = \varepsilon.
\end{aligned}
$$

Teorema 2.2.3. *Seja $U \subseteq \mathbb{R}^m$, sejam também $f, g : U \to \mathbb{R}^n$ e $\alpha : U \to \mathbb{R}$ aplicações contínuas, então as aplicações abaixo são contínuas.*

$$
\begin{aligned}
f + g : U \to \mathbb{R}^n, &\quad (f + g)(x) = f(x) + g(x) \\
\alpha \cdot f : U \to \mathbb{R}^n, &\quad (\alpha \cdot f)(x) = \alpha(x) \cdot f(x) \\
\langle f, g \rangle : U \to \mathbb{R}, &\quad \langle f, g \rangle(x) = \langle f(x), g(x) \rangle \\
\frac{1}{\alpha} : U \to \mathbb{R}, &\quad \left(\frac{1}{\alpha}\right)(x) = \frac{1}{\alpha(x)}, \text{ desde que } \alpha(U) \neq 0
\end{aligned}
$$

Demonstração. Considere $s : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n$ definida por $s(x, y) = x + y$, $\xi : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ definida por $\xi(x, y) = \langle x, y \rangle$, $\varphi : \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}^n$ definida por $\varphi(t, x) = t \cdot x$ e $\rho : \mathbb{R} - \{0\} \to \mathbb{R}$ definida por $\rho(t) = \dfrac{1}{t}$. Como foi visto nos exemplos 2.4, 2.5, 2.6 e 2.7, s, ξ, φ e ρ são contínuas. Como f, g e α são contínuas, então o Teorema 2.2.2 garante que $F : U \to \mathbb{R}^n \times \mathbb{R}^n$ com $F(x) = (f(x), g(x))$ e $G : U \to \mathbb{R} \times \mathbb{R}^n$ com $G(x) = (\alpha(x), f(x))$, são ambas contínuas, já que estas possuem coordenadas contínuas. Note que $(s \circ F)(x) = s(F(x)) = s(f(x), g(x)) = f(x) + g(x)$, analogamente $(\xi \circ F)(x) = \langle f(x), g(x) \rangle$, $(\varphi \circ G)(x) = \alpha(x) \cdot f(x)$, $(\rho \circ \alpha)(x) = \dfrac{1}{\alpha(x)}$ e o Teorema 2.2.1 garante que estas são todas aplicações contínuas, já que são composições de outras aplicações contínuas. □

Definição 2.14. Dado $U \subset \mathbb{R}^n$, uma *cisão* de U é uma decomposição disjunta de dois abertos em U, ou seja, uma decomposição $U = A \cup B$, em que A e B são abertos em U e $A \cap B = \varnothing$. Note que, o conjunto $\varnothing$ é aberto (LIMA, 1993, p. 64), portanto todo conjunto aberto $U \subset \mathbb{R}^n$ admite, no mínimo, a cisão $U = U \cup \varnothing$, a qual é dita *cisão trivial* de U.

Definição 2.15. Um conjunto $U \subset \mathbb{R}^n$ diz-se *conexo* quando não admite outra cisão que não seja a trivial, ou seja, se $U = A \cup B$, em que A e B são abertos e disjuntos, então $A = \varnothing$ ou $B = \varnothing$.

Tanto aqui quanto em contextos mais gerais, como por exemplo na Topologia dos Espaços Métricos, a noção de conjunto aberto é de extrema importância, pois esta garante que, a fim de que uma aplicação $f : U \subset \mathbb{R}^m \to V \subset \mathbb{R}^n$ seja contínua, é necessário e suficiente que a imagem inversa $f^{-1}(V')$ de todo subconjunto aberto $V' \in V$ seja um conjunto aberto em U (LIMA, 1993, p. 68). Outra consideração importante a ser feita antes de prosseguir, é que se $U = A_1 \cup A_2 \cup A_3 \cup \cdots \cup A_m$, em que cada $A_i \in \mathbb{R}^n$ é um conjunto aberto, então U é aberto, ou seja, uma reunião finita de subconjuntos abertos de $\mathbb{R}^n$ é um conjunto aberto (LIMA, 1993, p. 66). Estas duas afirmações serão utilizadas nos próximos resultados.

Teorema 2.2.4. *Seja $f : U \subset \mathbb{R}^m \to \mathbb{R}^n$ uma aplicação contínua, em que U é conexo. A imagem de U por f, $f(U)$, é um conjunto conexo.*

Demonstração. Se $f(U) = A \cup B$ é uma cisão, então A e B são abertos em $f(U)$, e mais, $f^{-1}(f(U)) = f^{-1}(A \cup B) = f^{-1}(A) \cup f^{-1}(B)$ implica $U = f^{-1}(A) \cup f^{-1}(B)$. Como f é contínua e os conjuntos A e B são abertos, então $f^{-1}(A)$ e $f^{-1}(B)$ são abertos em U, isto quer dizer que a união finita $f^{-1}(A) \cup f^{-1}(B)$ é, de fato, uma cisão em U, ou seja, $f^{-1}(A) = \varnothing$ (daí segue que $A = \varnothing$) ou $f^{-1}(B) = \varnothing$ (daí segue que $B = \varnothing$), logo a cisão $f(U) = A \cup B$ é a trivial e portanto $f(U)$ é um conexo. □

Teorema 2.2.5. *Todos os subconjuntos conexos da reta são intervalos.*

Demonstração. Tome $U \subseteq \mathbb{R}$ um conjunto conexo. Suponha que U não seja um intervalo, ou seja, tomando $a, b \in U$, então existe $c \notin U$ tal que $a < c < b$. Note que, ao tomar $A = \{x \in U | x < c\}$ e $B = \{x \in U | x > c\}$, segue que $U = A \cup B$. Note também que A e B são abertos, ou seja, $A \cup B$ é uma cisão em U, a qual não é a trivial, pois $a \in A$ e $b \in B$. Mas se esta cisão, que de fato existe, não for a trivial, então U não pode ser conexo e isto contradiz a hipótese inicial, portanto U é um intervalo. □

Teorema 2.2.6 (Teorema do Valor Intermediário)**.** *Seja $f : U \subset \mathbb{R}^n \to \mathbb{R}$ contínua, em que U é um conjunto conexo. Se existem $a, b \in U$ e $d \in \mathbb{R}$ tais que $f(a) < d < f(b)$, então existe $c \in U$ tal que $f(c) = d$.*

Demonstração. Como f é contínua e U é conexo, então pelo Teorema 2.2.4 $f(U)$ é conexo e pelo Teorema 2.2.5, $f(U)$ é um intervalo. □

2.3 Caminhos

Será feita agora uma extensão das funções de uma variável com imagem na reta para as funções de uma variável com imagem no espaço $\mathbb{R}^n$ e serão apresentados alguns resultados referentes a este tipo de aplicação, pois isto será de extrema importância no decorrer do trabalho.

Definição 2.16. Um caminho $f : I \to \mathbb{R}^n$ é uma aplicação definida em um intervalo real não-degenerado[1] I.

[1]Um intervalo de extremos a e b diz-de *não-degenerado* quando $a \neq b$, ou seja, contêm mais de um ponto da reta.

Semelhante as funções reais de uma variável, dado um caminho $f : I \to \mathbb{R}^n$ e um ponto $a \in I$, a derivada de f no ponto a é definida pelo limite

$$f'(a) = \lim_{t \to 0} \frac{f(a+t) - f(a)}{t},$$

quando existe. Se $f = (f_1, \ldots, f_n)$, observe que

$$\begin{aligned} \frac{f(a+t) - f(a)}{t} &= \frac{(f_1(a+t), \ldots, f_n(a+t)) - (f_1(a), \ldots, f_n(a))}{t} \\ &= \frac{(f_1(a+t) - f_1(a), \ldots, f_n(a+t) - f_n(a))}{t} \\ &= \left(\frac{f_1(a+t) - f_1(a)}{t}, \ldots, \frac{f_n(a+t) - f_n(a)}{t} \right), \end{aligned}$$

portanto, ao trabalhar com uma função que pega valores em um intervalo e caem no Espaço Euclidiano $\mathbb{R}^n$, é natural que derivar f signifique derivar cada uma de suas funções coordenadas $f = (f_1, f_2, f_3, \ldots, f_n)$, logo $f'(a) = (f_1'(a), f_2'(a), f_3'(a), \ldots, f_n'(a))$. De forma análoga se verifica que um caminho é contínuo se, e somente se, cada uma das suas funções coordenadas também o são. Neste caso de funções de uma variável, uma aplicação é dita *derivável* em um ponto a se possui derivada neste ponto, ou seja, dado $f : I \to \mathbb{R}^n$ e $a \in I$, se existe um vetor $v \in \mathbb{R}^n$ tal que, ao dar um incremento t no valor de a de tal forma que $a + t$ ainda esteja contido em I e

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

então f é *derivável* no ponto a, e mais, $v = f'(a)$. Intuitivamente, ao dar este incremento no valor de a e considerar a reta que intersecta o caminho nos pontos $f(a)$ e $f(a+t)$, quando o incremento tende a zero, estes dois pontos estão se aproximando ao máximo, a fim de conseguir um vetor tangente ao caminho e que o tangencie no ponto a e aponta para a direção em que uma partícula se move ao longo da imagem de f. Note que, como acabou de ser visto, é dito que um caminho $f = (f_1, f_2, f_3, \ldots, f_n) : I \to \mathbb{R}^n$ é derivável no ponto a, quando existe o vetor $f'(a) = (f_1'(a), f_2'(a), f_3'(a), \ldots, f_n'(a))$, ou seja, existe a derivada de cada $f_i : I \to \mathbb{R}$, mas dos conhecimentos da Análise Real, se uma função real de uma variável real é derivável, então ela é contínua, portanto se cada $f_i : I \to \mathbb{R}$ for derivável no ponto a, então cada uma é contínua em a e consequentemente $f(a) = (f_1(a), f_2(a), f_3(a), \ldots, f_n(a))$ é contínuo neste ponto, por ter coordenadas contínuas. Portanto, se um caminho é derivável em um ponto, então ele é contínuo neste ponto.

Teorema 2.3.1 (Regra da Cadeia). *Sejam I, J intervalos abertos, $\varphi : I \to J$ e $f : J \to \mathbb{R}^n$, deriváveis respectivamente nos pontos $a \in I$ e $b = \varphi(a) \in J$. Então $f \circ \varphi : I \to \mathbb{R}^n$ é um caminho derivável e $(f \circ \varphi)'(a) = \varphi'(a) \cdot f'(\varphi(a))$.*

Demonstração. Perceba que

$$\begin{aligned}(f\circ\varphi)(a) &= f(\varphi(a)) = (f_1(\varphi(a)), f_2(\varphi(a)), f_3(\varphi(a)), \dots, f_n(\varphi(a)))\\ &= ((f_1\circ\varphi)(a), (f_2\circ\varphi)(a), (f_3\circ\varphi)(a), \dots, (f_n\circ\varphi)(a))\end{aligned}$$

e neste sentido, basta provar a Regra da Cadeia para cada função coordenada do caminho $f\circ\varphi$.

Considere $f_i\circ\varphi : I \to \mathbb{R}$. Há dois casos a considerar sobre o valor de φ no ponto a :

CASO 1: Se $\varphi'(a) \neq 0$, sem perda de generalidade, suponha que seja $\varphi'(a) > 0$, como φ é contínua (pois é derivável), então existe um ε suficientemente pequeno tal que $\varphi'(x_0) > 0$ para todo $x_0 \in (a-\varepsilon, a+\varepsilon)$, logo f é crescente no intervalo $(a-\varepsilon, a+\varepsilon)$, portanto $\varphi(a) \neq \varphi(x_0)$ para todo $x_0 \in (a-\varepsilon, a+\varepsilon)$, isto quer dizer que $\varphi(x_0) - \varphi(a) \neq 0$, logo

$$\frac{f_i(\varphi(x_0)) - f_i(\varphi(a))}{x_0 - a} = \frac{\varphi(x_0) - \varphi(a)}{x_0 - a} \cdot \frac{f_i(\varphi(x_0)) - f_i(\varphi(a))}{\varphi(x_0) - \varphi(a)}, \tag{2.3}$$

passando o limite em ambos os lados e fazendo $x_0 \to a$ vale

$$(f_i\circ\varphi)'(a) = \varphi'(a)\cdot f_i'(\varphi(a)).$$

CASO 2: Se $\varphi'(a) = 0$, ou φ é constante numa vizinhança de a e para todo x_0 suficientemente próximo de a, vale $\varphi(x_0) = \varphi(a)$ e portanto

$$\lim_{x_0\to a} \frac{f_i(\varphi(x_0)) - f_i(\varphi(a))}{x_0 - a} = 0 = 0\cdot f_i'(\varphi(a)) = \varphi'(a)\cdot f_i'(\varphi(a)),$$

ou $\varphi(x_0) \neq \varphi(a)$ e portanto vale a equação (2.3) e

$$\begin{aligned}\lim_{x_0\to a} \frac{f(\varphi(x_0)) - f(\varphi(a))}{x_0 - a} &= \lim_{x_0\to a} \frac{\varphi(x_0) - \varphi(a)}{x_0 - a} \cdot \frac{f(\varphi(x_0)) - f(\varphi(a))}{\varphi(x_0) - \varphi(a)}\\ &\Rightarrow (f\circ\varphi)'(a) = \varphi'(a)\cdot f'(\varphi(a)) = 0.\end{aligned}$$

Como o resultado é válido em ambos os casos para cada função coordenada de $f\circ\varphi$, fica demonstrado o Teorema. $\square$

3 CÁLCULO DE VÁRIAS VARIÁVEIS

Deseja-se, a partir de agora, generalizar a noção de diferenciabilidade, as propriedades de derivadas do Cálculo de funções de uma variável real para o Cálculo de funções de n variáveis reais e alguns teoremas que serão importantes no decorrer deste trabalho para aplicações que têm algum subconjunto aberto de $\mathbb{R}^n$ como Domínio. Para isto, será de extrema importância a ideia de derivadas parciais, em que é feito algo parecido com o que se faz nas funções de uma variável real, porém dando um incremento apenas em uma das variáveis da aplicação.

3.1 Derivadas

Uma pergunta que pode surgir é "por quê é trabalhado com conjuntos abertos no domínio?", e a resposta é, apenas por simplicidade no sentido analítico, ao tomar um conjunto aberto, qualquer ponto tomado neste conjunto pode ser acrescido em qualquer direção para algum incremento t suficientemente pequeno de tal forma que o ponto mais o incremento ainda esteja contido no conjunto.

A ideia é a seguinte, seja $f : U \subseteq \mathbb{R}^2 \to \mathbb{R}$, U um conjunto aberto e $(x_0, y_0) \in U$. Como f é uma função de duas variáveis, então considere que seu domínio esteja contido no plano O_{xy}, logo ao fixar uma das coordenadas de f, obtêm-se duas novas funções $g(t) = f(x_0 + t, y_0)$ e $h(k) = f(x_0, y_0 + k)$, tais que, em g foi fixada a segunda coordenada e em h foi fixada a primeira coordenada. A partir de agora estas duas funções de uma variável real $g : (a, b) \to \mathbb{R}$ e $h : (c, d) \to \mathbb{R}$, podem ser derivadas como no Cálculo de uma variável, portanto

$$g'(x_0) = \lim_{t \to 0} \frac{g(x_0 + t) - g(x_0)}{t} = \lim_{t \to 0} \frac{f(x_0 + t, y_0) - f(x_0, y_0)}{t} \tag{3.1}$$

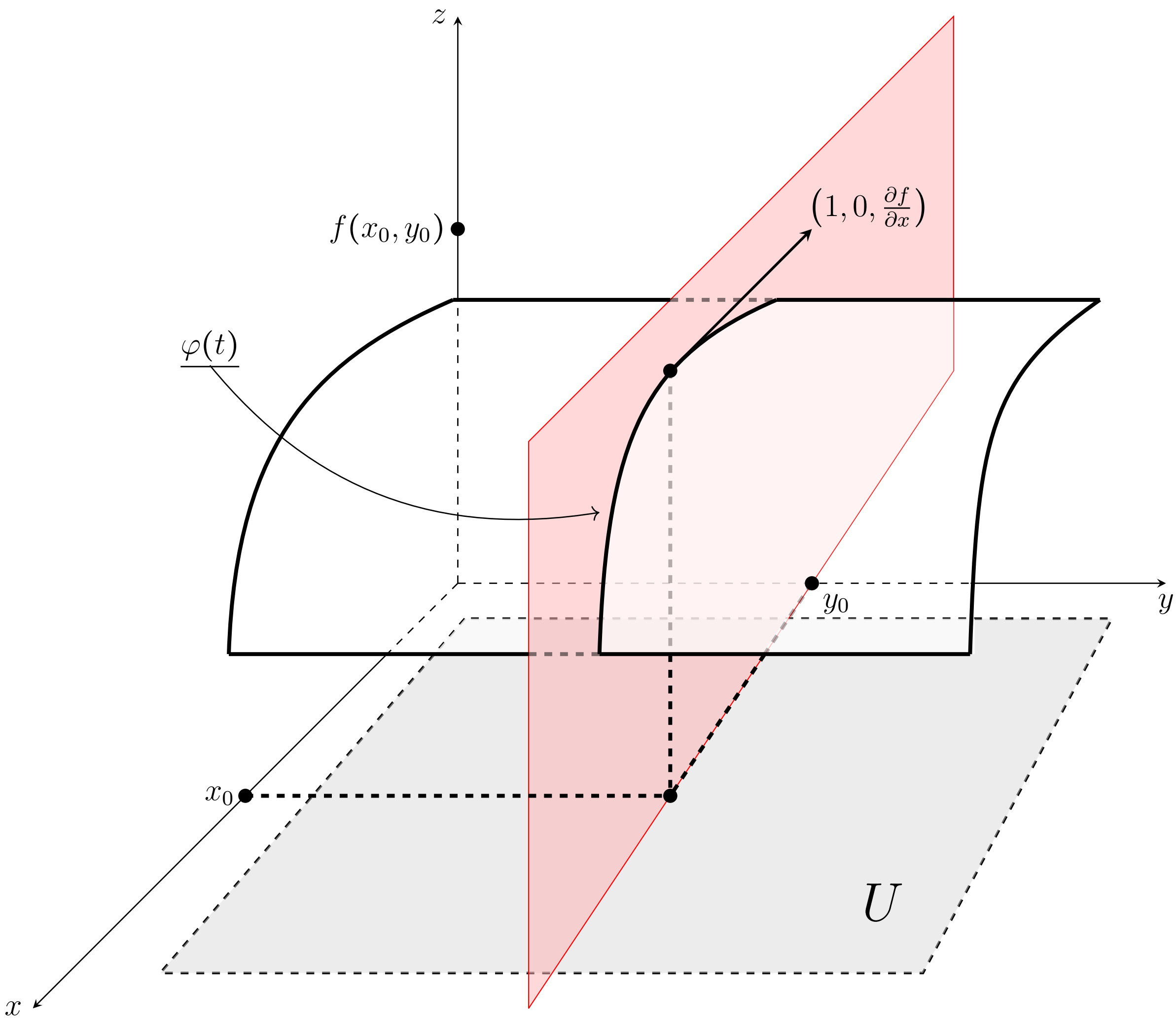
e

$$h'(y_0) = \lim_{k \to 0} \frac{h(y_0 + k) - h(y_0)}{k} = \lim_{k \to 0} \frac{f(x_0, y_0 + k) - f(x_0, y_0)}{k}. \tag{3.2}$$

O último limite de (3.1), quando existe, é definido como sendo a derivada parcial de f em relação à x, no ponto (x_0, y_0) e denotada por $\dfrac{\partial f}{\partial x}(x_0, y_0) = \lim\limits_{t \to 0} \dfrac{f(x_0 + t, y_0) - f(x_0, y_0)}{t}$, assim como o último limite de (3.2) é definido como sendo a derivada parcial de f em relação à y, no ponto (x_0, y_0) quando tal limite existe e denotada por $\dfrac{\partial f}{\partial y}(x_0, y_0) = \lim\limits_{k \to 0} \dfrac{f(x_0, y_0 + k) - f(x_0, y_0)}{k}$.

