

# Análise Matemática C

Diego Marcon

8 de Abril de 2020

Os conceitos de função escalar e de campo vetorial são dos mais fundamentais de matemática. Estão presentes nas formulações de leis da física, aplicações à engenharia e aparecem em qualquer campo da matemática.

O objetivo destas notas de aula da disciplina de Análise Matemática C da UFRGS é estudar propriedades de diferenciação e integração de funções de uma ou mais variáveis e com valores em espaços que também podem ter uma componente ou mais. As principais referências são [3, 11, 14, 15].

## Conteúdo

<b>1 Derivada primeira</b>	<b>3</b>
1.1 Funções reais de uma variável real . . . . .	5
1.2 Funções vetoriais de uma variável real – curvas em $\mathbb{R}^m$ . . . . .	5
1.3 Funções escalares de várias variáveis . . . . .	7
<b>2 Desigualdade do Valor Médio</b>	<b>8</b>
<b>3 Segunda derivada</b>	<b>10</b>
<b>4 Fórmula de Taylor</b>	<b>11</b>
4.1 Caso especial $k = 2$ ; Hessiana . . . . .	12
<b>5 Problemas de otimização - máximos e mínimos</b>	<b>13</b>
5.1 Comentário sobre pontos críticos não degenerados . . . . .	15
<b>6 Funções Convexas</b>	<b>15</b>
6.1 Regularidade de primeira ordem . . . . .	17
6.2 Regularidade de segunda ordem – o Teorema de Alexandrov . . . . .	19
<b>7 O Teorema da Função Inversa</b>	<b>19</b>
<b>8 O Teorema da Função Implícita</b>	<b>23</b>
<b>9 Formas canônicas locais</b>	<b>26</b>
<b>10 Mergulhos</b>	<b>31</b>
<b>11 Superfícies diferenciáveis em <math>\mathbb{R}^m</math></b>	<b>31</b>
11.1 Derivada de aplicações entre superfícies . . . . .	35
<b>12 Valores regulares</b>	<b>36</b>

<b>13 Superfícies Orientáveis</b>	<b>38</b>
13.1 Orientação de espaços vetoriais . . . . .	38
13.2 Orientação de superfícies . . . . .	39
<b>14 Multiplicadores de Lagrange</b>	<b>44</b>
<b>15 Integral de Riemann em <math>\mathbb{R}^n</math></b>	<b>46</b>
15.1 Conjuntos de medida zero e o Teorema de Lebesgue . . . . .	49
15.2 Conjuntos Jordan mensuráveis . . . . .	53
15.3 Cálculo de integrais de modo iterativo . . . . .	55
15.4 Teorema de Darboux – somas de Riemann mais gerais . . . . .	57
15.5 Fórmula de mudança de variáveis . . . . .	60
15.5.1 Coordenadas polares . . . . .	64
15.5.2 Coordenadas esféricas . . . . .	64
<b>16 Formas diferenciais</b>	<b>66</b>
<b>17 Produto exterior – motivação</b>	<b>71</b>
<b>18 Produto tensorial</b>	<b>73</b>
18.1 Álgebra tensorial . . . . .	77
<b>19 Produto exterior</b>	<b>78</b>
19.1 Álgebra exterior ou álgebra de Grassmann . . . . .	81
19.2 Identificação I; Aplicações bilineares alternadas . . . . .	81
19.3 Identificação II; Aplicações multilineares alternadas . . . . .	82
<b>20 Formas diferenciais de grau <math>p</math> e a derivada exterior</b>	<b>83</b>
20.1 O <i>pullback</i> de formas diferenciais . . . . .	86
<b>21 Formas exatas e fechadas</b>	<b>89</b>
<b>22 Variedades diferenciáveis</b>	<b>91</b>
22.1 Espaço tangente . . . . .	93
22.2 Diferencial . . . . .	97
<b>23 Partições da Unidade</b>	<b>98</b>

# 1 Derivada primeira

Vamos considerar os espaços vetoriais  $\mathbb{R}^n$  e  $\mathbb{R}^m$ . Naturalmente,  $m, n \in \{1, 2, 3, \dots\}$ . Seja  $U \subseteq \mathbb{R}^n$  um conjunto aberto. Dada uma aplicação  $f : U \subseteq \mathbb{R}^n \rightarrow \mathbb{R}^m$ , dizemos que  $f$  é diferenciável em  $x_0 \in \mathbb{R}^n$  quando existe uma transformação linear<sup>1</sup> que “aproxima”  $f$  perto de  $x_0$ . Mais precisamente,  $f$  é **diferenciável** em  $x_0 \in \mathbb{R}^n$  quando existe uma transformação linear  $T : \mathbb{R}^n \rightarrow \mathbb{R}^m$  tal que

$$f(x) = f(x_0) + T(x - x_0) + r(x - x_0), \quad \text{onde} \quad \lim_{x \rightarrow x_0} \frac{|r(x - x_0)|}{|x - x_0|} = 0.$$

Nesse caso, denotamos  $T = f'(x_0)$  ou  $T = Df(x_0)$  e dizemos que  $T$  é a **derivada** de  $f$  em  $x_0$ <sup>2</sup>.

Nós também dizemos que  $f$  é **diferenciável no conjunto aberto**  $U$ , ou simplesmente diferenciável em  $U$ , quando é diferenciável em todos os pontos de  $U$ .

Podemos reescrever a definição acima de várias maneiras equivalentes:

(i)  $f$  é diferenciável em  $x_0 \in \mathbb{R}^n$ .

(ii) existe uma transformação linear  $f'(x_0) : \mathbb{R}^n \rightarrow \mathbb{R}^m$  que satisfaz

$$f(x) = f(x_0) + f'(x_0) \cdot (x - x_0) + r(x - x_0), \quad \text{onde} \quad \lim_{x \rightarrow x_0} \frac{|r(x - x_0)|}{|x - x_0|} = 0.$$

(iii) existe uma transformação linear  $f'(x_0) : \mathbb{R}^n \rightarrow \mathbb{R}^m$  que satisfaz

$$f(x_0 + h) = f(x_0) + f'(x_0) \cdot h + r(h), \quad \text{onde} \quad \lim_{h \rightarrow 0} \frac{|r(h)|}{|h|} = 0.$$

(iv) existe uma transformação linear  $f'(x_0) : \mathbb{R}^n \rightarrow \mathbb{R}^m$  tal que

$$\lim_{h \rightarrow 0} \frac{|f(x_0 + h) - f(x_0) + f'(x_0) \cdot h|}{|h|} = 0.$$

Podemos também definir a **derivada direcional** de  $f : U \subseteq \mathbb{R}^n \rightarrow \mathbb{R}^m$ , na direção de um vetor  $v \in \mathbb{R}^n$  como

$$\frac{\partial f}{\partial v}(x) := \lim_{t \rightarrow 0} \frac{f(x + tv) - f(x)}{t} \in \mathbb{R}^m,$$

quando tal limite existe. É um exercício verificar que

$$\frac{\partial f}{\partial v}(x) = f'(x) \cdot v.$$

No caso dos vetores  $v = e_j$  da base canônica de  $\mathbb{R}^n$ , são bastante comum as notações

$$\frac{\partial f}{\partial x_j}(x) = \frac{\partial f}{\partial e_j}(x) = \partial_j f(x).$$

<sup>1</sup>Na verdade, o que aproxima  $f(x)$  é a transformação afim  $f(x_0) + T(x - x_0)$ . No entanto, a derivada de  $f$  é definida como a transformação linear  $T$ .

<sup>2</sup>Rigorosamente, deveríamos mostrar que, no caso de  $f$  ser diferenciável, a transformação  $T$  que, pela definição, existe, deve ser necessariamente única. Este é um exercício da nossa primeira lista de exercícios.

Também não é difícil verificar que, em componentes,

$$\frac{\partial f}{\partial x_j}(x) = \begin{bmatrix} \frac{\partial f_1}{\partial x_j}(x) \\ \frac{\partial f_2}{\partial x_j}(x) \\ \vdots \\ \frac{\partial f_m}{\partial x_j}(x) \end{bmatrix} \in \mathbb{R}^m.$$

Lembra de Álgebra Linear que a matriz canônica de uma transformação linear é àquela associada com as bases canônicas de  $\mathbb{R}^n$  e  $\mathbb{R}^m$ . Em outras palavras, é a matriz de ordem  $m \times n$  cujas colunas são os vetores  $f'(x) \cdot e_i \in \mathbb{R}^m$ :

$$[f'(x)] = \begin{bmatrix} | & | & & | \\ f'(x) \cdot e_1 & f'(x) \cdot e_2 & \cdots & f'(x) \cdot e_n \\ | & | & & | \end{bmatrix} = \begin{bmatrix} \frac{\partial f_1}{\partial x_1}(x) & \frac{\partial f_1}{\partial x_2}(x) & \cdots & \frac{\partial f_1}{\partial x_n}(x) \\ \frac{\partial f_2}{\partial x_1}(x) & \frac{\partial f_2}{\partial x_2}(x) & \cdots & \frac{\partial f_2}{\partial x_n}(x) \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial x_1}(x) & \frac{\partial f_m}{\partial x_2}(x) & \cdots & \frac{\partial f_m}{\partial x_n}(x) \end{bmatrix} = \left[ \frac{\partial f_i}{\partial x_j}(x) \right].$$

Esta matriz é, em alguns livros, conhecida como o **Jacobiano** de  $f$  em  $x$  e denotada por  $Jf(x)$ . Alguns autores reservam o nome Jacobiano não para a matriz, mas para o determinante da matriz acima:  $Jf(x) = \det [f'(x)]$ .

Se  $f$  é diferenciável em  $U$ , está bem definida a aplicação

$$\begin{aligned} f' : U &\longrightarrow \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m) \\ x &\longmapsto f'(x) \end{aligned}$$

onde  $\mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$  é o espaço vetorial das transformações lineares de  $\mathbb{R}^n$  em  $\mathbb{R}^m$ . Utilizando as matrizes canônicas associadas às transformações lineares, obtemos um isomorfismo (não canônico)  $\mathcal{L}(\mathbb{R}^n, \mathbb{R}^m) \simeq \mathbb{R}^{m \cdot n}$  e podemos falar em continuidade e diferenciabilidade de  $f'$  da mesma forma que vínhamos fazendo.

Quando  $f$  é diferenciável em  $U$  e  $f'$  é contínua, dizemos que  $f$  é **continuamente diferenciável** em  $U$  ou que  $f$  é de classe  $C^1$  em  $U$ . Escrevemos também  $f \in C^1(U, \mathbb{R}^m)$  ou, quando não houver confusão sobre o domínio ou contra-domínio, apenas  $f \in C^1$ .

**Teorema 1.** São equivalentes:

- (i)  $f \in C^1(U, \mathbb{R}^m)$ ;
- (ii) as funções coordenadas  $f_i : U \rightarrow \mathbb{R}$  possuem todas as derivadas parciais  $\frac{\partial f_i}{\partial x_j}$  contínuas;
- (iii) para todo  $x \in U$  e todo  $v \in \mathbb{R}^n$ , as derivadas direcionais  $\frac{\partial f}{\partial v} : U \rightarrow \mathbb{R}^m$  existem e são contínuas.

**Exercício 1.** Provar o Teorema 1. De qualquer maneira, a demonstração pode ser encontrada em várias referências, como por exemplo, [10, 11, 14, 15].

## 1.1 Funções reais de uma variável real

Este é o caso mais conhecido  $n = m = 1$ , da análise na reta, em que estudamos uma função do tipo  $f : \mathbb{R} \rightarrow \mathbb{R}$ . Quando  $f$  é diferenciável em  $a$ , nossa definição diz que existe a transformação linear  $f'(a) : \mathbb{R} \rightarrow \mathbb{R}$ , satisfazendo

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

Assim, nós podemos “confundir” a transformação linear  $f'(a)$  com o número associado pelo exercício abaixo, ou (o que é a mesma coisa) com a matriz canônica de ordem  $1 \times 1$  cujo único elemento é a derivada “parcial”:

$$[f'(a)] = \left[ \frac{df}{dx}(a) \right].$$

**Exercício 2.** Mostre que se  $T \in \mathcal{L}(\mathbb{R}; \mathbb{R})$ , então existe  $A \in \mathbb{R}$  tal que  $T(x) = Ax$ . Isto é, as transformações lineares em  $\mathbb{R}$  coincidem com as funções lineares usuais.

Sendo  $\mathbb{R}$  um corpo, é possível “dividir” por elementos de  $\mathbb{R}$ , isto é, dado  $h \in \mathbb{R}$ , existe o inverso multiplicativo  $h^{-1} \in \mathbb{R}$ . Neste caso, podemos escrever, como de costume,

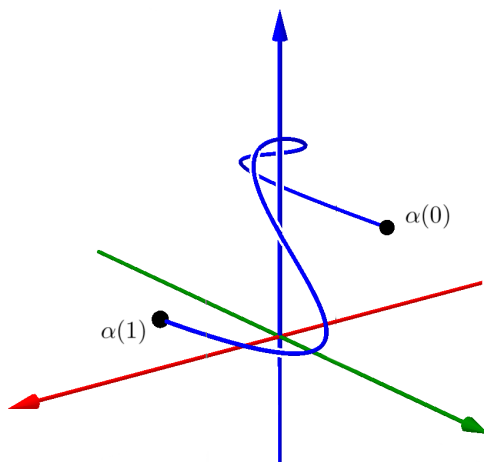
$$f'(a) = \lim_{h \rightarrow 0} \left( \frac{f(a+h) - f(a)}{h} - \frac{r(h)}{h} \right) = \lim_{h \rightarrow 0} \frac{f(a+h) - f(a)}{h}.$$

Em Cálculo I, se estuda derivada como o coeficiente da reta tangente ao gráfico da função. É justamente isto que nossa definição, com as devidas identificações, diz quando se escreve

$$f(x) \simeq f(a) + f'(a)(x - a) \quad \text{quando} \quad x \simeq a.$$

## 1.2 Funções vetoriais de uma variável real – curvas em $\mathbb{R}^m$

Quando  $n = 1$  e e.g.  $U = [0, 1]$  o gráfico da função  $\alpha : [0, 1] \rightarrow \mathbb{R}^m$  representa uma **curva** em  $\mathbb{R}^m$ .



É comum dizer que  $C = \alpha([0, 1])$  é a curva e que  $\alpha$  é uma **parametrização** da curva  $C$ . Ou ainda que  $\alpha$  é uma **curva parametrizada**. Isto é porque  $\alpha$ , conforme  $t$  cresce de 0 para 1, percorre a curva  $C$  em um sentido (ou orientação) fixado(a) e com certa velocidade, possuindo portanto mais informações. Para os nossos propósitos, vamos frequentemente confundir a terminologia e, por um abuso de linguagem, chamar  $\alpha$  de curva.

Em coordenadas (geralmente, na base canônica de  $\mathbb{R}^m$ ), podemos escrever

$$\alpha(t) = \begin{bmatrix} \alpha_1(t) \\ \alpha_2(t) \\ \vdots \\ \alpha_m(t) \end{bmatrix}.$$

No caso de ser  $\alpha$  diferenciável, segundo a nossa definição, a derivada de  $\alpha$  em um ponto  $t$  é a transformação linear  $\alpha'(t) : \mathbb{R} \rightarrow \mathbb{R}^m$ , cuja matriz canônica associada é de ordem  $m \times 1$ :

$$[\alpha'(t)] = \begin{bmatrix} \alpha'_1(t) \\ \alpha'_2(t) \\ \vdots \\ \alpha'_m(t) \end{bmatrix}. \quad (1)$$

Acima escrevemos, naturalmente, as derivadas “parciais” como

$$\alpha'_i(t) = \frac{d\alpha_i}{dt}(t).$$

Sendo  $h$  um número real (e não um vetor), é possível reescrever (1) como

$$[\alpha'(t)] = \lim_{h \rightarrow 0} \frac{\alpha(t+h) - \alpha(t)}{h}.$$

Identificando a transformação linear  $\alpha'(t)$  com a sua matriz canônica associada, podemos escrever, como de costume,

$$\alpha'(t) = \lim_{h \rightarrow 0} \frac{\alpha(t+h) - \alpha(t)}{h} \in \mathbb{R}^m. \quad (2)$$

Geometricamente, após estas identificações,  $\alpha'(t)$  é o **vetor tangente** a curva  $\alpha$  no ponto  $\alpha(t)$ . Faça um desenho dos quocientes em (2) para visualizar esta afirmação. Observe que, assim como no caso de uma função de uma variável com valores reais, a derivada de uma curva parametrizada pode ser vista como um objeto do mesmo tipo. A identificação que comentamos acima pode ser usada para considerar como equivalentes os objetos

$$\alpha' : [0, 1] \rightarrow \mathcal{L}([0, 1], \mathbb{R}^m) \quad \text{e} \quad \alpha' : [0, 1] \rightarrow \mathbb{R}^m.$$

**Exercício 3.** Observe que existe uma identificação  $\mathcal{L}(\mathbb{R}, \mathbb{R}^m) \simeq \mathbb{R}^m$  que é um isomorfismo canônico, isto é, que independe de escolha de bases para os espaços envolvidos: mostre que

$$\begin{aligned} \Phi : \mathbb{R}^m &\longrightarrow \mathcal{L}(\mathbb{R}, \mathbb{R}^m) \\ x &\longmapsto \Phi(x)(t) := tx \end{aligned}$$

é um isomorfismo de espaços vetoriais.

**Exercício 4.** Para  $\alpha : \mathbb{R} \rightarrow \mathbb{R}^m$  e  $\beta : \mathbb{R} \rightarrow \mathbb{R}^m$  curvas diferenciáveis, mostre que

$$\frac{d}{dt} \langle \alpha(t), \beta(t) \rangle = \langle \alpha'(t), \beta(t) \rangle + \langle \alpha(t), \beta'(t) \rangle \quad \forall t \in \mathbb{R}.$$

Em particular, prove que se  $\alpha$  tem velocidade constante, isto é,  $|\alpha'(t)| \equiv C$  para todo  $t$ , então

$$\alpha'(t) \perp \alpha''(t) \quad \forall t \in \mathbb{R}.$$

### 1.3 Funções escalares de várias variáveis

Este é o caso  $m = 1$ , de modo que nossos objetos de estudo são as funções da forma  $f : \mathbb{R}^n \rightarrow \mathbb{R}$ . Supondo  $f$  diferenciável, a derivada de  $f$  em  $a \in \mathbb{R}^n$  é a transformação linear  $f'(a) : \mathbb{R}^n \rightarrow \mathbb{R}$ . Observe que, neste caso particular, a derivada  $f'(a)$  é um **funcional linear**, ou seja, um elemento do espaço dual  $(\mathbb{R}^n)^*$ .

Em coordenadas canônicas, a matriz associada com  $f'(a)$  é de ordem  $1 \times n$  e dada por

$$[f'(a)] = \left[ \frac{\partial f}{\partial x_1}(a) \quad \frac{\partial f}{\partial x_2}(a) \quad \cdots \quad \frac{\partial f}{\partial x_n}(a) \right].$$

Desta forma, as derivadas direcionais são dadas pela fórmula

$$\frac{\partial f}{\partial v}(a) = f'(a) \cdot v = [f'(a)] [v] = \left[ \frac{\partial f}{\partial x_1}(a) \quad \frac{\partial f}{\partial x_2}(a) \quad \cdots \quad \frac{\partial f}{\partial x_n}(a) \right] \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} = \sum_{i=1}^n \frac{\partial f}{\partial x_i}(a) v_i. \quad (3)$$

**Observação 2 (Relação com o vetor gradiente).** Lembramos que, pelo Teorema da Representação de Riesz (ver, por exemplo, [2, Seção 8.5]), ao equiparmos  $\mathbb{R}^n$  com um produto interno  $(\cdot, \cdot)$ , existe um único vetor (que denotamos por)  $\nabla f(a) \in \mathbb{R}^n$  tal que

$$f'(a) \cdot v = (v, \nabla f(a)) \quad \forall v \in \mathbb{R}^n.$$

Uma observação fundamental é que o vetor  $\nabla f(a)$  depende do produto interno escolhido. Nós chamamos este vetor obtido pelo Teorema da Representação de Riesz de o **vetor gradiente** de  $f$  em  $a$ . No entanto, nas nossas notas, quando nos referirmos ao vetor gradiente, vamos *implicitamente assumir* que equipamos  $\mathbb{R}^n$  com o produto escalar usual:

$$\langle x, y \rangle = \sum_{i=1}^n x_i y_i,$$

onde  $x_i$  e  $y_i$  são, respectivamente, as coordenadas de  $x$  e  $y$  na base canônica de  $\mathbb{R}^n$ . Como pode ser visto por (3), o vetor  $\nabla f(a)$  em coordenadas canônicas é dado por

$$\nabla f(a) = \begin{bmatrix} \frac{\partial f}{\partial x_1}(a) \\ \frac{\partial f}{\partial x_2}(a) \\ \vdots \\ \frac{\partial f}{\partial x_n}(a) \end{bmatrix}$$

e podemos reescrever (3) como

$$f'(a) \cdot v = \langle v, \nabla f(a) \rangle. \quad (4)$$

Reforçamos que, enquanto  $f'(a) \in (\mathbb{R}^n)^*$  é um funcional linear que *independe* de qualquer estrutura geométrica,  $\nabla f(a) \in \mathbb{R}^n$  é um vetor que *depende* da escolha de produto interno utilizada.  $\triangleleft$

**Exercício 5.** Um campo vetorial em  $\mathbb{R}^3$  é uma função vetorial  $F : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ . Dizemos que o campo  $F$  é um **campo gradiente** quando existe uma função escalar  $\phi : \mathbb{R}^3 \rightarrow \mathbb{R}$  tal que  $F = -\nabla \phi$ .

Em um sistema mecânico clássico cujo campo de forças depende apenas da posição, uma trajetória é uma curva  $\alpha : I \rightarrow \mathbb{R}^3$  que satisfaz a Equação de Newton

$$m\alpha''(t) = F(\alpha(t)). \quad (5)$$

Mostre que, se  $F$  é um campo gradiente, então há conservação de energia, isto é, a função

$$E(t) = \frac{m|\alpha'(t)|^2}{2} + \phi(\alpha(t))$$

é constante. Por esta razão, campos gradiente também são conhecidos como **campos conservativos**.

Outra forma de enunciar este princípio de conservação de energia é o seguinte: o funcional energia definido como

$$E(x, v) = \frac{m|v|^2}{2} + \phi(x)$$

é constante *ao longo das trajetórias*  $t \mapsto (\alpha(t), \alpha'(t))$  do sistema.

*Dica:* Fazer o produto interno com  $\alpha'(t)$  em (5), utilizar o Exercício 4 e o Teorema Fundamental do Cálculo para funções reais de uma variável real.

No contexto de funções escalares, é comum a notação  $df(a) = f'(a) \in (\mathbb{R}^n)^*$ . É também comum chamar  $df(a)$  de o **diferencial** de  $f$  em  $a$ . Vamos escrever  $df(a)$  em termos da base canônica do espaço dual  $(\mathbb{R}^n)^*$ .

Lembramos da Álgebra Linear que, se  $\mathcal{B} = \{e_1, e_2, \dots, e_n\}$  é a base canônica de  $\mathbb{R}^n$ , podemos definir a **base dual**  $\mathcal{B}^* = \{dx_1, dx_2, \dots, dx_n\}$ . Naturalmente, cada um dos elementos da base dual é um funcional linear  $dx_i : \mathbb{R}^n \rightarrow \mathbb{R}$ , que é definido<sup>3</sup>, nos elementos de  $\mathcal{B}$ , por

$$dx_i(e_j) = \delta_{ij} = \begin{cases} 1, & \text{se } i = j \\ 0, & \text{se } i \neq j \end{cases} \quad (6)$$

Desta forma,

$$v = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} \rightsquigarrow dx_i(v) = v_i. \quad (7)$$

Então, relembando (3), temos que

$$df(a) \cdot v = \sum_{i=1}^n \frac{\partial f}{\partial x_i}(a) v_i = \sum_{i=1}^n \frac{\partial f}{\partial x_i}(a) dx_i(v) = \left( \sum_{i=1}^n \frac{\partial f}{\partial x_i}(a) dx_i \right) (v).$$

Portanto, o diferencial de  $f$  em  $a$  pode ser representado, na base dual de  $(\mathbb{R}^n)^*$ , por

$$df(a) = \sum_{i=1}^n \frac{\partial f}{\partial x_i}(a) dx_i.$$

## 2 Desigualdade do Valor Médio

Lembra que, em uma variável, vale o

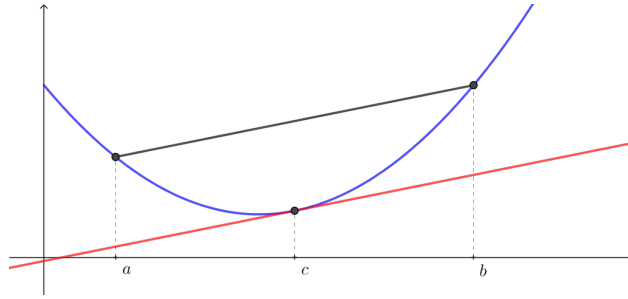
**Teorema 3 (Teorema do Valor Médio).** *Se  $f : U \rightarrow \mathbb{R}$  é diferenciável em  $(a, b) \subset U$ , então existe  $c \in (a, b)$  tal que*

$$f(b) - f(a) = f'(c)(b - a).$$

Geometricamente, o Teorema do Valor Médio afirma, dada a inclinação da reta que passa pelos pontos  $(a, f(a))$  e  $(b, f(b))$ , existe um valor  $c$  no interior do intervalo  $(a, b)$  cuja reta tangente tem essa inclinação.

<sup>3</sup>Nota, no entanto, que a notação utilizada é coerente com o seguinte: ao denotarmos por  $x_i : \mathbb{R}^n \rightarrow \mathbb{R}$  o funcional que associa com um vetor  $x$  a sua  $i$ -ésima coordenada na base canônica, temos que o seu diferencial é dado por (7).





A versão moderna deste resultado aparece no trabalho de Cauchy em 1823, embora versões preliminares tenham aparecido anteriormente (ver [wikipedia.org/meanvalue](http://wikipedia.org/meanvalue)).

A versão em dimensões maiores deste resultado é uma desigualdade, que, como veremos, segue da definição da norma de operadores (ou, se preferir, no caso  $n = 1$  é a desigualdade de Cauchy-Schwarz). O resultado seguinte segue a filosofia de passar para o caso de uma variável real por meio de um caminho unidimensional (por exemplo, um segmento).

**Teorema 4 (Desigualdade do Valor Médio).** *Seja  $f : U \subseteq \mathbb{R}^n \rightarrow \mathbb{R}^m$  uma aplicação diferenciável no aberto  $U$  e sejam  $x, y \in U$  de modo que o segmento  $[x, y]$  também está contido em  $U$ . Então, existe  $c \in [x, y]$  tal que*

$$|f(y) - f(x)| \leq \|f'(c)\| \cdot |y - x|. \quad (8)$$

*Demonstração.* Inicialmente mostramos o caso especial  $n = 1$ . Neste caso, podemos supor que  $U = [a, b] \subseteq \mathbb{R}$ . Definimos  $\phi : U \rightarrow \mathbb{R}$  por

$$\phi(t) = \langle f(t), f(y) - f(x) \rangle.$$

Pelo Teorema do Valor Médio (Teorema 3 acima) e pela Desigualdade de Cauchy-Schwarz, existe  $c \in (x, y)$  tal que

$$\begin{aligned} |f(y) - f(x)|^2 &= \phi(y) - \phi(x) \\ &= \phi'(c)(y - x) \\ &= \langle f'(c), f(y) - f(x) \rangle (y - x) \\ &\leq |f'(c)| \cdot |f(y) - f(x)| \cdot |y - x|, \end{aligned}$$

que é equivalente a (8) no caso  $n = 1$ .

Para o caso geral, consideramos uma parametrização do segmento entre  $x$  e  $y$  e recaímos no caso que já mostramos: definimos  $\alpha : [0, 1] \rightarrow \mathbb{R}^m$  por

$$\alpha(t) = f((1 - t)x + ty).$$

Assim, existe  $t_0 \in (0, 1)$  tal que  $|\alpha(1) - \alpha(0)| \leq |\alpha'(t_0)|$ . Pela regra da cadeia, isto é o mesmo que

$$|f(y) - f(x)| \leq |f'((1 - t_0)x + t_0y)| \cdot |y - x|.$$

O resultado segue da definição da norma de operadores, definindo  $c = (1 - t_0)x + t_0y$ . □

**Exercício 6.** Prove o Teorema 4 diretamente por integração. É necessário adicionar alguma hipótese para formalizar esta demonstração?

**Exercício 7.** Verifique que o Teorema do Valor Médio se estende (com igualdade) para funções escalares  $f : \mathbb{R}^n \rightarrow \mathbb{R}$ .

### 3 Segunda derivada

Considere  $f : U \rightarrow \mathbb{R}^m$  uma função diferenciável. Nós dizemos que  $f$  é **duas vezes diferenciável** em  $x_0$  quando  $f' : U \rightarrow \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$  é diferenciável em  $x_0$ . Mais explicitamente, quando existe uma transformação linear (que denotaremos por)

$$f''(x_0) : \mathbb{R}^n \longrightarrow \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m). \quad (9)$$

que satisfaz

$$\lim_{h \rightarrow 0} \frac{\|f'(x_0 + h) - f'(x_0) - f''(x_0) \cdot h\|}{|h|} = 0.$$

Observe que a norma do numerador acima é a norma de operadores em  $\mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$ :

$$T \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m) \rightsquigarrow \|T\| := \sup_{|x|=1} |T(x)|.$$

Por outro lado, poderíamos ter utilizado qualquer norma em  $\mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$  pois, em dimensão finita, são todas equivalentes.

**Exercício 8.** Defina

$$\mathcal{B}(\mathbb{R}^n; \mathbb{R}^m) = \{B : \mathbb{R}^n \times \mathbb{R}^n \longrightarrow \mathbb{R}^m; B \text{ é bilinear}\}$$

e mostre que

$$\begin{aligned} \Phi : \mathcal{L}(\mathbb{R}^n, \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)) &\longrightarrow \mathcal{B}(\mathbb{R}^n; \mathbb{R}^m) \\ T &\longmapsto \Phi(T)[x, y] := T(x)(y) \end{aligned}$$

é um isomorfismo de espaços vetoriais.

O exercício acima mostra que podemos considerar  $f''(x_0) \in \mathcal{B}(\mathbb{R}^n; \mathbb{R}^m)$ , ou seja,  $f''(x_0)$  pode tanto ser vista como uma aplicação linear do tipo (9) quanto como uma transformação bilinear do tipo

$$f''(x_0) : \mathbb{R}^n \times \mathbb{R}^n \longrightarrow \mathbb{R}^m.$$

Qualquer destas interpretações é chamada de **derivada segunda** de  $f$  em  $x_0$ . Tem-se, usando a identificação do Exercício 8,

$$f''(x_0)(v, w) = f''(x_0)(v)(w) = \frac{\partial}{\partial w} (f'(x_0)(v)) = \frac{\partial}{\partial w} \left( \frac{\partial f}{\partial v} \right) (x_0).$$

Uma notação mais comum talvez seja

$$f''(x_0)(v, w) = \frac{\partial^2 f}{\partial w \partial v} (x_0).$$

**Exercício 9.** Escrever  $f''(x_0)(v, w)$  em coordenadas. Dados  $v = \sum_{i=1}^n v^i e_i$  e  $w = \sum_{i=1}^n w^i e_i$ , obter as  $m$  componentes de  $f''(x_0)(v, w) \in \mathbb{R}^m$ .