Geometricamente, ao fixar y_0 na segunda coordenada de f, a função foi restrita ao segmento da reta que intersecta o eixo das ordenadas em y_0 e é paralelo ao eixo das abscissas. O gráfico de f restrito a este segmento equivale ao gráfico de g e tal gráfico é formado pela intersecção do plano que intersecta o eixo das ordenadas em y_0 e é paralelo ao plano O_{xz}.

Figura 1 – Derivada parcial em relação à x ($\mathbb{R}^2$).



Fonte: Elaborada pelo autor.

Considere o caminho $\varphi(t) = (x_0 + t, y_0, f(x_0 + t, y_0))$ com $t \in \mathbb{R}$, então quando $t = 0$ vale $\varphi(0) = (x_0, y_0, f(x_0, y_0))$. O gráfico da restrição de f citada acima equivale ao caminho φ, logo, derivando φ no ponto $t = 0$, segue que

$$
\begin{aligned}
\varphi'(0) &= \lim_{t\to 0} \frac{\varphi(t) - \varphi(0)}{t} = \lim_{t\to 0} \frac{(x_0 + t, y_0, f(x_0 + t, y_0)) - (x_0, y_0, f(x_0, y_0))}{t} \\
&= \lim_{t\to 0} \frac{(t, 0, f(x_0 + t, y_0) - f(x_0, y_0))}{t} = \lim_{t\to 0} \left(1, 0, \frac{f(x_0 + t, y_0) - f(x_0, y_0)}{t}\right) \\
&= \left(1, 0, \frac{\partial f}{\partial x}(x_0, y_0)\right).
\end{aligned}
$$

Então o vetor $\left(1, 0, \dfrac{\partial f}{\partial x}(x_0, y_0)\right)$ define uma direção tangente à superfície[1] de f no ponto

[1] Dada uma aplicação de duas variáveis $f(x, y)$, a superfície de f é o conjunto de todos os pontos (x, y, z) do espaço $\mathbb{R}^3$ tais que $z = f(x, y)$.

$(x_0, y_0, f(x_0, y_0))$ e $\dfrac{\partial f}{\partial x}(x_0, y_0)$ é o coeficiente angular da reta que tangencia a superfície de f relativamente ao plano O_{xz}.

Analogamente, ao fixar x_0 na primeira coordenada de f, a função foi restrita a um segmento de reta que intersecta o eixo das abscissas em x_0 e é paralelo ao eixo das ordenadas e portanto o gráfico de f nessa restrição equivale ao gráfico de h, em que tal gráfico é fomentado pela intersecção do plano que intersecta o eixo das abscissas no ponto x_0 e é paralelo ao plano O_{yz}. Considerando o caminho $\psi(h) = (x_0, y_0 + k, f(x_0, y_0 + k))$ com $k \in \mathbb{R}$, então $\psi(0) = (x_0, y_0, f(x_0, y_0))$ e portanto $\psi'(0) = \left(0, 1, \dfrac{\partial f}{\partial y}(x_0, y_0)\right)$ é mais uma direção tangente á superfície de f e $\dfrac{\partial f}{\partial y}(x_0, y_0)$ é o coeficiente angular da reta que tangencia a superfície de f relativamente ao plano O_{yz}.

Definição 3.1. Seja $f : U \subseteq \mathbb{R}^n \to \mathbb{R}$, U um conjunto aberto e um ponto $a \in U$. O limite

$$\lim_{t \to 0} \frac{f(a + te_i) - f(a)}{t} = \frac{\partial f}{\partial x_i}(a), \tag{3.3}$$

quando existe, é a i-ésima *derivada parcial* de f (ou a derivada parcial de f em relação à i-ésima coordenada), no ponto a.

Ao fazer $a + te_i$, é feito um incremento no ponto a apenas na sua i-ésima coordenada, ao calcular o limite (3.3) é feito algo semelhante ao caso de funções de uma variável para encontrar o coeficiente angular da reta que toca no gráfico da função epenas no ponto a, porém neste caso, a função está restrita a um segmento de reta que é paralelo ao i-ésimo eixo de $\mathbb{R}^n$, pois a variação da função é feita na direção do i-ésimo vetor da base canônica. Neste sentido, o cálculo de uma derivada parcial pode ser feito utilizando a regras de derivação para funções de uma variável, considerando todas as variáveis com exceção da de índice i, como constantes.

O conceito de derivadas parciais é muito bonito, tanto do ponto de vista geométrico quanto do ponto de vista analítico, uma vez que fornece informações extremamente importantes sobre o comportamento das funções, mas é um conceito um pouco limitado. Quando é feita uma restrição a uma reta paralela a algum dos eixos, no domínio de uma função, é feito basicamente cálculo da variação da função na direção do eixo em questão e portanto, em uma função de n variáveis, quando $n > 1$, é feito o cálculo da variação da função em n direções mas, existem infinitas direções que podem ser tomadas, nas quais podem ser calculadas as variações da função para cada ponto a do seu domínio. A fim de acabar com esta limitação a derivada direcional fornece informações sobre o comportamento de alguma função ao longo de um segmento de reta qualquer do seu domínio (em qualquer direção), mesmo que esta reta não seja paralela a algum dos eixos coordenados. Se ao invés de vetores da base canônica for tomado um vetor v qualquer, obtém-se então a derivada de f na direção do vetor v.

Seja $f : U \subseteq \mathbb{R}^n \to \mathbb{R}$ em que U é um conjunto aberto, $a \in U$ e $v \in \mathbb{R}^n$. Considere o caminho $\lambda : (-\varepsilon, \varepsilon) \to \mathbb{R}^n$ com $\lambda(t) = a + t \cdot v$, note que $\lambda(0) = a$ e $\lambda'(t) = v$. Como U é aberto, então todo o segmento de reta $\lambda(t) = a + t \cdot v$ com $t \in (-\varepsilon, \varepsilon)$, está contido em U para algum ε suficientemente pequeno e portanto a restrição de f ao caminho λ, ou seja, $f \circ \lambda : (-\varepsilon, \varepsilon) \to \mathbb{R}$, é uma função real de uma variável real, logo pode ser derivada no ponto a da seguinte forma:

$$(f \circ \lambda)'(0) = \lim_{t \to 0} \frac{(f \circ \lambda)(t) - (f \circ \lambda)(0)}{t} = \lim_{t \to 0} \frac{f(a + t \cdot v) - f(a)}{t}.$$

A última igualdade logo acima, é definida como sendo a *derivada direcional de* f no ponto a, na direção de v e denotada por

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

Observação 3.1. Deseja-se aqui, que a derivada direcional de f na direção de v dependa linearmente do vetor v e por isto, diferentemente das definições de derivada direcional dadas na maioria dos livros de Cálculo, aqui não será suposto que $\|v\| = 1$, será tomado v como um vetor arbitrário de $\mathbb{R}^n$.

Algumas considerações são importantes de serem feitas. A primeira é se a existência de derivadas parciais em todos os pontos assegura a continuidade da função. A segunda é se a derivada $\dfrac{\partial f}{\partial v}$ depende linearmente do vetor v.

Contraexemplo 3.1. Seja $f : \mathbb{R}^2 \to \mathbb{R}$ em que

$$f(x,y) = \begin{cases} \dfrac{xy^2}{x^2 + y^4}\,, & \text{se } (x,y) \neq (0,0) \\ 0\,, & \text{se } x = y = 0 \end{cases}$$

Para $(x,y) \neq (0,0)$:

$$\frac{\partial f}{\partial x}(x,y) = \frac{y^2 \cdot (x^2 + y^4) - xy^2 \cdot 2x}{(x^2 + y^4)^2} = \frac{x^2y^2 + y^6 - 2x^2y^2}{(x^2 + y^4)^2} = \frac{y^6 - x^2y^2}{(x^2 + y^4)^2},$$

$$\frac{\partial f}{\partial y}(x,y) = \frac{2xy \cdot (x^2 + y^4) - xy^2 \cdot 4y^3}{(x^2 + y^4)^2} = \frac{2x^3y + 2xy^5 - 4xy^5}{(x^2 + y^4)^2} = \frac{2x^3y - 2xy^5}{(x^2 + y^4)^2}.$$

Na origem vale:

$$\frac{\partial f}{\partial x}(0,0) = \lim_{t \to 0} \frac{f(0 + t, 0) - f(0,0)}{t} = \lim_{t \to 0} \frac{\frac{0}{t^2} - 0}{0} = 0,$$

$$\frac{\partial f}{\partial x}(0,0) = \lim_{t \to 0} \frac{f(0, 0 + t) - f(0,0)}{t} = \lim_{t \to 0} \frac{\frac{0}{t^4} - 0}{0} = 0.$$

Como acabou de ser visto, as derivadas parciais existem em todos os pontos do domínio, entretanto, a função não é contínua na origem, com efeito, para todo $t \neq 0$ segue que

$$f(t^2, t) = \frac{t^2 \cdot t}{(t^2)^2 + t^4} = \frac{t^4}{t^4 + t^4} = \frac{1}{2}$$

é constante, e isto quer dizer que ao considerar o caminho $\gamma(t) = (t^2, t)$ e calcular o limite de f ao longo de γ vale

$$\lim_{\substack{(x,y) \to (0,0) \\ (x,y) \in \gamma}} f(x,y) = \frac{1}{2} \neq f(0,0) = 0.$$

Como as derivadas parciais são casos particulares de derivada direcional. Considere um vetor $v = (\alpha, \beta)$ arbitrário de $\mathbb{R}^2$. Ao calcular a derivada de f na direção de v na origem pela definição, segue que

$$\frac{\partial f}{\partial v}(0,0) = \begin{cases} \dfrac{\beta^2}{\alpha} \ , & \text{se } \alpha \neq 0 \\ 0 \ , & \text{caso contrário} \end{cases}$$

e mesmo existindo a derivada direcional de f na direção de qualquer vetor, tal função ainda é descontínua na origem pelo mesmo motivo anterior.

Observação 3.2. Como foi tomado um caminho particular no contraexemplo 3.1, não se sabe se o limite da função existe na origem e muito menos se ele é igual a $\frac{1}{2}$, mas se tal limite existe ele é diferente de zero (o limite quando existe, é igual ao longo de qualquer caminho).

Contraexemplo 3.2. Considere a função $f : \mathbb{R}^2 \to \mathbb{R}$, definida por

$$f(x,y) = \begin{cases} \dfrac{|x|y}{\sqrt{x^2 + y^2}} \ , & \text{se } (x,y) \neq (0,0) \\ 0 \ , & \text{se } (x,y) = (0,0) \end{cases}.$$

Se for considerado $v = (\alpha, \beta)$ um vetor qualquer de $\mathbb{R}^2$, segue que

$$\begin{aligned} \frac{\partial f}{\partial v}(0,0) &= \lim_{t \to 0} \frac{f(0 + t \cdot \alpha, 0 + t \cdot \beta) - f(0,0)}{t} = \lim_{t \to 0} \frac{f(t\alpha, t\beta)}{t} = \lim_{t \to 0} \frac{\dfrac{|t\alpha| t\beta}{\sqrt{(t\alpha)^2 + (t\beta)^2}}}{t} \\ &= \lim_{t \to 0} \frac{|t||\alpha| t\beta}{\sqrt{t^2 \cdot \alpha^2 + t^2 \cdot \beta^2} \cdot t} = \lim_{t \to 0} \frac{|t||\alpha|\beta}{\sqrt{t^2 \cdot (\alpha^2 + \beta^2)}} = \lim_{t \to 0} \frac{|t||\alpha|\beta}{\sqrt{t^2} \cdot \sqrt{\alpha^2 + \beta^2}} \\ &= \lim_{t \to 0} \frac{|t||\alpha|\beta}{|t| \cdot \sqrt{\alpha^2 + \beta^2}} = \frac{|\alpha|\beta}{\sqrt{\alpha^2 + \beta^2}}. \end{aligned}$$

Perceba que, ao tomar $w = (\alpha', \beta')$, segue que

$$\begin{aligned}\frac{\partial f}{\partial(v+w)}(0,0) &= \frac{|\alpha+\alpha'|(\beta+\beta')}{\sqrt{(\alpha+\alpha')^2+(\beta+\beta')^2}} \\ &\neq \frac{|\alpha|\beta}{\sqrt{\alpha^2+\beta^2}} + \frac{|\alpha'|\beta'}{\sqrt{(\alpha')^2+(\beta')^2}} = \frac{\partial f}{\partial v}(0,0) + \frac{\partial f}{\partial w}(0,0)\end{aligned}$$

e isto quer dizer que a derivada de f não depende linearmente de v, porém mais frente será visto que isto de fato acontece se f é uma função *diferenciável.*

Como acabou de ser visto, a derivada que foi definida anteriormente sob as hipóteses mencionadas, não garante a continuidade da função, nem a dependência linear do vetor v. Isto será resolvido com a noção de *diferenciabilidade.* Será visto agora o Teorema do Valor Médio para funções de n variáveis.

Teorema 3.1.1 (Teorema do Valor Médio)**.** *Seja $f : U \subseteq \mathbb{R}^n \to \mathbb{R}$, em que U é um conjunto aberto que contenha todo o segmento de reta $[a, a+v]$ com $v \in \mathbb{R}^n$, f é contínua na restrição $f|_{[a,a+v]}$ e existe a derivada direcional $\frac{\partial f}{\partial v}(x)$ para todo $x \in (a, a+v)$. Então, existe $\theta \in (0,1)$ tal que*

$$f(a+v) - f(a) = \frac{\partial f}{\partial v}(a+\theta v).$$

Demonstração. Considere a função $\xi : [0,1] \to \mathbb{R}$, tal que $\xi(t) = f(a+tv)$. A continuidade de ξ segue da continuidade de f, e mais, $\xi(0) = f(a)$ e $\xi(1) = f(a+v)$. pelo Teorema do Valor Médio para funções de uma variável real, existe $\theta \in (0,1)$ tal que $\xi(1) - \xi(0) = \xi'(\theta)$ e vale

$$\begin{aligned}f(a+tv) - f(a) = \xi(1) - \xi(0) = \xi'(\theta) &= \lim_{t\to 0} \frac{\xi(\theta+t) - \xi(\theta)}{t} \\ &= \lim_{t\to 0} \frac{f(a+(\theta+t)v) - f(a+\theta v)}{t} \\ &= \lim_{t\to 0} \frac{f(a+\theta v+tv) - f(a+\theta v)}{t} = \frac{\partial f}{\partial v}(a+\theta v).\end{aligned}$$

□

3.2 Diferenciabilidade

No Cálculo de funções de uma variável real, é dito que uma função f é derivável em um ponto x_0 do seu domínio quando existe o limite que por definição é a derivada de f no ponto x_0 e a partir da existência da derivada são obtidas informações sobre o comportamento da f como por exemplo, a sua continuidade. No Cálculo de n variáveis, o conceito de *função derivável* é atribuído à existência das derivadas parciais, porém apenas a existência das derivadas parciais de uma função, mesmo que em todos os pontos do seu domínio, não garantem sequer a continuidade, como foi visto no exemplo 3.1.

Como muitos resultados do Cálculo dependem da continuidade das funções, então os matemáticos Fréchet[2] e Stolz[3] observaram a necessidade de construir uma extensão mais adequada de derivada para funções de n variáveis, o conceito de *função diferenciável*, este conceito é atribuído à existência de uma boa aproximação da função. Quando é falado que uma função f é *diferenciável* em um ponto x_0, está sendo afirmado basicamente que existe uma boa aproximação local de f em uma vizinhança do ponto x_0. Em outras palavras, no caso de uma variável, dada $f : I \subseteq \mathbb{R} \to \mathbb{R}$ com I aberto e um ponto $x_0 \in I$, ao dar um incremento $t \in (-\varepsilon, \varepsilon)$, com ε suficientemente pequeno de tal forma que $x_0 + t$ ainda esteja em I, então existe uma transformação linear T, tal que o valor do acréscimo $f(x_0+t)-f(x_0)$ é aproximadamente o valor de $T(t)$, isto quer dizer que para valores de t próximos de 0, o valor do acréscimo é exatamente o valor de uma transformação linear de t mais um resto (ou erro) em função de t, ou seja, $f(x_0+t)-f(x_0) = T(t)+r(t)$ e a única exigência pedida é que este resto seja infinitamente pequeno em relação ao comprimento de t. Mais à frente será explicado melhor esta exigência.

Seja $f : I \subseteq \mathbb{R} \to \mathbb{R}$, em que I é um intervalo aberto. Sejam também $x_0 \in I$ e $t \in (-\varepsilon, \varepsilon)$, com ε suficientemente pequeno tal que $x_0 + t \in I$. É dito que f é diferenciável em x_0 se existe um $a \in \mathbb{R}$ constante tal que o limite

$$\lim_{t\to 0} \frac{f(x_0+t) - f(x_0) - a \cdot t}{t} = 0 \tag{3.4}$$

exista. O numerador da fração acima é por definição o resto da aproximação em relação a t, ou seja, $r(t) = f(x_0+t) - f(x_0) - a \cdot t$, isto quer dizer que $f(x_0+t)-f(x_0) = a \cdot t + r(t)$, logo o acréscimo $f(x_0+t)-f(x_0)$ é igual a uma função linear de t mais um resto, a saber, $T(t) = a \cdot t$ é a função linear de t. Observe que quando existe um valor a que satisfaz o limite (3.4) então

$$\begin{aligned} \lim_{t\to 0} \frac{f(x_0+t) - f(x_0) - a \cdot t}{t} &= \lim_{t\to 0} \frac{f(x_0+t) - f(x_0)}{t} - a = 0 \\ &\Rightarrow f'(x_0) - a = 0 \Rightarrow f'(x_0) = a. \end{aligned}$$

Portanto, $T(t) = f'(x_0) \cdot t$, logo o acréscimo $f(x_0+t) - f(x_0) = f'(x_0) \cdot t + r(t)$, e mais $f(x_0+t)-f(x_0) = f'(x_0)\cdot t+r(t)$ implica $f(x_0+t) = f(x_0)+f'(x_0)\cdot t+r(t)$, logo, definindo uma aplicação $L(x_0+t) = f(x_0) + f'(x_0) \cdot t$, segue que o gráfico de L é justamente a f no ponto x_0 quando t varia em todo o conjunto $\mathbb{R}$. Observe que, quanto mais o valor de t se aproxima de 0, mais o valor de $f(x_0+t)$ se aproxima de $L(x_0+t)$ (veja a Figura 2).

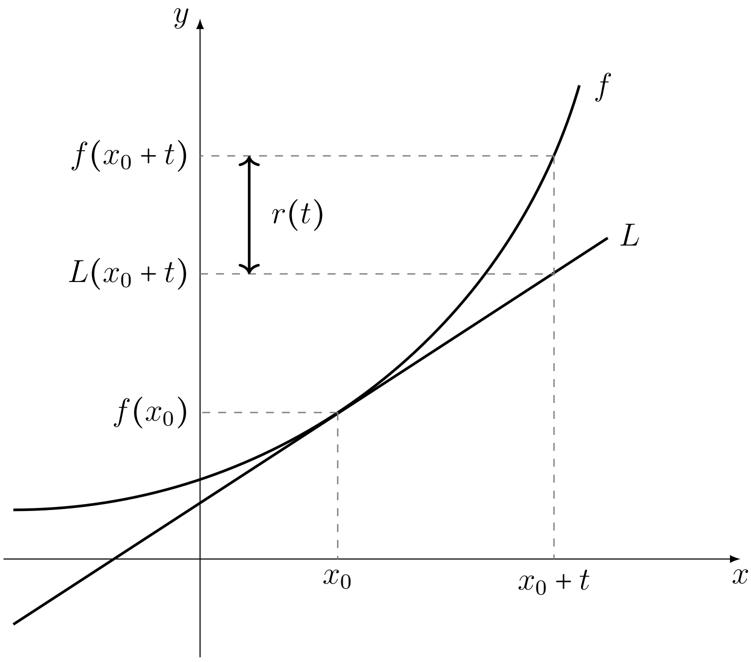
Em uma função de uma variável, a melhor aproximação local, em volta de um ponto x_0 do seu domínio, é por definição a reta tangente ao seu gráfico no ponto x_0 (GUIDORIZZI,

[2] René Maurice Fréchet (1878-1973) foi um matemático francês. Nascido em Maligny, na França, entrou para a École Normale Supérieure Paris em 1900, onde estudou matemática e doutorou-se em 1906.

[3] Otto Stolz (1842-1905), nascido em Hall, na Áustria, estudou na University of Innsbruck. Em 1864 o matemático austríaco doutorou-se na University of Vienna onde apresentou sua tese em geometria.

2001, P. 622-623). Esta informação aliada ao fato de que a é a derivada de f, permite afirmar que, no Cálculo de uma variável, dizer que uma função é diferenciável em um ponto equivale à dizer que ela é derivável no ponto em questão, e mais, a continuidade da função segue da existência de sua derivada, porém, quando f é uma função de mais de uma variável, estas noções se distanciam um pouco.

Figura 2 – Erro da aproximação pela reta tangente.



Fonte: Elaborada pelo autor.

Seja $f : U \subseteq \mathbb{R}^2 \to \mathbb{R}$, U aberto e $(x_0, y_0) \in U$, considere também um vetor $(h, k) \in \mathbb{R}^2$ de tal forma que $(x_0 + h, y_0 + k)$ ainda esteja contido em U. É dito que f é diferenciável em (x_0, y_0) quando existem constantes $a, b \in \mathbb{R}$ tais que o limite

$$\lim_{(h,k)\to(0,0)} \frac{f(x_0 + h, y_0 + k) - f(x_0, y_0) - a \cdot h - b \cdot k}{\|(h, k)\|} = 0 \tag{3.5}$$

exista. O numerador da fração (3.5) é por definição, o resto (ou erro) da aproximação, ou seja, $r(h, k) = f(x_0 + h, y_0 + k) - f(x_0, y_0) - a \cdot h - b \cdot k$, isto quer dizer que $f(x_0 + h, y_0 + k) - f(x_0, y_0) = a \cdot h + b \cdot k + r(h, k)$, logo o acréscimo $f(x_0 + h, y_0 + k) - f(x_0, y_0)$ é igual a uma função linear de (h, k), mais um resto em relação ao vetor (h, k), a saber, $a \cdot h + b \cdot k$ é a função linear. Na Definição 3.2 será visto que, quando os valores a e b existirem, então $a = \frac{\partial f}{\partial x}(x_0, y_0)$ e $b = \frac{\partial f}{\partial y}(x_0, y_0)$. Portanto, $f(x_0 + h, y_0 + k) - f(x_0, y_0) = a \cdot h + b \cdot k + r(h, k)$ implica $f(x_0 + h, y_0 + k) - f(x_0, y_0) = \frac{\partial f}{\partial x}(x_0, y_0) \cdot h + \frac{\partial f}{\partial y}(x_0, y_0) \cdot k + r(h, k)$ acarreta $f(x_0 + h, y_0 + k) = f(x_0, y_0) + \frac{\partial f}{\partial x}(x_0, y_0) \cdot h + \frac{\partial f}{\partial y}(x_0, y_0) \cdot k + r(h, k)$, logo definindo uma aplicação $L(x_0 + h, y_0 + k) = f(x_0, y_0) + \frac{\partial f}{\partial x}(x_0, y_0) \cdot h + \frac{\partial f}{\partial y}(x_0, y_0) \cdot k$, segue que o gráfico de L é justamente a equação do plano tangente[4] à superfície de f no ponto (x_0, y_0), basta fazer $h = x - x_0$ e $k = y - y_0$ (STEWART, 2013, p. 823).

Em uma função com domínio em $\mathbb{R}^2$, o plano tangente à sua superfície é uma boa aproximação em torno do ponto de tangência (STEWART, 2013, P. 826).

Mais geralmente, é dito que:

Definição 3.2. Dada uma função $f : U \subseteq \mathbb{R}^n \to \mathbb{R}$, com U aberto e $a \in U$, f é dita *diferenciável* no ponto a, quando para todo vetor $v = (\alpha_1, \alpha_2, \alpha_3, \ldots, \alpha_n) \in \mathbb{R}^n$ existem constantes $A_1, A_2, A_3, \ldots, A_n \in \mathbb{R}$ tais que o limite

$$\lim_{v \to 0} \frac{f(a+v) - f(a) - \sum_{i=1}^{n} A_i \cdot \alpha_i}{\|v\|} = 0$$

exista. Quando f é diferenciável em todos os pontos do domínio, é dito que f é *diferenciável.*

Por definição, $r(v) = f(a+v) - f(a) - \sum_{i=1}^{n} A_i \cdot \alpha_i$ logo, se for tomado $v = t \cdot e_i$ com $t \in \mathbb{R}$, como $\|e_i\| = 1$, segue que

$$\begin{aligned}
\lim_{t \to 0} \left| \frac{r(t \cdot e_i)}{\|t \cdot e_i\|} \right| &= \lim_{t \to 0} \left| \frac{r(t \cdot e_i)}{|t|} \right| \\
&= \lim_{t \to 0} \left| \frac{r(t \cdot e_i)}{t} \right| \\
&= \lim_{t \to 0} \left| \frac{f(a + t \cdot e_i) - f(a) - A_i \cdot t}{t} \right| \\
&= \lim_{t \to 0} \left| \frac{f(a + t \cdot e_i) - f(a)}{t} - A_i \right| \\
&= \left| \frac{\partial f}{\partial x_i}(a) - A_i \right| = 0 \Leftrightarrow \frac{\partial f}{\partial x_i}(a) - A_i = 0 \Leftrightarrow \frac{\partial f}{\partial x_i}(a) = A_i.
\end{aligned}$$

Observação 3.3. Perceba que, das normas que foram apresentadas, independentemente da norma que for adotada, sempre vai valer que a norma do vetor $t \cdot e_i$ é igual ao valor absoluto de t, pois o t é um escalar e o vetor e_i possui a i-ésima coordenada igual a 1 e as demais iguais a 0, isto quer dizer que a norma de e_i é igual a 1 para todo $i \in \{1, 2, 3, \ldots, n\}$, em particular, na norma euclidiana vale $\|t \cdot e_i\| = |t| \cdot \|e_i\| = |t| \cdot 1 = |t|$. Da definição do resto, $r(v)$ é uma função real, pois é uma soma de valores reais e portanto faz sentido $|r(v)|$.

Nas condições da Definição 3.2, dizer que o limite da fração tende a 0 quando v tende a 0, significa dizer que o numerador tende a 0 mais rapidamente que o denominador "$\|v\|$", em outras palavras, o resto $r(v)$ é infinitamente pequeno em relação à norma de v,

[4]Dada uma aplicação $f : \mathbb{R}^2 \to \mathbb{R}$ e um ponto (x_0, y_0) do domínio de f, o plano tangente à superfície de f no ponto (x_0, y_0) é o plano que contêm todas as retas tangentes à superfície de f, que passam pelo ponto (x_0, y_0).

portanto o limite do resto de v tende a 0 quando $v \to 0$, pois $\lim_{v\to 0} \dfrac{r(v)}{\|v\|} = \lim_{v\to 0} \dfrac{r(v)}{\|v\|} \cdot \|v\| = 0$, logo

$$\begin{aligned}
\lim_{v\to 0} f(a+v) - f(a) &= \left[\lim_{v\to 0} f(a+v) - f(a)\right] - 0 \\
&= \left[\lim_{v\to 0} f(a+v) - f(a)\right] - \lim_{v\to 0} \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(a)\cdot \alpha_i \\
&= \lim_{v\to 0} f(a+v) - f(a) - \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(a)\cdot \alpha_i = \lim_{v\to 0} r(v) = 0.
\end{aligned}$$

Destas igualdades, $\lim_{v\to 0} f(a+v) - f(a) = 0$ equivale à dizer que $f(a+v)$ tende à $f(a)$ quando v tende a 0, daí segue que se uma função é diferenciável em um ponto, então ela é contínua neste ponto. Perceba que $r(v) = f(a+v) - f(a) - \sum_{i=1}^{n} A_i \cdot \alpha_i$ se, e somente se, $f(a+v) - f(a) = \sum_{i=1}^{n} A_i \cdot \alpha_i + r(v)$, isto quer dizer que o acréscimo $f(a+v) - f(a)$ é igual à uma função linear de v mais um resto infinitamente pequeno em relação á norma de v. Não somente isto, esta definição de diferenciabilidade ainda garante que a derivada direcional de uma função f na direção de qualquer vetor v, depende linearmente do mesmo e vale

$$\frac{\partial f}{\partial v}(a) = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(a)\cdot \alpha_i.$$

Com efeito, seja $f : U \to \mathbb{R}$, a um ponto do aberto $U \subseteq \mathbb{R}^n$, $v \in \mathbb{R}^n$ e f diferenciável em a. Como U é aberto, então faz sentido considerar a imagem do ponto $a + tv$, para todo t suficientemente pequeno, pois $f(a+tv) \in U$, portanto:

$$\begin{aligned}
f(a+tv) - f(a) &= f(a+tv) - f(a) - \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(a)\cdot t\alpha_i + \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(a)\cdot t\alpha_i \\
\Leftrightarrow f(a+tv) - f(a) &= \frac{\left(f(a+tv) - f(a) - \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(a)\cdot t\alpha_i\right)\cdot \|v\|}{\|v\|} + \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(a)\cdot t\alpha_i \\
\Leftrightarrow \frac{f(a+tv) - f(a)}{t} &= \frac{\left(f(a+tv) - f(a) - \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(a)\cdot t\alpha_i\right)\cdot \|v\|}{t\cdot\|v\|} + \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(a)\cdot \alpha_i \\
\Leftrightarrow \frac{f(a+tv) - f(a)}{t} &= \pm\left(\frac{\left(f(a+tv) - f(a) - \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(a)\cdot t\alpha_i\right)}{|t|\cdot\|v\|}\cdot \|v\|\right) + \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(a)\cdot \alpha_i \\
\Leftrightarrow \frac{f(a+tv) - f(a)}{t} &= \pm\left(\frac{\left(f(a+tv) - f(a) - \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(a)\cdot t\alpha_i\right)}{\|t\cdot v\|}\cdot \|v\|\right) + \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(a)\cdot \alpha_i,
\end{aligned}$$

passando o limite em ambos os lados da última igualdade, segue que

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

e

$$\lim_{t\to 0}\frac{\left(f(a+tv)-f(a)-\sum_{i=1}^{n}\frac{\partial f}{\partial x_i}(a)\cdot t\alpha_i\right)}{\|t\cdot v\|}=\lim_{t\to 0}\frac{r(t\cdot v)}{\|t\cdot v\|}=0,$$

daí resulta a fórmula enunciada.

Estas últimas conclusões, da continuidade de f e da dependência linear da derivada direcional do vetor v são o ponto chave desda noção de diferenciabilidade, esta definição foi dada pelos matemáticos Fréchet e Stolz.

Algumas vezes, é conveniente considerar uma função $\rho(v)=\frac{r(v)}{\|v\|}$, em que ρ é contínua em $v=\overrightarrow{0}$ e se anula neste ponto, em outras palavras $\lim_{v\to 0}\rho(v)=0$.

Teorema 3.2.1 (Regra da Cadeia)**.** *Sejam $U\subseteq\mathbb{R}^m$ e $V\subseteq\mathbb{R}^n$ conjuntos abertos, $f=(f_1,f_2,f_3,\ldots,f_n):U\to V$, a um ponto de U tal $f(a)=b\in V$ e cada função coordenada $f_k:U\to\mathbb{R}$ é diferenciável no ponto a. Seja $g:V\to\mathbb{R}$ uma função diferenciável no ponto $b=f(a)$, então a composição $g\circ f:U\to\mathbb{R}$ é diferenciável em a e vale a fórmula*

$$\frac{\partial(g\circ f)}{\partial x_i}(a)=\sum_{k=1}^{n}\frac{\partial g}{\partial y_k}(b)\cdot\frac{\partial f_k}{\partial x_i}(a).$$

Demonstração. Considere U_0, o conjunto de todos os vetores de $\mathbb{R}^m$ tais que $a+v\in U$, considere também $v=(\alpha_1,\alpha_2,\alpha_3,\ldots,\alpha_m)$ e perceba que utilizando a definição de diferenciabilidade em f no ponto a, com a norma da soma, segue que

$$\begin{aligned}f(a+v)&=f(a)+\sum_{i=1}^{m}\frac{\partial f}{\partial x_i}(a)\cdot\alpha_i+\rho(v)\cdot\|v\|_S\\ \Rightarrow\quad f_k(a+v)&=f_k(a)+\sum_{i=1}^{m}\frac{\partial f_k}{\partial x_i}(a)\cdot\alpha_i+\rho_k(v)\cdot\|v\|_S\end{aligned}\tag{3.6}$$

em que $\rho(v)$, assim como cada uma de suas funções coordenadas $\rho_k(v)$, são contínuas em $v=0$ e são todas contínuas neste ponto. Perceba que, da forma como o conjunto U_0 foi tomado, nele pode ser definida uma aplicação $w:U_0\to\mathbb{R}^n$ com $w(v)=(\beta_1(v),\beta_2(v),\beta_3(v),\ldots,\beta_n(v))$ em que suas funções coordenada são definidas por $\beta_k(v)=\sum_{i=1}^{m}\frac{\partial f_k}{\partial x_i}(a)\cdot\alpha_i+\rho_k(v)\cdot\|v\|_S$. Definindo w desta forma, segue que

$$\begin{aligned}f_k(a+v)&=f_k(a)+\beta_k(v)\\ \Rightarrow\quad f(a+v)&=b+w\end{aligned}\tag{3.7}$$

e mais, como $\frac{|\alpha_i|}{\|v\|_S}\leq 1$, então cada $\frac{\|\beta_k(v)\|_S}{\|v\|_S}$ é limitado em uma vizinhança de 0 e portanto $\dfrac{\|w(v)\|_S}{\|v\|_S}$ também o é.

Utilizando agora a definição de diferenciabilidade em g no ponto $b = f(a)$, segue que para todo $v \in U_0$, vale

$$\begin{aligned}(g \circ f)(a+v) = g(f(a+v)) = g(b+w) &= g(b) + \sum_{k=1}^{n} \frac{\partial g}{\partial y_k}(b) \cdot \beta_k + \theta(w) \cdot \|w\|_S \\ &= (g \circ f)(a) + \sum_{k=1}^{n} \frac{\partial g}{\partial y_k}(b) \cdot \beta_k + \theta(w) \cdot \|w\|_S, \quad (3.8)\end{aligned}$$

em que $\theta(w)$ é contínua em 0 e se anula neste ponto. Perceba que $w(v)$ também se anula em 0, pela forma como foi definida, portanto $v \to 0$ implica $w(v) \to 0$ acarreta $\theta(w) \to 0$. Utilizando a definição de β_k em (3.8), vale

$$(g \circ f)(a+v) = (g \circ f)(a) + \sum_{k=1}^{n} \frac{\partial g}{\partial y_k}(b) \cdot \left(\sum_{i=1}^{m} \frac{\partial f_k}{\partial x_i}(a) \cdot \alpha_i + \rho_k(v) \cdot \|v\|_S \right) + \theta(w) \cdot \|w\|_S, \quad (3.9)$$

perceba que

$$\begin{aligned}&\sum_{k=1}^{n} \frac{\partial g}{\partial y_k}(b) \cdot \left(\sum_{i=1}^{m} \frac{\partial f_k}{\partial x_i}(a) \cdot \alpha_i + \rho_k(v) \cdot \|v\|_S \right) + \theta(w) \cdot \|w\|_S \\ &= \sum_{i=1}^{m} \left(\sum_{k=1}^{n} \frac{\partial g}{\partial y_k}(b) \cdot \frac{\partial f_k}{\partial x_i}(a) \right) \cdot \alpha_i + \sum_{k=1}^{n} \frac{\partial g}{\partial y_k}(b) \cdot \rho_k(v) \cdot \|v\|_S + \theta(w) \cdot \|w\|_S,\end{aligned}$$

logo ao chamar $\sum_{k=1}^{n} \frac{\partial g}{\partial y_k}(b) \cdot \frac{\partial f_k}{\partial x_i}(a) = A_i$ e $\sum_{k=1}^{n} \frac{\partial g}{\partial y_k}(b) \cdot \rho_k(v) \cdot \|v\|_S + \theta(w) \cdot \|w\|_S = R$, e substituir A_i e R em (3.9), vai valer

$$(g \circ f)(a+v) = (g \circ f)(a) + \sum_{i=1}^{m} A_i \cdot \alpha_i + R$$

em que, pelos resultados obtidos sobre $\theta(v)$ e sobre $\dfrac{\|w\|_S}{\|v\|_S}$, segue que $\lim\limits_{v \to 0} \dfrac{R}{\|v\|_S} = 0$. Portanto a composição $g \circ f$ é diferenciável no ponto a e suas derivadas parciais são os A_i. □

Quando uma função $f : U \to \mathbb{R}$ é contínua no aberto $U \subseteq \mathbb{R}^n$ é dito que f é de *classe* C^0; Quando existem e são contínuas as derivadas parciais $\dfrac{\partial f}{\partial x_i} : U \to \mathbb{R}$ em todos os pontos $x \in U$, é dito que f é de *classe* C^1; f diz-se de *classe* C^2 quando as suas n derivadas parciais são de *classe* C^1; Generalizando, é dito que f é de *classe* C^k para algum $k \in \mathbb{N}$, quando suas n derivadas parciais forem de *classe* C^{k-1}, se isto ocorre para todo $k \geq 0$, então f é de *classe* C^∞.