Prosseguindo como acima, define-se a terceira derivada em  $x_0$  como uma forma trilinear e assim por diante. Além disso, dizemos que

- $f$  é de classe  $C^0$  quando é contínua;
- $f$  é de classe  $C^1$  quando  $f$  é diferenciável e  $f'$  é contínua;
- $f$  é de classe  $C^2$  quando  $f$  e  $f'$  são diferenciáveis e  $f''$  é contínua;

⋮

- $f$  é de classe  $C^\infty$  quando existem as derivadas de todas as ordens de  $f$ .

A diferenciabilidade de composições de funções acontece de acordo com a **regra da cadeia**: para  $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$  e  $g : \mathbb{R}^m \rightarrow \mathbb{R}^p$  diferenciáveis,

$$(g \circ f)'(x_0) = g'(f(x_0)) \cdot f'(x_0),$$

ou, em componentes, para  $i = 1, 2, \dots, p$  e  $j = 1, 2, \dots, n$ ,

$$\frac{\partial}{\partial x_j} g_i(f(x_0)) = \sum_{k=1}^m \frac{\partial g_i}{\partial y_k}(f(x_0)) \frac{\partial f_k}{\partial x_j}(x_0).$$

## 4 Fórmula de Taylor

A Fórmula de Taylor unidimensional nos diz que se  $f : \mathbb{R} \rightarrow \mathbb{R}$  é uma função de classe  $C^{k+1}$ , então existe  $\xi$  no intervalo aberto entre  $a$  e  $x$  tal que

$$f(x) = f(a) + f'(a)(x-a) + \frac{f''(a)}{2!}(x-a)^2 + \dots + \frac{f^{(k)}(a)}{k!}(x-a)^k + \frac{f^{(k+1)}(\xi)}{(k+1)!}(x-a)^{k+1}.$$

Vamos obter uma versão multidimensional deste resultado.

**Teorema 5 (Fórmula de Taylor).** *Suponhamos que  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  é uma função de classe  $C^{k+1}$  em uma bola  $B_r(a)$ , então, para  $x \in B_r(a) \subseteq \mathbb{R}^n$ , existe  $\xi$  no segmento de  $a$  para  $x$  tal que*

$$f(x) = f(a) + f'(a) \cdot (x-a) + \dots + \frac{1}{k!} f^{(k)}(a) \cdot (x-a)^k + \frac{1}{(k+1)!} f^{(k+1)}(\xi) \cdot (x-a)^{k+1}. \quad (10)$$

A notação do Teorema deveria ser, mais precisamente,

$$f^{(j)}(a)(x-a, x-a, \dots, x-a) \text{ ao invés de } f^{(j)}(a) \cdot (x-a)^j,$$

pois estamos pensando em  $f^{(j)}(a)$  como uma aplicação  $j$ -linear.

*Demonstração.* Como já fizemos na prova da Desigualdade do Valor Médio, vamos nos restringir ao segmento de reta que une  $a$  a  $x$  e aplicar o caso conhecido unidimensional. Já que queremos considerar  $t = 0$  e  $t = 1$ , escolhemos  $\delta > 0$  adequado e definimos  $\phi : [-\delta, 1 + \delta] \rightarrow \mathbb{R}$  por

$$\phi(t) = f((1-t)a + tx) = f(a + t(x-a)).$$

A fórmula de Taylor unidimensional implica que existe  $\tau \in (0, 1)$  de modo que

$$\phi(1) = \phi(0) + \phi'(0) + \dots + \frac{1}{k!} \phi^{(k)}(0) + \frac{1}{(k+1)!} \phi^{(k+1)}(\tau). \quad (11)$$

O resultado segue da Regra da Cadeia, com  $\xi = a + \tau(x-a)$ . □

Para escrever mais explicitamente a fórmula de Taylor em termos de derivadas parciais de  $f$ , vamos introduzir uma notação adicional. Um **multi-índice** é um elemento  $\alpha \in \mathbb{N}^n$  que, para os nossos propósitos, representa o número de derivadas parciais com respeito à variável que está na respectiva posição. Por exemplo, para  $f : \mathbb{R}^5 \rightarrow \mathbb{R}$ , denotamos

$$\frac{\partial^5 f}{\partial x_3^3 \partial x_2 \partial x_4}(a) = D^\alpha f(a), \quad \text{onde } \alpha = (0, 3, 2, 1, 0).$$

Também é comum utilizar multi-índices para denotar produtos das componentes de um vetor:

$$x^\alpha = x_2^3 x_3^2 x_4, \quad \text{onde } \alpha = (0, 3, 2, 1, 0).$$

De maneira mais geral, escrevemos

$$x^\alpha = x_1^{\alpha_1} x_2^{\alpha_2} \cdots x_n^{\alpha_n} \quad \text{e} \quad D^\alpha f(a) = \frac{\partial^{\alpha_1}}{\partial x_1^{\alpha_1}} \frac{\partial^{\alpha_2}}{\partial x_2^{\alpha_2}} \cdots \frac{\partial^{\alpha_n}}{\partial x_n^{\alpha_n}} f(a) = \frac{\partial^{|\alpha|} f}{\partial x_1^{\alpha_1} \cdots \partial x_n^{\alpha_n}}(a).$$

É também comum escrever (como já o fizemos acima)

$$|\alpha| = \alpha_1 + \alpha_2 + \cdots + \alpha_n, \quad \alpha! = \alpha_1! \alpha_2! \cdots \alpha_n! \quad \text{e} \quad \binom{n}{\alpha} = \frac{n!}{\alpha!} = \frac{n!}{\alpha_1! \alpha_2! \cdots \alpha_n!}.$$

**Exercício 10 (Teorema Multinomial).** Mostre que, para  $N \in \mathbb{N}$  fixado e  $x_1, x_2, \dots, x_k \in \mathbb{R}$ , vale

$$(x_1 + x_2 + \cdots + x_k)^N = \sum_{\alpha_1 + \alpha_2 + \cdots + \alpha_k = N} \frac{N!}{\alpha_1! \alpha_2! \cdots \alpha_k!} x_1^{\alpha_1} x_2^{\alpha_2} \cdots x_k^{\alpha_k}.$$

Na notação introduzida acima, podemos escrever, mais sucintamente,

$$(x_1 + x_2 + \cdots + x_k)^N = \sum_{|\alpha|=N} \binom{N}{\alpha} x^\alpha, \quad \text{onde} \quad x = (x_1, x_2, \dots, x_k).$$

Dica: Mostrar o caso  $k = 2$ , conhecido como binômio de Newton, e depois fazer indução em  $k$ .

Agora, vamos utilizar o Teorema Multinomial para calcular (11) mais explicitamente. Iniciamos observando que

$$\phi'(t) = \frac{d}{dt} f(a + tv) = v_1 \frac{\partial f}{\partial x_1} + v_2 \frac{\partial f}{\partial x_2} + \cdots + v_n \frac{\partial f}{\partial x_n} = \left( v_1 \frac{\partial}{\partial x_1} + v_2 \frac{\partial}{\partial x_2} + \cdots + v_n \frac{\partial}{\partial x_n} \right) f.$$

Logo, utilizando o Teorema Multinomial,

$$\phi^{(j)}(t) = \left( v_1 \frac{\partial}{\partial x_1} + v_2 \frac{\partial}{\partial x_2} + \cdots + v_n \frac{\partial}{\partial x_n} \right)^j f = \sum_{|\alpha|=j} \frac{j!}{\alpha!} v^\alpha D^\alpha f(a + tv).$$

Portanto, (10) é o mesmo que

$$f(x) = \sum_{|\alpha| \leq k} \frac{1}{\alpha!} (x - a)^\alpha D^\alpha f(a) + \sum_{|\alpha|=k+1} \frac{1}{\alpha!} (x - a)^\alpha D^\alpha f(\xi).$$

#### 4.1 Caso especial $k = 2$ ; Hessiana

Para a maior parte das aplicações, é suficiente escrever a fórmula de Taylor até ordem 2 apenas. Além disso, observamos que, a partir da forma bilinear  $f''(a) : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$ , podemos definir uma forma quadrática  $Q(v) = f''(a)(v, v)$ . Esta forma quadrática é conhecida como a **Hessiana** de  $f$  em  $a$ . Vejamos qual é a matriz simétrica associada à Hessiana: escrevendo  $v$  na base canônica,

$$Q(v) = f''(a)(v, v) = \frac{\partial^2 f}{\partial v^2}(a) = \sum_{i=1}^n \frac{\partial^2 f}{\partial x_i \partial x_j}(a) v_i v_j = \langle \nabla^2 f(a) v, v \rangle = v^T \nabla^2 f(a) v,$$

onde  $\nabla^2 f(a)$  é a matriz das derivadas parciais de ordem 2 de  $f$ :

$$\nabla^2 f(a) = \left[ \frac{\partial^2 f}{\partial x_i \partial x_j}(a) \right]_{n \times n}.$$

Por abuso de linguagem,  $\nabla^2 f(a)$  é também conhecida como a Hessiana de  $f$  em  $a$ .

Com esta nova notação, enunciaremos um caso particular da Fórmula de Taylor que provamos:

**Corolário 6.** Suponhamos que  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  é uma função de classe  $C^3$  em uma bola  $B_r(a)$ , então, para  $x \in B_r(a) \subseteq \mathbb{R}^n$ , existe  $\xi$  no segmento de  $a$  para  $x$  tal que

$$f(x) = f(a) + \nabla f(a) \cdot (x - a) + \frac{1}{2}(x - a)^T \nabla^2 f(a)(x - a) + r(x - a), \quad (12)$$

onde

$$\lim_{x \rightarrow a} \frac{r(x - a)}{|x - a|^2} = 0.$$

**Exercício 11.** Verifique que é válido o corolário acima. É realmente necessário  $f \in C^3$ ?

## 5 Problemas de otimização - máximos e mínimos

Nesta seção, estudamos condições necessárias e suficientes para que um ponto  $a \in \mathbb{R}^n$  seja um minimizante (ou maximizante) para uma função escalar  $f : \mathbb{R}^n \rightarrow \mathbb{R}$ . Começamos com uma condição necessária de ordem um.

**Proposição 7.** Suponhamos que  $a \in \text{int } U$  é um ponto de mínimo local de  $f : U \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$  e que  $f$  é diferenciável em  $a$ . Então,  $df(a) = 0$ .

*Demonstração.* Utilizando que  $a$  é mínimo local e a definição de diferenciabilidade, temos que para todo  $v$  próximo da origem:

$$f(a) \leq f(a + v) = f(a) + df(a) \cdot v + r(v) \quad \text{e também} \quad \frac{r(v)}{|v|} \rightarrow 0.$$

Dividindo pelo comprimento de  $v$  e fazendo  $v \rightarrow 0$ , concluímos que  $df(a) \cdot w \geq 0$  para todo  $w$  unitário. Em particular, podemos trocar  $w$  por  $-w$  e segue que  $df(a) = 0$ .  $\square$

Um ponto  $a \in \mathbb{R}^n$  é dito **ponto crítico** de  $f$  quando  $df(a) = 0$ . Assim, o teorema acima afirma que todo extremizante de  $f$  é ponto crítico. A recíproca não é verdadeira:  $a = 0$  é ponto crítico de  $f(x) = x^3$  que não é de mínimo nem de máximo.

Lembramos que uma matriz simétrica  $A$  de ordem  $n \times n$  é dita **positiva definida** quando

$$\langle Av, v \rangle \geq 0 \quad \text{para todo } v \in \mathbb{R}^n.$$

É comum a notação  $A \geq 0$ . A matriz  $A$  é dita **estritamente positiva definida**, e se escreve  $A > 0$ , quando

$$\langle Av, v \rangle > 0 \quad \text{para todo } v \in \mathbb{R}^n.$$

**Proposição 8.** Suponhamos que  $a \in \text{int } U$  é um ponto de mínimo local de  $f : U \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$  e que  $f$  é duas vezes diferenciável em  $a$ . Então, a matriz Hessiana  $\nabla^2 f(a)$  é positiva definida:

$$\langle \nabla^2 f(a)v, v \rangle \geq 0 \quad \text{para todo } v \in \mathbb{R}^n.$$

*Demonstração.* Pela proposição anterior,  $\nabla f(a) = 0$ . Logo, pela fórmula de Taylor, temos que

$$f(a) \leq f(a + h) = f(a) + \frac{1}{2} \langle \nabla^2 f(a) \cdot h, h \rangle + r(h),$$

para todo  $h \in B_\delta(0)$  próximo de zero (já que  $a$  é um ponto de mínimo local). Segue que

$$\langle \nabla^2 f(a) \cdot h, h \rangle + 2r(h) \geq 0.$$

Dividindo por  $|h|^2$  e fazendo  $h \rightarrow 0$ , concluímos que

$$\langle \nabla^2 f(a)w, w \rangle \geq 0 \quad \text{para todo vetor unitário } w \in \mathbb{R}^n.$$

Para concluir a demonstração, basta fazer  $w = v/|v|$  e multiplicar por  $|v|^2$ .  $\square$

**Exercício 12.** A recíproca da Proposição 8 não é verdadeira; encontre um contra-exemplo. Enuncie e prove resultados análogos às duas proposições acima para pontos de máximo local.

Se tivermos uma matriz Hessiana estritamente positiva, daí sim a recíproca é verdadeira.

**Proposição 9.** Seja  $f : U \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$  de classe  $C^2$  em  $U$ . Suponhamos que

- $\nabla f(a) = 0$  e
- $\nabla^2 f(a) > 0$ .

Então,  $a$  é um ponto de mínimo local estrito para  $f$ .

*Demonstração.* Assim como na Proposição 8 acima, utilizamos a Fórmula de Taylor:

$$f(a+h) = f(a) + \frac{1}{2} \langle \nabla^2 f(a) \cdot h, h \rangle + r(h) \quad \text{onde} \quad \lim_{h \rightarrow 0} \frac{r(h)}{|h|^2} = 0.$$

Dividindo por  $|h|^2$  podemos escrever

$$\frac{f(a+h) - f(a)}{|h|^2} = \frac{1}{2} \left\langle \nabla^2 f(a) \cdot \frac{h}{|h|}, \frac{h}{|h|} \right\rangle + \rho(h) \quad \text{onde} \quad \lim_{h \rightarrow 0} \rho(h) = 0.$$

A aplicação  $w \mapsto \langle \nabla^2 f(a) \cdot w, w \rangle$  assume um mínimo positivo na esfera unitária, de modo que, para algum  $C > 0$ , vale que

$$\frac{f(a+h) - f(a)}{|h|^2} \geq C + \rho(h) \quad \text{onde} \quad \lim_{h \rightarrow 0} \rho(h) = 0.$$

Agora, existe  $\delta > 0$  tal que

$$h \in B_\delta(0) \implies C + \rho(h) > 0,$$

concluindo a demonstração. □

**Exercício 13.** Adaptar o raciocínio acima para pontos de máximo local de  $f$ .

**Proposição 10.** Seja  $f : U \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$  de classe  $C^2$  em  $U$ . Suponhamos que

- $\nabla f(a) = 0$  e
- $\nabla^2 f(a)$  é **indefinida**:  $\langle \nabla^2 f(a)v, v \rangle$  muda de sinal.

Então,  $a$  não é nem um ponto de mínimo local nem um ponto de máximo local para  $f$ .

*Demonstração.* O argumento utilizado nas demonstrações da Proposição 9 e do Exercício 13 mostra que, para  $h$  suficientemente próximo da origem,

$$\frac{f(a+h) - f(a)}{|h|^2} \text{ tem o mesmo sinal de } \left\langle \nabla^2 f(a) \cdot \frac{h}{|h|}, \frac{h}{|h|} \right\rangle.$$

Isto conclui a prova da proposição. □

É um exercício de Álgebra Linear mostrar que uma matriz  $A$  é positiva definida se, e somente se, todos os seus autovalores são positivos.

**Exercício 14.** Faça o exercício!

Mas pode ser bastante trabalhoso encontrar todos os autovalores de  $A$ . Um método indireto consiste em completar quadrados até se “livrar” dos termos mistos da forma quadrática  $\langle Av, v \rangle$ . Vejamos um exemplo.

**Exemplo 11.** Considere

$$A = \begin{bmatrix} 1 & 2 & 0 \\ 2 & 1 & 1 \\ 0 & 1 & -2 \end{bmatrix}$$

de modo que

$$\langle Ax, x \rangle = x_1^2 + x_2^2 - 2x_3^2 + 4x_1x_2 + 2x_2x_3.$$

Juntando os termos que têm  $x_1$ , obtemos

$$x_1^2 + 4x_1x_2 = (x_1 + 2x_2)^2 - 4x_2^2 =: a^2 - 4x_2^2 \quad \text{onde} \quad a = x_1 + 2x_2.$$

Logo,

$$\langle Ax, x \rangle = a^2 - 3x_2^2 - 2x_3^2 + 2x_2x_3.$$

Em seguida, juntando os termos com  $x_2$ :

$$-3x_2^2 + 2x_2x_3 = -3 \left( x_2 + \frac{2}{3}x_3 \right)^2 + \frac{x_3^2}{3} =: -3b^2 + \frac{x_3^2}{3} \quad \text{onde} \quad b = x_2 + \frac{x_3}{3}.$$

Denotando  $c = x_3$ , obtemos

$$\langle Ax, x \rangle = a^2 - 3b^2 + \frac{c^2}{3}.$$

Desta forma, é fácil ver que  $\langle Ax, x \rangle$  assume tanto valores positivos quanto negativos; basta escolher alguns valores baseados nos sinais que aparecem nos termos ao quadrado. Por exemplo, quando  $b = 0$ , temos valores positivos. Quando  $a = c = 0$ , temos valores negativos. Portanto,  $A$  é uma matriz indefinida.

## 5.1 Comentário sobre pontos críticos não degenerados

Seja  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  de classe  $C^2$ . Dizemos que um ponto crítico  $a$  de  $f$  é **não degenerado** quando a matriz Hessiana de  $f$  em  $a$  é invertível.

**Proposição 12.** *Seja  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  de classe  $C^2$ . O conjunto dos pontos críticos não degenerados de  $f$  é isolado em  $\mathbb{R}^n$ .*

*Demonstração.* Considere a aplicação “gradiente”  $g : \mathbb{R}^n \rightarrow \mathbb{R}^n$  dada por

$$g(x) = \nabla f(x), \quad x \in \mathbb{R}^n.$$

Se  $a \in \mathbb{R}^n$  é um ponto crítico não degenerado, temos  $g'(a) = \nabla^2 f(a)$  invertível, de modo que, pelo Teorema da Função Inversa (ou diretamente, o Passo 1 da prova do Teorema da Função Inversa que demonstraremos em breve já mostra que)  $g$  é injetiva em uma vizinhança de  $a$ . Em particular,  $g(x) \neq g(a) = 0$  nesta vizinhança de  $a$ . Logo,  $a$  é o único ponto crítico de  $f$  nesta vizinhança.  $\square$

## 6 Funções Convexas

Funções convexas são dos objetos mais importantes em matemática, principalmente em problemas de otimização e de cálculo de variações. A referência mais clássica para o assunto é Rockafellar [13]. O material desta seção é uma ampliação de Villani [16, Subseção 2.1.3], que já é bastante inspirado em [13]. No entanto, não vamos nos aprofundar nos tópicos que necessitam de Teoria da Medida, para não fugir dos pré-requisitos do nosso curso.

Uma função  $f : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$  é uma **função convexa própria**<sup>4</sup> quando  $f$  não é identicamente  $+\infty$  e, para todo  $x, y \in \mathbb{R}^n$  e para todo  $0 < \lambda < 1$ , tem-se

$$f[(1 - \lambda)x + \lambda y] \leq (1 - \lambda)f(x) + \lambda f(y). \quad (13)$$

<sup>4</sup>No contexto de Análise Convexa, a palavra *própria* indica que  $f > -\infty$  em todo ponto e que  $f(x) < +\infty$  para pelo menos um ponto [13, Seção 4].

Quando a desigualdade é estrita em (13), dizemos que  $f$  é **estritamente convexa**.

O **domínio efetivo** de uma função convexa  $f$  é o conjunto

$$\text{dom } f := \{x \in \mathbb{R}^n; f(x) < +\infty\}.$$

É fácil de ver que  $\text{dom } f$  é um conjunto convexo (exercício). É também fácil ver que:

**Exercício 15 (Desigualdade de Jensen - versão discreta).** A função  $f : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$  é convexa se, e somente se, para todo  $\lambda_1 \geq 0, \lambda_2 \geq 0, \dots, \lambda_m \geq 0$  com  $\lambda_1 + \lambda_2 + \dots + \lambda_m = 1$ , temos

$$f[\lambda_1 x_1 + \lambda_2 x_2 + \dots + \lambda_m x_m] \leq \lambda_1 f(x_1) + \lambda_2 f(x_2) + \dots + \lambda_m f(x_m).$$

Uma propriedade geométrica que segue da definição é que o gráfico de uma função convexa está “acima” do plano tangente em cada um dos pontos de diferenciabilidade:

**Proposição 13.** Se  $f : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$  é uma função convexa própria que é diferenciável em  $x \in \text{dom } f$ , então

$$f(y) \geq f(x) + \langle \nabla f(x), y - x \rangle,$$

para todo  $y \in \mathbb{R}^n$ . Em particular,  $\nabla f$  é monótono em pontos de diferenciabilidade:

$$\langle \nabla f(y) - \nabla f(x), y - x \rangle \geq 0.$$

*Demonstração.* Reescrevendo (13) como

$$\frac{f[x + \lambda(y - x)] - f(x)}{\lambda} \leq f(y) - f(x)$$

e mandando  $\lambda \rightarrow 0$ , obtemos o resultado. □

Na verdade, esta propriedade de o gráfico estar acima dos planos tangentes caracteriza as funções convexas diferenciáveis:

**Proposição 14.** Suponhamos que  $f : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$  é uma função de classe  $C^1$  que satisfaz

$$f(y) \geq f(x) + \langle \nabla f(x), y - x \rangle,$$

para todo  $x, y \in \mathbb{R}^n$ . Então  $f$  é uma função convexa.

*Demonstração.* É sempre verdade que (como se vê tomando  $x = y$ )

$$f(y) \leq \sup_{x \in \mathbb{R}^n} [f(x) + \langle \nabla f(x), y - x \rangle].$$

Logo, nossa hipótese garante que, na verdade,

$$f(y) = \sup_{x \in \mathbb{R}^n} [f(x) + \langle \nabla f(x), y - x \rangle].$$

O resultado segue da observação que o supremo pontual de uma família qualquer de funções convexas é também uma função convexa. □

**Exercício 16.** Mostre que o supremo pontual de uma família qualquer de funções convexas é também uma função convexa, isto é, dadas  $f_\lambda$  convexas, temos que

$$f(y) := \sup_{\lambda \in \Lambda} f_\lambda(y)$$

é convexa.

Em seguida, mostramos outra caracterização de funções convexas, mas com derivadas de ordem dois.



**Proposição 15.** *Seja  $f : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$  é uma função de classe  $C^2$ . Então  $f$  é convexa se, e somente se,  $\nabla^2 f(x)$  é positiva semi-definida, para todo  $x \in \text{dom } f$ , isto é,*

$$\langle \nabla^2 f(x)y, y \rangle \geq 0,$$

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

*Demonstração.*  $\boxed{\implies}$  Como  $x = \frac{x+ty}{2} + \frac{x-ty}{2}$ , a definição de função convexa implica que

$$2f(x) \leq f(x+ty) + f(x-ty).$$

Logo,

$$\frac{f(x+ty) + f(x-ty) - 2f(x)}{t^2} \geq 0.$$

Fazendo  $t \rightarrow 0$ , obtemos o que queríamos.

$\boxed{\impliedby}$  Pelo Teorema do Valor Médio, para todo  $x, y \in \mathbb{R}^n$ ,

$$f(y) - f(x) = \langle \nabla f(x), y - x \rangle + \frac{\langle \nabla^2 f(\xi)(y-x), y-x \rangle}{2} \geq \langle \nabla f(x), y - x \rangle.$$

Uma outra forma de provar (que também é a ideia da prova da Fórmula de Taylor com resto integral) é fazer duas aplicações do Teorema Fundamental do Cálculo:

$$\begin{aligned} f(y) - f(x) &= \int_0^1 \frac{d}{dt} f[x + t(y-x)] dt \\ &= \int_0^1 \langle \nabla f[x + t(y-x)], y-x \rangle dt \\ &= \langle \nabla f(x), y-x \rangle + \left\langle \int_0^1 [\nabla f[x + t(y-x)] - \nabla f(x)] dt, y-x \right\rangle \\ &= \langle \nabla f(x), y-x \rangle + \left\langle \int_0^1 \int_0^1 \frac{d}{ds} \nabla f[x + st(y-x)] ds dt, y-x \right\rangle \\ &= \langle \nabla f(x), y-x \rangle + \left\langle \int_0^1 \int_0^1 \nabla^2 f[x + st(y-x)] (t(y-x)) ds dt, y-x \right\rangle \\ &= \langle \nabla f(x), y-x \rangle + \int_0^1 \int_0^1 t \langle \nabla^2 f[x + st(y-x)](y-x), y-x \rangle ds dt \\ &\geq \langle \nabla f(x), y-x \rangle. \end{aligned} \quad \square$$

## 6.1 Regularidade de primeira ordem

A questão seguinte de nosso interesse é a análise da regularidade de funções convexas. Por exemplo, dada uma função convexa própria,

- será  $f$  contínua?
- será  $f$  diferenciável?

Passamos a estudar estas questões.

**Teorema 16.** *Suponhamos que  $f : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$  é uma função convexa própria. Então,  $f$  é localmente limitada e localmente Lipschitz em  $\text{int}(\text{dom } f)$ . Em particular<sup>5</sup>,  $f$  é diferenciável em quase todo ponto de  $\text{dom } f$ .*

<sup>5</sup>O Teorema de Rademacher (ver, por exemplo, [5, Section 3.1.2]) afirma que funções Lipschitz são diferenciáveis em quase todo ponto.

Vamos apresentar duas provas de que funções convexas são localmente de Lipschitz. A primeira vale apenas em dimensão um, mas é bastante instrutiva. A segunda é uma prova em dimensão  $n \geq 1$  que é também elementar.

*Demonstração em dimensão 1.* Se  $n = 1$ ,  $\text{int dom } f$  é um intervalo  $(a, b)$ . Em dimensão um, é possível caracterizar funções convexas como: dados  $a < x < b$ , temos

$$\frac{f(x) - f(a)}{x - a} \leq \frac{f(b) - f(a)}{b - a} \leq \frac{f(b) - f(x)}{b - x}. \quad (14)$$

Para ver isso, tomamos  $t = \frac{x - a}{b - a} \in (0, 1)$  de modo que  $x = (1 - t)a + tb$ . Logo,

$$f(x) \leq (1 - t)f(a) + tf(b) = f(a) + \frac{f(b) - f(a)}{b - a}(x - a).$$

Então,

$$\frac{f(x) - f(a)}{x - a} \leq \frac{f(b) - f(a)}{b - a}.$$

Considerando  $s \in (0, 1)$  tal que  $x = sa + (1 - s)b$ , conseguimos a segunda desigualdade em (14).

Considere  $(c, d) \subset\subset (a, b)$ . Iterando (14), temos que, para  $a < c < x < y < d < b$ ,

$$\frac{f(c) - f(a)}{c - a} \leq \frac{f(y) - f(x)}{y - x} \leq \frac{f(b) - f(d)}{b - d}.$$

Escolhendo

$$M := \max \left\{ \left| \frac{f(c) - f(a)}{c - a} \right|, \left| \frac{f(b) - f(d)}{b - d} \right| \right\},$$

temos  $|f(y) - f(x)| \leq M|x - y|$ , para quaisquer  $x, y \in (c, d)$ .  $\square$

Finalmente, apresentamos uma segunda prova de que  $f$  é localmente Lipschitz.

*Demonstração.* Considere  $x_0 \in \text{int dom } f$  e  $Q_r(x_0) \subset \text{int dom } f$  um cubo fechado centrado em  $x_0$  cujas arestas tem comprimento  $2r$  e são paralelas com os eixos coordenados. Observe que  $B_r(x_0) \subset Q_r(x_0)$ . Sendo  $f$  convexa, o máximo de  $f$  em  $Q_r(x_0)$  é atingido em um dos vértices  $x_k$  do cubo (exercício). Logo,

$$f(y) \leq \sup_k f(x_k) =: M < +\infty, \quad \text{para todo } y \in B_r(x_0).$$

Em seguida, vamos mostrar que  $f$  é também limitada por baixo em  $B_r(x_0)$ . De fato, dado  $x \in B_r(x_0)$ , temos que  $y := 2x_0 - x \in B_r(x_0)$ . Logo,

$$x_0 = \frac{x + y}{2} \implies f(x_0) \leq \frac{f(x) + f(y)}{2} \leq \frac{f(x) + M}{2} \implies f(x) \geq 2f(x_0) - M.$$

Logo,  $|f| \leq C$  em  $B_r(x_0)$ , o que mostra que  $f$  é localmente limitada em  $\text{int dom } f$ .

Agora, fixamos  $\delta \in (0, r)$ . Dados  $x, y \in B_{r-\delta}(x_0)$ , temos que

$$y = \frac{\delta}{\delta + \|y - x\|} x + \frac{\|y - x\|}{\delta + \|y - x\|} z, \quad \text{onde } z = y + \delta \frac{y - x}{\|y - x\|} \in B_r(x_0).$$

Usando convexidade,

$$f(y) \leq \frac{\delta}{\delta + \|y - x\|} f(x) + \frac{\|y - x\|}{\delta + \|y - x\|} f(z).$$

Logo,

$$f(y) - f(x) \leq \frac{f(z) - f(x)}{\delta + \|y - x\|} \|y - x\| \leq \frac{2C}{\delta} \|y - x\|.$$

Trocando os papéis de  $x, y \in B_{r-\delta}(x_0)$ , concluímos que  $f$  é de Lipschitz em  $B_{r-\delta}(x_0)$ .  $\square$

## 6.2 Regularidade de segunda ordem – o Teorema de Alexandrov

Em seguida, estudamos diferenciabilidade de segunda ordem para funções convexas. O Teorema de Alexandrov diz que funções convexas são duas vezes diferenciáveis para quase todo ponto, que é um resultado notável, já que diferenciabilidade de primeira ordem é apenas quase sempre. Um pouco mais explicitamente, o Teorema de Alexandrov afirma que, para quase todo ponto, existe uma matriz, que pode ser chamada de “Hessiana”, que de fato tem a ver (em um certo sentido) com derivadas de segunda ordem de  $f$ , e que cumpre seu papel no sentido da Fórmula de Taylor.

**Teorema de Alexandrov.** *Seja  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  uma função convexa. Então, existe um conjunto  $A$  de medida nula tal que, para todo  $x \in \mathbb{R}^n \setminus A$ , existe uma matriz (denotada por)  $\nabla^2 f(x)$  que satisfaz*

$$f(x+v) = f(x) + \langle \nabla f(x), v \rangle + \frac{\langle \nabla^2 f(x)v, v \rangle}{2} + o(|v|^2) \quad \text{quando } v \rightarrow 0.$$

*Demonstração.* Pode ser encontrada em Evans, Gariepy [5] ou Villani [17]. Necessita de conhecimentos de Teoria da Medida.  $\square$

## 7 O Teorema da Função Inversa

O Teorema da Função Inversa figura como um dos resultados mais importantes da análise matemática moderna. O teorema nos dá uma condição suficiente para que funções vetoriais sejam invertíveis na vizinhança de um ponto. Não é tão claro quem foi o primeiro a demonstrar este resultado, mas uma busca no Google<sup>6</sup> mostrou que o Teorema da Função Inversa aparece no trabalho de Ulisse Dini<sup>7</sup>, em “Lezioni di Analisi Infinitesimale”, de 1877.