Teorema 3.2.2. *Seja $f : U \subseteq \mathbb{R}^n \to \mathbb{R}$, em que U é um aberto. Se existem as derivadas parciais em todos os pontos de U e todas são contínuas em um ponto c, então f é diferenciável em c.*

Demonstração. Tomando $c = (c_1, c_2, c_3, \ldots, c_n) \in U$ tal que todas as $\dfrac{\partial f}{\partial x_i}$ são contínuas neste ponto e tomando $v = (\alpha_1, \alpha_2, \alpha_3, \ldots, \alpha_n) \in \mathbb{R}^n$ tal que $c + v \in U$. Utilizando a definição de diferenciabilidade em f no ponto c, segue que

$$
\begin{aligned}
r(v) = \quad & f(c+v) \quad -f(c) - \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(c) \cdot \alpha_i \\
= \quad & f(c+v) \quad -f\left(c + \sum_{k=2}^{n} e_k \cdot \alpha_k\right) + f\left(c + \sum_{k=2}^{n} e_k \cdot \alpha_k\right) \\
& \qquad\qquad\; -f\left(c + \sum_{k=3}^{n} e_k \cdot \alpha_k\right) + f\left(c + \sum_{k=3}^{n} e_k \cdot \alpha_k\right) \\
& \qquad\qquad\; -f\left(c + \sum_{k=4}^{n} e_k \cdot \alpha_k\right) + f\left(c + \sum_{k=4}^{n} e_k \cdot \alpha_k\right) - \cdots \\
& \qquad\qquad\; -f(c + e_n \cdot \alpha_n) + f(c + e_n \cdot \alpha_n) - f(c) - \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(c) \cdot \alpha_i.
\end{aligned}
$$

Tomar estes valores de f que foram adicionados e depois subtraídos possibilita a utilização do Teorema do Valor Médio para funções de uma variável real, segundo o qual existem $\theta_1, \theta_2, \theta_3, \ldots, \theta_n \in (0,1)$ tais que:

$$
\begin{aligned}
f(c+v) - f\left(c + \sum_{k=2}^{n} e_k \cdot \alpha_k\right) =& f(c_1 + \alpha_1, c_2 + \alpha_2, c_3 + \alpha_3, \ldots, c_n + \alpha_n) \\
& - f(c_1, c_2 + \alpha_2, c_3 + \alpha_3, \ldots, c_n + \alpha_n) \\
=& \frac{\partial f}{\partial x_1}(c + \theta_1 \cdot e_1 \cdot \alpha_1) \cdot \alpha_1
\end{aligned} \tag{3.10}
$$

$$
\begin{aligned}
f\left(c + \sum_{k=2}^{n} e_k \cdot \alpha_k\right) - f\left(c + \sum_{k=3}^{n} e_k \cdot \alpha_k\right) =& f(c_1, c_2 + \alpha_2, c_3 + \alpha_3, \ldots, c_n + \alpha_n) \\
& - f(c_1, c_2, c_3 + \alpha_3, \ldots, c_n + \alpha_n) \\
=& \frac{\partial f}{\partial x_2}(c + \theta_2 \cdot e_2 \cdot \alpha_2) \cdot \alpha_2
\end{aligned} \tag{3.11}
$$

$$
\begin{aligned}
f\left(c + \sum_{k=3}^{n} e_k \cdot \alpha_k\right) - f\left(c + \sum_{k=4}^{n} e_k \cdot \alpha_k\right) =& f(c_1, c_2, c_3 + \alpha_3, \ldots, c_n + \alpha_n) \\
& - f(c_1, c_2, c_3, c_4 + \alpha_4, \ldots, c_n + \alpha_n) \\
=& \frac{\partial f}{\partial x_3}(c + \theta_3 \cdot e_3 \cdot \alpha_3) \cdot \alpha_3
\end{aligned} \tag{3.12}
$$

$\vdots$

$$f(c+e_n\cdot\alpha_n)-f(c)=f(c_1,c_2,\dots,c_{n-1},c_n+\alpha_n)-f(c_1,c_2,\dots,c_{n-1},c_n+\alpha_n)$$
$$=\frac{\partial f}{\partial x_n}(c+\theta_n\cdot e_n\cdot\alpha_n)\cdot\alpha_n. \tag{3.13}$$

Portanto, substituindo (3.10), (3.11), (3.12) e (3.13) em $r(v)$, segue que

$$\begin{aligned}r(v)=&\frac{\partial f}{\partial x_1}(c+\theta_1\cdot e_1\cdot\alpha_1)\cdot\alpha_1+\frac{\partial f}{\partial x_2}(c+\theta_2\cdot e_2\cdot\alpha_2)\cdot\alpha_2+\frac{\partial f}{\partial x_3}(c+\theta_3\cdot e_3\cdot\alpha_3)\cdot\alpha_3\\&+\cdots+\frac{\partial f}{\partial x_n}(c+\theta_n\cdot e_n\cdot\alpha_n)\cdot\alpha_n-\sum_{i=1}^{n}\frac{\partial f}{\partial x_i}(c)\cdot\alpha_i\\=&\left[\frac{\partial f}{\partial x_1}(c+\theta_1\cdot e_1\cdot\alpha_1)-\frac{\partial f}{\partial x_1}(c)\right]\cdot\alpha_1+\cdots+\left[\frac{\partial f}{\partial x_n}(c+\theta_n\cdot e_n\cdot\alpha_n)-\frac{\partial f}{\partial x_n}(c)\right]\cdot\alpha_n.\end{aligned}$$

Como as derivadas parciais de f são contínuas em c e cada quociente $\frac{\alpha_i}{\|v\|}\le 1$, ou seja, é limitado independentemente da norma que for adotada (entre as normas que foram apresentadas), então $\lim_{v\to 0}\frac{r(v)}{\|v\|}=0$ e portanto f é diferenciável em c.

□

Corolário 3.1. *Para que f seja diferenciável é suficiente que $f\in C^1$.*

Perceba que

$f\in C^1\Rightarrow\frac{\partial f}{\partial x_i}\in C^0$ e $f\in C^0$,

$f\in C^2\Rightarrow\frac{\partial f}{\partial x_i}\in C^1\Rightarrow\frac{\partial f}{\partial x_i}\in C^0\Rightarrow f\in C^1\Rightarrow f\in C^0$,

$f\in C^3\Rightarrow\frac{\partial f}{\partial x_i}\in C^2\Rightarrow\frac{\partial f}{\partial x_i}\in C^1\Rightarrow f\in C^2\Rightarrow f\in C^1\Rightarrow f\in C^0$,

segue que

$$f\in C^k\Rightarrow f\in C^{k-1}\Rightarrow\cdots\Rightarrow f\in C^1\Rightarrow f\in C^0,$$

logo

$$C^0\supset C^1\supset C^2\supset\cdots\supset C^k.$$

Corolário 3.2. *Sejam $U\subseteq\mathbb{R}^m$ e $V\subseteq\mathbb{R}^n$ abertos, $f=(f_1,f_2,f_3,\dots,f_n):U\to\mathbb{R}^n$, com $f(U)\subset V$ e cada função coordenada $f_k:U\to\mathbb{R}$ é de classe C^k. Seja $g:V\to\mathbb{R}$ uma função de classe C^k. Então, a composição $g\circ f:U\to\mathbb{R}$ é de classe C^k.*

Demonstração. Para o caso $k=0$, o resultado segue como consequência do Teorema 2.2.1, então fica provado para $k=0$. Se $k\ge 1$, então as hipóteses acima permitem que seja utilizada a Regra da Cadeia na composição $g\circ f$, o que garante que tal composição possui todas as derivadas parciais, logo para todo $x\in U$ com $f(x)\in V$, vale

$$\frac{\partial f\circ g}{\partial x_i}=\sum_{j=1}^{n}\frac{\partial g}{\partial y_j}(f(x))\cdot\frac{\partial f_i}{\partial x_i}(x)=\sum_{j=1}^{n}\left(\frac{\partial g}{\partial y_j}\circ f\right)(x)\cdot\frac{\partial f_i}{\partial x_i}(x). \tag{3.14}$$

Por indução, suponha que o Teorema seja válido para a classe C^{k-1}, a partir daí será provado que vale para C^k. Com efeito, como $g \in C^k$, então todas as suas derivadas parciais são de classe C^{k-1}, como $f \in C^k$, em particular $f \in C^{k-1}$, pois $C^{k-1} \subset C^k$, então a composição $\dfrac{\partial g}{\partial x_i} \circ f$ ainda é de classe C^{k-1}, por hipótese de indução. Como cada $\dfrac{\partial f_j}{\partial x_i} \in C^k$, em particular $\dfrac{\partial f_j}{\partial x_i} \in C^{k-1}$, então $\left(\dfrac{\partial g}{\partial y_j} \circ f\right) \cdot \dfrac{\partial f_j}{\partial x_i} \in C^{k-1}$, pois será visto mais à frente que, o produto de funções de classe C^{k-1} ainda é desta classe, logo de (3.14) segue que todas as derivadas parciais da composição $g \circ f$ são de classe C^{k-1}, ou seja $g \circ f \in C^k$. □

Para funções diferenciáveis em $U \subset \mathbb{R}^n$, fixando um ponto de uma f, a derivada direcional de f neste ponto depende linearmente de v para qualquer $v \in \mathbb{R}^n$, e neste sentido, é conveniente que exista uma aplicação linear que, para cada ponto $a \in U$ fixado, a derivada de f no ponto a assuma valores reais quando aplicada nos vetores $v \in \mathbb{R}^n$. Será definida tal aplicação a seguir. Considere a aplicação $df : U \to (\mathbb{R}^n)^*$, em que tal aplicação associa a cada ponto a do domínio de f, um funcional linear $df(a) \in (\mathbb{R}^n)^*$.

Definição 3.3. Seja $f : U \subseteq \mathbb{R}^n \to \mathbb{R}$, em que U é um aberto, diferenciável em $a \in U$ e $v = (\alpha_1, \alpha_2, \alpha_3, \ldots, \alpha_n) \in \mathbb{R}^n$. O valor da *diferencial* de f no ponto a, aplicada no vetor v é dado por

$$df(a) : \mathbb{R}^n \to \mathbb{R}$$
$$df(a) \cdot v = \frac{\partial f}{\partial v}(a) = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(a) \cdot \alpha_i.$$

Para que df seja contínua é necessário e suficiente que $f \in C^1$ e além disso, a diferencial admite as fórmulas a seguir.

Teorema 3.2.3. *Sejam $f, g : U \to \mathbb{R}$ e a um ponto de U tal que f e g são diferenciáveis em a, segue*

1. $f + g$ *é diferenciável e* $d(f + g) = df + dg$*;*
2. $f \cdot g$ *é diferenciável e* $d(f \cdot g) = f \cdot dg + g \cdot df$*;*
3. $\dfrac{f}{g}$ *é diferenciável e* $d\left(\dfrac{f}{g}\right) = \dfrac{(g \cdot df - f \cdot dg)}{g^2}$*, desde que* $g(x) \neq 0$ *para todo* $x \in U$*.*

Demonstração. Considere as funções $s, m : \mathbb{R}^2 \to \mathbb{R}$ com $s(x, y) = x + y$, $m(x, y) = x \cdot y$ e a função $q : \mathbb{R} \times \mathbb{R} - \{0\} \to \mathbb{R}$ com $q(x, y) = \dfrac{x}{y}$. Utilizando as regras de derivação do Cálculo, as funções s, m e q possuem todas as derivadas parciais e todas são contínuas em seus respectivos domínios, logo as três funções são de classe C^1, ou seja, diferenciáveis. Sejam $f, g : U \subseteq \mathbb{R}^n \to \mathbb{R}$ funções diferenciáveis, então a aplicação $F : U \to \mathbb{R}^2$ definida por $F(x) = (f(x), g(x))$ possui coordenadas diferenciáveis, portanto a Regra da Cadeia garante que as composições $s \circ F$, $m \circ F$ e $q \circ F$ são diferenciáveis e ainda fornece uma fórmula para calcular suas derivadas parciais. Perceba que $s \circ F(x) = s(F(x)) = s(f(x), g(x)) = f(x) + g(x)$, com raciocínio análogo $m \circ F = f \cdot g$ e $q \circ F = \dfrac{f}{g}$, logo as funções $f+g$, $f \cdot g$ e $\dfrac{f}{g}$ são todas diferenciáveis, e mais

$$\frac{\partial(f+g)}{\partial x_i} = \frac{\partial(s \circ F)}{\partial x_i} = \frac{\partial s}{\partial f} \cdot \frac{\partial F_1}{\partial x_i} + \frac{\partial s}{\partial g} \cdot \frac{\partial F_2}{\partial x_i} = 1 \cdot \frac{\partial F_1}{\partial x_i} + 1 \cdot \frac{\partial F_2}{\partial x_i} = \frac{\partial f}{\partial x_i} + \frac{\partial g}{\partial x_i},$$

analogamente

$$\frac{\partial(f \cdot g)}{\partial x_i} = \frac{\partial(m \circ F)}{\partial x_i} = g \cdot \frac{\partial f}{\partial x_i} + f \cdot \frac{\partial g}{\partial x_i} \quad \text{e} \quad \frac{\partial(f/g)}{\partial x_i} = \frac{\partial(q \circ F)}{\partial x_i} = \frac{g \cdot \dfrac{\partial f}{\partial x_i} - f \cdot \dfrac{\partial g}{\partial x_i}}{g^2},$$

daí segue que

$$\begin{aligned} d(f+g) \cdot v &= \sum_{i=1}^{n} \frac{\partial(f+g)}{\partial x_i} \cdot \alpha_i = \sum_{i=1}^{n} \left(\frac{\partial f}{\partial x_i} + \frac{\partial g}{\partial x_i} \right) \cdot \alpha_i = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i} \cdot \alpha_i + \sum_{i=1}^{n} \frac{\partial g}{\partial x_i} \cdot \alpha_i \\ &= df \cdot v + dg \cdot v \;\Rightarrow\; d(f+g) = df + dg \end{aligned}$$

e com raciocínio análogo segue o restante das fórmulas enunciadas. □

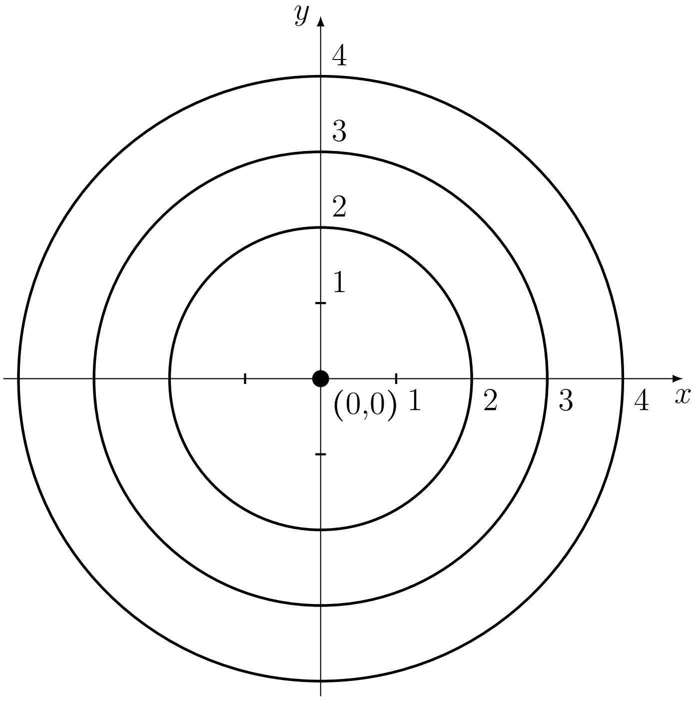
Uma consequência direta do Teorema 3.2.3 é que, se f e g são de classe C^1, então as derivadas parciais de ambas são contínuas, e mais, o Teorema 2.2.3 garante que a soma, o produto e a razão (em seu domínio) de funções contínuas ainda o são, logo do Teorema 3.2.3 segue que as derivadas parciais das funções $f+g$, $f \cdot g$ e f/g são todas contínuas, logo a soma, o produto e a razão de funções de classe C^1 ainda é C^1. Com raciocínio análogo, se f e g são de classe C^2, então suas derivadas parciais são todas de classe C^1, segue que que as derivadas parciais das funções $f+g$, $f \cdot g$ e f/g são todas de classe C^1, ou seja, $f+g$, $f \cdot g$ e f/g são todas de classe C^2. Generalizando, se f e g são de classe C^k, então $f+g$, $f \cdot g$ e f/g são todas de classe C^k.

Definição 3.4. Seja $f : U \subseteq \mathbb{R}^n \to \mathbb{R}$, em que U é um aberto, o conjunto dos pontos $x \in U$ tais que $f(x) = c$, ou seja, a imagem inversa de f no ponto c, $f^{-1}(c)$, é chamado o *conjunto de nível* c de $f(x)$.

Se $n = 2$, tal conjunto é chamado de *curva de nível* c de f, muito embora a imagem inversa de f no ponto c não seja necessariamente uma curva, já se $n = 3$, tal conjunto é chamado de *superfície de nível*.

Exemplo 3.1. Seja $f : \mathbb{R}^2 \to \mathbb{R}$, definida por $f(x,y) = x^2 + y^2$. Como f é a soma de dois quadrados, então a função nunca assumirá valores negativos, ou seja, sua imagem é o intervalo $[0, \infty)$. Portanto o conjunto de nível de f de nível c é o conjunto vazio para todo $c < 0$. Para $c = 0$, o único ponto de $\mathbb{R}^2$ que satisfaz a equação $x^2 + y^2 = 0$ é a origem. Já para $c > 0$, $x^2 + y^2 = c$ é a equação de uma circunferência de centro na origem e raio igual a $\sqrt{c}$, em outras palavras, o conjunto de nível de f de nível c para todo $c > 0$ é o conjunto de todos os pontos de $\mathbb{R}^2$ cuja distância até a origem mede $\sqrt{c}$.

Figura 3 – Curvas de nível 0, 4, 9 e 16 de $f(x,y) = x^2 + y^2$.



Fonte: Elaborada pelo autor.

Exemplo 3.2. Seja $k : \mathbb{R}^2 \to \mathbb{R}$, definida por $k(x,y) = \sqrt{x^2+y^2}$. Note que k tem as mesmas curvas de nível da aplicação f do exemplo 3.1, porém a curva de nível c de $k(x,y)$ equivale à curva de nível c^2 de $f(x,y)$.

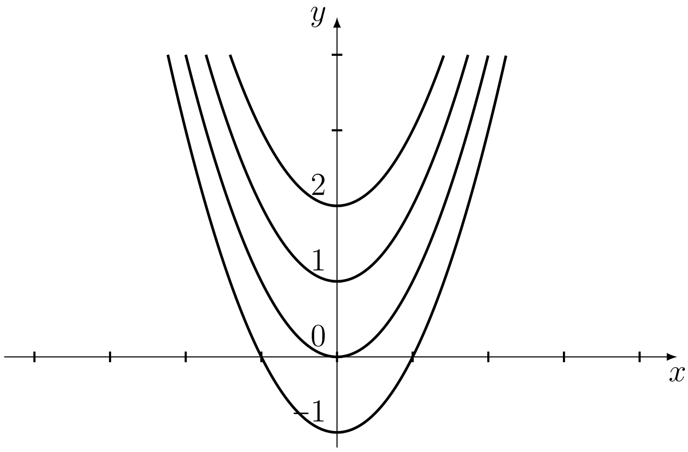
Exemplo 3.3. Seja $g : \mathbb{R}^2 \to \mathbb{R}$, definida por $g(x,y) = y - x^2$. Como não há nenhuma restrição na função, então a sua imagem é toda a reta. Para todo $c \in \mathbb{R}$, segue que

$$g(x,y) = c \;\Rightarrow\; y = x^2 + c, \tag{3.15}$$

conclui-se então que as curvas de nível de g são parábolas com concavidade para o lado positivo do eixo O_y, e mais, fazendo $x = 0$ na equação (3.15), segue que $y = 0^2 + c$ implica $y = c$, logo as parábolas intersectam o eixo O_y no ponto $(0,c)$ para cada $c \in \mathbb{R}$. Para os níveis $-1, 0, 1$ e 2, vale:

- $c = -1 \;\Rightarrow\; y = x^2 - 1$;
- $c = 0 \;\Rightarrow\; y = x^2$;
- $c = +1 \;\Rightarrow\; y = x^2 + 1$;
- $c = +2 \;\Rightarrow\; y = x^2 + 2$.

Figura 4 – Curvas de nível de $g(x,y) = y - x^2$.



Fonte: Elaborada pelo autor.

Exemplo 3.4. Seja $h : \mathbb{R}^2 \to \mathbb{R}$, definida por $h(x,y) = 4x^2 + 9y^2$. Note que esta função não está definida para valores reais negativos. Se $h(x,y) = 0$, então o único ponto que satisfaz a equação $4x^2 + 9y^2 = 0$ é a origem, logo o conjunto de nível de nível $c = 0$ de h é $\{(0,0)\}$. Dos conhecimentos do Cálculo Vetorial, uma equação do tipo $4x^2 + 9y^2 = c$ caracteriza uma elipse e portanto, para os níveis $1, 9$ e 16, segue que:

- $c = 1 \Rightarrow 4x^2 + 9y^2 = 1$:
  Fazendo $x = 0$ na equação acima, $4 \cdot 0^2 + 9y^2 = 1$ implica $y = \pm\dfrac{1}{3}$, já se $y = 0$, então $4x^2 + 9 \cdot 0^2 = 1$ acarreta $x = \pm\dfrac{1}{2}$, isto quer dizer que a elipse intersecta o eixo O_y nos pontos $(0, -\dfrac{1}{3})$, $(0, \dfrac{1}{3})$ e o eixo O_x nos pontos $(-\dfrac{1}{2}, 0)$, $(\dfrac{1}{2}, 0)$.
- $c = 9 \Rightarrow 4x^2 + 9y^2 = 9$:
  Para $x = 0$ na equação acima, $4 \cdot 0^2 + 9y^2 = 9$ implica $y = \pm 1$ e para $y = 0$, $4x^2 + 9 \cdot 0^2 = 9$ acarreta $x = \pm\dfrac{3}{2}$, logo a elipse intersecta o eixo O_y nos pontos $(0, -1)$, $(0, 1)$ e o eixo O_x nos pontos $(-\dfrac{3}{2}, 0)$, $(\dfrac{3}{2}, 0)$.
- $c = 16 \Rightarrow 4x^2 + 9y^2 = 16$:
  Se $x = 0$ na equação acima, $4 \cdot 0^2 + 9y^2 = 16$ implica $y = \pm\dfrac{4}{3}$ e se $y = 0$, $4x^2 + 9 \cdot 0^2 = 9$ acarreta $x = \pm 2$, portanto a elipse intersecta o eixo O_y nos pontos $(0, -\dfrac{4}{3})$, $(0, \dfrac{4}{3})$ e o eixo O_x nos pontos $(-2, 0)$, $(2, 0)$.

Exemplo 3.5. Seja $j : \mathbb{R}^3 \to \mathbb{R}$, com $j(x,y,z) = x + 2y + 4z$. Como esta função não possui nenhuma restrição, então sua imagem é toda a reta e portanto, as superfícies de nível de j são definidas pela equação $x + 2y + 4z = c$, para todo $c \in \mathbb{R}$. Bem, esta equação caracteriza planos, basta encontrar os pontos de intersecção do plano com os eixos O_x, O_y e O_z para poder esboçar seu gráfico. Fazendo $x = y = 0$ na equação $x + 2y + 4z = c$, segue que

$$0 + 2 \cdot 0 + 4z = c \;\Rightarrow\; 4z = c \;\Rightarrow\; z = \frac{c}{4}, \tag{3.16}$$

agora fazendo $x = z = 0$ na equação $x + 2y + 4z = c$, segue que

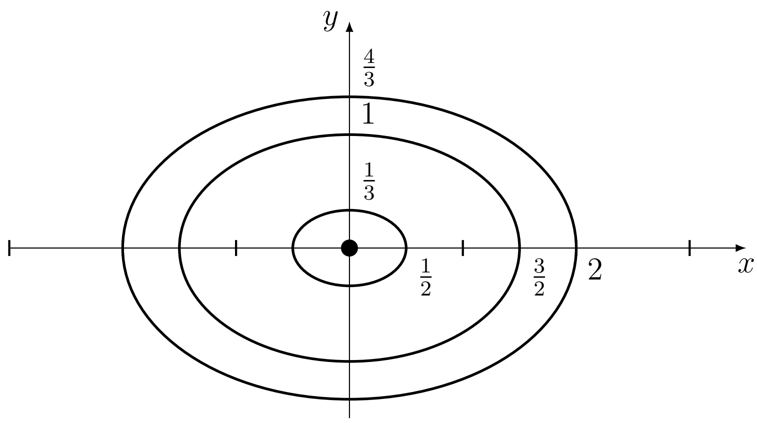
$$0 + 2y + 4 \cdot 0 = c \;\Rightarrow\; 2y = c \;\Rightarrow\; y = \frac{c}{2}, \tag{3.17}$$

finalmente fazendo $y = z = 0$ na equação $x + 2y + 4z = c$, segue que

$$x + 2 \cdot y + 4 \cdot z = c \;\Rightarrow\; x = c \tag{3.18}$$

e portanto as equações (3.16), (3.17) e (3.18) garantem que a superfície de nível c de j é o plano que intersecta respectivamente os eixos, O_z no ponto $(0, 0, \frac{c}{4})$, O_y no ponto $(0, \frac{c}{2}, 0)$ e O_x no ponto $(c, 0, 0)$.

Figura 5 – Curvas de nível de $h(x, y) = 4x^2 + 9y^2$.



Fonte: Elaborada pelo autor.

3.3 O Vetor Gradiente

Definição 3.5. Seja $f : U \subseteq \mathbb{R}^n \to \mathbb{R}$ uma função de classe C^1 e a um ponto do aberto U. Defina uma nova aplicação

$$\begin{aligned} \nabla f : U \subseteq \mathbb{R}^n &\longrightarrow \mathbb{R}^n \\ a &\longmapsto \nabla f(a) \end{aligned}$$

em que, para cada ponto a pertencente a U, é associado um vetor (uma direção) em $\mathbb{R}^n$, de tal forma que a i-ésima coordenada do vetor obtido equivale à i-ésima derivada parcial de f no ponto em questão, ou seja,

$$\nabla f(a) = \left(\frac{\partial f}{\partial x_1}(a), \frac{\partial f}{\partial x_2}(a), \frac{\partial f}{\partial x_3}(a), \ldots, \frac{\partial f}{\partial x_n}(a) \right).$$

Um ponto $x \in U$ diz-se *ponto crítico* de f quando $\nabla f(x) = \vec{0}$, ou seja, quando todas as derivadas parciais de f se anulam neste ponto.

A vantagem de se definir uma aplicação desta forma é que, como $f \in C^1$, ou seja, f é diferenciável, então tomando $v = (\alpha_1, \alpha_2, \alpha_3, \dots, \alpha_n)$ um vetor qualquer de $\mathbb{R}^n$ e $a \in U$, vale

$$\frac{\partial f}{\partial v}(a) = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(a) \cdot \alpha_i = \langle \nabla f(a), v \rangle$$

e enxergando a derivada direcional desta forma, como o produto interno do vetor gradiente de f no ponto a por um vetor $v \in \mathbb{R}^n$, podem ser obtidas algumas informações sobre o comportamento de f, como por exemplo, a direção para a qual a f cresce mais rapidamente, no sentido de que, dado um ponto a do domínio de f e um vetor v, para qual direção o vetor v deve apontar para que $f(a+v)$ tenha a maior variação possível em uma vizinhança do ponto a. E mais, da definição de diferenciabilidade, $r(v)$ pode ser definido como sendo

$$r(v) = f(a+v) - f(a) - \langle \nabla f(a), v \rangle.$$

Logo, dada $f : U \subseteq \mathbb{R}^n \to \mathbb{R}$ e $a \in U$ com U aberto, é dito que f é diferenciável no ponto a se, e somente se, f possui todas as derivadas parciais neste ponto e para todo vetor $v \in U$ valer $\lim\limits_{v \to 0} \dfrac{f(a+v) - f(a) - \langle \nabla f(a), v \rangle}{\|v\|} = 0$.

A notação da derivada parcial como um produto interno ainda proporciona uma notação mais simples para a Regra da Cadeia.

Sejam $U \subseteq \mathbb{R}^m$ e $V \subseteq \mathbb{R}^n$ abertos, $f = (f_1, f_2, f_3, \dots, f_n) : U \to V$, a um ponto de U tal $f(a) = b \in V$ e cada função coordenada $f_k : U \to V$ é diferenciável no ponto a. Seja $g : V \to \mathbb{R}$ uma função diferenciável no ponto $b = f(a)$, então a composição $g \circ f : U \to \mathbb{R}$ ainda é diferenciável em a e vale a fórmula

$$\frac{\partial (g \circ f)}{\partial x_i}(a) = \left\langle \nabla g(b), \frac{\partial f}{\partial x_i}(a) \right\rangle.$$

Perceba que esta é a mesma fórmula enunciada anteriormente, apenas foi melhorada a notação, deixando-a mais simples. Perceba também que se $m = 1$, ou seja, se $U \subseteq \mathbb{R}$ é um intervalo aberto e $f : U \subseteq \mathbb{R} \to V$ é um caminho diferenciável no ponto $a \in U$ tal que $f(a) = b \in V$, então a composição $(g \circ f)(a)$ é uma função real de uma variável real e a notação da sua derivada tornar-se-á ainda mais simples

$$(g \circ f)'(a) = \langle \nabla g(b), f'(a) \rangle.$$

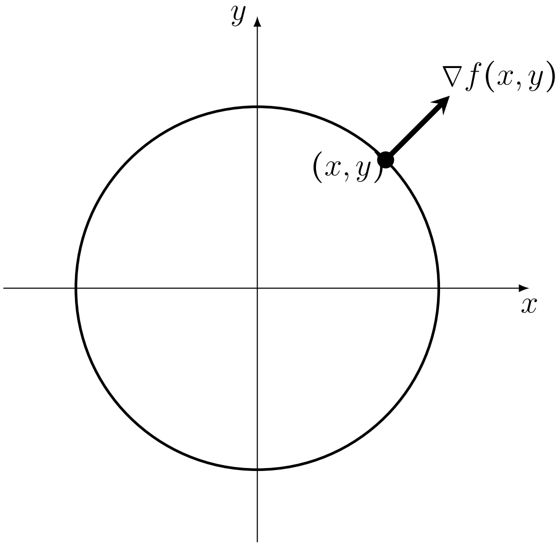
Exemplo 3.6. Considere as funções f, g e h dos exemplos 3.1, 3.3 e 3.4 respectivamente.

$f(x,y) = x^2 + y^2$ implica $\nabla f(x,y) = (2x, 2y)$, então para cada ponto $(x,y) \in \mathbb{R}^2$, o vetor $\nabla f(x,y)$ aponta para a direção (x,y), pois como o interesse aqui é apenas a direção, então ambas as coordenadas de ∇f podem ser divididas por 2.

$g(x,y) = y - x^2$ implica $\nabla g(x,y) = (-2x,1)$, logo para cada ponto $(x,y) \in \mathbb{R}^2$, o vetor $\nabla g(x,y)$ aponta para a direção $(-2x,1)$.

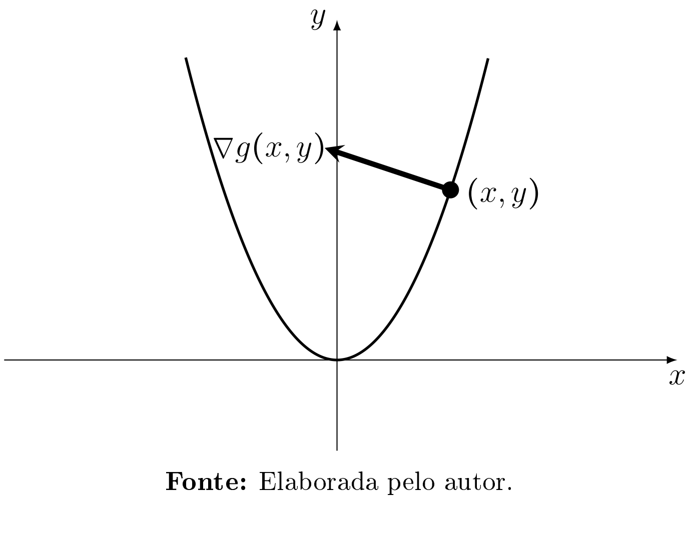
$h(x,y) = 4x^2 + 9y^2$ implica $\nabla h(x,y) = (8x,18y)$, novamente, como o interesse aqui é apenas na direção do vetor ∇h então ambas as coordenadas de $\nabla h(x,y)$ podem ser divididas por 2, logo para cada ponto (x,y) do domínio de h, o vetor $\nabla h(x,y)$ aponta para a direção $(4x,9y)$.

Figura 6 – Gradiente de $f(x,y)$.



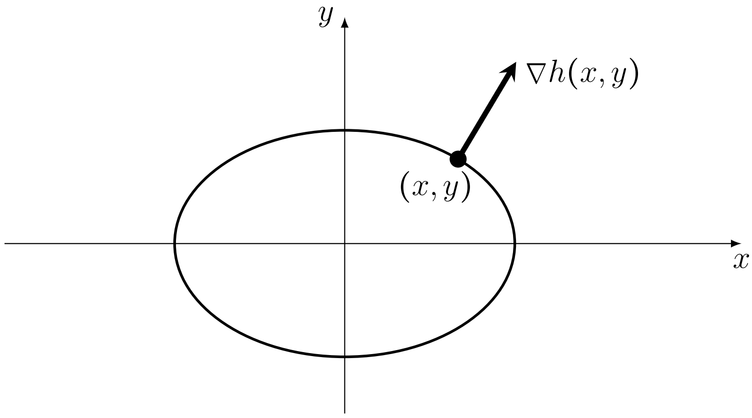
Fonte: Elaborada pelo autor.

Figura 7 – Gradiente de $g(x,y)$.



Fonte: Elaborada pelo autor.

Figura 8 – Gradiente de $h(x, y)$.



Fonte: Elaborada pelo autor.

4 O TEOREMA DA FUNÇÃO IMPLÍCITA

Uma função nada mais é que uma relação entre dois conjuntos, em que cada elemento do conjunto de partida (Domínio) se relaciona com um único elemento do conjunto de chegada (Contradomínio). Sabendo disto, é dito que uma equação do tipo $F : U \subseteq \mathbb{R}^n \to \mathbb{R}$ definida por $F(x) = c$, com $x = (x_1, x_2, x_3, \ldots, x_{k-1}, x_k, x_{k+1}, \ldots, x_n) \in \mathbb{R}^n$ e c uma constante real, *define implicitamente* uma função $x_k = \xi(x_1, x_2, x_3, \ldots, x_{k-1}, x_{k+1}, \ldots, x_n)$ quando para cada $(x_1, x_2, x_3, \ldots, x_{k-1}, x_{k+1}, \ldots, x_n)$ existir um único x_k de tal forma que $x = (x_1, x_2, x_3, \ldots, x_{k-1}, x_k, x_{k+1}, \ldots, x_n)$ seja solução de F. Perceba que considerar o nível c de $F(x)$ equivale à considerar o nível 0 de $f(x) = F(x) - c$. No caso $n = 2$, é dito que uma função $f : \mathbb{R}^2 \to \mathbb{R}$ com $f(x, y) = c$, define implicitamente uma função $y = \xi(x)$ quando para cada $x \in \mathbb{R}$ existe um único $y \in \mathbb{R}$ tais que $(x, y) = (x, \xi(x))$ é solução para f.

Em alguns casos, pode acontecer de uma equação definir mais de uma função implicitamente, ou de definir alguma função implicitamente apenas localmente, como por exemplo a equação $F(x, y) = x^2 + y^2 = 1$, ao tentar resolve-la em função de x, segue que $y = \pm\sqrt{1 - x^2}$, perceba que esta equação não define y em função de x, pois para cada $x \in (-1, 1)$ existem sempre dois valores de y que a resolvem. Analogamente, ao tentar resolve-la como função de y, segue que $x = \pm\sqrt{1 - y^2}$ e novamente a equação não define x como função de y pelo mesmo motivo. Entretanto, ao considerar os conjuntos $U_1 = \{(x, y) \in \mathbb{R}^2; y > 0\}$, $U_2 = \{(x, y) \in \mathbb{R}^2; y < 0\}$, $U_3 = \{(x, y) \in \mathbb{R}^2; x > 0\}$ e $U_4 = \{(x, y) \in \mathbb{R}^2; x < 0\}$, segue que se $(x, y) \in U_1$ então F equivale à $y = \sqrt{1 - x^2}$ e portanto a equação $F(x, y)$ restrita ao conjunto aberto U_1 define implicitamente y em função de x, apenas localmente. Analogamente F equivale à $y = -\sqrt{1 - x^2}$ quando $(x, y) \in U_2$, equivale à $x = \sqrt{1 - y^2}$ quando $(x, y) \in U_3$ e à $x = -\sqrt{1 - y^2}$ quando $(x, y) \in U_4$, logo, F restrita a algum destes dois últimos abertos também define x em função de y.

Logo, note que o conjunto de todas as soluções possíveis para $F(x, y) = 1$ estão contidas em algum dos abertos U_i, e mais, F restrita a algum destes abertos define implicitamente y em função de x ou x em função de y, isto quer dizer que o conjunto $F^{-1}(1) \cap U_i$ é o gráfico de alguma função $y = \xi(x)$ ou $x = \zeta(y)$.