<sup>6</sup><https://mathoverflow.net/questions/94323/who-was-the-first-to-formulate-the-inverse-function-theorem>, visualizado em 30 de Janeiro de 2018

<sup>7</sup>Matemático italiano que viveu de 14 de Novembro de 1845 até 28 de Outubro de 1918. Nascido em Pisa, Itália, estudou na Scuola Normale Superiore, sob orientação de Enrico Betti. Também estudou em Paris com Charles Hermite, através de bolsas de intercâmbio. É conhecido por seu trabalho em Análise Real, no período de formulação rigorosa de vários conceitos da teoria. Seu nome aparece no Critério de Dini (para convergência de séries de Fourier), Teorema de Dini (sobre convergência de sequências monótonas de funções), é creditado o Teorema da Função Implícita, que veremos mais adiante no curso, etc. Dini faleceu aos 72 anos em Pisa, Itália. Algumas destas informações foram retiradas de [https://en.wikipedia.org/w/index.php?title=Ulisse\\_Dini&oldid=812724987](https://en.wikipedia.org/w/index.php?title=Ulisse_Dini&oldid=812724987) em 03 de Janeiro de 2018. Segue um retrato de Dini (que também pode ser encontrado no link acima).

**Teorema 17 (Teorema da Função Inversa).** *Seja  $f : E \subseteq \mathbb{R}^n \rightarrow \mathbb{R}^n$  uma função de classe  $C^1(E, \mathbb{R}^n)$  e que, para algum  $a \in E$ , a transformação linear  $f'(a)$  é invertível. Então,  $f$  é invertível, com inversa diferenciável, perto do ponto  $a$ .*

*Mais precisamente, existe um aberto  $U$  que contém  $a$  e um aberto  $V$  que contém  $f(a)$  tais que*

- (i)  $f$  é injetiva em  $U$ ;
- (ii)  $f(U) = V$ , de modo que  $f : U \rightarrow V$  é uma bijeção;
- (iii) A inversa  $f^{-1}$  é de classe  $C^1(V; U)$  e vale que

$$(f^{-1})'(y) = [f'(f^{-1}(y))]^{-1}.$$

**Observação 18.** Em outras palavras, se  $f \in C^1$  e  $f'(a)$  é um isomorfismo, então  $f$  é um difeomorfismo de uma vizinhança do ponto  $a$  em uma vizinhança do ponto  $f(a)$ .

Vamos apresentar uma prova bastante direta, seguindo os passos de Rudin [14, pp. 193] ou Spivak [15, pp. 41] ou Lima [10, Capítulo 6]. Começamos com um resultado preliminar de Álgebra Linear, que será importante na demonstração.

**Lema 19.** *Seja  $M(n)$  o espaço vetorial das matrizes (quadradas) de ordem  $n \times n$  e  $GL(n) \subset M(n)$  o subconjunto das matrizes invertíveis. Então:*

- (i) Se  $A$  é invertível e  $B$  satisfaz

$$\|B - A\| < \frac{1}{\|A^{-1}\|}, \quad (15)$$

*então  $B$  também é invertível. Ou seja, a bola de centro em  $A \in GL(n)$  e de raio  $1/\|A^{-1}\|$  está contida em  $GL(n)$ . Em particular, o conjunto  $GL(n)$  é aberto em  $M(n)$ .*

- (ii) A aplicação  $\Phi : GL(n) \rightarrow GL(n)$  dada por  $\Phi(A) = A^{-1}$ , é um homeomorfismo<sup>8</sup>.

*Demonstração do lema.* Nas condições do item (i), consideramos  $x \in \mathbb{R}$  e vamos mostrar que existe uma constante  $k > 0$  tal que

$$k|x| \leq |Bx|, \quad (16)$$

o que implica que  $B$  é injetiva. Sendo uma matriz quadrada, concluímos que é invertível. Para verificar (16), calculamos

$$|x| = |A^{-1}Ax| \leq \|A^{-1}\| |Ax|$$

que implica

$$\frac{|x|}{\|A^{-1}\|} \leq |Ax| \leq |(B - A)x| + |Bx| \leq \|B - A\| |x| + |Bx|,$$

o que implica em (16) com

$$k := \frac{1}{\|A^{-1}\|} - \|(B - A)\| > 0.$$

Para provar (ii), notamos que

$$B^{-1} - A^{-1} = B^{-1}(A - B)A^{-1}.$$

Além disso, fazendo  $x = B^{-1}y$  em (16), obtemos

$$k|B^{-1}y| \leq |y|, \quad \text{o que implica} \quad \|B^{-1}\| \leq \frac{1}{k} = \frac{\|A^{-1}\|}{1 - \|A^{-1}\| \|B - A\|}.$$

Juntando estas informações, obtemos

$$\|\Phi(B) - \Phi(A)\| = \|B^{-1} - A^{-1}\| \leq \|B^{-1}\| \|A - B\| \|A^{-1}\| \leq \frac{\|A^{-1}\|^2}{1 - \|A^{-1}\| \|B - A\|} \|B - A\|. \quad (17)$$

<sup>8</sup>Isto é,  $\Phi$  é contínua, invertível e com inversa (que, neste caso, é a própria aplicação  $\Phi$ ) é contínua.

Logo, para  $A \in GL(n)$  fixado, o lado direito de (17) vai para zero quando  $\|B - A\| \rightarrow 0$ . Logo, dado  $\varepsilon > 0$ , existe  $\delta > 0$  tal que o lado direito de (17) fica menor do que  $\varepsilon$  quando  $\|B - A\| < \delta$ . Em particular,

$$\|B - A\| < \delta \implies \|\Phi(B) - \Phi(A)\| < \varepsilon. \quad \square$$

Passamos então à demonstração do teorema principal desta seção.

*Demonstração do Teorema da Função Inversa.* Dividimos a prova em três passos, todos presentes de uma maneira ou de outra em [10, 14, 15]. É um bom exercício comparar os argumentos para um melhor entendimento. Por exemplo, [10] separa a prova que apresentamos em vários lemas e teoremas independentes, e mesmo mais gerais, que podem esconder um pouco o trabalho da demonstração. Por outro lado, os teoremas independentes podem ser úteis em alguma outra situação.

Passo 1. Vamos provar que  $f$  é injetiva em uma vizinhança de  $a$ . Basta mostrar que, para  $|h|$  suficientemente próximo da origem e para algum  $\varepsilon > 0$ ,

$$|f(x+h) - f(x)| \geq \varepsilon|h|. \quad (18)$$

Sendo a desigualdade (18) válida para quaisquer  $x \in B_R(a)$  e  $x+h \in B_R(a)$ , temos que  $f$  é injetiva em  $B_R(a)$ . Comparar com (16) que é a versão linear do argumento de injetividade.

Devemos, assim, mostrar (18). Note que, pela desigualdade triangular,

$$|f(x+h) - f(x)| \geq |f'(a) \cdot h| - |f(x+h) - f(x) - f'(a) \cdot h|. \quad (19)$$

A ideia que vamos formalizar abaixo é a seguinte: sendo  $f$  continuamente diferenciável e  $x$  próximo de  $a$ , a derivada  $f'(a)$  nos dá alguma informação que seria dada por  $f'(x)$ . Assim, vamos conseguir mostrar que  $|f(x+h) - f(x) - f'(a) \cdot h|$  fica pequeno quando  $h$  está próximo da origem.

Fixamos um número  $R$  positivo que será escolhido posteriormente. Sejam  $x \in B_R(a)$  e  $x+h \in B_R(a)$  pontos quaisquer e definamos  $F : [0, 1] \rightarrow \mathbb{R}^n$  por

$$F(t) = f(x+th) - f(x) - tf'(a) \cdot h.$$

Pela Desigualdade do Valor Médio, existe  $t_0 \in (0, 1)$  que satisfaz

$$|f(x+h) - f(x) - f'(a) \cdot h| = |F(1) - F(0)| \leq |F'(t_0)|. \quad (20)$$

Temos ainda

$$|F'(t_0)| = |f'(x+t_0h) \cdot h - f'(a) \cdot h| \leq \|f'(x+t_0h) - f'(a)\| |h|. \quad (21)$$

Por hipótese,  $f \in C^1$ , de modo que  $f' : E \rightarrow \mathcal{L}(\mathbb{R}^n; \mathbb{R}^n)$  é contínua. Logo, dado  $\varepsilon > 0$ , existe  $R > 0$  tal que  $B_R(a) \subseteq E$  e

$$\|f'(x) - f'(a)\| < \varepsilon \quad \text{para todo } x \in B_R(a).$$

Além disso, pelo Lema 19 acima, podemos considerar  $\varepsilon > 0$  pequeno o suficiente, de tal maneira que  $f'(x)$  é invertível para *todo*  $x \in B_R(a)$ .

Podemos utilizar (20) e (21) para escrever

$$|f(x+h) - f(x) - f'(a) \cdot h| < \varepsilon|h|.$$

Já que  $f'(a)$  é invertível, possivelmente diminuindo  $\varepsilon > 0$  para que  $\varepsilon < 1/(2\|f'(a)^{-1}\|)$ , temos

$$\varepsilon|h| = \varepsilon|f'(a)^{-1}f'(a) \cdot h| \leq \varepsilon\|f'(a)^{-1}\| |f'(a) \cdot h| < \frac{|f'(a) \cdot h|}{2}. \quad (22)$$

Assim, (19) implica

$$|f(x+h) - f(x)| \geq \frac{|f'(a) \cdot h|}{2} \geq \varepsilon|h|.$$

onde, na última desigualdade, utilizamos (22) novamente.

Passo 2. Escrevemos  $U := B_R(a)$  e  $V := f(U)$  e temos, pelo passo 1, que  $f : U \rightarrow V$  é invertível. Vamos verificar que, nas hipóteses do teorema,  $f : U \rightarrow V$  é uma aplicação aberta. Lembramos que uma aplicação  $f$  é aberta quando a imagem de qualquer conjunto aberto por  $f$  é um conjunto aberto. Em particular, temos que  $V$  é aberto e que  $f^{-1}$  é uma aplicação contínua<sup>9</sup>.

<sup>9</sup>Uma aplicação  $g$  é contínua em  $\mathbb{R}^n$  se, e somente se,  $g^{-1}(A)$  é aberto para qualquer  $A$  aberto.

Para verificar que  $f$  é aberta, vamos provar que a imagem de qualquer bola (aberta) contida em  $U$  é um conjunto aberto. Consideramos então  $x_0 \in U$  e  $r > 0$  tais que  $\overline{B_r(x_0)} \subset U$ . Afirmando que  $B_{\varepsilon r/2}(f(x_0)) \subset f(B_r(x_0))$ . De fato, seja  $y \in B_{\varepsilon r/2}(f(x_0))$ . Tem-se

$$|y - f(x_0)| < \frac{\varepsilon r}{2}. \quad (23)$$

A função  $\phi(x) := |y - f(x)|$  é contínua no conjunto compacto  $\overline{B_r(x_0)}$ . Logo, assume um valor mínimo em  $x^* \in \overline{B_r(x_0)}$ . Se fosse  $|x^* - x_0| = r$ , teríamos que, usando (18) e (23),

$$\varepsilon r = \varepsilon |x^* - x_0| \leq |f(x^*) - f(x_0)| \leq |y - f(x^*)| + |y - f(x_0)| < \phi(x^*) + \frac{\varepsilon r}{2}.$$

Mas daí, utilizando (23) novamente,

$$\phi(x_0) < \frac{\varepsilon r}{2} < \phi(x^*),$$

o que contradiz ser  $x^*$  um ponto de mínimo. Segue que  $x^* \in B_r(x_0)$ .

Observamos ainda que  $x^*$  também minimiza a função  $\psi(x) = |f(x) - y|^2$  em  $\overline{B_r(x_0)}$ . Logo, sendo um ponto interior, temos  $\psi'(x^*) = 0$ , isto é,  $2(f(x^*) - y) = 0$ . Concluimos desta forma que  $y = f(x^*)$ , onde  $x^* \in B_r(x_0)$ . Logo,  $y \in f(B_r(x_0))$ , como havíamos afirmado. Segue que  $V$  é um conjunto aberto.

Passo 3. Neste último passo, vamos verificar (iii). Dados  $y \in V$  e  $y + k \in V$ , escrevemos

$$x = f^{-1}(y) \quad \text{e} \quad h = f^{-1}(y + k) - x.$$

Desta forma, temos as relações

$$k = f(x + h) - f(x) \quad \text{e} \quad h = f^{-1}(y + k) - f^{-1}(y). \quad (24)$$

Sendo  $f$  diferenciável em  $x$ , podemos escrever

$$k = f(x + h) - f(x) = f'(x) \cdot h + r(h) \quad \text{onde} \quad \lim_{h \rightarrow 0} \frac{r(h)}{h} = 0.$$

Agora, pela nossa escolha de  $\varepsilon > 0$ , o Lema 19 garante que  $f'(x)$  é invertível, para qualquer  $x \in U$ . Assim, aplicando  $f'(x)^{-1}$  em ambos os lados da identidade acima, obtemos

$$f'(x)^{-1} \cdot k = h + f'(x)^{-1} \cdot r(h).$$

Reescrevendo isto a partir das relações em (24), temos

$$f^{-1}(y + k) - f^{-1}(y) = h = f'(x)^{-1} \cdot k - f'(x)^{-1} \cdot r(h) =: f'(x)^{-1} \cdot k + \rho(k). \quad (25)$$

Por (18), temos que  $\lim_{k \rightarrow 0} |h(k)| = 0$  (em particular, isto mostra mais uma vez a continuidade da função inversa  $f^{-1}$ ). Finalmente, observamos que

$$\frac{|\rho(k)|}{|k|} = \frac{|f'(x)^{-1} \cdot r(h)|}{|k|} \leq \frac{\|f'(x)^{-1}\| |r(h)|}{|k|} = \frac{\|f'(x)^{-1}\| |h|}{|f(x + h) - f(x)|} \cdot \frac{|r(h)|}{|h|} \stackrel{(18)}{\leq} \frac{\|f'(x)^{-1}\|}{\varepsilon} \cdot \frac{|r(h)|}{|h|}.$$

Portanto,

$$\lim_{k \rightarrow 0} \frac{|\rho(k)|}{|k|} = 0$$

e (25) mostra que  $f^{-1}$  é diferenciável com  $(f^{-1})'(y) = f'(x)^{-1}$ .  $\square$

**Exemplo 20.** Considere  $f : \mathbb{R}^2 \rightarrow \mathbb{R}^2$  definida por  $f(x, y) = (e^x \cos y, e^x \sin y)$ . Temos,

$$f'(x, y) = \begin{bmatrix} e^x \cos y & e^x \sin y \\ -e^x \sin y & e^x \cos y \end{bmatrix} \implies \det f'(x, y) = e^{2x} > 0 \quad \forall (x, y) \in \mathbb{R}^2.$$

Pelo Teorema da Função Inversa, todo ponto  $(x, y) \in \mathbb{R}^2$  possui uma vizinhança onde  $f$  é um difeomorfismo. No entanto,  $f$  não é um difeomorfismo global, pois  $f$  nem é injetiva:  $f(x, y + 2\pi) = f(x, y)$ .

**Exercício 17.** Mostre que  $f : U \subseteq \mathbb{R}^n \rightarrow \mathbb{R}^n$ , de classe  $C^1$ , é um difeomorfismo local se, e somente se,  $f'(x) \in \mathcal{L}(\mathbb{R}^n; \mathbb{R}^n)$  é um isomorfismo para todo  $x \in U$ . Ou, equivalentemente, se, e somente se,  $\det f'(x) \neq 0$  para todo  $x \in U$ .

**Exercício 18.** Seja  $f : U \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$  uma função de classe  $C^2$ . Suponhamos

- $a \in U$  é um ponto crítico de  $f$
- a matriz Hessiana  $\nabla^2 f(a)$  é invertível.

Então, existe uma vizinhança de  $a$  onde  $a$  é o único ponto crítico de  $f$ .

## 8 O Teorema da Função Implícita

O Teorema da Função Implícita, historicamente, apareceu como uma condição necessária para resolução de sistemas não lineares, como, por exemplo,

$$\begin{cases} f_1(x_1, x_2, \dots, x_n, y_1, y_2, \dots, y_m) = 0 \\ f_2(x_1, x_2, \dots, x_n, y_1, y_2, \dots, y_m) = 0 \\ \vdots \\ f_n(x_1, x_2, \dots, x_n, y_1, y_2, \dots, y_m) = 0 \end{cases}.$$

A ideia é “isolar” as variáveis  $x_i$ , resolvendo o sistema para estas variáveis:

$$\begin{cases} x_1 = \xi_1(y_1, y_2, \dots, y_m) \\ x_2 = \xi_2(y_1, y_2, \dots, y_m) \\ \vdots \\ x_n = \xi_n(y_1, y_2, \dots, y_m) \end{cases}.$$

Vejam, recordando um exemplo dos cursos básicos de álgebra linear, como isto é feito no caso de um sistema linear de equações:

**Exemplo 21.** Vamos resolver o sistema linear  $Ax = 0$ , onde a matriz  $A$ , de ordem  $3 \times 5$ , é

$$A = \begin{bmatrix} 1 & 0 & 3 & 1 & 2 \\ -1 & 2 & -2 & 0 & 2 \\ 6 & -4 & 1 & 1 & -4 \end{bmatrix}.$$

Como função, tem-se  $A : \mathbb{R}^5 \rightarrow \mathbb{R}^3$ . Comparando com as funções  $f_i$  acima, as componentes de  $A$  são dadas pelas expressões

$$\begin{cases} f_1(x_1, x_2, x_3, y_1, y_2) = x_1 + 3x_3 + y_1 + 2y_2 \\ f_2(x_1, x_2, x_3, y_1, y_2) = -x_1 + 2x_2 - 2x_3 + y_2 \\ f_3(x_1, x_2, x_3, y_1, y_2) = 6x_1 - 4x_2 + x_3 + y_1 - 4y_2 \end{cases}.$$

Observe que o nome das variáveis é irrelevante e, em geral, apenas descobrimos quais podem ser isoladas em termos das demais *a posteriori*, isto é, depois de analisar a forma escalonada da matriz dada.

Por escalonamento (estamos utilizando o símbolo  $\sim$  para “equivalência por linhas”):

$$A \sim \begin{bmatrix} 1 & 0 & 3 & 1 & 2 \\ 0 & 2 & 1 & 1 & 4 \\ 0 & 4 & 17 & 5 & 16 \end{bmatrix} \sim \dots \sim \begin{bmatrix} 1 & 0 & 0 & 2/5 & 2/5 \\ 0 & 1 & 0 & 2/5 & 26/15 \\ 0 & 0 & 1 & 1/5 & 8/15 \end{bmatrix} \implies \begin{cases} x_1 = -\frac{2}{5}x_4 - \frac{2}{5}x_5 \\ x_2 = -\frac{2}{5}x_4 - \frac{26}{15}x_5 \\ x_3 = -\frac{1}{5}x_4 - \frac{8}{15}x_5 \end{cases}$$

Como já é conhecido da Álgebra Linear

$$\dim \ker A + \dim \operatorname{Im} A = 5.$$

A dimensão do núcleo de  $A$  representa o número de variáveis livres do sistema. As demais podem ser escritas em termos das variáveis livres.

Mais geralmente, temos o seguinte exercício de Álgebra Linear.

**Exercício 19.** Seja  $T \in \mathcal{L}(\mathbb{R}^{n+k}, \mathbb{R}^n)$  tal que

$$T(x, 0) = 0 \implies x = 0. \quad (26)$$

Acima, estamos escrevendo  $(x, y) \in \mathbb{R}^{n+k}$  e pensando em  $x \in \mathbb{R}^n$  e  $y \in \mathbb{R}^k$ . Prove que, para  $y \in \mathbb{R}^k$  fixado, existe único  $x = \xi(y) \in \mathbb{R}^n$  tal que

$$T(\xi(y), y) = 0.$$

Além disso, a aplicação  $y \mapsto \xi(y)$  é linear. Isto significa que  $x$  está definido implicitamente em termos de  $y$  pela equação  $T(x, y) = 0$ . Este problema pode ser visto como a “versão linear” do Teorema da Função Implícita que veremos abaixo.

Passamos a analisar o Teorema da Função Implícita. A ideia geral (da seção anterior) continua válida aqui: se a derivada de  $f$  satisfaz uma condição do tipo do Exercício 19, então a conclusão também é válida, localmente, para a função  $f$ , possivelmente não linear. Aparentemente, na Itália, o Teorema da Função Implícita também conhecido como Teorema de Dini.

Antes de enunciar e provar o teorema, deixamos mais um exercício para o leitor.

**Exercício 20.** Seja  $T \in \mathcal{L}(\mathbb{R}^{n+k}, \mathbb{R}^n)$ . São equivalentes:

- (i)  $T(x, 0) = 0 \implies x = 0$ .
- (ii) As  $n$  primeiras colunas da matriz associada a  $T$  são linearmente independentes.
- (iii) A matriz de ordem  $n \times n$ , obtida ao considerar apenas as  $n$  primeiras colunas da matriz canônica associada a  $T$ , é invertível.
- (iv) Qualquer dos itens do Exercício 4 da Lista Zero para a matriz do item anterior.

Reforçamos que a posição das colunas não é importante. O exercício acima poderia ser enunciado (mas seria mais chato de escrever) como: se  $n$  das colunas são linearmente independentes (tipo o item (ii)), então teríamos versões análogas para os outros itens.

**Teorema 22 (Teorema da Função Implícita).** Seja  $f : E \subseteq \mathbb{R}^{n+k} \rightarrow \mathbb{R}^n$  uma função de classe  $C^1$ . Suponhamos que  $(a, b) \in E$  é tal que

- $f(a, b) = 0$ .



- $f'(a, b) : \mathbb{R}^{n+k} \rightarrow \mathbb{R}^n$  satisfaz (qualquer) um dos itens do Exercício 20.

Então, existe uma vizinhança  $W$  de  $b \in \mathbb{R}^k$  e única função  $\xi : W \subseteq \mathbb{R}^k \rightarrow \mathbb{R}^n$  de classe  $C^1$  tal que

$$a = \xi(b) \quad e \quad f(\xi(y), y) = 0 \quad \text{para todo } y \in W.$$

*Demonstração.* A ideia é estender  $f$  para uma função de  $\mathbb{R}^{n+k}$  em  $\mathbb{R}^{n+k}$  e aplicar o Teorema da Função Inversa. Definimos  $F : \mathbb{R}^{n+k} \rightarrow \mathbb{R}^{n+k}$  por

$$F(x, y) = (f(x, y), y).$$

Dessa forma,

$$F'(x, y) \cdot (h, k) = (f'(x, y)(h, k), k).$$

Precisamos ver que  $F'(a, b)$  é invertível. Como está definida em espaços vetoriais de mesma dimensão, basta verificar que é injetiva:

$$F'(a, b) \cdot (h, k) = 0 \iff (f'(a, b) \cdot (h, k), k) = 0 \iff k = 0 \quad e \quad f'(a, b) \cdot (h, 0) = 0.$$

Nossa hipótese garante que  $h = 0$  e, logo,  $F'(a, b)$  é invertível. Pelo Teorema da Função Inversa, existe uma vizinhança  $U$  de  $(a, b) \in \mathbb{R}^{n+k}$  e  $V$  de  $(0, b) \in \mathbb{R}^{n+k}$  onde  $F$  admite uma inversa  $G$ . Observe que devemos ter

$$G(u, v) = (g(u, v), v), \quad (u, v) \in V,$$

para alguma função  $g : V \subseteq \mathbb{R}^{n+k} \rightarrow \mathbb{R}^n$  de classe  $C^1$ . Notamos que

$$\rightsquigarrow F(a, b) = (0, b) \implies G(0, b) = (a, b) \implies g(0, b) = a;$$

$$\rightsquigarrow F(G(x, y)) = (x, y) \implies f(g(x, y), y) = x, \quad (x, y) \in V.$$

Concluimos a demonstração ao considerar uma vizinhança  $W \subseteq \mathbb{R}^k$  do ponto  $b$  tal que  $(0, b) \in \{0\} \times W \subseteq V$  e definirmos  $\xi : W \rightarrow \mathbb{R}^n$  por

$$\xi(y) = g(0, y), \quad y \in W.$$

A unicidade de  $\xi$  segue de ser  $F : U \rightarrow V$  injetiva. □

**Exercício 21.** Pela Regra da Cadeia, escrever uma equação diferencial que a aplicação  $\xi$  dada pelo Teorema da Função Implícita deve satisfazer, a saber, para  $y \in W$ ,

$$\xi'(y) = -\partial_x f(\xi(y), y)^{-1} \partial_y f(\xi(y), y).$$

*Dica:* Diferenciar a identidade  $f(\xi(y), y) = 0$ .

**Exemplo 23.** A equação  $x^2 + y^2 + z^2 = 1$  define implicitamente  $z = \xi(x, y)$  em vizinhanças de pontos que não sejam da forma  $(x, y, 0)$  (pense geometricamente!). Analogamente, define implicitamente  $y = h(x, z)$  em vizinhanças de pontos que não sejam da forma  $(x, 0, z)$ .

O Teorema da Função Implícita pode, por exemplo, ser aplicado da seguinte forma: defina  $f : \mathbb{R}^3 \rightarrow \mathbb{R}$  por  $f(x, y, z) = x^2 + y^2 + z^2 - 1$  e queremos entender como identificar funções implícitas a partir da equação

$$f(x, y, z) = 0.$$

A (matriz da) derivada de  $f$  em um ponto  $(x, y, z)$  qualquer é dada por

$$[2x \quad 2y \quad 2z].$$

Qualquer uma coluna dessa matriz forma um conjunto linearmente independente se, e somente se, sua entrada for não nula.

Um caso especial do Teorema da Função Implícita que generaliza o exemplo acima é o de funções escalares:

**Corolário 24.** *Seja  $f : E \subseteq \mathbb{R}^{k+1} \rightarrow \mathbb{R}$  uma função de classe  $C^1$ . Suponhamos que  $(a_1, a_2, \dots, a_k, a_{k+1}) \in E$  é tal que*

$$f(a_1, a_2, \dots, a_k, a_{k+1}) = 0 \text{ e que } \frac{\partial f}{\partial x_{k+1}}(a_1, a_2, \dots, a_k, a_{k+1}) \neq 0.$$

*Então, existe uma vizinhança  $W$  de  $(a_1, a_2, \dots, a_k) \in \mathbb{R}^k$  e única função  $\xi : W \subseteq \mathbb{R}^k \rightarrow \mathbb{R}$  de classe  $C^1$  tal que*

$$a_{k+1} = \xi(a_1, a_2, \dots, a_k) \text{ e } f(a_1, a_2, \dots, a_k, \xi(a_1, a_2, \dots, a_k)) = 0 \text{ para todo } (a_1, a_2, \dots, a_k) \in W.$$

Note que, com relação ao enunciado mais geral do Teorema da Função Implícita, invertemos a ordem das variáveis:  $a_{k+1}$  faz o papel de  $a$  e as demais variáveis fazem o papel de  $b$ . Isto não deve causar confusão. Além do mais, poderíamos enunciar o corolário para qualquer uma das variáveis da função, substituindo a hipótese da derivada pela que é natural. Por exemplo, para uma função  $f : \mathbb{R}^4 \rightarrow \mathbb{R}$ , a equação

$$f(x, y, z, w) = 0$$

define  $y$  como função das demais variáveis em uma vizinhança do ponto  $(a, b, c, d)$  quando

$$\frac{\partial f}{\partial y}(a, b, c, d) \neq 0.$$

Vejamos um exemplo explícito de aplicações com valores vetoriais.

**Exemplo 25 (Adaptação de Rudin [14]).** Seja  $f : \mathbb{R}^5 \rightarrow \mathbb{R}^2$  dada por

$$f(x_1, x_2, x_3, x_4, x_5) = (2e^{x_1} + x_2x_3 - 4x_4 + 3, x_2 \cos x_1 - 6x_1 + 2x_3 - x_5).$$

Observamos que  $f(0, 1, 3, 2, 7) = (0, 0)$ . A (matriz da) derivada de  $f$  na base canônica é

$$\left[ \begin{array}{ccccc} 2e^{x_1} & x_3 & x_2 & -4 & 0 \\ -x_2 \operatorname{sen} x_1 - 6 & \cos x_1 & 2 & 0 & -1 \end{array} \right] \Big|_{(0,1,3,2,7)} = \left[ \begin{array}{ccccc} 2 & 3 & 1 & -4 & 0 \\ -6 & 1 & 2 & 0 & -1 \end{array} \right].$$

Nesta matriz duas colunas quaisquer são linearmente independentes, de modo que é possível escrever duas variáveis quaisquer como função das demais. Por exemplo, a matriz menor referente às colunas 1 e 4 é

$$\left[ \begin{array}{cc} 2 & -4 \\ -6 & 0 \end{array} \right],$$

que tem colunas linearmente independentes. Pelo Teorema da Aplicação Implícita, existe vizinhança  $U$  de  $(1, 3, 7) \in \mathbb{R}^3$  e uma função  $\xi : U \rightarrow \mathbb{R}^2$  de classe  $C^\infty$  tais que

$$(x_1, x_4) = \xi(x_2, x_3, x_5)$$

Devemos ainda ter as condições (conferir com o Exercício 21)

$$\begin{bmatrix} \partial_{x_5} \xi_1 & \partial_{x_3} \xi_1 & \partial_{x_2} \xi_1 \\ \partial_{x_5} \xi_2 & \partial_{x_3} \xi_2 & \partial_{x_2} \xi_2 \end{bmatrix} = - \begin{bmatrix} 0 & 1/3 \\ 1/2 & 1/12 \end{bmatrix} \begin{bmatrix} 3 & 1 & 0 \\ 1 & 2 & -1 \end{bmatrix} = \begin{bmatrix} -1/3 & -2/3 & 1/3 \\ -3/4 & -2/3 & 1/12 \end{bmatrix},$$

ou ainda,

$$\begin{aligned} \partial_{x_5} \xi_1 &= -1/3, & \partial_{x_3} \xi_1 &= -2/3, & \partial_{x_2} \xi_1 &= 1/3 \\ \partial_{x_5} \xi_2 &= -3/4, & \partial_{x_3} \xi_2 &= -2/3, & \partial_{x_2} \xi_2 &= 1/12. \end{aligned}$$

## 9 Formas canônicas locais

A ideia desta seção é entender localmente alguns tipos de aplicações de classe  $C^1$ . O leitor pode encontrar mais detalhes de conceitos desta seção em [6, 10, 18]. Consideramos  $f : U \subseteq \mathbb{R}^n \rightarrow \mathbb{R}^m$  uma aplicação de classe  $C^1$ . Nós dizemos que:

- (i)  $f$  é uma **imersão** quando, para todo  $x \in U$ ,  $f'(x) : \mathbb{R}^n \rightarrow \mathbb{R}^m$  é injetiva.
- (ii)  $f$  é uma **submersão** quando, para todo  $x \in U$ ,  $f'(x) : \mathbb{R}^n \rightarrow \mathbb{R}^m$  é sobrejetiva.

Por propriedades básicas das transformações lineares, só é possível  $f$  ser uma imersão no caso  $n \leq m$ . De forma parecida, só é possível  $f$  ser uma submersão no caso  $m \leq n$ . O caso  $m = n$  já foi caracterizado pelo Teorema da Função Inversa, de modo que não precisamos dar um nome especial a este caso.

**Exercício 22.** A aplicação  $f$  é um difeomorfismo local se, e somente se, para todo  $x \in U$ ,  $f'(x) : \mathbb{R}^n \rightarrow \mathbb{R}^m$  é um isomorfismo.

Começamos estudando a forma local de uma submersão, pois esta análise tem relação direta com o Teorema da Função Implícita ( $n \geq m$ ) que demonstramos na seção anterior.

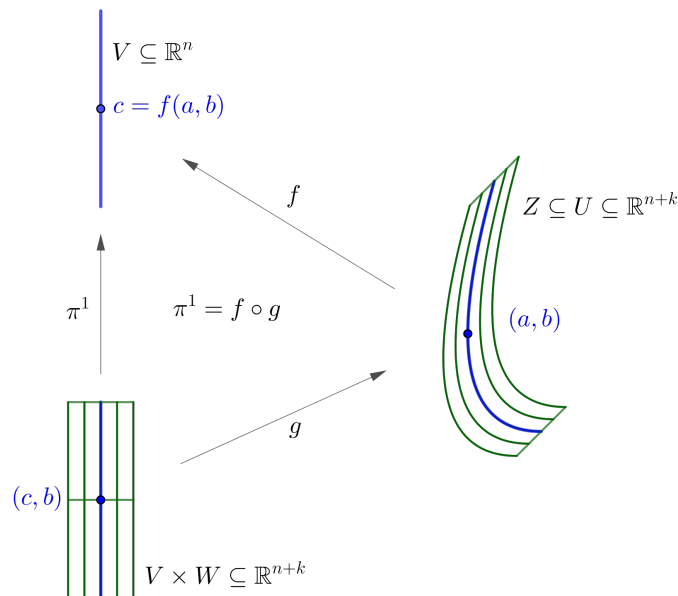
**Teorema 26 (Forma Local de Submersões).** *Seja  $f : U \subseteq \mathbb{R}^{n+k} \rightarrow \mathbb{R}^n$  uma submersão de classe  $C^k(U; \mathbb{R}^n)$  e  $(a, b) \in U \subseteq \mathbb{R}^{n+k}$ . Então, existem*

- (i) uma vizinhança  $W$  de  $b \in \mathbb{R}^k$ ;
- (ii) uma vizinhança  $V$  de  $c = f(a, b) \in \mathbb{R}^n$ ;
- (iii) uma vizinhança  $Z$  de  $(a, b) \in \mathbb{R}^{n+k}$  e;
- (iv) um difeomorfismo  $g : V \times W \rightarrow Z$  de classe  $C^k$

tais que, para todo  $(u, v) \in \mathbb{R}^{n+k}$ , tem-se

$$f(g(u, v)) = u.$$

Em outras palavras, a menos de uma mudança de coordenadas representada pelo difeomorfismo  $g$ , a aplicação  $f$  localmente se parece com a projeção na primeira coordenada  $\pi^1(x, y) = x$ .



*Demonstração.* A prova deste resultado é essencialmente a mesma do Teorema da Função Implícita que apresentamos na seção anterior. Repetimos a demonstração para fixar bem as ideias.

Vamos estender a aplicação  $f$  para uma função de  $\mathbb{R}^{n+k}$  da forma mais intuitiva possível em  $\mathbb{R}^{n+k}$  e aplicar o Teorema da Função Inversa. Definimos  $F : \mathbb{R}^{n+k} \rightarrow \mathbb{R}^{n+k}$  por

$$F(x, y) = (f(x, y), y).$$

Dessa forma,

$$F'(x, y) \cdot (h, k) = (f'(x, y)(h, k), k).$$

Sendo  $f$  uma submersão, sabemos que a dimensão da imagem de  $f'(a, b)$  é  $n$  e, logo, existem  $n$  colunas linearmente independentes na matriz canônica associada. Podemos supor, sem perda de generalidade (como o enunciado já faz), que estas são as  $n$  primeiras colunas. Para verificar que  $F'(a, b)$  é invertível, utilizamos (como anteriormente) o Exercício 20:

$$F'(a, b) \cdot (h, k) = 0 \iff (f'(a, b) \cdot (h, k), k) = 0 \iff k = 0 \text{ e } f'(a, b) \cdot (h, 0) = 0 \implies h = 0.$$

Pelo Teorema da Função inversa, existe uma vizinhança  $Z$  de  $(a, b) \in \mathbb{R}^{n+k}$  e uma vizinhança (que podemos tomar da forma)  $V \times W$  de  $(c, b) \in \mathbb{R}^{n+k}$  onde  $F$  admite uma inversa  $g : V \times W \rightarrow Z$  que é um difeomorfismo de classe  $C^k$ . Observe que devemos ter

$$g(u, v) = (g_1(u, v), v), \quad (u, v) \in V \times W,$$

isto é, a segunda componente de  $g$  é simplesmente  $v \mapsto v$ . Tem-se

$$F(g(u, v)) = (u, v) \iff (f(g(u, v)), v) = (u, v) \iff f(g(u, v)) = u. \quad \square$$

**Exemplo 27.** A função  $\pi^1 : \mathbb{R}^{n+k} \rightarrow \mathbb{R}^n$  definida por  $\pi^1(x, y) = x$  é uma submersão de classe  $C^1$ . A forma local das submersões afirma que toda submersão é localmente dessa forma.

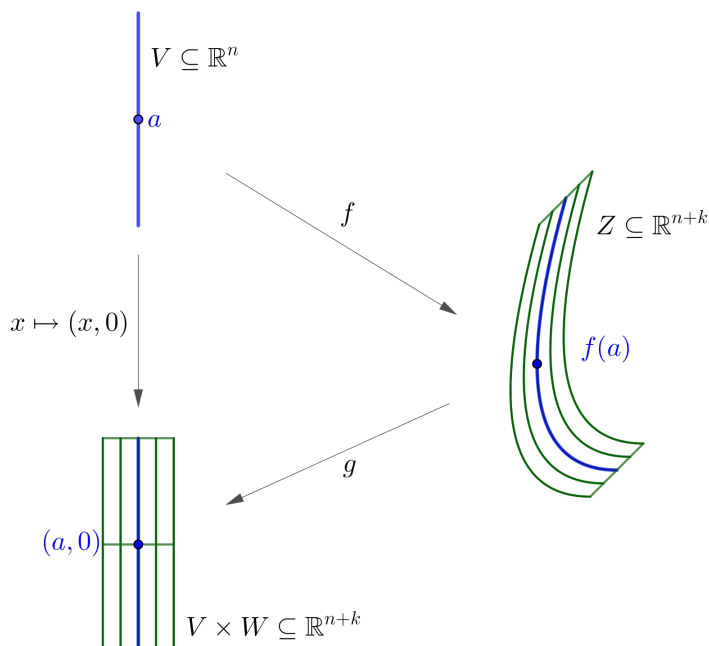
**Teorema 28 (Forma Local de Imersões).** Consideramos  $f : U \subseteq \mathbb{R}^n \rightarrow \mathbb{R}^{n+k}$  uma imersão de classe  $C^k(U; \mathbb{R}^n)$  e  $a \in U \subseteq \mathbb{R}^n$ . Então, existem

- (i) uma vizinhança  $W$  de  $0 \in \mathbb{R}^k$ ;
- (ii) uma vizinhança  $V$  de  $a \in \mathbb{R}^n$ ;
- (iii) uma vizinhança  $Z$  de  $f(a) \in \mathbb{R}^{n+k}$  e;
- (iv) um difeomorfismo  $g : Z \rightarrow V \times W$  de classe  $C^k$

tais que, para todo  $u \in \mathbb{R}^n$ , tem-se

$$g(f(u)) = (u, 0).$$

Em outras palavras, a menos de uma mudança de coordenadas representada pelo difeomorfismo  $g$ , a aplicação  $f$  localmente se parece com a aplicação de "inclusão"  $x \mapsto (x, 0)$ .



*Demonstração.* Este teorema é também uma consequência do Teorema da Função Inversa. Sendo  $f'(a)$  uma aplicação injetiva, temos que sua imagem  $f'(a)(\mathbb{R}^n)$  tem dimensão  $n$ . Por exemplo,

$$\{f'(a) \cdot e_i\}_{i=1}^n = \left\{ \frac{\partial f}{\partial x_i}(a) \right\}_{i=1}^n \quad \text{é linearmente independente em } \mathbb{R}^{n+k}.$$

Completamos este conjunto até uma base, isto é, consideramos

$$\left\{ \frac{\partial f}{\partial x_1}(a), \frac{\partial f}{\partial x_2}(a), \dots, \frac{\partial f}{\partial x_n}(a), v_1, v_2, \dots, v_k \right\} \quad \text{base de } \mathbb{R}^{n+k}. \quad (27)$$

Definimos  $F : \mathbb{R}^{n+k} \rightarrow \mathbb{R}^{n+k}$  como

$$F(x, y) = f(x) + \sum_{i=1}^k y^i v_i, \quad \text{onde } y = \sum_{i=1}^k y^i e_i.$$

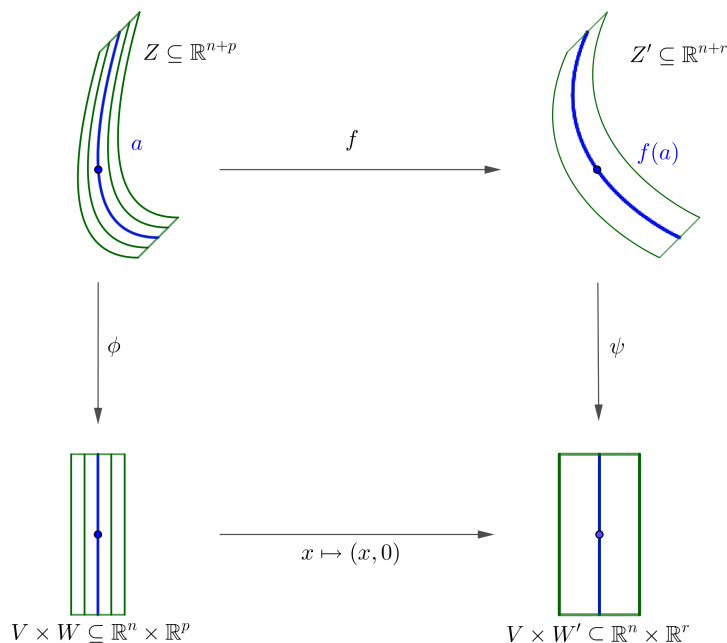
Assim, a expressão para a derivada de  $F$  em  $(a, 0)$  é

$$F'(a, 0) \cdot (h, k) = f'(a) \cdot h + \sum_{i=1}^k k^i v_i, \quad \text{onde } k = \sum_{i=1}^k k^i e_i.$$

Segue de (27) que  $F'(a, 0) \cdot (h, k) = 0 \implies (h, k) = 0$ , de modo que  $F'(a, 0)$  é um isomorfismo. Pelo Teorema da Aplicação Inversa, existe uma vizinhança (que podemos tomar da forma)  $V \times W$  de  $(0, a)$  e uma vizinhança  $Z$  de  $f(a)$  tais que  $F$  admite uma inversa  $g : Z \rightarrow V \times W$  que é um difeomorfismo de classe  $C^k$ . Temos

$$g(f(u)) = g(F(u, 0)) = (u, 0). \quad \square$$

Nós vamos estudar imersões mais adiante como parametrizações ou “cartas” para superfícies no espaço euclidiano. Antes disso, enunciaremos um teorema que engloba todos os anteriores como casos particulares:



**Teorema 29 (Teorema do Posto).** *Seja  $f : U \subseteq \mathbb{R}^{n+p} \rightarrow \mathbb{R}^{n+r}$  uma aplicação de classe  $C^1$  e  $a \in U$ . Suponhamos que o posto da aplicação derivada  $f'(x) : \mathbb{R}^{n+p} \rightarrow \mathbb{R}^{n+r}$  é constante igual a  $n$  para todo  $x$  numa vizinhança de  $a$ . Então existem difeomorfismos de classe  $C^k$*

(i)  $\phi : Z \rightarrow V \times W$  de uma vizinhança de  $a$  em um aberto de  $\mathbb{R}^n \times \mathbb{R}^p$  e

(ii)  $\psi : Z' \rightarrow V \times W'$  de uma vizinhança de  $f(a)$  em um aberto de  $\mathbb{R}^n \times \mathbb{R}^r$

tais que

$$(\psi \circ f \circ \phi^{-1})(x, y) = (x, 0).$$

Em [10], é apresentada uma demonstração deste teorema que utiliza as formas locais de imersões e submersões da seção anterior. Apresentamos uma prova independente (nas mesmas linhas de [6]) que, em particular, implica nas formas locais acima (de fato, a demonstração é a mesma das formas locais, com poucas mudanças).

*Demonstração.* Reordenando as coordenadas, se necessário, nós podemos assumir que a matriz (lembra que o conjunto das matrizes invertíveis é aberto!)

$$\left[ \frac{\partial f}{\partial x_i}(x) \right]_{i,j=1}^n$$

é invertível para todo  $x \in U$ . Definindo  $\phi : \mathbb{R}^{n+p} \rightarrow \mathbb{R}^{n+p}$  por

$$\phi(x, y) = (f_1(x, y), f_2(x, y), \dots, f_n(x, y), y_1, y_2, \dots, y_p),$$

o Teorema da Função Inversa implica que  $\phi$  é um difeomorfismo de uma vizinhança  $Z$  de  $a \in U$  em uma vizinhança de  $\phi(a) \in \mathbb{R}^{n+p}$ . Logo,  $\phi^{-1}$  em componentes é da forma

$$\phi^{-1}(x, y) = (\varphi(x, y), y).$$

Pela definição de inversa, temos

$$(x, y) = \phi(\phi^{-1}(x, y)) = \phi(\varphi(x, y), y) = ((f_1, f_2, \dots, f_n)(\varphi(x, y), y), y)$$

Assim,

$$f(\phi^{-1}(x, y)) = ((f_1, f_2, \dots, f_n)(\varphi(x, y), y), (f_{n+1}, f_{n+2}, \dots, f_{n+p})(\varphi(x, y), y)) = (x, (f_j \circ \phi^{-1})(x, y)).$$

Agora, a função  $f \circ \phi^{-1} : \mathbb{R}^{n+p} \rightarrow \mathbb{R}^{n+r}$  tem derivada com matriz com matriz canônica

$$(f \circ \phi^{-1})'(x, y) = \begin{bmatrix} I_r & 0 \\ * & \partial_{y_i}(f_j \circ \phi^{-1}) \end{bmatrix},$$

onde no canto inferior direito temos  $j > r$ . Sendo o posto de  $(f \circ \phi^{-1})'(x, y)$  exatamente igual a  $r$ , temos que necessariamente  $\partial_{y_i}(f_j \circ \phi^{-1}) = 0$  para todo  $j > r$ . Logo,

$$j > r \implies f_j \circ \phi^{-1}(x, y) = f_j \circ \phi^{-1}(x), \quad (28)$$

isto, é,  $f_j \circ \phi^{-1}$  não depende de  $y$  quando  $j > r$  (mais rigorosamente, deveríamos tomar o cuidado de que o domínio em  $y$ , para cada  $x$  fixado, é um conjunto conexo. Como escolhemos a vizinhança livremente, é possível fazer isso).

Agora, definindo  $\psi : V \times W' \subseteq \mathbb{R}^{n+r} \rightarrow \mathbb{R}^{n+r}$  por

$$\psi(w_1, w_2, \dots, w_{n+r}) := (w_1, \dots, w_n, w_{n+1} - (f_{n+1} \circ \phi^{-1})(w), \dots, w_{n+r} - (f_{n+r} \circ \phi^{-1})(w))$$

Temos, por (28), que

$$\psi'(a) = \begin{bmatrix} I_n & 0 \\ * & I_r \end{bmatrix},$$

de modo que  $\psi$  é um difeomorfismo em uma vizinhança de  $a$ . Além disso,

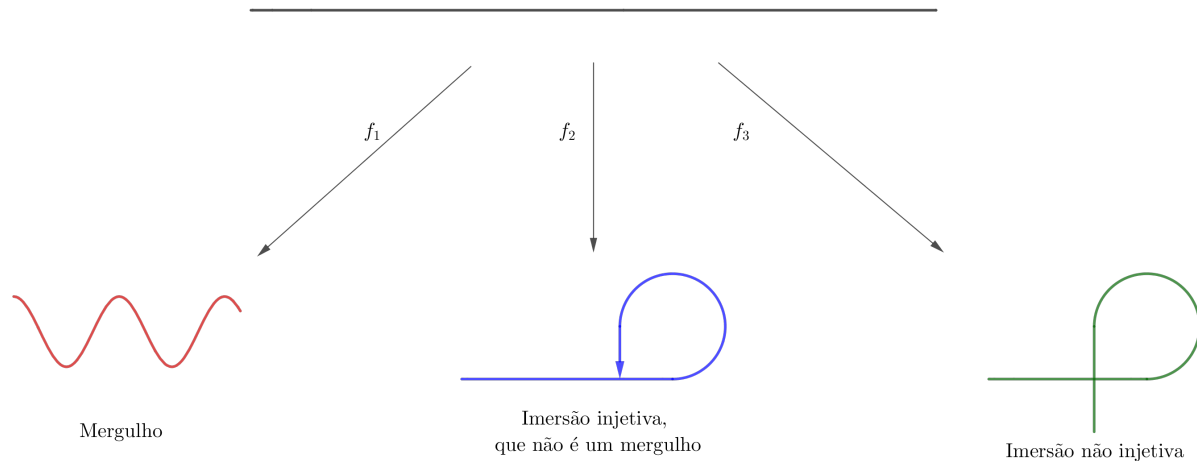
$$(\psi \circ f \circ \phi^{-1})(x, y) = \psi(x, (f_j \circ \phi^{-1})(x, y)) = (x, 0). \quad \square$$

**Exercício 23.** Obter as formas canônicas locais como corolários do Teorema do Posto.

## 10 Mergulhos

Dizemos que  $\phi : U \subseteq \mathbb{R}^n \rightarrow \phi(U) \subseteq \mathbb{R}^{n+k}$  é um **mergulho** quando, além de ser uma imersão,  $\phi$  for um homeomorfismo sobre a sua imagem.

Geometricamente, podemos pensar como nos seguintes exemplos, como protótipos:

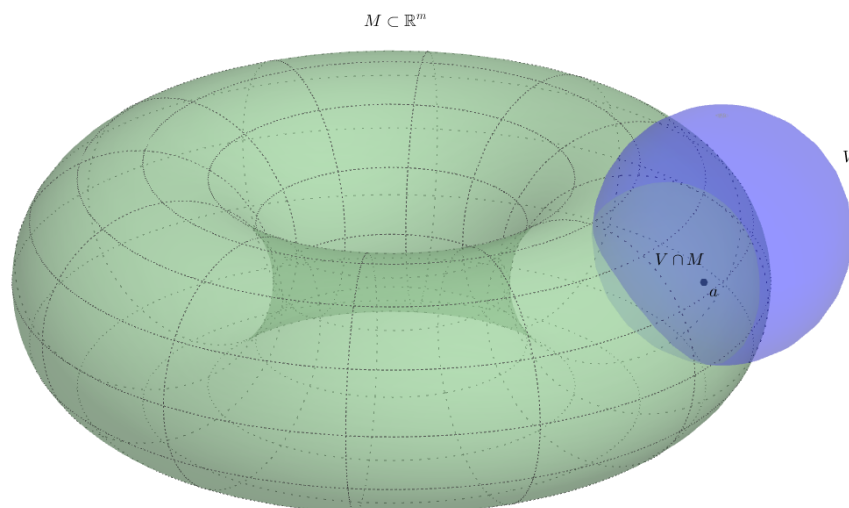


Em muitos contextos, é comum identificar  $U$  com a sua imagem  $f(U)$  e dizer que  $U$  está **mergulhado** em  $\mathbb{R}^{n+k}$ . Sendo imersões, podemos expressar  $f$  em forma canônica, como nas seções anteriores.

## 11 Superfícies diferenciáveis em $\mathbb{R}^m$

Intuitivamente, uma superfície de dimensão  $n$  é um subconjunto do espaço euclidiano que localmente se parece com um subconjunto aberto de  $\mathbb{R}^n$ . É neste sentido que podemos pensar na superfície da Terra como uma superfície bidimensional em  $\mathbb{R}^3$  (talvez algum outro espaço? ☺), já que localmente se parece com um plano. Existem várias maneiras de formalizar este conceito. Faremos de pelo menos duas formas diferentes ao longo do nosso curso. Segue a primeira.

Uma **superfície** de **dimensão**  $n$  e de classe  $C^k$  em  $\mathbb{R}^m$  é um conjunto  $M \subseteq \mathbb{R}^m$  tal que todo ponto  $a \in M$  possui vizinhança aberta  $V \subseteq \mathbb{R}^m$  que pode ser **parametrizada** por um aberto  $U \subseteq \mathbb{R}^n$ , isto é, existe um mergulho  $\phi : U \rightarrow V \cap M$  de classe  $C^k$ . Nós também dizemos que  $\phi$  é uma **parametrização** ou uma **carta** (de classe  $C^k$ ). Também dizemos que  $M$  tem **codimensão**  $m - n$ .



A ideia, em uma superfície de classe  $C^k$ , é que cada vizinhança  $V \cap M$  do desenho acima pode ser vista como o mergulho de um aberto  $U \subseteq \mathbb{R}^n$ .

**Exemplo 30.** O exemplo que talvez seja o mais familiar é o gráfico de uma função de classe  $C^k$ . Mais explicitamente, se  $U \subseteq \mathbb{R}^m$  um aberto e  $f : U \rightarrow \mathbb{R}^n$  uma função de classe  $C^k$ , vamos mostrar que o gráfico de  $f$

$$M = \{(x, f(x)) \in \mathbb{R}^{m+n}; x \in U\}$$

é uma superfície de dimensão  $n$  e classe  $C^k$  em  $\mathbb{R}^{m+n}$ . De fato, considere  $\phi : U \rightarrow M \subseteq \mathbb{R}^{m+n}$  dada por

$$\phi(x) = (x, f(x))$$

e observe que  $\phi$  é uma parametrização de todo ponto de  $M$ :

- $\phi(x) = \phi(y) \implies x = y$ ;
- $\phi^{-1} : M \rightarrow U$  é dada por  $\phi^{-1}(x, y) = x$  e é a restrição a  $M$  de uma função contínua;
- $\phi'(x)(h) = (h, f'(x) \cdot h)$  é injetiva.

Um dos resultados desta seção diz que, *localmente*, toda superfície diferenciável  $M$  de dimensão  $n$  e de classe  $C^k$  é o gráfico de uma função  $f : U \subseteq \mathbb{R}^n \rightarrow \mathbb{R}^{m-n}$  de classe  $C^k$ . Esta propriedade não é verdade globalmente, como mostra o exemplo do toro acima ou mesmo da esfera  $x^2 + y^2 + z^2 = 1$ .

**Proposição 31.** *Toda superfície diferenciável  $M$  de dimensão  $n$  e de classe  $C^k$  em  $\mathbb{R}^m$  é, localmente, o gráfico de uma função  $f : U_0 \subseteq \mathbb{R}^n \rightarrow \mathbb{R}^{m-n}$  de classe  $C^k$ .*

*Demonstração.* Consideramos uma parametrização qualquer  $\phi : U \rightarrow M$  e seja  $a \in \phi(U)$ . Consideramos  $x_0 \in U$  tal que  $\phi(x_0) = a$ . Sendo  $\phi$  uma imersão, as colunas da matriz associada com a transformação linear  $\phi'(x_0) : \mathbb{R}^n \rightarrow \mathbb{R}^m$  são linearmente independentes. Denotamos por  $\pi$  a projeção sobre as  $n$  primeiras coordenadas, isto é, para  $(h, k) \in \mathbb{R}^n \times \mathbb{R}^{m-n}$ ,  $\pi(h, k) = h$ . Logo, pelo Teorema da Aplicação Inversa, existem vizinhanças  $U_0$  de  $x_0$  e  $V$  de  $\pi(a)$  tais que  $\pi \circ \phi : U_0 \rightarrow V$  é um difeomorfismo. Observe agora que podemos definir uma nova parametrização da vizinhança  $\phi(U_0) \subseteq M$  por  $\psi := \phi \circ (\pi \circ \phi)^{-1}$ . Trivialmente, temos  $(\pi \circ \psi)(v) = v$  para todo  $v \in V$ . Logo, para  $v \in V$ , o vetor  $\psi(v) \in \mathbb{R}^m$  pode ser escrito como

$$\psi(v) = (v, f(v)).$$

Como  $\psi$  é de classe  $C^k$ , também  $f$  o é. □

**Exemplo 32.** A esfera  $\mathbb{S}^n = \{(x_1, x_2, \dots, x_{n+1}) \in \mathbb{R}^{n+1}; x_1^2 + x_2^2 + \dots + x_{n+1}^2 = 1\} \subseteq \mathbb{R}^{n+1}$  é uma superfície  $n$ -dimensional e de classe  $C^\infty$  de  $\mathbb{R}^{n+1}$ .

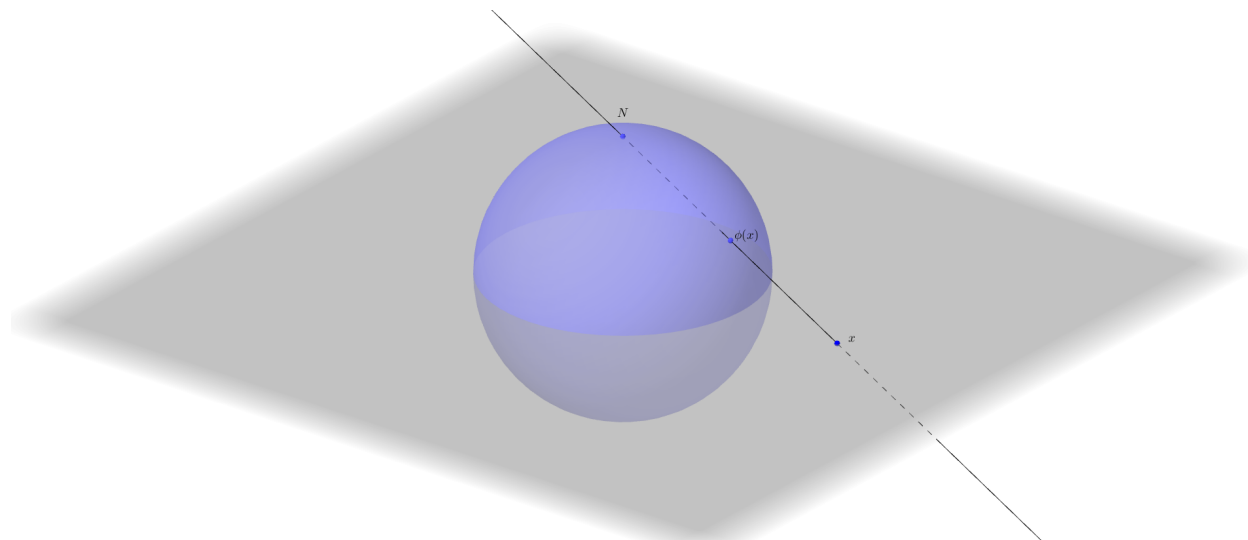
Para justificar esta afirmação, notamos que qualquer ponto do polo norte  $x_{n+1} > 0$  da esfera pode ser parametrizada por  $\phi : B_1(0) \rightarrow \{x_{n+1} > 0\} \subseteq \mathbb{R}^{n+1}$  dada por

$$\phi(x) = \left( x_1, x_2, \dots, x_n, \sqrt{1 - x_1^2 - x_2^2 - \dots - x_n^2} \right).$$

Como exercício, verifique que  $\phi$  é um mergulho. Fazendo o análogo para o polo sul (sinal negativo na fórmula acima), já conseguimos cobrir todos os pontos da esfera, exceto o equador  $x_{n+1} = 0$ . Podemos, no entanto, repetir o argumento isolando cada uma das variáveis e cobrir toda a esfera  $\mathbb{S}^n$  com  $2n + 2$  parametrizações.

Uma outra possibilidade, que permite cobrir a esfera com apenas duas parametrizações ou cartas, é usar projeções estereográficas:





Denotamos  $N = (0, 0, \dots, 0, 1)$  o polo norte da esfera. Para  $x = (x_1, x_2, \dots, x_n) \in \mathbb{R}^n$ , associamos o ponto  $\phi(x) \in \mathbb{S}^n \subset \mathbb{R}^{n+1}$ , como na figura. Uma fórmula pode ser obtida pela parametrização do segmento da figura é

$$\phi(x) = (1-t)N + t(x, 0) = (tx_1, tx_2, \dots, tx_n, 1-t).$$

O parâmetro  $t$  é escolhido de modo que se tenha  $\phi(x) \in \mathbb{S}^n$ :

$$t^2x_1^2 + \dots + t^2x_n^2 + (1-t)^2 = 1 \iff t = \frac{2}{1+|x|^2}.$$

Logo,

$$\phi(x) = \left( \frac{2x_1}{1+|x|^2}, \frac{2x_2}{1+|x|^2}, \dots, \frac{2x_n}{1+|x|^2}, \frac{-1+|x|^2}{1+|x|^2} \right) = \left( \frac{2x}{1+|x|^2}, \frac{-1+|x|^2}{1+|x|^2} \right).$$

Também olhando para o desenho, podemos escrever a inversa de  $\phi$  e obter que  $\phi$  é uma parametrização de classe  $C^\infty$  para o conjunto aberto  $\mathbb{S}^n \setminus \{S\}$ , onde  $S$  é o polo sul da esfera.

Ainda, um terceiro modo é a utilização das conhecidas coordenadas esféricas, que já aparecem nos cursos de cálculo. Por simplicidade, consideramos o caso  $n = 2$  da esfera  $\mathbb{S}^2$  em  $\mathbb{R}^3$ : uma parametrização é dada por:  $\varphi : (0, 2\pi) \times (0, \pi) \subset \mathbb{R}^2 \rightarrow \mathbb{S}^2$  dada por

$$\varphi(\theta, \phi) = (\cos \theta \sin \phi, \sin \theta \sin \phi, \cos \phi).$$

Quais pontos esta parametrização não cobre. Como remediar isto?

**Exercício 24.** Obter uma parametrização do toro (superfície que parece uma rosquinha - ver figura da página 31) em  $\mathbb{R}^3$ .

**Exercício 25 (Produtos cartesianos).** Verifique que o produto cartesiano  $M \times N \subseteq \mathbb{R}^m \times \mathbb{R}^k$  de duas (ou de qualquer número finito de) superfícies

- $M \subseteq \mathbb{R}^m$  de dimensão  $p$  em  $\mathbb{R}^m$  e
- $N \subseteq \mathbb{R}^k$  de dimensão  $q$  em  $\mathbb{R}^k$

é uma superfície de dimensão  $p+q$  em  $\mathbb{R}^m \times \mathbb{R}^k$ , cuja classe de diferenciabilidade é menor dentre as duas.

A partir de agora, dada uma superfície diferenciável  $M \subseteq \mathbb{R}^m$ , nós gostaríamos de dar sentido à noção de diferenciabilidade da aplicações da forma

$$f : M \rightarrow \mathbb{R} \quad \text{ou} \quad f : M \rightarrow \mathbb{R}^k \quad \text{ou mesmo} \quad f : M \rightarrow N.$$

A menos que a dimensão de  $M$  seja igual a  $m$ , a superfície  $M$  tem interior vazio em  $\mathbb{R}^m$  e, portanto, não é um conjunto aberto. Por isto este é um conceito novo, que não consequência imediata da nossa definição usual de diferenciabilidade.

O seguinte lema é de importância fundamental neste sentido.

**Teorema 33.** *Seja  $M \subseteq \mathbb{R}^m$  uma superfície de dimensão  $n$  e de classe  $C^k$  e  $\phi : U \rightarrow \phi(U) \subseteq M$  uma parametrização de classe  $C^k$ . Então*

(i)  $f : V \subseteq \mathbb{R}^p \rightarrow \mathbb{R}^m$  de classe  $C^k$  com  $f(V) \subset \phi(U) \implies \phi^{-1} \circ f : V \rightarrow U$  é de classe  $C^k$ .