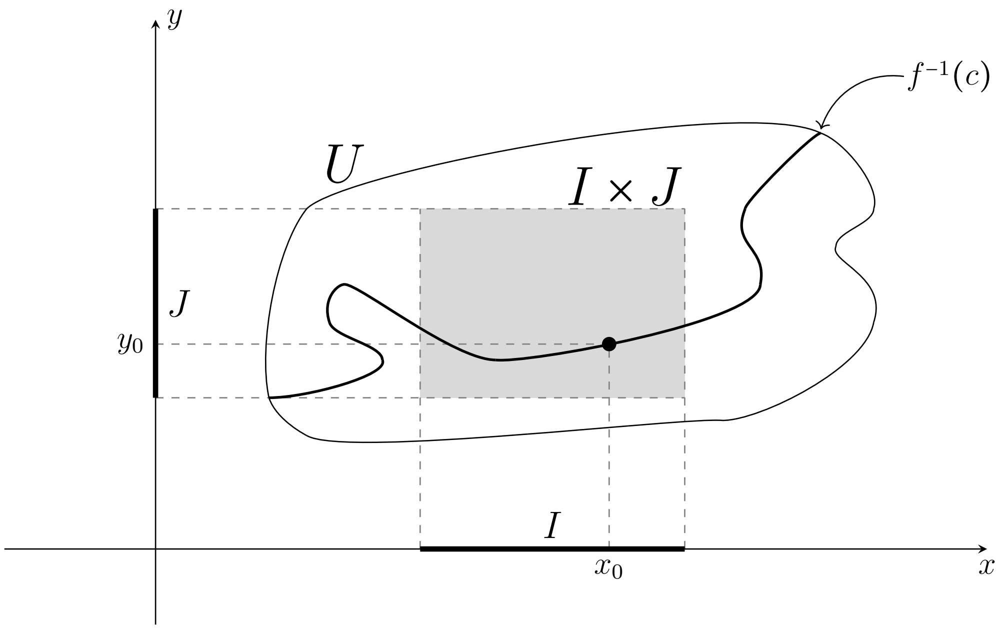
Acontece que, no geral, existem muitas equações em que é muito difícil ou impossível explicitar alguma das suas variáveis em função das demais e fazer esta análise que acabou de ser feita com a equação $x^2 + y^2 = 1$, em alguns casos é difícil ou impossível de até mesmo esboçar o gráfico de algumas equações, mas apesar disto, ainda é desejado obter delas algumas informações como por exemplo, se a equação define implicitamente alguma função mesmo que não seja possível explicitá-la e caso defina, é desejado obter informações desta função como a sua continuidade e derivada, ou derivadas parciais se for o caso. Será visto agora que é possível obter estas informações desde que a equação atenda a algumas condições.

Teorema 4.0.1 (Teorema da Função Implícita (duas variáveis))**.** *Sejam* $f : U \subseteq \mathbb{R}^2 \to \mathbb{R}$ *com* U *aberto e* $(x_0, y_0) \in U$*, tais que* $f(x_0, y_0) = c$*,* $f \in C^k$ *, com* $k \geq 1$*, e* $\dfrac{\partial f}{\partial y}(x_0, y_0) \neq 0$*. Então, existem intervalos abertos* I *de centro em* x_0 *e* J *de centro em* y_0 *e uma única função* $\xi : I \to J$ *tal que* $f^{-1}(c) \cap (I \times J)$ *é o gráfico de* ξ*, em que* $(I \times J)$ *é um retângulo aberto de centro em* (x_0, y_0)*. A função* ξ *é de classe* C^k *e para todo* $x \in I$ *vale*

$$\xi'(x) = -\frac{\dfrac{\partial f}{\partial x}(x, \xi(x))}{\dfrac{\partial f}{\partial y}(x, \xi(x))}.$$

Dizer que o gráfico de $\xi : I \to J$ é dado por $f^{-1}(c) \cap (I \times J)$, significa que para cada $x \in I$ existe um único $y \in J$ tal que $f(x, y) = c$ e portanto, y é uma função de x, a saber $\xi(x) = y$.

Figura 1 – Função Implícita ($\mathbb{R}^2$).



Fonte: Elaborada pelo autor.

Nas hipóteses do Teorema 4.0.1, ao supor $\dfrac{\partial f}{\partial y}(x_0, y_0) \neq 0$, conclui-se de que f define implicitamente uma função $\xi(x) = y$, y em função de x, e mais, obtém-se ainda uma fórmula de derivação para ξ, porém, o raciocínio é análogo quando $\dfrac{\partial f}{\partial x}(x_0, y_0) \neq 0$, quando isto ocorre, então f define implicitamente uma função $x = \zeta(y)$, x em função de y e é obtida a fórmula

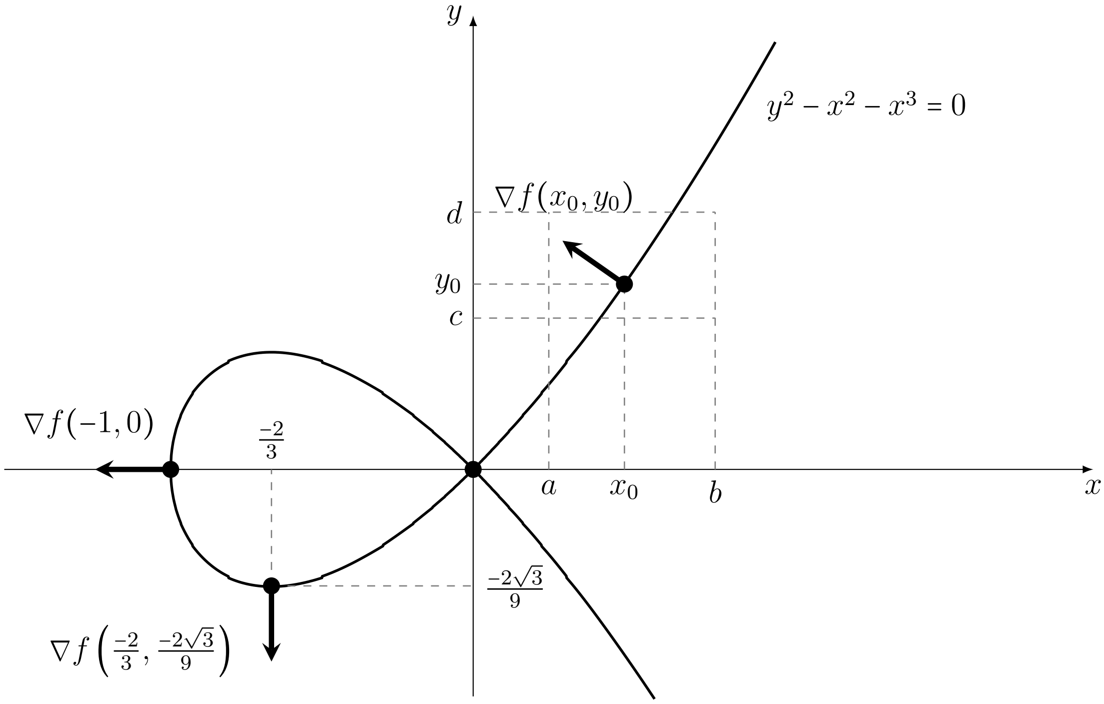
$$\zeta'(y) = -\frac{\dfrac{\partial f}{\partial y}(\zeta(y), y)}{\dfrac{\partial f}{\partial x}(\zeta(y), y)}.$$

Então, para que uma equação $f(x, y)$ defina implicitamente alguma função, pedir que

$\nabla f(x_0, y_0) \neq \overrightarrow{0}$ seria uma hipótese mais geral, pois caso isto ocorra, obrigatoriamente alguma das derivadas parciais de f no ponto (x_0, y_0) tem que ser diferente de 0 e então f define implicitamente uma função em que a variável dependente é aquela em que a referente coordenada do vetor gradiente de f no ponto (x_0, y_0) é diferente de 0.

Exemplo 4.1. Considere o nível 0 da equação $f(x,y) = y^2 - x^2 - x^3$, considere também os pontos $(0,0)$, $(-1,0)$, $\left(-\frac{2}{3}, -\frac{2\sqrt{3}}{9}\right)$, um ponto (x_0, y_0) em que $x_0 > 0$ e um retângulo aberto $V = (a,b) \times (c,d)$ centrado em (x_0, y_0).

Figura 2 – $f(x,y) = 0$



Fonte: Elaborada pelo autor.

Perceba que f é contínua, pois é um polinômio e todos os pontos mencionados acima estão no nível 0 de f. Calculando as derivadas parciais em um ponto qualquer de f, utilizando as regras de derivação, segue que

$$\frac{\partial f}{\partial x}(x,y) = -3x^2 - 2x \quad \text{e} \quad \frac{\partial f}{\partial y}(x,y) = 2y.$$

No ponto $(-1,0)$, $\frac{\partial f}{\partial y}$ é igual a 0, ou seja, a reta que tangencia f no ponto $(-1,0)$ é paralela ao eixo das abcissas, isto quer dizer que para algum x no interior de algum retângulo aberto centrado em $(-1,0)$ vai existir dois valores de y que são soluções da equação e portanto f não define nenhuma função $\xi(x) = y$ em uma vizinhança do ponto $(-1,0)$.

No ponto $\left(-\frac{2}{3}, -\frac{2\sqrt{3}}{9}\right)$, $\frac{\partial f}{\partial x}$ vale 0 e portanto a reta tangente à f neste ponto é paralela ao eixo das ordenadas e para algum y no interior de qualquer retângulo aberto centrado em $\left(-\frac{2}{3}, -\frac{2\sqrt{3}}{9}\right)$ vai existir dois valores de x que são soluções de $f(x,y) = 0$, isto quer dizer que f não define nenhuma função $\zeta(y) = x$ em uma vizinhança do ponto $\left(-\frac{2}{3}, -\frac{2\sqrt{3}}{9}\right)$.

Em um ponto qualquer (x_0, y_0) com $x_0 > 0$, tanto $\frac{\partial f}{\partial x}$ quanto $\frac{\partial f}{\partial y}$ assumirão valores diferentes de 0, ou seja, o vetor $\nabla f(x_0, y_0)$ não possui nenhuma coordenada nula, logo, às luzes do Teorema da Função implícita, f define tanto uma função $\xi(x) = y$ quanto uma função $\zeta(y) = x$ dentro de algum retângulo aberto $V = (a,b) \times (c,d)$ de centro em (x_0, y_0). Em outras palavras, para cada x no interior de V existe um único y que soluciona f, assim como para cada y, existe um único x que a soluciona. Por fim, na origem, o vetor $\nabla f(0,0) = \vec{0}$ (ponto crítico). Para quaisquer ε e δ, com $K = (-\varepsilon, \varepsilon) \times (-\delta, \delta)$, o conjunto $f^{-1}(0,0) \cap K$ não é gráfico de nenhuma função, pois para cada x no interior de algum retângulo aberto K centrado na origem, existem dois valores de y e para cada $y \in K$ existem dois valores de x que solucionam $f(x,y)$.

Será necessário demonstrar o seguinte Lema para poder demonstrar o Teorema da Função Implícita:

Lema 4.1. *Seja $f : X \times K \to \mathbb{R}^p$ uma função contínua, em que $X \subset \mathbb{R}^m$ e $K \subset \mathbb{R}^k$ é um compacto. Se $f^{-1}(c)$ é o gráfico de uma aplicação $\xi : X \to K$, com $c \in \mathbb{R}^p$, então ξ é contínua. Dizer que $f^{-1}(c)$ é o gráfico de uma aplicação $\xi : X \to K$, com $c \in \mathbb{R}^p$, significa que para cada $x \in X$ existe $y = \xi(x)$ único, tal que $f(x, \xi(x)) = c$.*

Demonstração. Sejam $(x_n) \in X$ uma sequência de pontos, com $\lim_{n\to\infty} x_n = x_0 \in X$ e $y_0 = \xi(x_0)$. Como ξ é uma aplicação de X em K e $x_n \in X$ para todo $n \in \mathbb{N}$, então $\xi(x_n) \in K$ para todo n e como K é compacto, em particular fechado, então ele contêm todos os seus pontos aderentes, portanto a sequência $(\xi(x_n))$ tem limite em K, ou seja, $\lim_{n\to\infty} \xi(x_n) = y \in K$, e mais, como K é limitado, então $(\xi(x_n))$ também o é. Para que ξ seja contínua em um ponto x_0 basta mostrar que para toda sequência de pontos $(x_n) \in X$, com $\lim_{n\to\infty} x_n = x_0 \in X$, tem-se $\lim_{n\to\infty} \xi(x_n) = \xi(x_0)$, como a sequência (x_k) foi tomada arbitrariamente, é suficiente mostrar que $\lim_{n\to\infty} \xi(x_n) = \xi(x_0) = y_0$, para isto, pelo Teorema 2.1.4, basta mostrar que a sequência $(\xi(x_n))$ possui um único valor de aderência, ou seja, toda subsequência $(\xi(x'_n))$ convergente converge para y_0. Tomando uma subsequência (x'_n) arbitrária de (x_n), como $\lim_{n\to\infty} f(x'_n, \xi(x'_n)) = c$ para todo n, pois f define implicitamente ξ, então $\lim_{n\to\infty} f(x'_n, \xi(x'_n)) = \lim_{n\to\infty} f(x_0, y) = c$, logo $y = \xi(x_0) = y_0$ pela unicidade de y_0, portanto $\lim_{n\to\infty} \xi(x'_n) = \xi(x_0) = y_0$ implica $\lim_{n\to\infty} \xi(x_n) = \xi(x_0) = y_0$, pois (x'_n) foi tomada arbitrariamente e isto quer dizer que ξ é contínua em x_0. Como x_0 foi tomado arbitrariamente então ξ é contínua. □

Será enunciado agora o Teorema da Função Implícita para funções de um espaço de dimensão $n+1$ qualquer ($\mathbb{R}^{n+1}$) e em seguida, demonstrado. Na verdade será feita a generalização do Teorema 4.0.1, tal qual equivale ao caso em que $n = 1$, ou seja, o caso $\mathbb{R}^2$.

Teorema 4.0.2 (Teorema da Função Implícita)**.** *Sejam $f : U \subseteq \mathbb{R}^{n+1} \to \mathbb{R}$ com U aberto e $p = (x_0, y_0) \in U$ tais que $f(p) = c$, $f \in C^k (k \geq 1)$ e $\dfrac{\partial f}{\partial y}(p) \neq 0$. Então, existem uma bola aberta $B = B(x_0; \delta)$ e um intervalo aberto $J = (y_0 - \varepsilon, y_0 + \varepsilon)$ e uma única função $\xi : B \to J$ tal que $f^{-1}(c) \cap (B \times J)$ é o gráfico de ξ. A função ξ é de classe C^k e para todo $x \in B$ vale*

$$\frac{\partial \xi}{\partial x_i}(x) = -\frac{\dfrac{\partial f}{\partial x_i}(x, \xi(x))}{\dfrac{\partial f}{\partial y}(x, \xi(x))} \quad , \quad (i = 1, 2, 3, \ldots, n).$$

Dizer que o gráfico de $\xi : B \to J \in C^k$ é dado por $f^{-1}(c) \cap (B \times J)$, significa que para cada $x \in B$ existe um único $y \in J$ tal que $f(x, y) = c$ e portanto y é uma função de x, a saber $\xi(x) = y$.

Demonstração. Sem perda de generalidade, suponha $\dfrac{\partial f}{\partial y}(x_0, y_0) > 0$. Como $f \in C^k$ com $k \geq 1$ e em particular $f \in C^1$, segue todas as derivadas parciais de f são contínuas, em particular $\dfrac{\partial f}{\partial y}$ é contínua. Da continuidade de $\dfrac{\partial f}{\partial y}$, segue que existem δ e ε estritamente positivos e suficientemente pequenos tais que, tomando δ e ε como raios da bola aberta de centro em x_0, $B = B(x_0; \delta)$, e do intervalo aberto de centro em y_0, $J = (y_0 - \varepsilon, y_0 + \varepsilon)$, respectivamente, tais que $B \times \overline{J} \subset U$ (basta tomar um ε_0 tal que $B \times J \subset U$, escolhendo $\varepsilon < \varepsilon_0$, então $B \times \overline{J} \subset U$) e para todo ponto $(x, y) \in B \times \overline{J}$, segue que $\dfrac{\partial f}{\partial y} > 0$. A vantagem de tomar um ε suficientemente pequeno tal que o conjunto "B cartesiano *fecho de* J"esteja contido em U é que isto permite analisar o comportamento de f nos extremos do intervalo compacto $\overline{J} = [y_0 - \varepsilon, y_0 + \varepsilon]$. Nesta condições, f é estritamente crescente no intervalo $\overline{J}$ para todo $x \in B$ e consequentemente, para qualquer incremento no valor de y_0, o valor de f é diferente de c, ou seja, $f(x_0, y_0 - \varepsilon) < c$ e $f(x_0, y_0 + \varepsilon) > c$. Suponha agora δ tão pequeno que $f(x, y_0 - \varepsilon) < c$ e $f(x, y_0 + \varepsilon) > c$ para todo $x \in B$. Como f é contínua e estritamente crescente em $\overline{J}$, pelo Teorema do Valor Intermediário, para todo $x \in B$ fixado, existe $y = \xi(x)$, único, tal que $f(x, \xi(x)) = c$. Como $f(x, y_0 \pm \varepsilon) \neq c$ então $y \in (y_0 - \varepsilon, y_0 + \varepsilon)$, ou seja, $f^{-1}(c) \cap (B \times \overline{J}) = f^{-1}(c) \cap (B \times J)$, tal qual é o gráfico de uma função $\xi : B \to J$ e tal função é contínua pelo Lema 4.1.

Será mostrando agora que existem todas as derivadas parciais de ξ e que vale a fórmula enunciada. Pois bem, note que, como $B = B(x_0; \delta)$ é um conjunto aberto, então para cada $x \in B$ existe um δ' tal que $B(x; \delta') \subset B$. Em particular, para cada $x \in B$, segue que $x + te_i \in B$ para todo $t \in (-\delta', \delta')$, isto quer dizer que $\xi(x + te_i) \in J$ e portanto $f(x + te_i, \xi(x + te_i)) = c$, pois ξ é definida implicitamente por f. Defina $k(t) = \xi(x + te_i) - \xi(x)$ para todo $t \in (-\delta', \delta')$. Note que, definindo k desta forma, $\lim_{t \to 0} k(t) = 0$ se, e somente se,

$\lim_{t\to 0}[\xi(x+te_i)-\xi(x)] = 0$ se, e somente se, $\lim_{t\to 0}\xi(x+te_i) = \xi(x)$ se e somente se, ξ é contínua em x, o que de fato acontece pelo 4.1. Note também que, $k(t) = \xi(x+te_i)-\xi(x)$ implica $\xi(x+te_i) = \xi(x)+k(t)$, logo $c = f(x+te_i, \xi(x+te_i)) = f(x+te_i, \xi(x)+k(t))$ e portanto

$$f(x+te_i, \xi(x)+k(t)) - f(x, \xi(x)) = c - c = 0. \tag{4.1}$$

Utilizando o Teorema do Valor Médio em (4.1), para cada $t \in (-\delta', \delta')$ existe $\theta = \theta(t) \in (0,1)$ de tal forma que tomando $v = (te_i, k(t))$,

$$\begin{aligned} 0 = f(x+te_i, \xi(x)+k(t)) - f(x, \xi(x)) =& \frac{\partial f}{\partial v}(x+\theta\cdot te_i, \xi(x)+\theta\cdot k(t)) \\ =& \frac{\partial f}{\partial x_i}(x+\theta\cdot te_i, \xi(x)+\theta\cdot k(t))\cdot t \\ &+ \frac{\partial f}{\partial y}(x+\theta\cdot te_i, \xi(x)+\theta\cdot k(t))\cdot k(t) \end{aligned}$$

$$\begin{aligned} &\Rightarrow \frac{\partial f}{\partial x_i}(x+\theta\cdot te_i, \xi(x)+\theta\cdot k(t))\cdot t = -\frac{\partial f}{\partial y}(x+\theta\cdot te_i, \xi(x)+\theta\cdot k(t))\cdot k(t) \\ &\Rightarrow \frac{\xi(x+te_i)-\xi(x)}{t} = \frac{k(t)}{t} = -\frac{\frac{\partial f}{\partial x_i}(x+\theta\cdot te_i, \xi(x)+\theta\cdot k(t))}{\frac{\partial f}{\partial y}(x+\theta\cdot te_i, \xi(x)+\theta\cdot k(t))} \\ &\Rightarrow \frac{\xi(x+te_i)-\xi(x)}{t} = -\frac{\frac{\partial f}{\partial x_i}(x+\theta\cdot te_i, \xi(x)+\theta\cdot k(t))}{\frac{\partial f}{\partial y}(x+\theta\cdot te_i, \xi(x)+\theta\cdot k(t))}. \end{aligned} \tag{4.2}$$

Passando o limite em (4.2) com t tendendo a 0, segue que

$$\lim_{t\to 0}\frac{\xi(x+te_i)-\xi(x)}{t} = \lim_{t\to 0} -\frac{\frac{\partial f}{\partial x_i}(x+\theta\cdot te_i, \xi(x)+\theta\cdot k(t))}{\frac{\partial f}{\partial y}(x+\theta\cdot te_i, \xi(x)+\theta\cdot k(t))}$$

$$\Rightarrow \frac{\partial \xi}{\partial x_i}(x) = -\frac{\frac{\partial f}{\partial x_i}(x, \xi(x))}{\frac{\partial f}{\partial y}(x, \xi(x))}. \tag{4.3}$$

Note que, como f é de classe C^1, ou seja, suas derivadas parciais são todas contínuas, então a existência do limite da esquerda de (4.3) segue da existência do limite da direita, tal qual é por definição a i-ésima derivada parcial de ξ. Analisando a expressão (4.3), se $f \in C^k$, então todas as suas derivadas parciais são de classe C^{k-1} e portando $\frac{\partial \xi}{\partial x_i} \in C^{k-1}$, pois da fórmula acima segue que elas são de mesma classe, logo se $\frac{\partial \xi}{\partial x_i} \in C^{k-1}$ para todo

$(n = 1, 2, 3, \ldots, n)$ então $\xi \in C^k$. □

Observação 4.1. Utilize o Lema 4.1 no fecho de J, $\overline{J}$, no Teorema 4.0.2.

Exemplo 4.2. Dada a equação $f(x,y) = e^{2y} + yx + x^5 - 2x^2 - 2$ e o ponto $(1,0)$, considere o nível 0 de f.

O ponto $(1,0)$ está no nível 0 de f, pois $f(1,0) = e^{2\cdot 0} + 0\cdot 1 + 1^5 - 2\cdot 1^2 = 1 + 1 - 2 = 0$. f é de classe C^∞, pois a função exponencial é de classe C^∞ assim como polinômios também o são, e mais, a composição de funções de classe C^∞ ainda é C^∞. Como

$$\frac{\partial f}{\partial y}(x,y) = 2e^{2y} + x \Rightarrow \frac{\partial f}{\partial y}(1,0) = 2e^{2\cdot 0} + 1 = 3 \neq 0$$

e

$$\frac{\partial f}{\partial x}(x,y) = y + 5x^4 - 4x \Rightarrow \frac{\partial f}{\partial x}(1,0) = 0 + 5\cdot 1^4 - 4\cdot 1 = 1 \neq 0,$$

então o Teorema da Função Implícita garante a existência de uma única função $\xi(x) = y$ e uma única função $\zeta(y) = x$, ambas de classe C^∞ em uma vizinhança do ponto $(1,0)$. E mais, o Teorema ainda fornece uma fórmula para calcular as derivadas destas funções no ponto $(1,0)$, com $\xi(1) = 0$ e $\zeta(0) = 1$:

$$\xi'(x) = -\frac{\dfrac{\partial f}{\partial x}(x,\xi(x))}{\dfrac{\partial f}{\partial y}(x,\xi(x))} \Rightarrow \xi'(1) = -\frac{\dfrac{\partial f}{\partial x}(1,0)}{\dfrac{\partial f}{\partial y}(1,0)} = -\frac{0 + 5\cdot 1^4 - 4\cdot 1}{2e^{2\cdot 0} + 1} = -\frac{1}{3},$$

$$\zeta'(y) = -\frac{\dfrac{\partial f}{\partial y}(\zeta(y),y)}{\dfrac{\partial f}{\partial x}(\zeta(y),y)} \Rightarrow \zeta'(0) = -\frac{\dfrac{\partial f}{\partial y}(1,0)}{\dfrac{\partial f}{\partial x}(1,0)} = -\frac{2e^{2\cdot 0} + 1}{0 + 5\cdot 1^4 - 4\cdot 1} = -3.$$

Exemplo 4.3. Considere o nível 0 da equação $g(x,y,z) = y\ln(x) + 2ye^z + \mathrm{sen}(z)$ e o ponto $(1,1,0)$.

Note que, o ponto $(1,1,0)$ está no nível 0 de g, pois $g(1,1,0) = 1\cdot \ln(1) + 2\cdot 1\cdot e^0 + \sin(0) - 2 = 0$. Como polinômios são de classe C^∞ assim como a composição de funções C^∞, então g é de classe C^∞ em uma vizinhança do ponto $(1,1,0)$.

Agora, calculando as derivadas parciais de f no ponto dado, segue que

$$\frac{\partial g}{\partial x}(x,y,z) = y\cdot\frac{1}{x} \Rightarrow \frac{\partial g}{\partial x}(1,1,0) = 1 \neq 0,$$

$$\frac{\partial g}{\partial y}(x,y,z) = \ln(x) + 2e^z \Rightarrow \frac{\partial g}{\partial y}(1,1,0) = 2 \neq 0$$

e

$$\frac{\partial g}{\partial z}(x,y,z) = 2ye^z + \cos(z) \Rightarrow \frac{\partial g}{\partial z}(1,1,0) = 2\cdot 1\cdot e^0 + \cos(0) = 3 \neq 0.$$

Isto quer dizer que o vetor $\nabla g(1,1,0)$ não possui nenhuma coordenada nula, logo o Teorema da Função implícita garante que g define implicitamente uma função $\xi(x,z) = y$, uma função $\zeta(y,z) = x$ e uma função $\gamma(x,y) = z$ no interior de alguma bola aberta de centro em $(1,1,0)$. O Teorema fornece ainda uma fórmula para poder calcular as derivadas destas funções:

$$\frac{\partial \xi}{\partial x}(1,0) = -\frac{\frac{\partial g}{\partial x}(1,1,0)}{\frac{\partial g}{\partial y}(1,1,0)} = -\frac{1}{2}\ , \quad \frac{\partial \xi}{\partial z}(1,0) = -\frac{\frac{\partial g}{\partial z}(1,1,0)}{\frac{\partial g}{\partial y}(1,1,0)} = -\frac{3}{2}\ ,$$

$$\frac{\partial \zeta}{\partial y}(1,0) = -\frac{\frac{\partial g}{\partial y}(1,1,0)}{\frac{\partial g}{\partial x}(1,1,0)} = -3\ , \quad \frac{\partial \zeta}{\partial z}(1,0) = -\frac{\frac{\partial g}{\partial z}(1,1,0)}{\frac{\partial g}{\partial x}(1,1,0)} = -2\ ,$$

$$\frac{\partial \gamma}{\partial x}(1,1) = -\frac{\frac{\partial g}{\partial x}(1,1,0)}{\frac{\partial g}{\partial z}(1,1,0)} = -\frac{1}{3}\quad , \frac{\partial \gamma}{\partial y}(1,0) = -\frac{\frac{\partial g}{\partial y}(1,1,0)}{\frac{\partial g}{\partial z}(1,1,0)} = -\frac{2}{3}.$$

5 APLICAÇÕES

5.1 O Vetor Gradiente e as curvas de nível

Dentre as várias aplicações possíveis, também importantíssimas, para o Teorema da Função Implícita, este permite provar que o Vetor Gradiente é perpendicular aos conjuntos de nível, mais precisamente, dada $f : U \subseteq \mathbb{R}^{n+1} \to \mathbb{R}$ de classe C^k $(k \geq 1)$, seja $M \subset \mathbb{R}^{n+1}$ o conjunto de nível c de f, ou seja, $M = f^{-1}(c) = \{(x, y) \in \mathbb{R}^{n+1}; x \in \mathbb{R}^n, y \in \mathbb{R} \text{ e } f(x, y) = c\}$, seja também $(x_0, y_0) \in M$, então $\nabla f(x_0, y_0)$ é perpendicular a M em (x_0, y_0), quando não for o vetor nulo.

Teorema 5.1.1. *Seja $f : U \subseteq \mathbb{R}^{n+1} \to \mathbb{R}$ de classe C^k $(k \geq 1)$, U aberto e $(x_0, y_0) \in U$ tais que $x_0 \in \mathbb{R}^n, y_0 \in \mathbb{R}, f(x_0, y_0) = c$ e $\nabla f(x_0, y_0) \neq \vec{0}$. Se $M \subset \mathbb{R}^{n+1}$ é o conjunto de nível c de f então, $\nabla f(x_0, y_0) \perp M$ em (x_0, y_0).*

Demonstração. O Teorema da Função Implícita garante que f define localmente uma função de n variáveis. Considere o caminho

$$\begin{aligned} \lambda : (-\varepsilon, \varepsilon) &\longrightarrow M \\ t &\longmapsto \lambda(t) = (x, y) \end{aligned}$$

tal que $\lambda(0) = (x_0, y_0)$, perceba agora que para todo $t \in (-\varepsilon, \varepsilon)$, segue que $\lambda(t) \in M$ implica $f(\lambda(t)) = f \circ \lambda(t) = c$, logo, se λ é diferenciável em 0 então, como já foi visto, a derivada de um caminho resulta em um vetor tangente a tal caminho, no ponto em questão, portanto, utilizando a regra da cadeia na composição $f \circ \lambda$, sabendo que tal composição é constante igual a c, segue que

$$0 = (f \circ \lambda)'(0) = \langle \nabla f(\lambda(0)), \lambda'(0) \rangle = \langle \nabla f(x_0, y_0), \lambda'(0) \rangle.$$

Isto quer dizer que o vetor $\nabla f(x_0, y_0)$ é ortogonal ao vetor velocidade $\lambda'(0)$, como foi tomado um caminho λ arbitrário, então $\nabla f(x_0, y_0)$ é ortogonal a todo vetor velocidade de M no ponto (x_0, y_0), daí segue o resultado. □

Mais precisamente, dizer que ∇f é ortogonal á todo vetor velocidade de M no ponto (x_0, y_0) significa dizer que $\nabla f(x_0, y_0)$ é perpendicular a qualquer reta que tangencie o conjunto M no ponto (x_0, y_0), ou seja $\nabla f(x_0, y_0) \perp M$ em (x_0, y_0).

No caso $f : U \subseteq \mathbb{R}^2 \to \mathbb{R}$, por exemplo, sabendo que o vetor gradiente é perpendicular às curvas de nível, o motivo pelo qual tal vetor precisa ser diferente do vetor nulo para que f defina alguma função implicitamente é o seguinte. Quando $\dfrac{\partial f}{\partial y}(x_0, y_0) = 0$ por exemplo, isto quer dizer que o vetor ∇f neste ponto não tem componente em $\vec{j}$ neste ponto e obrigatoriamente a reta que tangencia f neste ponto possui apenas componente em $\vec{j}$, por ser perpendicular a $\nabla f(x_0, y_0)$, ou seja, tal reta é paralela ao eixo das ordenadas, isto quer dizer que, como f é contínua, então para qualquer vizinhança tomada do ponto (x_0, y_0), para algum x dentro desta vizinhança vai existir mais de um y que seja solução de f e portanto f não define nenhuma função $\xi(x) = y$ em nenhuma vizinhança de (x_0, y_0). Analogamente, se $\frac{\partial f}{\partial x}(x_0, y_0) = 0$, então o vetor ∇f neste ponto não tem componente em $\vec{i}$ neste ponto e obrigatoriamente a reta tangente à f em (x_0, y_0) possui apenas componente em $\vec{i}$, logo para algum y em qualquer vizinhança do ponto (x_0, y_0) vai existir mais de um x que seja solução de f.

5.2 O Teorema da Função Inversa

Uma outra importantíssimo aplicação para o Teorema da Função Implícita é utilizá-lo para demonstrar o Teorema da Função Inversa. Por simplicidade, será denotado por $\partial_i f$ a i-ésima derivada parcial de f. Denota-se por $\mathcal{L}(\mathbb{R}^n; \mathbb{R}^m)$, o conjunto das transformações lineares $T : \mathbb{R}^n \to \mathbb{R}^m$, tal qual é um espaço vetorial.

Definição 5.1. Sejam $U \subset \mathbb{R}^n$ e $V \subset \mathbb{R}^m$. Um *isomorfismo* de U em V é uma aplicação linear bijetora $f : U \to V$.

Definição 5.2. Sejam $U, V \subset \mathbb{R}^n$ conjuntos abertos. Um *difeomorfismo* de U em V é uma bijeção diferenciável $f : u \to V$, cuja inversa f^{-1} também é diferenciável. f diz-se um difeomorfismo de classe C^k, quando f e f^{-1} são ambas diferenciáveis e de classe C^k.

Teorema 5.2.1 (Teorema da Função Inversa)**.** *Sejam $f : U \subseteq \mathbb{R}^n \to \mathbb{R}^m$, em que U é aberto e f é de classe $C^k (k \geq 1)$, tal que $f'(y_0) \in \mathcal{L}(\mathbb{R}^n; \mathbb{R}^m)$ é um isomorfismo. Então f é um difeomorfismo de classe C^k de uma vizinhança $B(y_0; \delta)$ de y_0 sobre uma vizinhança $B(f(y_0); \varepsilon)$ de $f(y_0)$, ou seja, existem abertos $W = B(y_0; \delta)$, $V = B(f(y_0); \varepsilon)$ e uma única função $g : W \to V$, a qual é de classe C^k, tal que $g = f^{-1}$.*

Demonstração. Considere o caminho

$$\begin{aligned} \varphi : \mathbb{R}^m \times U &\longrightarrow U \\ (x, y) &\longmapsto f(y) - x. \end{aligned}$$

Note que φ satisfaz todas as hipóteses do Teorema da Função Implícita:

1. $f \in C^k \;\Rightarrow\; \varphi \in C^k$;

2. $\varphi(f(y_0), y_0) = f(y_0) - f(y_0) = 0$;

3. $\partial_2 \varphi(f(y_0), y_0) = f'(y_0)$ é um isomorfismo.

Aplicando em φ o Teorema da Função Implícita, é obtida a existência de bolas abertas $W = B(y_0; \delta)$, $V = B(f(y_0); \varepsilon)$ e uma única função $g : W \to V$, a qual é de classe C^k, tais que $g(f(y_0)) = y_0$ e as únicas soluções em $W \times V$ para

$$\varphi(x,y) = 0$$

são da forma

$$(x, g(x)),$$

isto quer dizer que

$$0 = \varphi(x, g(x)) = f(g(x)) - x \;\Rightarrow\; f(g(x)) = x$$

para todo $x \in W$, ou seja, $g = f^{-1}$ em uma vizinhança de $(f(y_0), y_0) \in W \times V \subset \mathbb{R}^m \times U$. $\square$

Observação 5.1. Como $f' : U \to \mathbb{R}^m$ é um vetor, supor que $f'(y_0)$ seja um isomorfismo, em vez de supor $f'(y_0) \neq 0$, terá o mesmo efeito.

Existem várias versões tanto do Teorema da Função Implícita quanto do Teorema da Função Inversa, as que foram apresentadas aqui são as mais simples delas, existem versões mais gerais, com hipóteses mais gerais, como por exemplo, em Espaços de Banach. Estes dois Teoremas são equivalentes, ou seja, aceitando um, pode-se demonstrar o outro e vice-versa. Aqui foi demonstrado o Teorema da Função Inversa como consequência do Teorema da Função Implícita, o caso contrário pode ser encontrado em Rudin (1964, p. 224).

5.3 Bilhares

Definição 5.3. Seja $f : U \subset \mathbb{R}^m \to \mathbb{R}^n$, U aberto e $a \in U$. Chama-se *matriz jacobiana* de f no ponto a, e denota-se por $Jf(a)$, a matriz $n \times m$

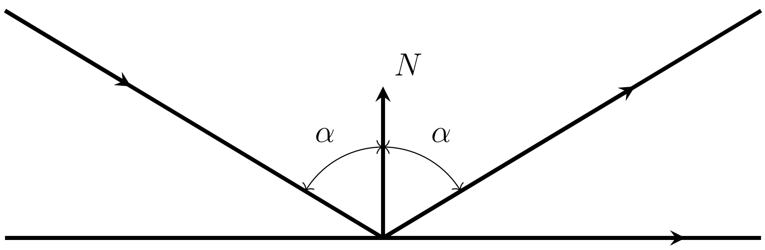
$$\begin{bmatrix} \frac{\partial f_1}{\partial x_1}(a) & \frac{\partial f_1}{\partial x_2}(a) & \frac{\partial f_1}{\partial x_3}(a) & \cdots & \frac{\partial f_1}{\partial x_n}(a) \\ \frac{\partial f_2}{\partial x_1}(a) & \frac{\partial f_2}{\partial x_2}(a) & \frac{\partial f_2}{\partial x_3}(a) & \cdots & \frac{\partial f_2}{\partial x_n}(a) \\ \frac{\partial f_3}{\partial x_1}(a) & \frac{\partial f_3}{\partial x_2}(a) & \frac{\partial f_3}{\partial x_3}(a) & \cdots & \frac{\partial f_3}{\partial x_n}(a) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial x_1}(a) & \frac{\partial f_m}{\partial x_2}(a) & \frac{\partial f_m}{\partial x_3}(a) & \cdots & \frac{\partial f_m}{\partial x_n}(a) \end{bmatrix}.$$

A fim de que $f : U \subset \mathbb{R}^m \to \mathbb{R}^n$ seja um isomorfismo, basta que $\det Jf(x) \neq 0$ para todo $x \in U$.

Seja $\gamma : [a,b] \to \Omega \subset \mathbb{R}^2$ um caminho (ou *arco*) em $\mathbb{R}^2$, que tem como domínio um intervalo fechado. Sua imagem $\gamma([a,b]) = \{\gamma(t) \in \Omega \mid t \in [a,b]\}$ é dita uma *curva* em Ω. Um ponto (x_0, y_0) do arco γ diz-se *ponto múltiplo* do arco quando a equação $\gamma(t) = (x_0, y_0)$ é satisfeita para mais de um $t \in [a,b]$. Quando um arco γ possui apenas um *ponto duplo* $\gamma(a) = \gamma(b)$, correspondendo aos pontos final e inicial, diz-se que γ é um *arco fechado simples* e sua imagem, $\gamma([a,b])$, é dita uma *curva fechada simples*. Uma regiam $\Omega \subset \mathbb{R}^2$ é dita *simplesmente conexa* quando, em toda curva fechada simples em Ω, os pontos dentro desta curva também pertencem à Ω.