(ii)  $\psi : W \rightarrow \phi(U)$  é uma parametrização  $\iff \xi := \psi^{-1} \circ \phi : U \rightarrow W$  é um difeomorfismo.

*Demonstração.* Começamos por provar (i). Note inicialmente que este item não segue da regra da cadeia, pois a inversa  $\phi^{-1}$  não está definida em um subconjunto aberto de  $\mathbb{R}^m$ . Mas sabemos que  $\phi$  é um mergulho. Dado um ponto  $a \in V$ , vamos mostrar que  $\phi^{-1} \circ f$  é de classe  $C^k$  em uma vizinhança de  $a$ . Sendo  $\phi$  imersão, existe um sistema de coordenadas onde  $\phi$  se comporta como uma inclusão. Ou seja, existe um difeomorfismo  $g : Z \rightarrow \tilde{V} \times W$  de classe  $C^k$  tal que

$$g(\phi(v)) = (v, 0).$$

Temos que  $Z$  é uma vizinhança de  $\phi(x_0) = f(a)$  e  $\tilde{V}$  é uma vizinhança de  $x_0$ . Assim, escrevendo por  $\pi$  a projeção na primeira coordenada, podemos escrever

$$(\pi \circ g \circ \phi)(v) = v, \quad \text{ou ainda,} \quad \pi \circ g|_{\phi(\tilde{V})} = \phi^{-1}.$$

Logo,  $\phi^{-1}$  é a restrição da aplicação  $\pi \circ g$  de classe  $C^k$  ao conjunto  $\phi(\tilde{V}) \subseteq M$ . Sendo  $f$  contínua,  $V_1 := f^{-1}(\phi(\tilde{V}))$  é uma vizinhança aberta de  $a$  em  $\mathbb{R}^p$  e tem-se

$$\phi^{-1} \circ f = \pi \circ g \circ f \quad \text{em } V_1.$$

Logo,  $\phi^{-1} \circ f$  é de classe  $C^k$  em  $V_1$ , que é o que tínhamos afirmado.

Para provar (ii), observamos que o item (i) implica que ambas as funções

$$\xi = \psi^{-1} \circ \phi \quad \text{e} \quad \xi^{-1} = \phi^{-1} \circ \psi$$

são diferenciáveis. Logo  $\xi$  é um difeomorfismo. □

Equipados com o lema acima, nós podemos definir diferenciabilidade para funções definidas em superfícies (e não necessariamente um aberto do espaço euclidiano).

Uma função  $f : M \rightarrow \mathbb{R}^d$  é dita **diferenciável** quando, para toda carta  $\phi : U \subseteq \mathbb{R}^n \rightarrow M$ , tem-se que  $f \circ \phi : U \rightarrow \mathbb{R}^d$  é uma função diferenciável. Em outras palavras, uma aplicação  $f$  é dita diferenciável na superfície  $M$  se for diferenciável quando vista em coordenadas locais. Note que a noção de diferenciabilidade deve ser uma propriedade intrínseca de  $M$ , ou seja, não pode depender do sistema de coordenadas particular utilizado. Neste ponto que o lema acima é fundamental. Se eu utilizo o sistema de coordenadas  $\phi$  e o leitor um outro  $\psi$ , o lema acima implica que

$$f \circ \phi \text{ é diferenciável} \iff f \circ \psi = (f \circ \phi) \circ (\phi^{-1} \circ \psi) \text{ é diferenciável.}$$

O **espaço tangente**  $T_x M \subseteq \mathbb{R}^m$  a uma superfície  $M \subseteq \mathbb{R}^m$  no ponto  $x$  pode ser definido como a imagem  $\phi'(x_0)(\mathbb{R}^n)$ , onde  $\phi : U \rightarrow M$  é qualquer parametrização com  $x = \phi(x_0) \in \phi(U)$ . Como  $\phi$  é uma imersão,  $T_x M$  tem dimensão  $n$ , para todo  $x \in M$ . Uma base pode ser obtida com os vetores  $\phi(x_0) \cdot e_i$ , isto é,

$$T_x M = \text{Span} \left\{ \frac{\partial \phi}{\partial x_1}(x_0), \frac{\partial \phi}{\partial x_2}(x_0), \dots, \frac{\partial \phi}{\partial x_n}(x_0) \right\} \subset \mathbb{R}^m.$$

**Proposição 34.** *Temos a seguinte caracterização para o espaço tangente definido acima:*

$$T_x M = \{v \in \mathbb{R}^m; v = \gamma'(0) \text{ para alguma curva diferenciável } \gamma : (-\delta, \delta) \rightarrow M \text{ com } \gamma(0) = x\}.$$

Em particular,  $T_x M$  é independente de qualquer parametrização de  $M$ .

*Demonstração.* Sejam  $\gamma : (-\delta, \delta) \rightarrow M$  uma curva diferenciável em  $M \subseteq \mathbb{R}^m$  com  $\gamma(0) = x$  e  $\phi$  uma parametrização de uma vizinhança de  $x$  que contém toda a curva  $\gamma$  (isto é possível pois podemos, se necessário, diminuir  $\delta > 0$ ). Trazemos a curva de “volta” para  $\mathbb{R}^n$  pela parametrização:  $\mu(t) = \phi^{-1}(\gamma(t))$  satisfaz

$$\gamma'(0) = (\phi \circ \phi^{-1} \circ \gamma)'(0) = \left. \frac{d}{dt} \phi(\mu(t)) \right|_{t=0} = \phi'(\mu(t)) \cdot \mu'(t) \Big|_{t=0} = \phi'(x_0) \cdot \mu'(0);$$

logo,  $\gamma'(0) \in \phi'(x_0)(\mathbb{R}^n) = T_x M$ .

Reciprocamente, consideramos  $v \in \phi'(x_0)(\mathbb{R}^n)$ , isto é, suponhamos que exista  $u \in \mathbb{R}^n$  tal que  $v = \phi'(x_0) \cdot u$ . Para todo  $t \ll 1$ , temos  $x_0 + tu \in U$  e definimos  $\gamma(t) = \phi(x_0 + tu)$ . Trivialmente  $\gamma(0) = x$  e temos

$$\gamma'(0) = \phi'(x_0) \cdot u = v. \quad \square$$

Quando desenhamos o espaço tangente, estamos pensando em um espaço linear que aproxima a superfície  $M$ . No entanto, não estamos nos importando que a aproximação perto de  $x \in M$  ocorre perto da origem  $0 \in T_x M$ . Em verdade, é mais natural indexar os espaços vetoriais  $T_x M$  pensar na união disjunta

$$TM := \bigcup_{x \in M} \{x\} \times T_x M = \{(x, v); x \in M, v \in T_x M\} \subset \mathbb{R}^{2m},$$

que é conhecido como o **fibrado tangente** à superfície  $M$ . Este fibrado possui uma estrutura de superfície diferenciável em  $\mathbb{R}^{2m}$  (de classe  $C^{k-1}$ ), de dimensão  $2n$ , induzida pelas parametrizações da superfície  $M$ .

**Exercício 26.** Prove esta última afirmação.

## 11.1 Derivada de aplicações entre superfícies

Assim como na seção anterior, podemos falar de aplicações diferenciáveis entre superfícies. A ideia é que uma aplicação  $f$  com valores em uma superfície  $N$  é diferenciável na superfície  $M$  se for diferenciável quando vista em coordenadas locais de  $M$  e de  $N$ .

Sejam  $M$  superfície de dimensão  $n$  e  $N$  superfície de dimensão  $p$ , ambas de classe  $C^k$ . Uma função  $f : M \rightarrow N$  é dita **diferenciável** quando<sup>10</sup>, para toda carta  $\phi : U \subseteq \mathbb{R}^n \rightarrow M$  e para toda carta  $\psi : V \subseteq \mathbb{R}^p \rightarrow N$ , tem-se que

$$\psi^{-1} \circ f \circ \phi : U \rightarrow V$$

é uma função diferenciável (entre espaços euclidianos). Esta definição, como deveria ser, é independente das parametrizações escolhidas, pois, dadas outras cartas nas mesmas vizinhanças de cada superfície, temos

$$\psi_1^{-1} \circ f \circ \phi_1 = (\psi_1^{-1} \circ \psi) \circ (\psi^{-1} \circ f \circ \phi) \circ (\phi^{-1} \circ \phi_1)$$

e as mudanças de parâmetros são todas diferenciáveis.

A **derivada** de uma aplicação diferenciável  $f : M \rightarrow N$  é a transformação linear

$$f'(x) : T_x M \rightarrow T_{f(x)} N$$

dada pela seguinte lei: se  $v \in T_x M$  e  $\gamma$  é uma curva diferenciável com  $\gamma(0) = x$  e  $\gamma'(0) = v$ , então

$$f'(x) \cdot v = \left. \frac{d}{dt} f(\gamma(t)) \right|_{t=0}.$$

Note que  $f'(x) \cdot v \in T_{f(x)} N$  pois  $f \circ \gamma$  é uma curva em  $N$  com  $(f \circ \gamma)(0) = f(x)$ .

Na notação da seção anterior, uma aplicação  $f : M \rightarrow N$  induz uma aplicação  $f_* : TM \rightarrow TN$  entre os respectivos fibrados tangentes:

$$f_*(x, v) = (f(x), f'(x) \cdot v).$$

Esta aplicação  $f_*$  é às vezes chamada de **push-forward** de  $f$ .

<sup>10</sup>Para as considerações desta seção, de superfícies mergulhadas em espaços euclidianos, não é necessária a parametrização de  $N$  nesta definição. Voltaremos a esta observação quando discutirmos variedades diferenciáveis.

## 12 Valores regulares

Seja  $f : U \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$  uma aplicação diferenciável. Nós dizemos que  $c \in \mathbb{R}^m$  é um **valor regular** de  $f$  quando, para todo  $x \in f^{-1}(c)$ , tem-se  $f'(x) : \mathbb{R}^n \rightarrow \mathbb{R}^m$  sobrejetiva. Em particular, deve-se ter  $n \geq m$ .

**Teorema 35 (Superfície de nível de um valor regular).** *Seja  $f : U \subset \mathbb{R}^{n+k} \rightarrow \mathbb{R}^k$  uma função de classe  $C^k$ . Se  $c \in \mathbb{R}^k$  é um valor regular de  $f$ , então  $M = f^{-1}(c)$  é uma superfície de dimensão  $n$  e classe  $C^k$  em  $\mathbb{R}^{n+k}$ . Além disso,*

$$T_x M = \ker f'(x) \quad \text{para todo } x \in M.$$

*Demonstração.* Este teorema é um rephraseamento do Teorema da Função Implícita. Sendo  $f'(w)$  sobrejetiva, sua matriz canônica possui  $n$  colunas linearmente independentes. Reordenando as coordenadas, se necessário, podemos supor que são as  $n$  primeiras. Escrevemos  $w = (x, y) \in \mathbb{R}^{n+k}$ . As primeiras colunas linearmente independentes implicam diretamente que

$$f'(x, y) \cdot (h, 0) = 0 \implies h = 0.$$

Logo, localmente,  $x = \xi(y)$  de classe  $C^k$ , de modo que, localmente,  $M$  é o gráfico de uma função de classe  $C^k$  e, portanto, uma superfície diferenciável, como enunciado.

Para a segunda afirmação, seja  $v \in T_x M$ . Logo,  $v = \gamma'(0)$  para alguma curva diferenciável  $\gamma : (-\delta, \delta) \rightarrow M$  com  $\gamma(0) = x$ . Pela definição de  $M$ , temos

$$f(\gamma(t)) = c \implies f'(x)[v] = 0.$$

Logo,  $T_x M \subseteq \ker f'(x)$ . Sendo  $\dim \text{Im } f'(x) = k$ , devemos ter  $\dim \ker f'(x) = n$ . Portanto, só pode ser que  $T_x M = \ker f'(x)$ .  $\square$

Considere uma aplicação  $f : \mathbb{R}^{n+1} \rightarrow \mathbb{R}$ . Um valor regular neste caso é um número  $c \in \mathbb{R}$  tal que, para todo  $x \in f^{-1}(c)$ , temos  $f'(x) \neq 0 \in (\mathbb{R}^{n+1})^*$ . Ou, equivalentemente,  $\nabla f(x) \neq 0 \in \mathbb{R}^{n+1}$ . Neste caso,  $M = f^{-1}(c)$  é uma superfície de dimensão  $n$  (ou hiperfície<sup>11</sup>) em  $\mathbb{R}^{n+1}$  e o espaço tangente é

$$T_x M = \ker df(x) = \{\nabla f(x)\}^\perp = \{v \in \mathbb{R}^{n+1}; \nabla f(x) \cdot v = 0\}.$$

**Exemplo 36.** A esfera  $\mathbb{S}^n \subset \mathbb{R}^{n+1}$  pode ser vista como a imagem inversa de  $c = 1$  pela função

$$f(x) = x_1^2 + x_2^2 + \dots + x_{n+1}^2 = |x|^2.$$

Temos  $\nabla f(x) = 2x$ , de modo que todo  $c \neq 0$  é um valor regular de  $f$  (apesar de somente  $c > 0$  ser interessante). Em particular, para  $c = 1$ , temos que  $\mathbb{S}^n = f^{-1}(1)$  é uma hiperfície de classe  $C^\infty$  cujo espaço tangente é

$$T_x M = \{x\}^\perp = \{v \in \mathbb{R}^{n+1}; x \cdot v = 0\}.$$

**Exemplo 37.** Seja  $A \in M(n \times n)$  uma matriz real simétrica e consideramos a forma quadrática  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  dada por

$$f(x) = \langle Ax, x \rangle = \sum_{i,j=1}^n a_{ij} x_i x_j.$$

Temos  $\nabla f(x) = 2Ax$ . Para  $c \neq 0$ , temos

$$x \in f^{-1}(c) \iff \langle Ax, x \rangle = c \implies Ax \neq 0 \iff \nabla f(x) \neq 0.$$

Segue que  $M = f^{-1}(c)$  é uma hiperfície de classe  $C^\infty$  em  $\mathbb{R}^n$ . Seu espaço tangente em um ponto  $x \in M$  pode ser caracterizado por

$$T_x M = \{A(x)\}^\perp.$$

A esfera é o caso especial em que  $A = I$ .

<sup>11</sup>Uma **hiperfície** é, por definição, uma superfície de codimensão um, isto é, uma superfície cuja dimensão é um a menos que a dimensão do espaço ambiente.

**Exemplo 38.** Consideramos o grupo especial linear

$$SL(\mathbb{R}^n) = \{A \in M(n \times n); \det A = 1\} \subset M(n \times n) \simeq \mathbb{R}^{n^2}$$

e a aplicação  $f : \mathbb{R}^{n^2} \rightarrow \mathbb{R}$  dada por

$$f(A) = \det A.$$

Expandindo a linha  $i$  em cofatores, obtemos

$$f(A) = \sum_{j=1}^n (-1)^{i+j} a_{ij} \det A_{ij};$$

logo,

$$\frac{\partial f}{\partial a_{ij}}(A) = (-1)^{i+j} \det A_{ij}.$$

Segue que os pontos críticos de  $f$  são as matrizes com todos os determinantes menores nulos. Em particular, pela expansão em cofatores, tem-se que todo ponto crítico tem determinante nulo. Logo, toda matriz em  $SL(\mathbb{R}^n) = f^{-1}(1)$  não é um ponto crítico de  $f$ . Equivalentemente,  $c = 1$  é valor regular de  $f$ . Segue que  $SL(\mathbb{R}^n)$  é uma hipersuperfície de classe  $C^\infty$  em  $M(n \times n)$ .

Vamos analisar o espaço tangente  $T_I(SL(\mathbb{R}^n))$ .<sup>12</sup> Note que

$$\frac{\partial f}{\partial a_{ij}}(I) = \delta_{ij} = \begin{cases} 1, & \text{se } i = j \\ 0, & \text{se } i \neq j \end{cases}.$$

Logo,

$$f'(I) \cdot H = 0 \iff \sum_{i,j=1}^n a_{ij} \delta_{ij} = 0 = \sum_{i=1}^n a_{ii} = \text{tr } A.$$

Portanto, o espaço tangente a  $SL(\mathbb{R}^n)$  em  $I$  é o espaço das matrizes de traço nulo:

$$SL(\mathbb{R}^n) = \{A \in M(n \times n); \text{tr } A = 0\}.$$

**Exemplo 39.** Considere o grupo ortogonal e o conjunto das matrizes simétricas:

$$\mathcal{O}(n) = \{A \in M(n \times n); AA^T = I\} \subset M(n \times n) \simeq \mathbb{R}^{n^2} \text{ e } \mathcal{S}(n) = \{A \in M(n \times n); A = A^T\} \simeq \mathbb{R}^{n(n+1)/2}.$$

Seja  $f : \mathbb{R}^{n^2} \rightarrow \mathbb{R}^{n(n+1)/2}$  a aplicação

$$f(A) = AA^T$$

Assim,  $\mathcal{O}(n) = f^{-1}(I)$ . Temos (exercício)

$$f'(A) \cdot H = HA^T + AH^T.$$

Mostramos que, para  $A \in \mathcal{O}(n)$ , a transformação  $f'(A)$  é sobrejetiva: para  $S \in \mathcal{S}(n)$  simétrica, temos

$$f'(A) \cdot \frac{SA}{2} = \frac{SA}{2} A^T + A \frac{A^T S^T}{2} = S \text{ quando } AA^T = I.$$

Segue que  $\mathcal{O}(n)$  é uma superfície  $C^\infty$  e de dimensão

$$\dim \mathcal{O}(n) = n^2 - \frac{n(n+1)}{2} = \frac{n(n-1)}{2}$$

em  $\mathcal{S}(n)$ . O espaço tangente fica caracterizado como

$$T_I(\mathcal{O}(n)) = \{H \in M(n \times n); H + H^T = 0\},$$

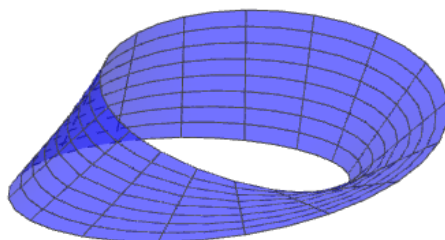
que é o conjunto das matrizes anti-simétricas e que, necessariamente, deve ter a dimensão correta (embora obter esta dimensão seja muito mais elementar com métodos puramente algébricos).

<sup>12</sup> $SL(\mathbb{R}^n)$  possui uma estrutura de grupo com o produto de matrizes. Uma superfície diferenciável que também é um grupo é conhecida como um **grupo de Lie**. Para grupos de Lie, basta conhecer o espaço tangente em seu elemento neutro para conhecer o espaço tangente em qualquer outro ponto. Ver, por exemplo, [3, 18].

**Exemplo 40.** Nem toda superfície é imagem inversa de valor regular. Se uma hiperfície da forma  $M = f^{-1}(c)$  para  $c$  valor regular e  $f \in C^1$ , então

$$x \in M \mapsto \frac{\nabla f(x)}{\|\nabla f(x)\|} \in \mathbb{R}^{n+1}$$

é um campo de vetores contínuo, não nulo (unitário) e normal à  $M$  (isto é,  $\langle \nabla f(x), v \rangle = 0$  para todo  $v \in T_x M$ ). Nem toda superfície admite um campo de vetores com estas propriedades. Um exemplo é a faixa de Möbius do desenho abaixo.



**Proposição 41.** Toda superfície diferenciável é, localmente, a imagem inversa de um valor regular.

*Demonstração.* Já sabemos que toda superfície  $M \subset \mathbb{R}^m$  de dimensão  $n$  é, localmente, o gráfico de uma aplicação diferenciável  $\xi : U \subset \mathbb{R}^n \rightarrow \mathbb{R}^{m-n}$ . Definimos  $f : U \times \mathbb{R}^{m-n} \rightarrow \mathbb{R}^m$  por

$$f(x, y) = y - \xi(x) = (y_i - \xi_i(x)).$$

Temos que o gráfico de  $\xi$  é a superfície de nível  $f^{-1}(0)$  e

$$[f'(x, y)] = [\xi'(x) \quad I_{(m-n) \times (m-n)}]$$

é certamente sobrejetiva. □

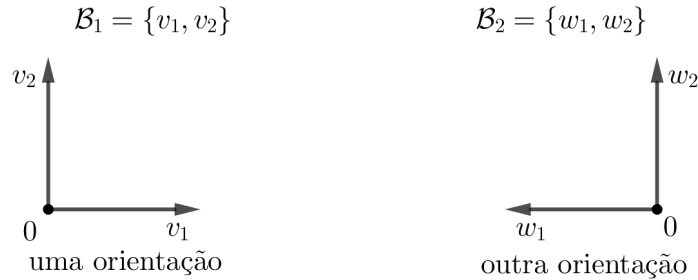
**Observação 42.** Note que nem toda superfície que é imagem inversa de valor regular é o gráfico de uma função. Considere, por exemplo, a esfera. Mais geralmente, o gráfico de uma função é homeomorfo ao seu domínio e é, por consequência, aberto. Dessa maneira, nenhuma superfície compacta pode ser o gráfico de uma função (apenas localmente, o que já demonstramos).

## 13 Superfícies Orientáveis

Começamos discutindo orientação de espaços vetoriais e em seguida definimos superfícies orientáveis. Principais referências consultadas: [3, 7, 10, 11].

### 13.1 Orientação de espaços vetoriais

Intuitivamente, em um espaço vetorial, uma orientação é a forma particular que escolhemos para representar os vetores. De forma (bem pouco) mais rigorosa, é a forma que escolhemos para representar os vetores da base do espaço vetorial. Por exemplo, no espaço  $\mathbb{R}^2$ , as duas bases da figura abaixo têm orientações diferentes.



Nenhuma dessas orientações é mais natural do que a outra (inclusive, se olharmos para a figura pelo lado avesso da página, uma base se transforma na outra). É apenas uma questão de gosto e, claro, muitas vezes do nosso costume de utilizar bases como a da esquerda. Um outro exemplo é a nossa escolha de orientação para o espaço  $\mathbb{R}^3$  a partir da regra da mão direita. Esta escolha é certamente arbitrária; no entanto, é interessante o uso desta convenção para facilitar a leitura de textos diferentes (por exemplo, é difícil encontrar um texto de cálculo onde o produto vetorial é definido a partir da regra da mão esquerda!).

Sejam  $\mathcal{B}_1 = \{v_1, v_2, \dots, v_n\}$  e  $\mathcal{B}_2 = \{w_1, w_2, \dots, w_n\}$  bases ordenadas de um espaço vetorial  $V$ . Nós dizemos que as duas  $\mathcal{B}_1$  e  $\mathcal{B}_2$  têm a **mesma orientação** quando o determinante da matriz de mudança de bases é positivo. Mais explicitamente, escrevendo

$$v_i = \sum_{j=1}^n a_{ij} w_j,$$

$\mathcal{B}_1$  e  $\mathcal{B}_2$  têm a mesma orientação se, e somente se,

$$\det [a_{ij}] > 0.$$

Note que (exercício!) a propriedade “ter mesma orientação” define uma relação de equivalência que possui apenas duas classes de equivalência. Uma **orientação** para  $V$  é a escolha de uma dessas classes. Em geral, depois de feita a escolha por uma das classes de equivalência, chamamos qualquer base desta orientação de **base positiva**. As bases da outra classe de equivalência possível são então chamadas de **bases negativas**.

Notamos que esta definição captura as propriedades que falamos no início da seção. Se trocamos o sinal de algum dos vetores da base, nós mudamos a orientação; se trocamos dois vetores da base ordenada de lugar, mudamos a orientação. Mais geralmente, se fizermos uma permutação dos vetores da base, uma permutação par corresponde a não trocar orientação enquanto que uma permutação ímpar corresponde a trocar de orientação.

Quando  $V = \mathbb{R}^n$ , nós fixamos a orientação que deixa a base canônica positiva. Assim, dada uma outra base  $\mathcal{B} = \{v_1, v_2, \dots, v_n\}$  de  $\mathbb{R}^n$ , tem-se que  $\mathcal{B}$  é uma base positiva (isto é, tem a mesma orientação da base canônica) quando

$$\det [v_1 \ v_2 \ \dots \ v_n] > 0.$$

A matriz acima é formada colocando os vetores  $v_i$  nas colunas.

## 13.2 Orientação de superfícies

A ideia por trás de uma superfície orientada é a possibilidade de fixar orientações para os espaços tangentes que são compatíveis quando se faz mudanças de parâmetros.

Seja  $M \subset \mathbb{R}^m$  uma superfície de classe  $C^k$  e dimensão  $n$  e  $\phi : U \rightarrow \phi(U) \subset M$  uma parametrização. Uma orientação para os espaços tangentes  $T_x M$ , para  $x \in U$ , é obtida a partir da parametrização  $\phi$ , a saber, a orientação das bases

$$\left\{ \frac{\partial \phi}{\partial x_1}(x), \frac{\partial \phi}{\partial x_2}(x), \dots, \frac{\partial \phi}{\partial x_n}(x) \right\} \subset T_x M.$$

Se tivermos uma outra parametrização  $\psi : V \rightarrow \psi(V) \subset M$  com  $W := \phi(U) \cap \psi(V) \neq \emptyset$ , então a orientação induzida na vizinhança  $W$  por cada uma das parametrizações  $\phi$  e  $\psi$  pode coincidir ou não (dãã!). É possível decidir qual é o caso olhando para a matriz de mudança de bases. Pela regra da cadeia,

$$\frac{\partial \phi}{\partial x_j}(x) = \frac{\partial}{\partial x_j}(\psi \circ \psi^{-1} \circ \phi)(x) = \sum_{i=1}^n \frac{\partial \psi}{\partial y_i}((\psi^{-1} \circ \phi)(x)) \frac{\partial (\psi^{-1} \circ \phi)_i}{\partial x_j}(x).$$

Escrevendo<sup>13</sup>  $y = (\psi^{-1} \circ \phi)(x)$ , obtemos a fórmula de mudança de base em sua forma mais clássica:

$$\frac{\partial \phi}{\partial x_j}(x) = \sum_{i=1}^n \frac{\partial y_i}{\partial x_j}(x) \frac{\partial \psi}{\partial y_i}(y).$$

Segue da discussão acima que a orientação dos espaços tangentes  $T_x M$  obtidas a partir de duas parametrizações, para  $x \in W$ , é a mesma quando

$$\det \left[ \frac{\partial (\psi^{-1} \circ \phi)_i}{\partial x_j}(x) \right] > 0.$$

Notamos que o determinante acima é o determinante do Jacobiano da mudança de parâmetros.

Nós vamos dizer que  $\phi : U \rightarrow \phi(U) \subset M$  e  $\psi : V \rightarrow \psi(V) \subset M$  são duas cartas **coerentes** ou **compatíveis** quando

$$\det J(\psi^{-1} \circ \phi) > 0.$$

Por questões técnicas, não pedimos que  $W = \phi(U) \cap \psi(V) \neq \emptyset$  e cartas em vizinhanças coordenadas disjuntas serão sempre consideradas compatíveis. Quando é possível cobrir a superfície  $M$  por parametrizações duas a duas coerentes, nós dizemos que  $M$  é uma superfície **orientável**. Uma coleção de parametrizações coerentes  $\mathcal{A}$  que cobrem toda a superfície  $M$  é chamado de um **atlas coerente**. Adicionando cartas coerentes, se necessário, é fácil de se convencer que todo atlas coerente está contido em um atlas coerente maximal<sup>14</sup>. Um atlas coerente maximal é dito uma **orientação** de  $M$  e, no caso de fixarmos um atlas coerente maximal, dizemos que a superfície  $M$  está **orientada**. Adiantamos que nem toda superfície admite um atlas coerente, isto é, nem toda superfície é orientável. Antes de analisar esta questão, vejamos alguns exemplos.

**Exemplo 43.** Qualquer superfície que possa ser coberta por uma vizinhança coordenada só, como gráficos de funções, são orientáveis.

**Exemplo 44.** Subconjuntos abertos de superfícies orientáveis são também orientáveis; basta considerar a restrição das cartas do atlas.

**Exemplo 45.** Qualquer superfície que possa ser coberta por apenas duas parametrizações e cuja interseção das vizinhanças coordenadas for conexa, é necessariamente orientável.

Isto segue da seguinte observação: se  $M = \phi(U) \cup \psi(V)$  para duas parametrizações  $\phi$  e  $\psi$ , então

$$\text{ou } \det J(\psi^{-1} \circ \phi) > 0 \text{ sempre ou } \det J(\psi^{-1} \circ \phi) < 0 \text{ sempre,}$$

pois  $\psi^{-1} \circ \phi$  é um difeomorfismo e  $\phi(U) \cap \psi(V)$  é conexo. No caso de ser positivo, pela definição de orientabilidade,  $M$  é orientável. No caso de ser negativo, nós podemos “trocar o sinal” de uma das parametrizações, como segue. Definimos  $\tilde{\phi} : \tilde{U} \rightarrow \phi(U)$  por

$$\tilde{\phi}(x_1, x_2, \dots, x_n) = \phi(-x_1, x_2, \dots, x_n),$$

onde  $\tilde{U} = \{(-x_1, x_2, \dots, x_n) \in \mathbb{R}^n; x = (x_1, x_2, \dots, x_n) \in U\}$ . Daí temos

$$J(\psi^{-1} \circ \tilde{\phi}) = -J(\psi^{-1} \circ \phi).$$

<sup>13</sup>É comum denotar as coordenadas por  $x = (x_1, x_2, \dots, x_n)$  em uma vizinhança (por exemplo,  $U$ ) e por  $y = (y_1, y_2, \dots, y_n)$  em outra. Ao considerar uma mudança de parâmetros, escrevemos ainda  $y_i(x) = (\psi^{-1} \circ \phi)_i(x)$ .

<sup>14</sup>Um atlas  $\mathcal{A}$  é dito maximal se, para  $\phi$  compatível com toda parametrização de  $\mathcal{A}$ , tem-se  $\phi \in \mathcal{A}$ . Para provar que todo atlas coerente está contido em um atlas coerente maximal pode-se utilizar o Lema de Zorn (ou Axioma da Escolha). No entanto, isto não é absolutamente necessário; ver, por exemplo, <https://math.stackexchange.com/questions/123456/existence-of-a-maximal-atlas> (acessado em 08 de Abril de 2018).



**Exemplo 46.** Devido ao exemplo anterior,  $\mathbb{S}^n$  é orientável, para todo  $n > 1$ ; ver projeção estero-gráfica da página 32. Por que excluimos o caso  $n = 1$ ?

**Exercício 27.** Mostre que  $\mathbb{S}^1$  é orientável.

**Exercício 28.** Mostre que  $M \times N$  é orientável se, e somente se, ambas  $M$  e  $N$  são orientáveis.

Seja  $M \subseteq \mathbb{R}^m$  uma superfície de dimensão  $n$  e classe  $C^k$ . Uma função do tipo  $v : M \rightarrow \mathbb{R}^m$  tal que  $v(x) \in T_x M$  é chamada de um **campo de vetores** em  $M$ . É também comum pensar em um campo vetorial como uma função que associa uma “seta” para cada  $x \in M$ . Claro que esta seta é um elemento de  $T_x M$ , pois são as únicas direções disponíveis para quem enxerga a superfície “pelo lado de dentro”. As vezes um campo vetorial é definido como uma função  $v : M \rightarrow TM$  que satisfaz  $\pi \circ v = id$ . Neste contexto,  $\pi : TM \rightarrow M$  é a projeção  $\pi(x, v) = x$  e tem-se que  $\pi \circ v = id \iff v(x) \in \{x\} \times T_x M$ .