Definição 5.4. Seja Ω uma região simplesmente conexa do plano $\mathbb{R}^2$, denomina-se *bilhar* toda curva fechada simples γ, de classe C^k $(k \geq 2)$, orientada positivamente[1], com curvatura estritamente positiva, contida em Ω. Chama-se *trajetória de bilhar*, a trajetória descrita por uma partícula (a bola de bilhar) se movendo em velocidade constante, satisfazendo a Lei da Reflexão[2] com colisão elástica na fronteira γ.

Figura 1 – Lei da Reflexão.



Fonte: Elaborada pelo autor.

Sabendo que, entre duas colisões consecutivas, o movimento descrito pela partícula é retilíneo, então a dinâmica de movimento de uma partícula em um bilhar pode ser caracterizada por uma sequência de pontos $P_n = (s_n, \alpha_n) \in \gamma$, nos momentos de colisão, em que $s_n \in [0, l)$, l é comprimento de γ, e α_n é o ângulo orientado entre o vetor tangente à γ (no ponto P_n) e o segmento de trajetória de saída da partícula. O objetivo aqui será estudar apenas o caso em que $\alpha_n \in (0, \pi)$ para todo $n \in \mathbb{N}$. Tomando o eixo das abscissas (O_x) como direção fixada, denotar-se-á por φ, o Ângulo orientado entre o eixo das abscissas e o vetor tangente à γ.

Note que, dados dois pontos (s_0, α_0) e (s_1, α_1) em um bilhar (veja a Figura 2), obtém-se

$$\varphi_1 - \alpha_1 = \varphi_0 + \alpha_0 \;\Rightarrow\; \alpha_1 = \varphi_1 - \varphi_0 - \alpha_0, \tag{5.1}$$

chamando de θ o ângulo orientado entre o eixo das abscissas e o segmento de trajetória de saída, segue que

$$\theta_0 = \alpha_0 - \varphi_0 \;\Rightarrow\; \alpha_0 = \theta_0 - \varphi_0. \tag{5.2}$$

Substituindo (5.2) em (5.1), segue que

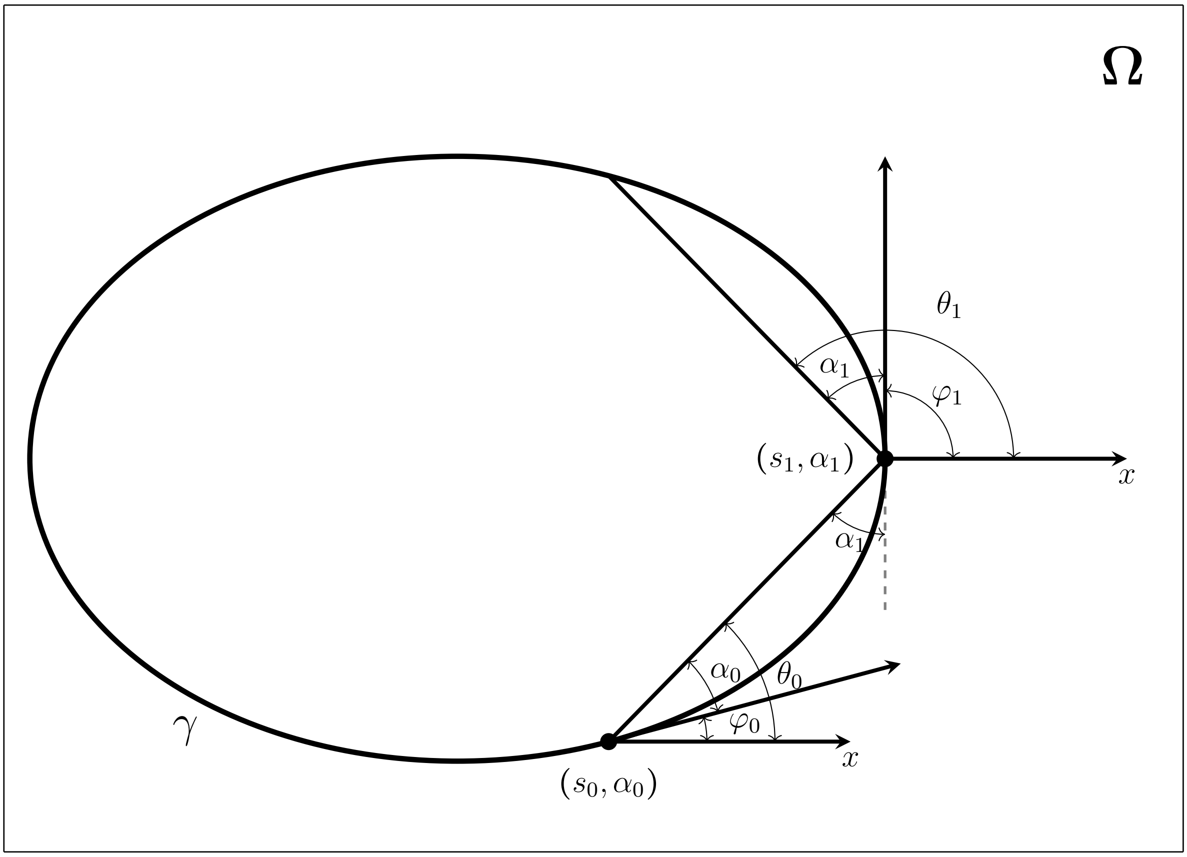
$$\alpha_1 = \varphi_1 - \theta_0. \tag{5.3}$$

Definição 5.5. Seja $T : [0, l) \times (0, \pi) \to [0, l) \times (0, \pi)$, definida por $T(s_0, \alpha_0) = (s_1, \alpha_1)$, em que l é o comprimento de uma bilhar γ. T diz-se a *aplicação bilhar* da curva γ.

[1]Uma curva está orientada positivamente (ou no sentido anti-horário) quando, ao caminhar sobre ela, a região interior à curva fica a sua esquerda.

[2]Lei da Reflexão: O ângulo de incidência é igual ao ângulo de reflexão.

Figura 2 – Bilhar.



Fonte: Elaborada pelo autor.

Teorema 5.3.1. *Seja γ uma curva fechada simples, de classe C^k ($k \geq 2$), de comprimento l, com curvatura estritamente positiva, que limita uma região simplesmente conexa $\Omega \subset \mathbb{R}^2$. Seja $T : [0, l) \times (0, \pi) \to [0, l) \times (0, \pi)$ a aplicação bilhar de γ. T é um difeomorfismo de classe C^{k-1}.*

Demonstração. Tomando as equações paramétricas de γ no parâmetro s, obtêm-se $x = f(s)$, $y = g(s)$ e $\gamma(s) = (f(s), g(s))$, em que $s \in [0, l)$. Como a curva γ agora está parametrizada por comprimento de arco, então o vetor unitário tangente à γ, $\gamma'(s) = (\frac{dx}{ds}, \frac{dy}{ds}) = (\cos(\varphi(s)), \mathrm{sen}(\varphi(s)))$, é tal que $f'(s) = \cos\varphi$ e $g'(s) = \mathrm{sen}(\varphi)$.

Tomando agora dois pontos $\tilde{P}_0 = (f(\tilde{s}_0), g(\tilde{s}_0)) \in \gamma$, $\tilde{P}_1 = (f(\tilde{s}_1), g(\tilde{s}_1)) \in \gamma$ e dois pontos $P_0 = (f(s_0), g(s_0))$, $P_1 = (f(s_1), g(s_1))$ que pertencem, respectivamente, a vizinhanças de $\tilde{P}_0$ e de $\tilde{P}_1$. Tome também, tais vizinhanças disjuntas. Suponha $f'(\tilde{s}_0) \neq 0$ e $f'(\tilde{s}_1) \neq 0$, a fim de que isto sempre ocorra, basta girar a região Ω no plano, se necessário. Suponha também $f'(s_0) \neq 0$ e $f'(s_1) \neq 0$ em vizinhanças suficientemente pequenas de $\tilde{s}_0$ e $\tilde{s}_1$. Das equações (5.2) e (5.3), segue que

$$\tan\varphi_0 = \frac{\text{sen}\varphi_0}{\cos\varphi_0} = \frac{g'(s_0)}{f'(s_0)} \quad \Rightarrow \quad \varphi_0 = \arctan\left(\frac{g'(s_0)}{f'(s_0)}\right),$$
$$\tan\varphi_0 = \frac{\text{sen}\varphi_1}{\cos\varphi_1} = \frac{g'(s_1)}{f'(s_1)} \quad \Rightarrow \quad \varphi_1 = \arctan\left(\frac{g'(s_1)}{f'(s_1)}\right)$$

e

$$\tan\theta_0 = \tan(\varphi_0 + \alpha_0) = \frac{g(s_1) - g(s_0)}{f(s_1) - f(s_0)} \quad \Rightarrow \quad \theta_0 = \arctan\left(\frac{g(s_1) - g(s_0)}{f(s_1) - f(s_0)}\right)$$

logo, ao substituir estes valores em $\alpha_0 = \theta_0 - \varphi_0$ e $\alpha_1 = \varphi_1 - \theta_0$, obtêm-se

$$\alpha_0 = \arctan\left(\frac{g(s_1) - g(s_0)}{f(s_1) - f(s_0)}\right) - \arctan\left(\frac{g'(s_0)}{f'(s_0)}\right) := L(s_0, s_1)$$

e

$$\alpha_1 = \arctan\left(\frac{g'(s_1)}{f'(s_1)}\right) - \arctan\left(\frac{g(s_1) - g(s_0)}{f(s_1) - f(s_0)}\right) := M(s_0, s_1).$$

Ao definir uma função $F : \mathbb{R}^4 \to \mathbb{R}^2$, definida da seguinte forma, $F(s_0, \alpha_0, s_1, \alpha_1) = (U(s_0, \alpha_0, s_1, \alpha_1), V(s_0, \alpha_0, s_1, \alpha_1))$, em que

$$F(s_0, \alpha_0, s_1, \alpha_1) = \begin{cases} U(s_0, \alpha_0, s_1, \alpha_1) := L(s_0, s_1) - \alpha_0 = 0 \\ V(s_0, \alpha_0, s_1, \alpha_1) := M(s_0, s_1) - \alpha_1 = 0 \\ (s_1(\tilde{s}_0, \tilde{\alpha}_0), \alpha_1(\tilde{s}_0, \tilde{\alpha}_0)) = (\tilde{s}_1, \tilde{\alpha}_1) \end{cases},$$

o Teorema da Função Implícita garante que, a fim de que existam funções $s_1 = s_1(\tilde{s}_0, \tilde{\alpha}_0)$ e $\alpha_1 = \alpha_1(\tilde{s}_0, \tilde{\alpha}_0)$, ambas de classe C^{k-1} (pois $\gamma \in C^k$ implica $f', g' \in C^{k-1}$ acarreta $U, V \in C^{k-1}$) em uma vizinhança de $(\tilde{s}_1, \tilde{\alpha}_1)$, é suficiente que a derivada de F restrita às duas últimas coordenadas, $[\partial F(\tilde{s}_0, \tilde{\alpha}_0, \tilde{s}_1, \tilde{\alpha}_1)]|_{(s_1, \alpha_1)}$, seja um isomorfismo, mais precisamente, é suficiente que seja

$$\det[JF(\tilde{s}_0, \tilde{\alpha}_0, \tilde{s}_1, \tilde{\alpha}_1)]|_{(s_1, \alpha_1)} = \begin{vmatrix} \frac{\partial U}{\partial s_1} & \frac{\partial U}{\partial \alpha_1} \\ \frac{\partial V}{\partial s_1} & \frac{\partial V}{\partial \alpha_1} \end{vmatrix} \neq 0. \tag{5.4}$$

Note que $\dfrac{\partial U}{\partial s_1} = \dfrac{\partial L}{\partial s_1}$, $\dfrac{\partial U}{\partial \alpha_1} = 0$, $\dfrac{\partial V}{\partial s_1} = \dfrac{\partial M}{\partial s_1}$ e $\dfrac{\partial V}{\partial \alpha_1} = -1$, logo, calcular o determinante (5.4) equivale à calcular o determinante

$$\det\begin{bmatrix} \frac{\partial L}{\partial s_1} & 0 \\ \frac{\partial M}{\partial s_1} & -1 \end{bmatrix} = -\frac{\partial L}{\partial s_1}.$$

Logo, para que o determinante (5.4) seja diferente de 0, deve acontecer $-\dfrac{\partial L}{\partial s_1}(\tilde{s_0},\tilde{s_1}) \neq 0$. Note que, derivando L no ponto $(\tilde{s_0},\tilde{s_1})$ em relação à s_1, segue que

$$\begin{aligned}
\frac{\partial L}{\partial s_1}(\tilde{s_0},\tilde{s_1}) = \frac{\partial}{\partial s_1}\left[\arctan\left(\frac{g(\tilde{s_1})-g(\tilde{s_0})}{f(\tilde{s_1})-f(\tilde{s_0})}\right)\right] &= \frac{\frac{g'(\tilde{s_1})\cdot[f(\tilde{s_1})-f(\tilde{s_0})]-f'(\tilde{s_1})\cdot[g(\tilde{s_1})-g(\tilde{s_0})]}{[f(\tilde{s_1})-f(\tilde{s_0})]^2}}{1+\frac{[g(\tilde{s_1})-g(\tilde{s_0})]^2}{[f(\tilde{s_1})-f(\tilde{s_0})]^2}} \\
&= \frac{\frac{g'(\tilde{s_1})f(\tilde{s_1})-g'(\tilde{s_1})f(\tilde{s_0})-f'(\tilde{s_1})g(\tilde{s_1})+f'(\tilde{s_1})g(\tilde{s_0})}{[f(\tilde{s_1})-f(\tilde{s_0})]^2}}{\frac{[f(\tilde{s_1})-f(\tilde{s_0})]^2+[g(\tilde{s_1})-g(\tilde{s_0})]^2}{[f(\tilde{s_1})-f(\tilde{s_0})]^2}} \\
&= \frac{\det\begin{bmatrix} f'(\tilde{s_1}) & g'(\tilde{s_1}) \\ f(\tilde{s_0})-f(\tilde{s_1}) & g(\tilde{s_0})-g(\tilde{s_1}) \end{bmatrix}}{[f(\tilde{s_1})-f(\tilde{s_0})]^2+[g(\tilde{s_1})-g(\tilde{s_0})]^2}. \qquad (5.5)
\end{aligned}$$

Note que, como $\tilde{P}_0 \neq \tilde{P}_1$ e o denominador de (5.5) é justamente a Distância Euclidiana entre estes dois pontos ($\|\tilde{P}_0 - \tilde{P}_1\| > 0$), então ele nunca se anula portanto, para que o determinante (5.4) de fato seja não-nulo, basta que o numerador de (5.5) seja não-nulo, ou seja

$$\det\begin{bmatrix} f'(\tilde{s_1}) & g'(\tilde{s_1}) \\ f(\tilde{s_0})-f(\tilde{s_1}) & g(\tilde{s_0})-g(\tilde{s_1}) \end{bmatrix} \neq 0,$$

mas para que ista condição se verifique, basta que os vetores $(f(\tilde{s_0})-f(\tilde{s_1}), g(\tilde{s_0})-g(\tilde{s_1}))$ e $(f'(\tilde{s_1}), g'(\tilde{s_1}))$ não sejam paralelos, o que de fato acontece, pois $\tilde{\alpha}_1 \neq 0$ e $\tilde{\alpha}_1 \neq \pi$, lembrando que $\tilde{\alpha}_1$ é o ângulo orientado entre o vetor tangente à γ no ponto $\tilde{P}_1 = (f(\tilde{s_1}), g(\tilde{s_1}))$ e o segmento que liga $\tilde{P}_0$ a $\tilde{P}_1$. Com isto, é satisfeita a hipótese do Teorema da Função Implícita a respeito do isomorfismo. Logo T é de classe C^{k-1} em uma vizinhança de $(\tilde{s}_0, \tilde{\alpha}_0)$.

A aplicação bilhar T é inversível (ANUNCIAÇÃO, 2012, P. 9), tomando $T^{-1}(s,\alpha) = T(s, \pi-\alpha)$, o caminho que foi feito para demonstrar que T é de classe C^{k-1} é exatamente o mesmo para mostrar que T^{-1} também é desta classe em uma vizinhança de $(\tilde{s}_0, \tilde{\alpha}_0)$. Como $T, T^{-1} \in C^{k-1}$ $(k \geq 2)$, em particular ambas são de classe C^1, ou seja, são diferenciáveis, logo T é um difeomorfismo de classe C^{k-1} em uma vizinhança de $(\tilde{s}_0, \tilde{\alpha}_0)$. $\square$

Estes são apenas alguns dos diversos e importantes resultados que seguem do Teorema da Função Implícita. Foram apresentados aqui, duas aplicações do Teorema dentro da Matemática e um fora da Matemática, a sua aplicabilidade é bem ampla e fornece resultados importantíssimos. Este é um Teorema que merece ser estudado mais profundamente, a pesar de que aqui foi feito apenas um estudo introdutório a seu respeito.

REFERÊNCIAS

AMORIM, Cláudia Rabelo Oliveira. **Teorema da função implícita e suas aplicações**. 2016. Dissertação (Mestrado em Matemática) - Universidade Federal de Minas Gerais, Belo Horizonte, 2016.

ANUNCIAÇÃO, Monique Rafaella. **Órbitas Periódicas em Bilhares Convexos**. 2012. Dissertação (Mestrado em Matemática) - Universidade Federal de Minas Gerais, Belo Horizonte, 2012.

GUIDORIZZI, Hamilton Luiz. **Um curso de cálculo, vol. 1**. 5. ed. Rio de Janeiro: LTC, 2013.

LIMA, Elon Lages. **Curso de análise vol. 2**. 11. ed. Rio de Janeiro: IMPA, 2010.

LIMA, Elon Lages. **Elementos de topologia geral**. 11. ed. Rio de Janeiro: Editora da Universidade de São Paulo, 1970.

LIMA, Elon Lages. **Espaços Métricos**. 3. ed. Rio de Janeiro: Instituto de Matemática Pura e Aplicada, CNPq, 1993.

MACIEL, Aldo Bezerra; LIMA, Osmundo Alves. **Introdução à Análise Real**. 2. ed. Campina Grande: EDUEPB, 2005.

O'CONNOR, J. J.; ROBERTSON, E. F. **Otto Stolz**. [*S. I.: s. n.*], 2007. Disponível em: https://mathshistory.st-andrews.ac.uk/Biographies/Stolz/. Acesso em: 19 ago. 2021.

O'CONNOR, J. J.; ROBERTSON, E. F. **René Maurice Fréchet**. [*S. I.: s. n.*], 2005. Disponível em: https://mathshistory.st-andrews.ac.uk/Biographies/Frechet/. Acesso em: 19 ago. 2021.

RUDIN, Walter. **Principles of Mathematical Analysis**. 3. ed. New York: McGraw-hill, 1964.

STEWART, James. **Cálculo, volume 2**. São Paulo: Cengage Learning, 2013.

WINTERLE, Paulo. **Vetores e Geometria Analítica**. São Paulo: Pearson Makron Books, 2000.