Um **campo normal** a  $M$  é uma função  $v : M \rightarrow \mathbb{R}^m$  tal que  $v(x) \in T_x M^\perp$  para todo  $x \in M$ . Desta maneira,  $v$  é normal quando  $v(x)$  é ortogonal a todo  $v \in T_x M$ .

Vejam agora que, no caso de hiperfícies, há uma caracterização de orientabilidade em termos de campos normais a  $M$ .

**Teorema 47.** *Seja  $M \subset \mathbb{R}^{n+1}$  uma superfície de dimensão  $n$  (aqui é importante que  $M$  tenha codimensão 1), então  $M$  é orientável se, e somente se, existe um campo normal à superfície  $M$  que é unitário e contínuo.*

*Demonstração.*  $\implies$  Suponhamos que  $M$  é orientável e seja  $\mathcal{A}$  um atlas positivo. Seja  $\phi \in \mathcal{A}$  uma carta positiva. Definimos  $u : M \rightarrow \mathbb{R}^{n+1}$  como segue: para  $x = \phi(x_0) \in M$ ,  $u(x)$  é o único vetor unitário em  $\mathbb{R}^{n+1}$  que satisfaz<sup>15</sup>

$$\det \begin{bmatrix} \left| \frac{\partial \phi}{\partial x_1}(x_0) \right| & \left| \frac{\partial \phi}{\partial x_2}(x_0) \right| & \cdots & \left| \frac{\partial \phi}{\partial x_n}(x_0) \right| & \left| u(x) \right| \\ \left| \frac{\partial \phi}{\partial x_1}(x_0) \right| & \left| \frac{\partial \phi}{\partial x_2}(x_0) \right| & \cdots & \left| \frac{\partial \phi}{\partial x_n}(x_0) \right| & \left| u(x) \right| \\ \left| \frac{\partial \phi}{\partial x_1}(x_0) \right| & \left| \frac{\partial \phi}{\partial x_2}(x_0) \right| & \cdots & \left| \frac{\partial \phi}{\partial x_n}(x_0) \right| & \left| u(x) \right| \\ \left| \frac{\partial \phi}{\partial x_1}(x_0) \right| & \left| \frac{\partial \phi}{\partial x_2}(x_0) \right| & \cdots & \left| \frac{\partial \phi}{\partial x_n}(x_0) \right| & \left| u(x) \right| \end{bmatrix} > 0.$$

Esta definição é independente da parametrização *positiva* escolhida, pois a matriz de mudança de parâmetros tem determinante positivo e podemos escrever

$$\begin{bmatrix} \left| \frac{\partial \phi}{\partial x_1}(x_0) \right| & \cdots & \left| \frac{\partial \phi}{\partial x_n}(x_0) \right| & \left| u(x) \right| \\ \left| \frac{\partial \phi}{\partial x_1}(x_0) \right| & \cdots & \left| \frac{\partial \phi}{\partial x_n}(x_0) \right| & \left| u(x) \right| \\ \left| \frac{\partial \phi}{\partial x_1}(x_0) \right| & \cdots & \left| \frac{\partial \phi}{\partial x_n}(x_0) \right| & \left| u(x) \right| \\ \left| \frac{\partial \phi}{\partial x_1}(x_0) \right| & \cdots & \left| \frac{\partial \phi}{\partial x_n}(x_0) \right| & \left| u(x) \right| \end{bmatrix}^T = \begin{bmatrix} \left( \frac{\partial y_i}{\partial x_j}(x_0) \right)_{n \times n} & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \left| \frac{\partial \psi}{\partial x_1}(y_0) \right| & \cdots & \left| \frac{\partial \psi}{\partial x_n}(y_0) \right| & \left| u(x) \right| \\ \left| \frac{\partial \psi}{\partial x_1}(y_0) \right| & \cdots & \left| \frac{\partial \psi}{\partial x_n}(y_0) \right| & \left| u(x) \right| \\ \left| \frac{\partial \psi}{\partial x_1}(y_0) \right| & \cdots & \left| \frac{\partial \psi}{\partial x_n}(y_0) \right| & \left| u(x) \right| \\ \left| \frac{\partial \psi}{\partial x_1}(y_0) \right| & \cdots & \left| \frac{\partial \psi}{\partial x_n}(y_0) \right| & \left| u(x) \right| \end{bmatrix}^T. \quad (29)$$

Vejam que  $u$  é um campo contínuo. Localmente,  $M$  é a imagem inversa de um valor regular de uma função do tipo  $g : A \subset \mathbb{R}^{n+1} \rightarrow \mathbb{R}$ . Temos que, para todo  $x \in U := g^{-1}(c) \subseteq M$ , o espaço tangente é  $T_x M = \ker g'(x)$ . Assim, temos

$$v(x) := \frac{\nabla g(x)}{\|\nabla g(x)\|} \in T_x M^\perp$$

é um campo contínuo de vetores normais à  $M$ , definido em  $U$ . Segue da continuidade e de ser  $T_x M^\perp$  unidimensional que  $v \equiv u$  em  $U$  ou  $v \equiv -u$  em  $U$ . Em qualquer caso,  $u$  é uma função contínua em cada vizinhança (que pode ser tomada conexa)  $U$ .

<sup>15</sup>Ao completar a base de  $T_x M$  a uma base de  $\mathbb{R}^{n+1}$  com um vetor unitário, existem apenas duas possibilidades; escolhamos àquela que resulta em uma base positiva de  $\mathbb{R}^{n+1}$ .

◀ Seja  $\mathcal{A}$  o conjunto das parametrizações  $\phi : U \rightarrow \phi(U) \subset M$  com  $U$  conexo e que satisfazem

$$\det \begin{bmatrix} \frac{\partial \phi}{\partial x_1}(x_0) & \frac{\partial \phi}{\partial x_2}(x_0) & \cdots & \frac{\partial \phi}{\partial x_n}(x_0) & u(x) \end{bmatrix} > 0.$$

Pela nossa hipótese, a matriz acima depende continuamente de  $x$  e, logo, não muda de sinal no conexo  $U$ . Observamos que, dada uma parametrização  $\psi : V \rightarrow \psi(V)$  qualquer, se  $\psi$  satisfizer

$$\det \begin{bmatrix} \frac{\partial \psi}{\partial y_1}(y_0) & \frac{\partial \psi}{\partial y_2}(y_0) & \cdots & \frac{\partial \psi}{\partial y_n}(y_0) & u(x) \end{bmatrix} < 0$$

em seu domínio conexo  $V$ , podemos “trocar o sinal” de  $\psi$  e obter uma carta na mesma vizinhança coordenada  $\psi(V)$  que pertence a  $\mathcal{A}$ . Segue daí que  $\mathcal{A}$  é um atlas de  $M$ .

Resta mostrar que  $\mathcal{A}$  é um atlas coerente. Mas isto é imediato da definição de  $\mathcal{A}$  e da fórmula de mudança de parâmetros (29).  $\square$

**Exemplo 48.** Vamos construir parametrizações para a **Faixa de Möbius**<sup>16</sup> e em seguida mostrar que é uma superfície não orientável. A faixa de Möbius (ver figura da página 38) pode ser obtida da seguinte maneira: para  $x_0 > 0$  e  $a \in (0, x_0)$ , consideramos em  $\mathbb{R}^3$  um segmento da forma  $(x_0 - a, x_0 + a)$  que está sobre o eixo dos  $x$ . O ponto médio  $x_0$  do segmento percorre uma volta sobre um círculo de raio  $x_0$  no plano  $xy$  enquanto que o segmento faz uma rotação de  $180^\circ$ , retornando ao segmento original, mas “invertido”.

Fixado um parâmetro  $t \in (0, 2\pi)$ , o centro do segmento se desloca para o ponto  $x_0 \cdot u(t)$ , onde  $u(t) = (\cos t, \sin t, 0)$ . Consideramos o plano em  $\mathbb{R}^3$  gerado pelos vetores unitários  $u(t)$  e  $e_3$ ; o semi-círculo de raio  $s$  centrado no ponto  $x_0 \cdot u(t)$  deste plano pode ser parametrizado por

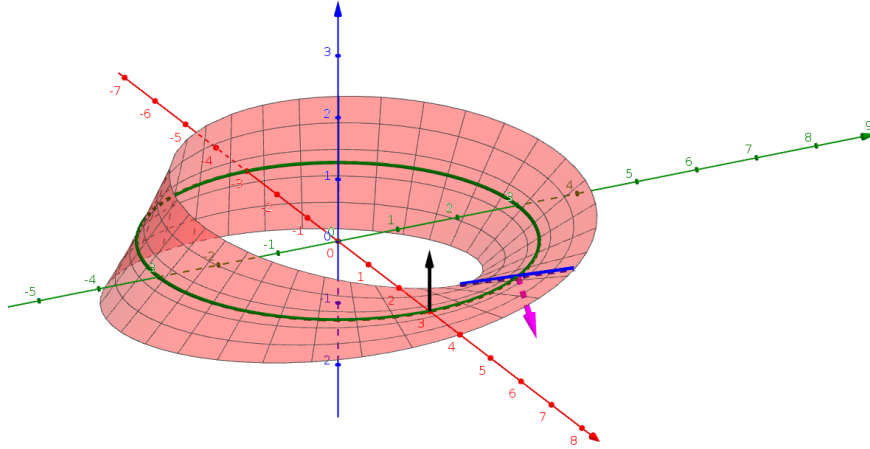
$$t \in (0, 2\pi) \mapsto x_0 \cdot u(t) + s \cos(t/2)u(t) + s \sin(t/2)e_3.$$

Uma parametrização para a faixa de Möbius pode ser obtida com essas ideias. Definimos uma carta  $\phi : (0, 2\pi) \times (-a, a) \rightarrow \mathbb{R}^3$  por

$$\phi(t, s) = \left( (x_0 + s \cos(t/2)) \cos t, (x_0 + s \cos(t/2)) \sin t, s \sin(t/2) \right).$$

Note que, para cada  $s \in (-a, a)$  fixado,  $(x_0 + s, 0, 0)$  representa um ponto do segmento inicial que vai ser rotacionado por um ângulo  $\pi$  sobre um círculo de raio  $s$  enquanto  $t$  variar de 0 a  $2\pi$ .

<sup>16</sup>A Faixa de Möbius apareceu por volta de 1860 em trabalhos independentes de dois matemáticos e astrônomos alemães: Johann Benedict Listing (25 de Julho de 1808 – 24 de Dezembro de 1882) e August Ferdinand Möbius (17 de Novembro de 1790 – 26 de Setembro de 1868). Este foi um primeiro exemplo de superfície não orientável e que “tem apenas um lado”.



Esta parametrização não cobre o segmento inicial que está sobre o plano  $xy$ . Fazendo uma parametrização análoga começando em outro segmento (como o azul da figura), podemos cobrir a faixa de Möbius com duas parametrizações<sup>17</sup>.

Observe que qualquer campo contínuo de vetores normais deve ser, ao longo da curva  $t \mapsto \phi(t, 0)$ , da forma

$$v(t) = \frac{\phi_t(t, 0) \times \phi_s(t, 0)}{\|\phi_t(t, 0) \times \phi_s(t, 0)\|}$$

ou o negativo deste. No entanto, tem-se  $v(0) = -v(2\pi)$ . Isto mostra que a faixa de Möbius não é orientável. As contas ficam como exercício.

Vejamus que uma das implicações do Teorema 47 segue válida para superfícies de codimensão maior. A prova é bastante semelhante.

**Teorema 49.** *Seja  $M \subset \mathbb{R}^m$  uma superfície de dimensão  $n$ . Se existirem  $m - n$  campos normais, unitários e contínuos  $v_j : M \rightarrow \mathbb{R}^m$ , à superfície  $M$  e tais que  $\{v_j(x)\}$  é linearmente independente para todo  $x \in M$ , então  $M$  é orientável.*

*Demonstração.* Seja  $\mathcal{A}$  o conjunto das parametrizações  $\phi : U \rightarrow \phi(U) \subset M$  com  $U$  conexo e que satisfazem

$$\det \begin{bmatrix} \frac{\partial \phi}{\partial x_1}(x_0) & \frac{\partial \phi}{\partial x_2}(x_0) & \cdots & \frac{\partial \phi}{\partial x_n}(x_0) & v_1(x) & v_2(x) & \cdots & v_{m-n}(x) \end{bmatrix} > 0.$$

Pela nossa hipótese, a matriz acima depende continuamente de  $x$  e, logo, não muda de sinal no conexo  $U$ . Observamos que, dada uma parametrização  $\psi : V \rightarrow \psi(V)$  qualquer, se  $\psi$  satisfizer

$$\det \begin{bmatrix} \frac{\partial \psi}{\partial y_1}(y_0) & \frac{\partial \psi}{\partial y_2}(y_0) & \cdots & \frac{\partial \psi}{\partial y_n}(y_0) & v_1(x) & v_2(x) & \cdots & v_{m-n}(x) \end{bmatrix} < 0$$

em seu domínio conexo  $V$ , podemos “trocar o sinal” de  $\psi$  e obter uma carta na mesma vizinhança coordenada  $\psi(V)$  que pertence a  $\mathcal{A}$ . Segue daí que  $\mathcal{A}$  é um atlas de  $M$ .

<sup>17</sup>Isto não contradiz o exemplo 45, pois a união das vizinhanças parametrizadas não é conexa.

Resta mostrar que  $\mathcal{A}$  é um atlas coerente. Mas isto é imediato da definição de  $\mathcal{A}$  e da fórmula de mudança de parâmetros equivalente a que exibimos em (29), trocando a matriz central por

$$\begin{bmatrix} \left( \frac{\partial y_i}{\partial x_j}(x_0) \right)_{n \times n} & 0 \\ 0 & I_{(m-n) \times (m-n)} \end{bmatrix} \quad \square$$

**Corolário 50.** *Seja  $f : \mathbb{R}^{n+k} \rightarrow \mathbb{R}^k$  uma função de classe  $C^1$ . Se  $M = f^{-1}(c)$  é a imagem inversa de um valor regular  $c \in \mathbb{R}^k$ , então  $M$  é orientável.*

*Demonstração.* Neste caso,  $T_x M = \ker f'(x)$  e podemos construir  $k$  campos contínuos normais unitários por

$$v_1(x) = \frac{\nabla f_1(x)}{\|\nabla f_1(x)\|}, \quad v_2(x) = \frac{\nabla f_2(x)}{\|\nabla f_2(x)\|}, \quad \dots, \quad v_k(x) = \frac{\nabla f_k(x)}{\|\nabla f_k(x)\|}.$$

São contínuos porque  $f \in C^1$ . São unitários por construção. São normais pois todo  $v \in T_x M$  satisfaz

$$0 = f'(x) \cdot v = \begin{bmatrix} - & \nabla f_1(x) \cdot v & - \\ - & \nabla f_2(x) \cdot v & - \\ & \vdots & \\ - & \nabla f_k(x) \cdot v & - \end{bmatrix} \quad \square$$

## 14 Multiplicadores de Lagrange

O método dos multiplicadores de Lagrange, como já se sabe desde disciplinas de Cálculo, é utilizado com o seguinte objetivo:

minimizar  $f(x)$  sujeito às restrições (ou vínculos) da forma  $g(x) = c$ .

Com a linguagem das seções anteriores, isto equivale a minimizar  $f : M \subset \mathbb{R}^n \rightarrow \mathbb{R}$  onde  $M = g^{-1}(c)$  é a imagem inversa de um valor  $c$ , que vamos assumir ser valor regular de  $g$ .

Sejam  $U \subset \mathbb{R}^{n+1}$  um aberto e  $f : U \rightarrow \mathbb{R}$  uma função de classe  $C^k$ . Um **ponto crítico** da restrição  $f|_M : M \rightarrow \mathbb{R}$  é um ponto  $x \in M$  tal que

$$f'(x) : T_x M \rightarrow \mathbb{R}$$

é a transformação linear nula. Isto é equivalente a dizer que

$$\frac{\partial f}{\partial v}(x) = 0 \quad \text{para todo } v \in T_x M.$$

**Teorema 51 (Multiplicadores de Lagrange).** *Seja  $g : \mathbb{R}^{n+1} \rightarrow \mathbb{R}$  uma função de classe  $C^1$  e  $M = g^{-1}(c)$  é a imagem inversa de um valor regular  $c \in \mathbb{R}$ . Seja  $U \subset \mathbb{R}^{n+1}$  um aberto e  $f : U \rightarrow \mathbb{R}$  uma função de classe  $C^1$ . Então, um ponto  $x \in M$  é um ponto crítico da restrição  $f|_M$  se, e somente se, existe uma constante  $\lambda \in \mathbb{R}$  tal que*

$$\nabla f(x) = \lambda \nabla g(x). \quad (30)$$

*Em particular, se a restrição  $f|_M$  tem um mínimo ou máximo em  $x_0$ , então vale (30).*

*Demonstração.* No caso de ser  $M = g^{-1}(c)$  é a imagem inversa de um valor regular, sabemos que

$$T_x M = \ker g'(x) = \{\nabla g(x)\}^\perp.$$

Se, além disso,  $x \in M$  é um ponto crítico de  $f$  tem-se que  $\nabla f(x) \in T_x M^\perp$ , pois  $\langle \nabla f(x), v \rangle$  é a derivada de  $f$  na direção tangente  $v$ . Segue que  $\nabla g(x)$  e  $\nabla f(x)$  estão no mesmo espaço unidimensional.  $\square$

A constante  $\lambda \in \mathbb{R}$  acima é chamada um multiplicador de Lagrange e é, geralmente encontrada para encontrar pontos críticos de funções com restrição. Observamos que, no teorema acima, não há nada especial no fato de a codimensão da superfície de nível ser um, como passamos a analisar.

**Teorema 52 (Multiplicadores de Lagrange).** *Seja  $g : \mathbb{R}^{n+k} \rightarrow \mathbb{R}^k$  uma função de classe  $C^1$  e  $M = g^{-1}(c)$  é a imagem inversa de um valor regular  $c \in \mathbb{R}^k$ . Sejam  $U \subset \mathbb{R}^{n+k}$  um aberto e  $f : U \rightarrow \mathbb{R}$  uma função de classe  $C^1$ . Então, um ponto  $x \in M$  é um ponto crítico da restrição  $f|_M$  se, e somente se, existe uma constante  $\lambda \in \mathbb{R}^k$  tal que*

$$\nabla f(x) = \lambda^T g'(x). \quad (31)$$

Mais explicitamente,

$$\nabla f(x) = \lambda_1 \nabla g_1(x) + \lambda_2 \nabla g_2(x) + \cdots + \lambda_k \nabla g_k(x). \quad (32)$$

As constantes  $\lambda_1, \lambda_2, \dots, \lambda_n$  são conhecidas como **multiplicadores de Lagrange**.

*Demonstração.* Basta notar que ser um ponto crítico é equivalente a  $\nabla f(x) \in (\ker A)^\perp$ , onde  $A = g'(x)$ . Isto por que

$$(\ker A)^\perp = \text{Im } A^T = \text{espaço gerado pelas linhas de } A = g'(x). \quad \square$$

**Exemplo 53.** Seja  $M \subset \mathbb{R}^m$  uma superfície de dimensão  $n$  e seja  $a \in \mathbb{R}^m \setminus M$ . Queremos encontrar  $x_0 \in M$  tal que a distância de  $x_0$  até a superfície  $M$  seja mínima. A restrição da função suave  $f(x) = |x - a|^2$  à  $M$  deve ter um ponto crítico em  $x_0$ . Assim, deve ser  $\nabla f(x_0) \perp T_{x_0}M$

**Exemplo 54.** É comum em textos de Álgebra Linear, como aplicação do estudo de autovalores e autovetores, otimizar formas quadráticas em esferas. Consideramos  $A$  uma matriz simétrica de ordem  $n \times n$  e definimos  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  a sua forma quadrática associada:

$$f(x) = \langle Ax, x \rangle.$$

Seja  $\mathbb{S}^{-1} = g^{-1}(1)$  a esfera unitária de  $\mathbb{R}^n$ , onde  $g(x) = |x|^2$ . Pelo Teorema dos Multiplicadores de Lagrange, pontos de mínimo ou máximo de  $f|_{\mathbb{S}^{n-1}}$  (na verdade, quaisquer pontos críticos) devem satisfazer

$$\nabla f(x) = \lambda \nabla g(x) \quad \text{ou, equivalentemente,} \quad Ax = \lambda x.$$

Matrizes simétricas possuem autovalores reais. Se  $\lambda_{\max}$  é o maior autovalor de  $A$ , com autovetor unitário associado  $u_{\max}$ , então

$$f(u_{\max}) = \lambda_{\max} |u_{\max}|^2 = \lambda_{\max}$$

é o valor máximo atingido por  $f$  em  $\mathbb{S}^{n-1}$ .

Uma observação interessante é que o hiperplano ortogonal a  $u_{\max}$  é invariante por  $A$ , isto é,  $\langle x, u_{\max} \rangle = 0 \implies \langle Ax, u_{\max} \rangle = 0$ . Dessa forma, obtemos uma aplicação  $\tilde{A} : \{u_{\max}\}^T \rightarrow \{u_{\max}\}^T$  e podemos fazer o raciocínio anterior, para obter mais um autovetor com autovalor maximal. Iteração deste raciocínio fornece uma demonstração do Teorema Espectral de Álgebra Linear.

**Exercício 29.** Nas condições do Teorema dos Multiplicadores de Lagrange, definimos a função de Lagrange por

$$L(x, \lambda) := f(x) - \lambda^T (g(x) - c).$$

Mostre que  $(x_0, \lambda_0)$  é ponto crítico de  $L$  se, e somente se,  $x_0$  é ponto crítico da restrição de  $f$  ao conjunto  $g(x) = c$ .

Mostre também que, se  $x_0$  é um mínimo local para o problema com restrição e  $\lambda_0$  é o multiplicador de Lagrange associado, então, para todo  $v \in \ker g'(x_0)$ , temos

$$\langle \nabla_x^2 L(x_0, \lambda_0) v, v \rangle \geq 0.$$

## 15 Integral de Riemann em $\mathbb{R}^n$

Nos primeiros cursos de Cálculo ou de Introdução à Análise Matemática, se define a integral de uma função  $f$  real de uma variável real em um intervalo  $[a, b]$ , denotada por

$$\int_a^b f(x) dx \quad \text{ou simplesmente} \quad \int_a^b f,$$

como um limite de somas de Riemann. Vamos ver que a definição se estende facilmente para funções limitadas definidas em **retângulos**  $n$ -dimensionais da forma

$$R = [a_1, b_1] \times [a_2, b_2] \times \cdots \times [a_n, b_n] \subseteq \mathbb{R}^n.$$

Vamos chamar este retângulo de “semi-aberto”, por serem os intervalos que o definem todos semi-abertos. Definimos o volume do retângulo  $R$  como

$$\text{vol}(R) = (b_1 - a_1) \cdot (b_2 - a_2) \cdot \cdots \cdot (b_n - a_n). \quad (33)$$

A definição de um retângulo aberto ou de um retângulo fechado é clara e o volume é definido pela mesma fórmula (33) acima.

Seja  $f : R \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$  uma função limitada. Consideramos uma partição de intervalo  $[a_i, b_i]$  em um número finito de subintervalos disjuntos  $\{I_{\lambda_i}\}_{\lambda_i \in \Lambda}$ , fechados na primeira extremidade e abertos na outra, e os respectivos subretângulos  $R_\lambda = I_{\lambda_1} \times I_{\lambda_2} \times \cdots \times I_{\lambda_n}$ , que formam uma partição (finita)  $\mathcal{P}$  de  $R$ . Para

$$m_\lambda := \inf_{x \in R_\lambda} f(x) \quad \text{e} \quad M_\lambda := \sup_{x \in R_\lambda} f(x), \quad (34)$$

definimos a **soma de Riemann inferior** e a **soma de Riemann superior**, respectivamente, por

$$I(f; \mathcal{P}) := \sum_{R_\lambda \in \mathcal{P}} m_\lambda \cdot \text{vol}(R_\lambda) \quad \text{e} \quad S(f; \mathcal{P}) := \sum_{R_\lambda \in \mathcal{P}} M_\lambda \cdot \text{vol}(R_\lambda).$$

Além disso, definimos a **integral inferior** de  $f$  em  $R$  como

$$\int_R f(x) dx := \sup_{\mathcal{P}} I(f; \mathcal{P}) = \sup_{\mathcal{P}} \sum_{R_\lambda \in \mathcal{P}} m_\lambda \cdot \text{vol}(R_\lambda).$$

Analogamente, definimos a **integral superior** de  $f$  em  $R$  como

$$\overline{\int}_R f(x) dx := \inf_{\mathcal{P}} S(f; \mathcal{P}) = \inf_{\mathcal{P}} \sum_{R_\lambda \in \mathcal{P}} M_\lambda \cdot \text{vol}(R_\lambda).$$

A função  $f$  é dita **integrável** no retângulo  $R$  quando a integral inferior e a superior coincidem. Neste caso, definimos a **integral** de  $f$  por

$$\int_R f(x) dx := \int_R f(x) dx = \overline{\int}_R f(x) dx.$$

A definição de integral não é boa para cálculos. Depois de estudar algumas propriedades da integral, vamos mostrar alguns métodos para efetivamente calcular uma integral como em cursos de cálculo. No entanto, vejamos um primeiro exemplo que indica que a integral faz bem seu papel de calcular volumes de regiões.

Definimos a **função característica**  $\chi_A : S \rightarrow \mathbb{R}$  de uma região  $A$  contida em um retângulo  $S$  como

$$\chi_A(x) = \begin{cases} 1, & \text{se } x \in A \\ 0, & \text{se } x \in S \setminus A \end{cases}.$$

**Exemplo 55.** Considere dois retângulos  $R$  e  $S$  tais que  $R \subset \text{int } S$ . Vamos mostrar que a função característica  $\chi_R : S \rightarrow \mathbb{R}$  é integrável e tem-se

$$\int_S \chi_R(x) dx = \text{vol}(R).$$

Considerando uma partição  $\mathcal{P}$  de  $S$  que contem  $R$  como um de seus elementos, obtemos que, para qualquer subpartição  $\tilde{\mathcal{P}}$  de  $\mathcal{P}$ ,

$$I(\chi_R; \tilde{\mathcal{P}}) = \text{vol}(R).$$

Segue daí que

$$\int_S \chi_R(x) dx = \text{vol}(R).$$

Por outro lado, dado qualquer retângulo  $\tilde{R}$  entre  $R$  e  $S$ , isto é, tal que  $R \subset \tilde{R} \subset S$ , podemos escolher uma partição  $\mathcal{P}$  de  $S$  que contém  $\tilde{R}$  e obter

$$S(\chi_R; \mathcal{P}) = \text{vol}(\tilde{R}) \quad \text{e logo} \quad \text{vol}(R) = \int_S \chi_R(x) dx \leq \overline{\int_S \chi_R(x) dx} \leq \text{vol}(\tilde{R}).$$

Como  $\text{vol}(\tilde{R})$  pode ser tomado tão próximo quanto se queira de  $\text{vol}(R)$ , nossa afirmação fica demonstrada.

**Exemplo 56.** A função  $f : [0, 1] \rightarrow \mathbb{R}$  definida por

$$f(x) = \begin{cases} 1, & \text{se } x \in \mathbb{Q} \cap [0, 1] \\ 0, & \text{se } x \in [0, 1] \setminus \mathbb{Q} \end{cases}$$

não é integrável, pois, dada qualquer partição  $\mathcal{P}$ , tem-se

$$I(f; \mathcal{P}) = 0 \quad \text{e} \quad S(f; \mathcal{P}) = 1.$$

Este exemplo é facilmente estendido para a função que vale 1 nos pontos de  $[0, 1]^n$  de coordenadas racionais e zero nos demais pontos de  $[0, 1]^n$ .

Um critério simples para que uma função  $f : R \rightarrow \mathbb{R}$  seja integrável, que é conhecido como **Crítério de Riemann**, segue diretamente das definições e é o seguinte:

**Proposição 57.** Uma função  $f : R \rightarrow \mathbb{R}$  é integrável se, e somente se, para todo  $\varepsilon > 0$ , existe uma partição  $\mathcal{P}$  do retângulo  $R$  tal que

$$0 \leq S(f, \mathcal{P}) - I(f, \mathcal{P}) < \varepsilon.$$

*Demonstração.*  $\boxed{\implies}$  Se  $f$  é integrável, tem-se

$$\int_R f(x) dx = \sup_{\mathcal{P}} I(f; \mathcal{P}) = \inf_{\mathcal{P}} S(f; \mathcal{P})$$

Usando a definição de ínfimo, temos que, dado  $\varepsilon > 0$ , existe  $\mathcal{P}_\varepsilon^1$  tal que

$$S(f; \mathcal{P}_\varepsilon^1) < \int_R f(x) dx + \frac{\varepsilon}{2}.$$

Pela definição de supremo, existe  $\mathcal{P}_\varepsilon^2$  tal que

$$\int_R f(x) dx - \frac{\varepsilon}{2} < I(f; \mathcal{P}_\varepsilon^2).$$

Agora, basta notar que para  $\mathcal{P}$  qualquer partição mais fina que ambas  $\mathcal{P}_\varepsilon^1$  e  $\mathcal{P}_\varepsilon^2$ , temos

$$\int_R f(x) dx - \frac{\varepsilon}{2} < I(f; \mathcal{P}_\varepsilon^2) \leq I(f; \mathcal{P}) \leq S(f; \mathcal{P}) \leq S(f; \mathcal{P}_\varepsilon^1) < \int_R f(x) dx + \frac{\varepsilon}{2}$$

Isto implica que

$$0 \leq S(f, \mathcal{P}) - I(f, \mathcal{P}) < \varepsilon.$$

$\Rightarrow$  Reciprocamente, sempre vale que

$$I(f, \mathcal{P}) \leq \int_R f(x) dx \leq \overline{\int_R f(x) dx} \leq S(f, \mathcal{P}).$$

Nossa hipótese garante que, para todo  $\varepsilon > 0$ , tem-se

$$0 \leq \overline{\int_R f(x) dx} - \int_R f(x) dx < \varepsilon.$$

Logo, as integrais inferior e superior coincidem e  $f$  é integrável.  $\square$

**Corolário 58.** Se  $f : R \rightarrow \mathbb{R}$  é uma função contínua limitada, então  $f$  é integrável.

*Demonstração.* Sendo limitada,  $f$  pode ser estendida continuamente para  $\bar{R}$ . Sendo contínua em um compacto, é uniformemente contínua: dado  $\varepsilon > 0$ , existe  $\delta > 0$  tal que

$$|y - x| < \delta \implies |f(y) - f(x)| < \frac{\varepsilon}{\text{vol}(A)}.$$

Considerando uma partição  $\mathcal{P} = \{R_\lambda\}$  de  $A$  tal que os diâmetros dos retângulos é menor do que  $\delta > 0$ , temos

$$\overline{\int_R f(x) dx} - \int_R f(x) dx \leq \sum_\lambda [M_\lambda - m_\lambda] \text{vol}(R_\lambda) < \frac{\varepsilon}{\text{vol}(A)} \sum_\lambda \text{vol}(R_\lambda) = \varepsilon.$$

A conclusão segue do Critério de Riemann.  $\square$

Ainda outra maneira de enunciar o Critério de Riemann é a partir da oscilação da função  $f$  nos retângulos de uma partição. Dado um conjunto  $A \subseteq \mathbb{R}^n$ , definimos a **oscilação** de  $f$  em  $A$  como

$$\omega_A(f) = \text{osc}_A(f) := \sup_A f - \inf_A f = \sup_{x, y \in A} [f(y) - f(x)] = \sup_{x, y \in A} |f(y) - f(x)|.$$

Notando que

$$S(f, \mathcal{P}) - I(f, \mathcal{P}) = \sum_\lambda \text{osc}_{R_\lambda}(f) \cdot \text{vol}(R_\lambda)$$

é apenas uma mudança de notação, obtemos a seguinte caracterização.

**Proposição 59.** Uma função  $f : R \rightarrow \mathbb{R}$  é integrável se, e somente se, para todo  $\varepsilon > 0$ , existe uma partição  $\mathcal{P}$  do retângulo  $R$  tal que

$$0 \leq \sum_\lambda \text{osc}_{R_\lambda}(f) \cdot \text{vol}(R_\lambda) < \varepsilon.$$

Passamos ao estudo de algumas das propriedades básicas de integração.

**Proposição 60 (Propriedades da integral).** As propriedades abaixo são satisfeitas pela integral de Riemann:



(i) Linearidade: se  $f, g : R \rightarrow \mathbb{R}$  são funções integráveis e  $c \in \mathbb{R}$ , então

$$\int_R [f(x) + cg(x)] dx = \int_R f(x) dx + c \int_R g(x) dx.$$

(ii) Positividade: se  $f : R \rightarrow \mathbb{R}$  é integrável e  $f(x) \geq 0$  para todo  $x \in R$ , então

$$\int_R f(x) dx \geq 0.$$

Em particular,

$$f \leq g \text{ integráveis} \implies \int_R f(x) dx \leq \int_R g(x) dx$$

(iii) Se  $f : R \rightarrow \mathbb{R}$  é integrável, então  $|f| : R \rightarrow \mathbb{R}$  é integrável e tem-se

$$\left| \int_R f(x) dx \right| \leq \int_R |f(x)| dx. \quad (35)$$

(iv) Valor médio: se  $f : R \rightarrow \mathbb{R}$  é contínua e limitada, então existe  $x_0 \in R$  tal que

$$\int_R f(x) dx = f(x_0) \cdot \text{vol}(R).$$

*Demonstração.* (i) e (ii) ficam com exercício para o leitor.

(iii) Da desigualdade elementar  $||a| - |b|| \leq |a - b|$ , segue que a oscilação de  $|f|$  é menor do que ou igual à oscilação de  $f$ . Logo, podemos escrever, para alguma partição  $\mathcal{P}$  de  $R$ ,

$$\overline{\int_R |f(x)| dx} - \underline{\int_R |f(x)| dx} \leq \sum_{\lambda} \omega_{R_{\lambda}}(|f|) \cdot \text{vol}(R_{\lambda}) \leq \sum_{\lambda} \omega_{R_{\lambda}}(f) \cdot \text{vol}(R_{\lambda}) < \varepsilon.$$

A última desigualdade segue do Critério de Riemann, pois  $f$  é integrável. Novamente pelo Critério de Riemann, concluímos que  $|f|$  é integrável. Agora, pelo item (ii), temos

$$-|f(x)| \leq f(x) \leq |f(x)| \implies -\int_R |f(x)| dx \leq \int_R f(x) dx \leq \int_R |f(x)| dx,$$

que é equivalente a (35).

(iv) Sendo  $f$  limitada, podemos escrever  $f(x) \in [m, M]$  para todo  $x \in R$ . Pelo item (ii) e pelo Exemplo 55, temos

$$m \text{ vol}(R) \leq \int_R f(x) dx \leq M \text{ vol}(R), \quad \text{que implica} \quad \frac{\int_R f(x) dx}{\text{vol}(R)} \in [m, M].$$

Como  $f$  é contínua, o Teorema do Valor Intermediário implica que existe  $x_0 \in R$  tal que

$$f(x_0) = \frac{\int_R f(x) dx}{\text{vol}(R)}. \quad \square$$

## 15.1 Conjuntos de medida zero e o Teorema de Lebesgue

Na tentativa de definir a integral para conjuntos mais gerais do que retângulos, aparece naturalmente a questão de poder ou não falar em “volume” desse conjunto mais geral. É possível falar em uma noção de volume de alguns subconjuntos de  $\mathbb{R}^n$ , chamada de medida de Lebesgue. Este tópico é estudado em cursos de Teoria da Medida. Por simplicidade, vamos restringir nossa atenção para conjuntos cuja medida (de Lebesgue) é nula.

Um conjunto  $A \subset \mathbb{R}^n$  é dito de **medida nula** em  $\mathbb{R}^n$  quando, para qualquer  $\varepsilon > 0$ , podemos encontrar uma cobertura de  $A$  por retângulos cuja soma dos volumes é menor do que  $\varepsilon$ , isto é,

$$A \subseteq \bigcup_{i=1}^{+\infty} R_i \quad \text{e} \quad \sum_{i=1}^{+\infty} \text{vol}(R_i) < \varepsilon.$$

**Exercício 30.** Prove que, na definição de conjunto de medida nula, é equivalente considerar que a cobertura de  $A$  é dada por:

- (i) retângulos abertos;
- (ii) retângulos fechados;
- (iii) cubos abertos;
- (iv) cubos fechados.

É imediato da definição acima que subconjuntos de um conjunto de medida zero também são de medida zero.

**Exemplo 61.** Se  $A = \{x^i\}_{i=1}^{+\infty} \subset \mathbb{R}^n$  é um conjunto enumerável, então  $A$  tem medida zero. De fato, dado  $\varepsilon > 0$ , consideramos retângulos (cubos) da forma

$$R_i = \left( x_1^i - \frac{\varepsilon^{1/n}}{2^{1+i/n}}, x_1^i + \frac{\varepsilon^{1/n}}{2^{1+i/n}} \right) \times \cdots \times \left( x_n^i - \frac{\varepsilon^{1/n}}{2^{1+i/n}}, x_n^i + \frac{\varepsilon^{1/n}}{2^{1+i/n}} \right).$$

Dessa forma,

$$\text{vol}(R_i) = \frac{\varepsilon}{2^i} \quad \text{e} \quad \sum_{i=1}^{+\infty} \text{vol}(R_i) = \varepsilon.$$

A técnica do Exemplo 61 acima pode ser usada para verificar o seguinte exercício.

**Exercício 31.** Mostre que se  $A \subseteq \bigcup_{i \in \mathbb{N}} A_i$  e todo  $A_i$  tem medida nula, então  $A$  tem medida nula.

**Exercício 32.** Mostre que se, para todo  $\varepsilon > 0$ , existe um número enumerável de retângulos  $R_i$  tais que

$$A \subseteq B \cup \bigcup_{i \in \mathbb{N}} R_i, \quad \sum_{i=1}^{+\infty} \text{vol}(R_i) < \varepsilon$$

e  $B$  tem medida nula, então  $A$  tem medida nula.

**Exemplo 62.** Retângulos não degenerados (isto é, um produto cartesiano de intervalos não vazios) **não** são conjuntos de medida zero. De fato, seja  $R$  um retângulo fechado. Dada qualquer cobertura de  $R$ , podemos considerar uma subcobertura finita do compacto  $\bar{R}$  por retângulos e escrever

$$R \subset \bar{R} \subset \bigcup_{i=1}^N R_i.$$

Considerando um retângulo  $S$  que contém todos os retângulos  $R_i$ , onde estão definidas as funções características  $\chi_{R_i}$ , temos

$$\chi_R \leq \chi_{\cup_i R_i} \leq \chi_{R_1} + \chi_{R_2} + \cdots + \chi_{R_N}.$$

Assim, utilizando que funções características de retângulos são integráveis, como no Exemplo 55, e a propriedade (ii) da Proposição 60 acima, obtemos

$$0 < \text{vol}(R) = \int_S \chi_R(x) dx \leq \int_S (\chi_{R_1} + \chi_{R_2} + \cdots + \chi_{R_N})(x) dx = \text{vol}(R_1) + \text{vol}(R_2) + \cdots + \text{vol}(R_N).$$

Logo, dado qualquer  $\varepsilon < \text{vol}(R)$ , não é possível encontrar uma cobertura cuja soma dos volumes seja menor do que  $\varepsilon$ .  $\triangleleft$

**Exemplo 63.** Uma consequência imediata do exemplo anterior é que conjuntos cujo interior é não vazio não podem ter medida nula.

**Teorema 64.** Se  $U \subset \mathbb{R}^n$  tem medida nula e  $f : U \rightarrow \mathbb{R}^n$  é localmente de Lipschitz, então  $f(U)$  tem medida nula.

*Demonstração.* Iniciamos considerando o caso em que  $f$  é globalmente de Lipschitz. Consideramos uma cobertura enumerável de  $U$  por cubos  $R_i$ , de arestas  $\ell_i$ , tais que

$$\sum_{i=1}^{+\infty} \text{vol}(R_i) < \frac{\varepsilon}{M^n},$$

onde  $M$  é a constante de Lipschitz de  $f$ . Note que

$$|f(y) - f(x)| \leq M|y - x| \text{ para todo } x, y \implies \text{diam}[f(U \cap R_i)] \leq M \text{diam}(R_i).$$

Na verdade, mais do que isso, as componentes  $f^j$  de  $f$  satisfazem

$$|f^j(y) - f^j(x)| \leq |f(y) - f(x)| \leq M|y - x| \text{ para todo } x, y \implies \text{diam}[f^j(U \cap R_i)] \leq M \text{diam}(R_i).$$

Em outras palavras, as projeções da imagem de  $U \cap R_i$  por  $f$  nos vetores da base canônica estão contidas em um intervalo de comprimento  $M \text{diam}(R_i)$ . Consideramos  $\tilde{R}_i$  um cubo que contém  $f(U \cap R_i)$  e cujas arestas medem  $M \text{diam}(R_i)$ . Temos  $\text{vol}(\tilde{R}_i) = M^n \text{diam}(R_i)^n$ . Logo,

$$f(U) = \bigcup_{i=1}^{+\infty} f(U \cap R_i) \subseteq \bigcup_{i=1}^{+\infty} \tilde{R}_i \quad \text{e} \quad \sum_{i=1}^{+\infty} \text{vol} \tilde{R}_i \leq M^n \sum_{i=1}^{+\infty} \text{vol}(R_i) < \varepsilon.$$

No caso de ser  $f$  localmente Lipschitz, tomamos uma cobertura enumerável de  $U$  por abertos onde  $f$  é Lipschitz; logo, a imagem de  $U$  por  $f$  é uma reunião enumerável de conjuntos de medida nula e, portanto, tem medida nula.  $\square$

**Proposição 65.** Para  $n < m$ , se  $f : U \subseteq \mathbb{R}^n \rightarrow \mathbb{R}^m$  é localmente de Lipschitz, então  $f(U)$  tem medida nula.

*Demonstração.* Consideramos  $U \times \{0\} \subset \mathbb{R}^n \times \mathbb{R}^{m-n}$  que é um conjunto de medida nula em  $\mathbb{R}^m$  (por quê?). Definimos  $g : U \times \mathbb{R}^{m-n} \rightarrow \mathbb{R}^m$  por  $g(x, y) = f(x)$ . Claramente,  $g$  também é localmente de Lipschitz. Pela proposição anterior,  $g(U \times \{0\}) = f(U)$  tem medida nula em  $\mathbb{R}^m$ .  $\square$

**Corolário 66.** Se  $M \subset \mathbb{R}^m$  é uma superfície de dimensão  $n < m$  e classe  $C^1$ , então  $M$  é um conjunto de medida nula em  $\mathbb{R}^m$ .

**Corolário 67.** Se  $M \subset \mathbb{R}^m$  é uma reunião enumerável de gráficos de aplicações de Lipschitz da forma  $f_i : A_i \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$ , com  $n < m$ , então  $M$  é de medida nula em  $\mathbb{R}^m$ . Estes conjuntos são conhecidos como  **$n$ -contavelmente retificáveis**.

**Teorema 68 (Lebesgue).** Seja  $f : R \subset \mathbb{R}^n \rightarrow \mathbb{R}$  uma função limitada. Então,  $f$  é integrável (a la Riemann) se, e somente se, o conjunto dos pontos onde  $f$  é descontínua tem medida zero em  $\mathbb{R}^n$ .

*Demonstração.* A prova utiliza a oscilação pontual da função  $f$ : é a função  $\text{osc}(f) : R \rightarrow \mathbb{R}$  dada por

$$\text{osc}(f)[x] := \inf_{\delta > 0} \left[ \text{osc}_{R \cap B_\delta(x)}(f) \right].$$

Em palavras, a oscilação definida acima mede a como que se comporta a oscilação em bolas cada vez “mais perto” do ponto  $x \in R$ . Ou ainda, mede quão ruim é a descontinuidade de  $f$  em  $x$ . Notamos que (prove como exercício)

$$f \text{ é contínua no ponto } x \iff \text{osc}(f)[x] = 0.$$

$\implies$  Suponhamos que  $f$  é integrável. Pelas definições acima, o conjunto  $D_f$  dos pontos onde  $f$  é descontínua pode ser escrito como

$$D_f = \bigcup_{k \in \mathbb{N}} D_k \quad \text{onde} \quad D_k = \left\{ x \in R; \text{osc}(f)[x] \geq \frac{1}{k} \right\}.$$

Basta mostrar que cada  $D_k$  tem medida nula. Pela Critério de Riemann, existe uma partição  $\mathcal{P}$  tal que

$$\sum_{\lambda} \text{osc}_{R_{\lambda}}(f) \cdot \text{vol}(R_{\lambda}) < \varepsilon.$$

Agora, podemos escrever

$$D_k = \{x \in D_k; x \in \partial R_{\lambda} \text{ para algum } \lambda\} \cup \{x \in D_k; x \in \text{int } R_{\lambda} \text{ para algum } \lambda\}$$

O primeiro desses conjuntos tem medida nula, pois as “faces” dos retângulos têm medida nula. Resta mostrar que o segundo, que denotamos por  $\tilde{D}$ , também tem medida nula.

Seja  $\mathcal{Q}$  a coleção dos retângulos de  $\mathcal{P}$  que contém algum ponto de  $D_k$  em seu interior. Logo,  $\mathcal{Q}$  é uma cobertura de  $\tilde{D}$ . Além disso, lembrando que a oscilação pontual é um ínfimo, para  $R \in \mathcal{Q}$ , temos

$$\text{osc}_R(f) \geq \text{osc}(f)[x] \geq \frac{1}{k}$$

de modo que

$$\sum_{R \in \mathcal{Q}} 1 \cdot \text{vol}(R) \leq k \sum_{R \in \mathcal{Q}} \text{osc}_R(f) \cdot \text{vol}(R) \leq k \sum_{R \in \mathcal{P}} \text{osc}_R(f) \cdot \text{vol}(R) < k \cdot \varepsilon.$$

Segue que  $\mathcal{Q}$  é uma cobertura de  $\tilde{D}$  por retângulos que têm soma dos volumes “pequena”. Segue que  $\tilde{D}$  tem medida zero, como queríamos.

$\Leftarrow$  Reciprocamente, suponhamos que o conjunto  $D_f$  dos pontos onde  $f$  é descontínua tem medida nula. O conjunto

$$D_{\varepsilon} := \{x \in R; \text{osc}(f)[x] \geq \varepsilon\}$$

está contido em  $D_f$  e, logo, é limitado e tem medida nula.

Vejamus que  $D_{\varepsilon}$  é fechado (e logo compacto): suponhamos que  $x_n \in D_{\varepsilon}$  e  $x_n \rightarrow x$ . Temos  $\text{osc}(f)[x_n] \geq \varepsilon$  para todo  $n$ . Se fosse  $\text{osc}(f)[x] < \varepsilon$ , existiria  $\delta > 0$  tal que

$$\text{osc}_{R \cap B_{\delta}(x)}(f) < \varepsilon.$$

No entanto, para  $n$  suficientemente grande,  $x_n \in B_{\delta}(x)$  e, neste caso, todo  $\eta$  suficientemente pequeno satisfaz  $B_{\eta}(x_n) \subset B_{\delta}(x)$ . Daí seria

$$\text{osc}(f)[x_n] \leq \text{osc}_{R \cap B_{\eta}(x_n)}(f) \leq \text{osc}_{R \cap B_{\delta}(x)}(f) < \varepsilon.$$

Vamos subdividir o retângulo  $R$  em uma partição com dois tipos de retângulos. Um tipo é dos retângulos que fazem parte de uma cobertura do conjunto de medida nula  $D_f$ . No outro tipo, a oscilação de  $f$  deve ser pequena.

Dada uma cobertura de  $D_{\varepsilon}$  por retângulos abertos cuja soma dos volumes é menor do que  $\varepsilon$ , podemos extrair uma subcobertura finita e escrever:

$$D_{\varepsilon} \subset \bigcup_{k=1}^N R_k \quad \text{com} \quad \sum_{k=1}^N \text{vol}(R_k) < \varepsilon.$$

Agora,  $R \setminus D_{\varepsilon}$  é aberto; logo, dado  $x \in R \setminus D_{\varepsilon}$ , existe um retângulo  $S_x$  aberto que contém  $x$  e está inteiramente contido em  $R \setminus D_{\varepsilon}$ . Podemos ainda escolher  $S_x$ , pela definição de ínfimo, de modo que  $\text{osc}_{S_x}(f) < \varepsilon$ . Sendo  $R$  compacto, a cobertura

$$R \subset \left( \bigcup_{k=1}^N R_k \right) \cup \bigcup_{x \in R \setminus D_{\varepsilon}} S_x$$

admite uma subcobertura finita:

$$R \subset \left( \bigcup_{k=1}^N R_k \right) \cup \bigcup_{j=1}^{\tilde{N}} S_j$$

Considere  $\mathcal{P}$  uma partição fina o suficiente de modo que todo elemento esteja contido em pelo menos um dos retângulos  $R_k$  ou  $S_j$ . Logo,

$$\begin{aligned} \sum_{\mathcal{P}} \text{osc}_{R_\lambda}(f) \cdot \text{vol}(R_\lambda) &\leq \sum_{R_\lambda \subseteq R_k} \text{osc}_{R_\lambda}(f) \cdot \text{vol}(R_\lambda) + \sum_{R_\lambda \subseteq S_j} \text{osc}_{R_\lambda}(f) \cdot \text{vol}(R_\lambda) \\ &\leq \text{osc}_R(f) \sum_{k=1}^N \text{vol}(R_k) + \sum_{R_\lambda \subseteq S_j} \text{osc}_{R_\lambda}(f) \cdot \text{vol}(R_\lambda) \\ &< \varepsilon [\text{osc}_R(f) + \text{vol}(R)] =: C \varepsilon. \end{aligned}$$

Portanto, pelo Critério de Riemann,  $f$  é integrável.  $\square$

## 15.2 Conjuntos Jordan mensuráveis

A questão de definir a integral de uma função limitada em um subconjunto limitado  $A \subset \mathbb{R}^n$  mais geral do que um cubo ou retângulo está intimamente ligada à possibilidade de se calcular o “volume” de  $A$ . Intuitivamente, o volume está associado ao cálculo de integrais de funções constantes.

Dizemos que um conjunto limitado  $A \subset \mathbb{R}^n$  é **Jordan mensurável** quando a função característica  $\chi_A : R \rightarrow \mathbb{R}$ , definida em um retângulo  $R$  que contém  $A$ , é integrável. Também se diz que  $A$  **possui conteúdo de Jordan**. Quando este é o caso, definimos o **volume** de  $A$  ou o **conteúdo de Jordan** por

$$\text{vol } A = \int_R \chi_A(x) \, dx.$$

Uma consequência imediata do Teorema de Lebesgue é o seguinte.

**Proposição 69.** *Um conjunto limitado  $A \subset \mathbb{R}^n$  é Jordan mensurável se, e somente se,  $\partial A$  tem medida (de Lebesgue) nula.*

*Demonstração.* Suponhamos, por simplicidade que  $A \subset \text{int } R$ . Basta então observar que o conjunto dos pontos de descontinuidade de  $\chi_A$  são justamente os pontos da fronteira de  $A$ . Dado qualquer conjunto  $A$ , podemos escrever a união disjunta

$$\mathbb{R}^n = \text{int}(A) \cup \partial A \cup \text{int}(\mathbb{R}^n \setminus A)$$

e  $\chi_A$  é constante em  $\text{int}(A)$  e em  $\text{int}(\mathbb{R}^n \setminus A)$ .  $\square$

**Exemplo 70.** Qualquer conjunto limitado cuja fronteira seja uma superfície de codimensão pelo menos um é Jordan mensurável. Deste modo, as bolas abertas ou fechadas, o toro sólido, elipsoides sólidos, etc, são conjuntos Jordan mensuráveis.

**Exemplo 71.** Mais geralmente, ver Corolário 67. Se a fronteira de um conjunto limitado for um conjunto  $n$ -contavelmente retificável, então este conjunto é Jordan mensurável.

**Exercício 33.** Seja  $A$  Jordan mensurável. Mostre que são equivalentes:

(i)  $\text{vol}(A) = 0$ ;

(ii) Dado  $\varepsilon > 0$ , existe uma cobertura finita por retângulos abertos  $R_k$  tal que

$$\sum_{k=1}^N \text{vol}(R_k) < \varepsilon.$$

**Exercício 34.** Seja  $K$  um conjunto compacto. Mostre que são equivalentes:

(i)  $\text{vol}(K) = 0$ ;

(ii)  $K$  tem medida zero.

**Exercício 35.** Discuta a equivalência das seguintes afirmações:

(i)  $A$  é Jordan mensurável;

(ii)  $\partial A$  tem medida nula;

(iii)  $\partial A$  tem volume zero.

Observamos que ser  $A$  de medida nula (segundo Lebesgue) não implica em Jordan mensurabilidade, ver Exemplo 56. Na linguagem desta seção, mostramos no Exemplo 56 que  $\mathbb{Q}^n \cap [0, 1]^n$  não é mensurável, apesar de ser enumerável. Uma outra prova deste fato consiste em notar que a fronteira  $\partial(\mathbb{Q}^n \cap [0, 1]^n) = [0, 1]^n$  não é de medida nula.

**Proposição 72.** *Seja  $A$  um conjunto Jordan mensurável. Então,  $\text{vol}(A) = 0 \iff \text{int}(A) = \emptyset$ .*

*Demonstração.* Dizer que  $A$  tem interior vazio é equivalente a dizer que nenhum retângulo (não degenerado) pode estar contido em  $A$ . Desta maneira, para qualquer partição  $\mathcal{P}$ , devemos ter  $I(\chi_A, \mathcal{P}) = 0$  e segue que

$$\int_R \chi_A(x) \, dx = 0.$$

Se  $A$  é Jordan mensurável, isto é equivalente a

$$\text{vol}(A) = \int_R \chi_A(x) \, dx = \int_R \chi_A(x) \, dx = 0. \quad \square$$

Passamos à definição de integral de uma função limitada  $f : A \rightarrow \mathbb{R}$  definida em um conjunto Jordan mensurável  $A$ . Consideramos  $R$  um retângulo que contém  $A$  e, por abuso de notação<sup>18</sup>, consideramos também a extensão de  $f$  como zero fora de  $A$ : definimos  $f\chi_A : R \rightarrow \mathbb{R}$  definida como

$$(f\chi_A)(x) = \begin{cases} f(x), & \text{se } x \in A \\ 0, & \text{se } x \in R \setminus A \end{cases}$$

Definimos a integral de  $f$  em  $A$  como

$$\int_A f(x) \, dx := \int_R (f\chi_A)(x) \, dx.$$

O análogo do Teorema de Lebesgue – Teorema 68 – vale trivialmente, pois ao considerar a extensão  $f\chi_A$ , estamos possivelmente adicionando descontinuidades na fronteira de  $A$ , que tem medida nula (devido à hipótese de  $A$  ser Jordan mensurável).

Além disso, todas as propriedades de integração – Proposição 60 – que vimos em domínios retangulares também se estendem trivialmente para integrais em domínios Jordan mensuráveis.

**Exercício 36.** Verificar estas últimas afirmações!

**Proposição 73.** *Sejam  $A$  e  $B$  são Jordan mensuráveis. Então a função  $f : A \cup B \rightarrow \mathbb{R}$  é integrável se, e somente se, as restrições  $f|_A : A \rightarrow \mathbb{R}$  e  $f|_B : B \rightarrow \mathbb{R}$  são integráveis. Alé disso, se  $\text{int}(A) \cap \text{int}(B) = \emptyset$ , temos*

$$\int_{A \cup B} f(x) \, dx = \int_A f(x) \, dx + \int_B f(x) \, dx.$$

<sup>18</sup>As funções  $f : A \rightarrow \mathbb{R}$  e  $\chi_A : R \rightarrow \mathbb{R}$  têm domínios diferentes, de modo que este não é um produto usual de funções.

*Demonstração.* Basta analisar os conjuntos de descontinuidade de  $f$ : temos

$$D_A \cup D_B \subset D_f \subset D_A \cup D_B \cup \partial A \cup \partial B.$$

Por hipótese,  $\partial A$  e  $\partial B$  tem medida nula. Logo,

$$D_f \text{ tem medida nula} \iff D_A \text{ e } D_B \text{ têm medida nula,}$$

o que prova a primeira afirmação. Agora, temos

$$\chi_{A \cup B} + \chi_{A \cap B} = \chi_A + \chi_B$$

Além disso, se  $\text{int}(A) \cap \text{int}(B) = \emptyset$ , então  $A \cap B$  tem interior vazio e  $\text{vol}(A \cap B) = 0$ . Sendo  $f$  limitada (digamos, por  $M$ ), concluímos que

$$\left| \int_{A \cap B} f(x) \, dx \right| = \int_R |f \chi_{A \cap B}(x)| \, dx \leq M \text{vol}(A \cap B) = 0.$$

Portanto,

$$\int_R f \chi_{A \cup B}(x) \, dx = \int_R f \chi_A(x) \, dx + \int_R f \chi_B(x) \, dx. \quad \square$$

### 15.3 Cálculo de integrais de modo iterativo

Para de fato calcular integrais multidimensionais, é comum considerarmos iterativamente integrais unidimensionais. Nesta seção, vamos provar o resultado que diz isto ser possível.

**Teorema 74.** *Consideramos dois retângulos  $R \subset \mathbb{R}^n$  e  $Q \subset \mathbb{R}^m$  e seja  $f : R \times Q \rightarrow \mathbb{R}$  uma função integrável. Então, as funções*

$$x \in R \mapsto \int_{\underline{Q}} f(x, y) \, dy \quad \text{e} \quad x \in R \mapsto \int_{\overline{Q}} f(x, y) \, dy$$

são integráveis e vale que

$$\int_{R \times Q} f(x, y) \, d(x, y) = \int_R \left[ \int_{\underline{Q}} f(x, y) \, dy \right] dx = \int_R \left[ \int_{\overline{Q}} f(x, y) \, dy \right] dx. \quad (36)$$

*Enunciado semelhante vale trocando os papéis de  $x$  e de  $y$ .*

*Demonstração.* Vamos provar apenas a primeira igualdade, pois a outra é completamente análoga. Toda partição de  $R \times Q$  é da forma  $\mathcal{P} = \mathcal{P}_1 \times \mathcal{P}_2$ , onde  $\mathcal{P}_1$  é uma partição de  $R$  e  $\mathcal{P}_2$  é uma partição de  $Q$ . Vamos mostrar inicialmente a seguinte relação entre somas inferiores:

$$I(f; \mathcal{P}) \leq I\left(\int_{\underline{Q}} f(\cdot, y) \, dy; \mathcal{P}_1\right).$$

Pela nossa definição de volume para retângulos, é claro que, para  $R_1 \times Q_1 \in \mathcal{P}$ ,

$$\text{vol}(R_1 \times Q_1) = \text{vol}(R_1) \cdot \text{vol}(Q_1)$$

Ainda, se  $R_1 \times Q_1 \in \mathcal{P}$ , temos que, para qualquer  $x \in R_1$  fixado, vale

$$m_{R_1 \times Q_1}(f) \leq m_{Q_1}(f(x, \cdot)),$$

pois o ínfimo na direita é tomado sobre um conjunto menor (estamos utilizando a notação estabelecida em (34)). Observamos ainda que, para todo  $x \in R_1$ ,

$$\sum_{Q_1 \in \mathcal{P}_2} m_{R_1 \times Q_1}(f) \cdot \text{vol}(Q_1) \leq \sum_{Q_1 \in \mathcal{P}_2} m_{Q_1}(f(x, \cdot)) \cdot \text{vol}(Q_1) \leq \int_{\underline{Q}} f(x, y) \, dy.$$

Logo, tomando o ínfimo para  $x \in R_1$ , temos

$$\sum_{Q_1 \in \mathcal{P}_2} m_{R_1 \times Q_1}(f) \cdot \text{vol}(Q_1) \leq m_{R_1} \left( \int_Q f(\cdot, y) \, dy \right).$$

Segue que

$$\begin{aligned} I(f; \mathcal{P}) &= \sum_{R_1 \times Q_1 \in \mathcal{P}} m_{R_1 \times Q_1}(f) \cdot \text{vol}(R_1 \times Q_1) \\ &= \sum_{R_1 \in \mathcal{P}_1} \left( \sum_{Q_1 \in \mathcal{P}_2} m_{R_1 \times Q_1}(f) \cdot \text{vol}(Q_1) \right) \text{vol}(R_1) \\ &\leq \sum_{R_1 \in \mathcal{P}_1} \left( \sum_{Q_1 \in \mathcal{P}_2} m_{Q_1}(f(x, \cdot)) \cdot \text{vol}(Q_1) \right) \text{vol}(R_1) \\ &\leq \sum_{R_1 \in \mathcal{P}_1} m_{R_1} \left( \int_Q f(\cdot, y) \, dy \right) \cdot \text{vol}(R_1) \\ &= I \left( \int_Q f(\cdot, y) \, dy; \mathcal{P}_1 \right), \end{aligned}$$

como havíamos afirmado. De modo análogo, provamos que

$$S \left( \int_Q f(\cdot, y) \, dy; \mathcal{P}_1 \right) \leq S(f; \mathcal{P}),$$

de modo que obtemos a cadeia de desigualdades

$$I(f; \mathcal{P}) \leq I \left( \int_Q f(\cdot, y) \, dy; \mathcal{P}_1 \right) \leq S \left( \int_Q f(\cdot, y) \, dy; \mathcal{P}_1 \right) \leq S(f; \mathcal{P}).$$

Sendo  $f$  integrável, todas estas desigualdade devem ser, em verdade, igualdades.  $\square$

Observamos que o teorema acima é mais geral do que o que estamos acostumados a aplicar em cursos de cálculo, pois dada uma função integrável  $f$  como acima, não é necessariamente verdade que, para  $x$  fixado, a função

$$y \in Q \mapsto f(x, y)$$

é integrável. De fato, o conjunto  $\{x\} \times Q$  tem medida nula em  $\mathbb{R}^{n+m}$  e alterar o valor da função  $f$  (de qualquer maneira imaginável) neste conjunto não mudaria o fato de que  $f$  é integrável. *Anyways*, no caso (usual dos cursos de Cálculo) em que todas as funções da forma

$$y \in Q \mapsto f(x, y) \quad \text{e} \quad x \in R \mapsto f(x, y)$$

são integráveis (por exemplo, quando  $f$  é contínua), temos como consequência direta as fórmulas usuais de integração iterada e mudança na ordem de integração

$$\int_{R \times Q} f(x, y) \, d(x, y) = \int_R \int_Q f(x, y) \, dy \, dx = \int_Q \int_R f(x, y) \, dx \, dy.$$

De forma relacionada está outro caso típico de cursos de Cálculo.

**Proposição 75.** *Considere o conjunto*

$$A = \{(x, y) \in [a, b] \times \mathbb{R}; \varphi(x) \leq y \leq \psi(x)\},$$

onde  $\varphi$  e  $\psi$  são funções contínuas (e logo limitadas) no intervalo compacto  $[a, b]$ . Seja  $f : A \rightarrow \mathbb{R}$  uma função contínua. Então, o conjunto  $A$  é Jordan mensurável e

$$\int_A f(x, y) \, d(x, y) = \int_a^b \int_{\varphi(x)}^{\psi(x)} f(x, y) \, dy \, dx.$$



*Demonstração.* Observe que a fronteira de  $A$  consiste de quatro “partes”, todas de medida nula:

$$\partial A = (\{a\} \times [\varphi(a), \psi(a)]) \cup \{(x, \varphi(x)); x \in [a, b]\} \cup (\{b\} \times [\varphi(b), \psi(b)]) \cup \{(x, \psi(x)); x \in [a, b]\}$$

Segue que  $A$  é Jordan mensurável. Além disso, considerando um retângulo  $R \times Q$  que contém  $A$  (note novamente que as funções  $\varphi$  e  $\psi$  são limitadas), temos que as funções  $y \mapsto f(x, y)\chi_{[\varphi(x), \psi(x)]}$  são contínuas, exceto possivelmente nos pontos  $\varphi(x)$  e  $\psi(x)$ . Logo, são integráveis e o resultado segue do teorema anterior.  $\square$

## 15.4 Teorema de Darboux – somas de Riemann mais gerais

Seja  $A \subseteq \mathbb{R}^n$  um conjunto Jordan mensurável. Vamos chamar a coleção de subconjuntos de  $A$

$$\mathcal{D} = \{X_1, X_2, \dots, X_k\}$$

de uma **decomposição** de  $A$  quando  $X_1, X_2, \dots, X_k$  são todos Jordan mensuráveis

$$A = \bigcup_{i=1}^k X_i \quad \text{e} \quad \text{int}(X_i \cap X_j) \neq \emptyset.$$

Denotamos também

$$|\mathcal{D}| := \max_{1 \leq i \leq k} \text{diam}(X_i).$$

Além disso, em consonância com a notação introduzida para partições de retângulos, escrevemos

$$m_{X_i} := \inf_{x \in X_i} f(x), \quad I(f; \mathcal{D}) := \sum_{i=1}^k m_{X_i} \cdot \text{vol}(X_i) \quad \text{e}$$

$$M_{X_i} := \sup_{x \in X_i} f(x), \quad S(f; \mathcal{D}) := \sum_{i=1}^k M_{X_i} \cdot \text{vol}(X_i).$$

Podemos ter escolhido um ponto arbitrário  $x_i \in X_i$  e definir uma soma de Riemann mais geral:

$$\sum_{i=1}^k f(x_i) \cdot \text{vol}(X_i).$$

Vamos mostrar nesta seção que todas estas formas de considerar somas de Riemann dão origem à mesma noção de integral. Decomposições mais gerais serão úteis para provar a Fórmula de Mudança de Variáveis para integrais múltiplas que apresentaremos na Subseção 15.5 abaixo.

O objetivo desta subseção é mostrar o seguinte.

**Teorema 76 (Critério de Darboux).** *Seja  $f : A \rightarrow \mathbb{R}$  uma função limitada em um conjunto Jordan mensurável  $A \subset \mathbb{R}^n$ . Então são equivalentes:*

(i)  $f$  é integrável em  $A$  o valor da integral é  $I$ ;

(ii) Existe o limite

$$I = \lim_{|\mathcal{D}| \rightarrow 0} \sum_{i=1}^k f(x_i) \cdot \text{vol}(X_i),$$

qualquer que seja a escolha de pontos  $x_i \in X_i$ .

Iniciamos a prova com um lema que afirma o seguinte: para um conjunto  $B \subset A$  de volume zero e para decomposições de  $A$  com diâmetros suficientemente pequenos, a soma dos volumes dos conjuntos da decomposição que estão “próximos” de  $B$  fica tão pequena quanto se queira.

Denotamos a “distância” entre dois conjuntos  $C$  e  $D$  por

$$d(C, D) := \inf \{|x - y|; x \in C, y \in D\}.$$

A palavra distância está entre aspas pois este objeto que definimos não é uma métrica propriamente dita, como se estuda em cursos de topologia geral (aqui pode-se ter  $d(C, D) = 0$  sem que necessariamente  $C = D$ , pois basta que  $C \cap D \neq \emptyset$ ). No entanto, serve para nossos propósitos.

**Lema 77.** *Seja  $B \subset A$  com  $\text{vol}(B) = 0$ . Temos que, dado  $\varepsilon > 0$ , existe  $\delta > 0$  tal que*

$$\mathcal{D} \text{ decomposição de } A \text{ com } |c\mathcal{D}| < \delta \implies \sum_{d(X_i, B) < \delta} \text{vol}(X_i) < \varepsilon.$$

*Demonstração.* Dado  $\varepsilon > 0$ , existe cobertura por retângulos  $R_1, R_2, \dots, R_N$  tais que

$$B \subseteq \bigcup_{i=1}^N R_i \quad \text{e} \quad \sum_{i=1}^N \text{vol}(R_i) < \varepsilon.$$

Nota que podemos “aumentar” um pouco os retângulos de modo a ainda ter a soma dos volumes menor do que  $\varepsilon$ . Mais precisamente, consideramos

$$R = [a_1, b_1] \times \dots \times [a_n, b_n] \rightsquigarrow R^\delta = [a_1 - \delta, b_1 + \delta] \times \dots \times [a_n - \delta, b_n + \delta].$$

Para  $\delta \ll 1$  suficientemente pequeno, temos

$$\sum_{i=1}^N \text{vol}(R_i^{2\delta}) < \varepsilon.$$

Assim, se  $\mathcal{D} = \{X_1, X_2, \dots, X_k\}$  é uma decomposição de  $A$  com  $|\mathcal{D}| < \delta$ , afirmamos que

$$d(X_i, B) < \delta \implies X_i \subseteq R_j^{2\delta} \text{ para algum } j.$$

De fato, da condição  $d(X_i, B) < \delta$ , sabemos que existem  $x \in X_i$  e  $b \in B$  tais que  $|x - b| < \delta$ . Além disso,  $\text{diam } X_i < \delta$  implica que todo ponto de  $X_i$  dista, no máximo  $2\delta$  de  $b$ :

$$y \in X_i \implies |y - b| \leq |x - y| + |x - b| < 2\delta.$$

Logo,  $b \in R_j$  para algum  $j$  implica  $y \in R_j^{2\delta}$  para todo  $y \in X_i$  como havíamos afirmado. Concluimos que

$$\mathcal{D} \text{ decomposição de } A \text{ com } |c\mathcal{D}| < \delta \implies \sum_{d(X_i, B) < \delta} \text{vol}(X_i) \leq \sum_{i=1}^N \text{vol}(R_i^{2\delta}) < \varepsilon \quad \square$$

*Demonstração do Teorema 76.* Sem perda de generalidade, podemos supor  $f \geq 0$ , pois  $f$  é limitada e somar uma constante  $C$  a  $f$  altera todos os termos do teorema pela mesma constante  $C \cdot \text{vol}(A)$ . Além disso, vamos considerar  $A$  um retângulo, pois o caso geral é uma consequência deste ao considerar  $\tilde{A}$  um retângulo que contem  $A$  e  $\tilde{f} := f \chi_A$ . Iniciamos mostrando que as somas de Riemann que consideramos aqui são compatíveis com as integrais inferior e superior que já havíamos definido.

Passo 1. Vale que

$$\overline{\int}_A f(x) dx = \lim_{|\mathcal{D}| \rightarrow 0} S(f; \mathcal{D}) = \inf_{\mathcal{D}} S(f; \mathcal{D}).$$

Dado  $\varepsilon > 0$ , existe uma partição  $\mathcal{P}$  de  $A$  por retângulos tal que

$$S(f; \mathcal{P}) < \overline{\int}_A f(x) dx + \frac{\varepsilon}{2}.$$

Por um lado, seja  $\mathcal{D} = \{X_1, X_2, \dots, X_k\}$  uma decomposição qualquer por subconjuntos fechados de  $A$  e definamos

$$B := \bigcup_{i=1}^k \partial X_i.$$

Temos que  $\text{vol}(B) = 0$ , pois cada  $X_i$  é compacto; ver Exercícios 33, 34 e 35. Logo, o Lema 77 acima implica que existe  $\delta > 0$  tal que toda partição  $\mathcal{P}$  com  $|\mathcal{P}| < \delta$  satisfaz

$$\sum_{d(R_i, B) < \delta} \text{vol}(R_i) < \varepsilon.$$

Observamos que, como  $B$  contém todas as fronteiras dos conjuntos  $X_i$ , a única possibilidade de se ter  $d(R_j, B) \geq \delta$  é quando  $R_j \subset \text{int } X_i$ . Logo, podemos separar a soma em duas parcelas, uma cuja soma dos volumes é pequena e a outra que é controlada pela soma superior de  $f$ :

$$\overline{\int}_A f(x) dx \leq S(f; \mathcal{P}) = \sum_{R \subset \text{int } X_i} M_R \cdot \text{vol}(R) + \sum_{R \cap B \neq \emptyset} M_R \cdot \text{vol}(R) \leq S(f; \mathcal{D}) + (\max_A f) \cdot \varepsilon.$$

Observe que a decomposição  $\mathcal{D}$  é qualquer; logo, sendo  $\varepsilon > 0$  arbitrário, temos

$$\overline{\int}_A f(x) dx \leq S(f; \mathcal{D}). \quad (37)$$

Por outro lado, considerando  $B$  como a união das faces dos retângulos da partição  $\mathcal{P}$ , tem-se também  $\text{vol}(B) = 0$  e, pelo Lema 77, existe  $\delta > 0$  tal que toda decomposição  $\mathcal{D}$  com  $|\mathcal{D}| < \delta$  satisfaz

$$\sum_{d(X_i, B) < \delta} \text{vol}(X_i) < \varepsilon.$$

Argumentando como acima, obtemos

$$\begin{aligned} S(f; \mathcal{D}) &= \sum_{X_i \subset \text{int } R} M_{X_i} \cdot \text{vol}(X_i) + \sum_{X_i \cap B \neq \emptyset} M_{X_i} \cdot \text{vol}(X_i) \\ &\leq \sum_{R \in \mathcal{P}} M_R \cdot \text{vol}(R) + (\max_A f) \sum_{X_i \cap B \neq \emptyset} \text{vol}(X_i) \\ &< \overline{\int}_A f(x) dx + (1 + \max_A f) \cdot \varepsilon. \end{aligned} \quad (38)$$

Juntando (37) e (38), concluímos que, dado  $\varepsilon > 0$ , existe  $\delta > 0$  tal que

$$|\mathcal{D}| < \delta \implies S(f; \mathcal{D}) - C\varepsilon \leq \overline{\int}_A f(x) dx \leq S(f; \mathcal{D}),$$

como queríamos.

Passo 2. Analogamente, vale que

$$\underline{\int}_A f(x) dx = \lim_{|\mathcal{D}| \rightarrow 0} I(f; \mathcal{D}) = \sup_{\mathcal{D}} I(f; \mathcal{D}).$$

Passo 3. (i)  $\implies$  (ii). Trivialmente, para qualquer decomposição  $\mathcal{D}$  e para qualquer escolha de pontos  $x_i$ , vale que

$$I(f; \mathcal{D}) \leq \sum_{i=1}^k f(x_i) \cdot \text{vol}(X_i) \leq S(f; \mathcal{D})$$

Se for  $f$  integrável, os Passos 1 e 2, em conjunto com o Teorema do Sanduiche, implicam (ii).

**Passo 4.** (ii)  $\implies$  (i). Dado  $\varepsilon > 0$ , existe uma decomposição  $\mathcal{D} = \{X_1, X_2, \dots, X_N\}$  tal que

$$\left| \sum_{i=1}^k f(x_i) \cdot \text{vol}(X_i) - I \right| < \varepsilon$$

qualquer que seja a escolha de pontos  $x_i \in X_i$ . Escolhemos  $y_i \in X_i$  e  $z_i \in X_i$  tais que, respectivamente,

$$f(y_i) \leq m_{X_i} + \frac{\varepsilon}{\text{vol}(X_i)} \quad \text{e} \quad f(z_i) \geq M_{X_i} - \frac{\varepsilon}{\text{vol}(X_i)}.$$

Logo – vá lendo as desigualdades desde a trivial do meio até as pontas –

$$I - N\varepsilon \leq \sum_{i=1}^N f(y_i) \cdot \text{vol}(X_i) - N\varepsilon \leq I(f; \mathcal{D}) \leq S(f; \mathcal{D}) \leq \sum_{i=1}^N f(z_i) \cdot \text{vol}(X_i) + N\varepsilon \leq I + N\varepsilon.$$

Portanto, dado  $\varepsilon > 0$ , existe  $\mathcal{D}$  decomposição tal que

$$\left| S(f; \mathcal{D}) - I(f; \mathcal{D}) \right| < 2N\varepsilon. \quad \square$$

## 15.5 Fórmula de mudança de variáveis

Para integrais unidimensionais, temos a seguinte fórmula de mudança de coordenadas (também conhecido como integrar por substituição): para  $I \subseteq \mathbb{R}$  um intervalo e  $\phi : [a, b] \rightarrow I$  é diferenciável com derivada contínua e  $f : I \rightarrow \mathbb{R}$  é uma função contínua, então

$$\int_{\phi(a)}^{\phi(b)} f(x) dx = \int_a^b f(\phi(t)) \phi'(t) dt.$$

Intuitivamente, pensamos na substituição  $x = \phi(t) \implies dx = \phi'(t) dt$ . Observamos, no entanto, que poderia ser  $\phi(b) < \phi(a)$ , de modo que, na notação de conjuntos, devemos escrever

$$\int_{\phi([a,b])} f(x) dx = \int_{[a,b]} f(\phi(t)) \cdot |\phi'(t)| dt.$$

Um tópico que nem sempre é abordado em cursos de Cálculo é a fórmula de mudança de variáveis para integrais múltiplas, onde o determinante Jacobiano aparece de forma fundamental.

**Teorema 78 (Fórmula de mudança de variáveis).** *Sejam  $U, V \subset \mathbb{R}^n$  conjuntos abertos e  $\phi : U \rightarrow V$  um difeomorfismo de classe  $C^1$ . Seja  $K \subset U$  um conjunto compacto Jordan mensurável. Então, para qualquer  $f : \phi(K) \rightarrow \mathbb{R}$  integrável, vale a fórmula de mudança de variáveis:*

$$\int_{\phi(K)} f(x) dx = \int_K f(\phi(y)) \cdot |\det \phi'(y)| dy.$$

Em subseções seguintes, veremos algumas aplicações clássicas deste resultado em coordenadas polares e esféricas. Passamos à prova, que consiste em vários passos.

*Demonstração.* Vamos demonstrar o teorema em passos que vão aumentando a generalidade.

**Passo 1.** Vale para dimensão  $n = 1$ ,  $\phi : \mathbb{R} \rightarrow \mathbb{R}$  afim invertível e  $K = [a, b]$  um intervalo compacto.

Neste caso, podemos escrever  $\phi(x) = cx + d$ , com  $c \neq 0$ . Seja  $\mathcal{P} = \{t_1 < t_2 < \dots < t_N\}$  uma partição de  $\phi(K)$ , que é um intervalo. A soma de Riemann superior pode ser escrita como

$$S(f; \mathcal{P}) = \sum_{i=1}^N M_{[t_i, t_{i+1}]}(f) \cdot (t_i - t_{i-1})$$

Observe, sendo  $\phi$  invertível, que  $t_i = \phi(s_i) = cs_i + d$  implica que  $\mathcal{Q} = \{s_i\}$  é uma partição de  $K$  e temos que

$$S(f \circ \phi; \mathcal{Q}) = \sum_{i=1}^N M_{[s_i, s_{i+1}]}(f \circ \phi) \cdot |s_i - s_{i-1}| = \frac{1}{|c|} \sum_{i=1}^N M_{[t_i, t_{i+1}]}(f) \cdot (t_i - t_{i-1}) = \frac{S(f; \mathcal{P})}{|c|}.$$

Como toda partição de  $K$  pode ser obtida desta maneira, concluímos que

$$\overline{\int_{\phi(K)} f(x) dx} = |c| \overline{\int_K f(\phi(y)) dy} = |\det \phi'(y)| \overline{\int_K f(\phi(y)) dy}. \quad (39)$$

A igualdade entre as integrais inferiores é obtida analogamente.

**Passo 2.** Vale para  $\phi: \mathbb{R}^n \rightarrow \mathbb{R}^n$  linear invertível.

Na linguagem da Álgebra Linear, pode-se dizer que “ $\phi$  é equivalente por linhas à transformação identidade”. Na verdade, é assim que se calcula inversas de matrizes em Álgebra Linear. Um pouco mais precisamente, escrevendo a matriz canônica associada a  $\phi$ , podemos chegar à matriz identidade fazendo uma sequência de dois tipos de operações elementares (que são escritas em termos de transformações invertíveis):

- Trocar duas linhas de lugar, que corresponde à transformação linear

$$T(x) = (x_1, x_2, \dots, x_{i-1}, x_j, x_{i+1}, \dots, x_{j-1}, x_i, x_{j+1}, \dots, x_{n-1}, x_n),$$

cujas matriz associada é (trocar linha  $i$  e  $j$  de lugar na matriz identidade)

$$\begin{bmatrix} 1 & 0 & \cdots & 0 & \cdots & 0 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & \cdots & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & & \vdots & & \vdots & & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & \cdots & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & & \vdots & & \vdots & & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & \cdots & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & & \vdots & & \vdots & & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & \cdots & 0 & \cdots & 1 & 0 \\ 0 & 0 & \cdots & 0 & \cdots & 0 & \cdots & 0 & 1 \end{bmatrix}$$

- Adicionar um múltiplo de uma linha à outra linha:

$$S(x) = (x_1, x_2, \dots, ax_i + bx_j, \dots, x_{n-1}, x_n),$$

cujas matriz associada é da forma

$$\begin{bmatrix} 1 & 0 & \cdots & 0 & \cdots & 0 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & \cdots & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & & \vdots & & \vdots & & \vdots & \vdots \\ 0 & 0 & \cdots & a & \cdots & b & \cdots & 0 & 0 \\ \vdots & \vdots & & \vdots & & \vdots & & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & \cdots & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & & \vdots & & \vdots & & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & \cdots & 0 & \cdots & 1 & 0 \\ 0 & 0 & \cdots & 0 & \cdots & 0 & \cdots & 0 & 1 \end{bmatrix}$$

Dessa maneira, temos que  $T_1 T_2 T_3 \cdots T_N \phi = I$ , onde cada  $T_i$  é uma transformação do tipo de  $T$  ou do tipo de  $S$  acima. Como as inversas são também elementares e do mesmo tipo, podemos escrever  $\phi$  como um produto de transformações destes dois tipos.

É importante a observação que, se a fórmula de mudança de variáveis for válida para duas transformações  $T_1$  e  $T_2$ , então vale também para  $\psi = T_1 T_2$ :

$$\int_{\psi(K)} f = \int_{T_2(K)} (f \circ T_1) |\det T_1| = \int_K f(T_1 T_2(y)) |\det T_1| |\det T_2| dy = \int_K f(\psi(y)) |\det \psi'(y)| dy.$$

Resta mostrar que a fórmula vale para  $T$  e  $S$ , como acima. A prova para  $T$  é direta: temos  $\det T = -1$  e qualquer retângulo satisfaz

$$R_k = [a_1, b_1] \times \cdots \times [a_i, b_i] \times \cdots \times [a_j, b_j] \times \cdots \times [a_n, b_n] \iff$$

$$T(R_k) = [a_1, b_1] \times \cdots \times [a_j, b_j] \times \cdots \times [a_i, b_i] \times \cdots \times [a_n, b_n]$$

Em outras palavras, partições por retângulos de (um retângulo que contém  $K$ , denotado por)  $R$  estão em relação biunívoca com partições por retângulos de  $T(R)$  (que também é um retângulo). Além disso, o máximo e o mínimo nos retângulos das partições são os mesmos. Logo,

$$\int_{T(R)} f \chi_K = \int_R (f \circ T) \chi_K \quad \text{e também} \quad \int_{T(R)} f \chi_K = \int_R (f \circ T) \chi_K,$$

como queríamos.

Provamos para também vale a mudança de variáveis para transformações do tipo de  $S$  acima: sem perda de generalidade, suponhamos

$$S(x) = (ax_1 + bx_2, x_2, x_3, \dots, x_{n-1}, x_n).$$

Note que  $\det S = a$ . Escrevemos  $y = (x_2, x_3, \dots, x_n)$  e  $x = (x_1, y)$ . Vamos aplicar integração repetida:

$$\begin{aligned} \int_{S(K)} f(x) dx &= \int_{\pi^2 S(K)} \left( \int_{\pi^1 S(K)} f(x_1, y) dx_1 \right) dy \\ &= \int_{\pi^2(K)} \left( |a| \int_{\pi^1(K)} f(ax_1 + bx_2, y) dx_1 \right) dy \\ &= \int_K f(S(x)) |\det S| dx. \end{aligned}$$

**Passo 3. Vale para  $\phi : \mathbb{R}^n \rightarrow \mathbb{R}^n$  geral.**

Note que  $D_{f \circ \phi} = \phi^{-1}(D_f)$ . Sendo  $\phi$  e  $\phi^{-1}$  ambas localmente de Lipschitz, temos que

$$f \text{ é integrável em } \phi(K) \iff f \circ \phi \text{ é integrável em } K.$$

Vamos utilizar que

$$\int_{\phi(K)} f(x) dx = \lim_{|\mathcal{P}| \rightarrow 0} \sum_{Y_i \in \mathcal{D}} f(y_i) \cdot \text{vol}(Y_i),$$

onde  $\mathcal{D}$  é uma decomposição (não necessariamente por retângulos) de  $h(K)$  e  $y_i \in Y_i$  é qualquer. Sendo  $\phi$  um difeomorfismo, decomposições deste tipo estão em correspondência biunívoca com decomposições  $\{X_i\}$  de  $K$ . Desta forma, podemos considerar  $Y_i = \phi(X_i)$  e  $y_i = \phi(x_i)$  acima.

A aplicação  $\phi'(x_i)$  é linear invertível. Pelo Passo 2, podemos escrever

$$\text{vol}(\phi(X_i)) = \text{vol} \left[ \phi'(x_i) \phi'(x_i)^{-1}(\phi(X_i)) \right] = |\det \phi'(x_i)| \cdot \text{vol} \left[ \phi'(x_i)^{-1}(\phi(X_i)) \right].$$

**Afirmção:**  $\text{vol} \left[ \phi'(x_i)^{-1}(\phi(X_i)) \right] \leq M_i^n \cdot \text{vol}(X_i)$ , onde  $M_i := \sup_{x \in X_i} |\phi'(x_i)^{-1}(\phi'(x))|$ .

De fato, vamos denotar temporariamente  $g = \phi'(x_i)^{-1} \circ \phi$ . Em verdade, nossa afirmação vale para o caso mais geral em que  $g$  é um difeomorfismo qualquer:

$$M := \sup_{x \in X} |g'(x)| \implies \text{vol}[g(X)] \leq M^n \text{vol}(X).$$

Se for  $X$  um cubo da forma

$$X = \{x \in \mathbb{R}^n; |x - x_0|_\infty \leq \ell/2\},$$

então, pela Desigualdade do Valor Médio, temos

$$|g(x) - g(x_0)| \leq |g'(c)| \cdot |x - x_0| \leq M|x - x_0|,$$

de modo que  $g(X) \subseteq \{y \in \mathbb{R}^n; |y - g(x_0)|_\infty \leq M\ell/2\}$ . Segue que

$$\text{vol}(g(X)) \leq M^n \ell^n = M^n \text{vol}(X).$$

O caso geral segue deste, considerando uma cobertura de  $X_i$  por cubos e utilizando a continuidade da derivada  $g'$ : dado  $\varepsilon > 0$ , existe  $A$  aberto com  $X \subset A \subset U$  e  $|g'(x)| \leq M^n + \varepsilon$  para todo  $x \in A$ . Sendo  $X$  Jordan mensurável, tomamos uma cobertura finita por cubos  $C_k$  de modo que

$$X \subset \bigcup_k C_k \subset A \quad \text{e} \quad \sum_k \text{vol}(C_k) < \text{vol}(X) + \varepsilon.$$

Como os cubos estão em  $A$ , temos

$$g(X) \subset \bigcup_k g(C_k) \quad \text{e} \quad \sup_{x \in C_k} |g'(x)| \leq M.$$

Pelo que fizemos para cubos

$$\text{vol}(g(X)) \leq \sum_k \text{vol}(g(C_k)) \leq (M^n + \varepsilon) \sum_k \text{vol}(C_k) < (M^n + \varepsilon)(\text{vol}(X) + \varepsilon)$$

Sendo  $\varepsilon > 0$  arbitrário, a afirmação está provada.

De volta à prova da Fórmula de Mudança de Variáveis, podemos escrever

$$\begin{aligned} \int_{\phi(K)} f(x) dx &= \lim_{|\mathcal{D}| \rightarrow 0} \sum_{X_i \in \mathcal{D}} f(h(x_i)) \cdot \text{vol}(h(X_i)) \\ &\leq \lim_{|\mathcal{D}| \rightarrow 0} \sum_{X_i \in \mathcal{D}} f(h(x_i)) |\det \phi'(x_i)| M_i^n \cdot \text{vol}(X_i) \\ &= \lim_{|\mathcal{D}| \rightarrow 0} \sum_{X_i \in \mathcal{D}} f(h(x_i)) |\det \phi'(x_i)| (1 + (M_i - 1))^n \cdot \text{vol}(X_i) \\ &= \int_K f(\phi(y)) |\det \phi'(y)| dy. \end{aligned} \tag{40}$$

A última igualdade seria imediata, não fosse o termo  $(1 + (M_i - 1))^n$ . No entanto, pelo Lema 79 abaixo, aplicado para a função  $\varphi(x, y) := |\phi'(y)^{-1} \cdot \phi'(x)|$  em  $X_i$ , temos que, dado  $\varepsilon > 0$ , existe  $\delta > 0$  tal que

$$|\mathcal{D}| < \delta \implies |M_i - 1| < \varepsilon.$$

Deste modo, temos que  $(1 + (M_i - 1))^n \rightarrow 1$  quando  $|\mathcal{D}| \rightarrow 0$  (uniformemente em  $i$ ). Assim, dado  $\varepsilon > 0$ , existe  $\delta > 0$  tal que, para todo  $i = 1, 2, \dots, N$ ,

$$|\mathcal{D}| < \delta \implies |(1 + (M_i - 1))^n - 1| < \frac{\varepsilon}{f(h(x_i)) |\det \phi'(x_i)| \text{vol}(K)},$$

o que mostra que (faça a diferença entre os os somatórios abaixo e veja que fica menor do que  $\varepsilon$ )

$$\lim_{|\mathcal{D}| \rightarrow 0} \sum_{X_i \in \mathcal{D}} f(h(x_i)) |\det \phi'(x_i)| (1 + (M_i - 1))^n \text{vol}(X_i) = \lim_{|\mathcal{D}| \rightarrow 0} \sum_{X_i \in \mathcal{D}} f(h(x_i)) |\det \phi'(x_i)| \text{vol}(X_i).$$

Finalmente, a desigualdade em (40) pode ser aplicada para  $\phi^{-1}$  na última integral de (40), obtendo

$$\int_{\phi^{-1}(\phi(K))} f(\phi(y)) |\det \phi'(y)| dy \leq \int_{\phi(K)} f(x) |\det \phi'(\phi^{-1}(x))| |\det(\phi^{-1})'(x)| dx = \int_{\phi(K)} f(x) dx. \quad \square$$

Encerramos a seção com a prova do lema que foi utilizado na demonstração acima.

**Lema 79.** *Seja  $X$  um conjunto compacto e Jordan mensurável e  $\varphi : X \times X \rightarrow \mathbb{R}$  uma função contínua tal que  $\phi(x, x) = 1$ . Então, dado  $\varepsilon > 0$ , existe  $\delta > 0$  tal que*

$$|x - y| < \delta \implies |\phi(x, y) - 1| < \varepsilon.$$

*Demonstração.* De fato, do contrário seria que, para todo  $k \in \mathbb{N}$ , conseguimos encontrar  $x_k \in X$  e  $y_k \in X$  tal que

$$|x_k - y_k| \leq \frac{1}{k} \quad \text{e} \quad |\varphi(x_k, y_k) - 1| \geq \varepsilon.$$

Sendo  $X$  compacto, a menos de extrair subsequências,  $x_k$  e  $y_k$  convergem para um mesmo limite  $a$ . Mas, sendo  $\varphi$  contínua, isto contradiz ser  $\varphi(a, a) = 1$ .  $\square$

### 15.5.1 Coordenadas polares

A função  $\phi$  que muda de coordenadas polares para Cartesianas pode ser escrita como

$$\phi(r, \theta) = (r \cos \theta, r \sin \theta), \quad \text{onde } (r, \theta) \in (0, +\infty) \times (0, 2\pi).$$

Assim, a matriz das derivadas parciais é

$$\begin{bmatrix} \frac{\partial \phi_1}{\partial r} & \frac{\partial \phi_1}{\partial \theta} \\ \frac{\partial \phi_2}{\partial r} & \frac{\partial \phi_2}{\partial \theta} \end{bmatrix} = \begin{bmatrix} \cos \theta & -r \sin \theta \\ \sin \theta & r \cos \theta \end{bmatrix} \quad \text{que satisfaz } |\det \phi'(r, \theta)| = |r \cos^2 \theta + r \sin^2 \theta| = |r| = r.$$

Portanto, a fórmula de mudança de variáveis implica que, em coordenadas polares:

$$\iint_K f(x, y) d(x, y) = \iint_{\phi^{-1}(R)} f(r \cos \theta, r \sin \theta) r d(r, \theta).$$

### 15.5.2 Coordenadas esféricas

Vamos verificar como fazer para escrever uma integral tripla em coordenadas esféricas. O raciocínio segue as mesmas linhas da subseção anterior para coordenadas polares. Sabemos que. A função  $\phi$  que muda de coordenadas esféricas para Cartesianas pode ser escrita como

$$\phi(\rho, \theta, \phi) = (\rho \sin \phi \cos \theta, \rho \sin \phi \sin \theta, \rho \cos \phi), \quad \text{onde } (\rho, \theta, \phi) \in U := (0, +\infty) \times (0, 2\pi) \times (0, \pi).$$

Assim, a matriz das derivadas parciais é

$$\begin{bmatrix} \frac{\partial \phi_1}{\partial \rho} & \frac{\partial \phi_1}{\partial \theta} & \frac{\partial \phi_1}{\partial \phi} \\ \frac{\partial \phi_2}{\partial \rho} & \frac{\partial \phi_2}{\partial \theta} & \frac{\partial \phi_2}{\partial \phi} \\ \frac{\partial \phi_3}{\partial \rho} & \frac{\partial \phi_3}{\partial \theta} & \frac{\partial \phi_3}{\partial \phi} \end{bmatrix} = \begin{bmatrix} \sin \phi \cos \theta & -\rho \sin \phi \sin \theta & \rho \cos \phi \cos \theta \\ \sin \phi \sin \theta & \rho \sin \phi \cos \theta & \rho \cos \phi \sin \theta \\ \cos \phi & 0 & -\rho \sin \phi \end{bmatrix}.$$



cujo Jacobiano é (usando, por exemplo a segunda coluna para expandir em cofatores)

$$\begin{aligned}
 J\phi &= \left| \rho \operatorname{sen} \phi \operatorname{sen} \theta \begin{vmatrix} \operatorname{sen} \phi \operatorname{sen} \theta & \rho \cos \phi \operatorname{sen} \theta \\ \cos \phi & -\rho \operatorname{sen} \phi \end{vmatrix} + \rho \operatorname{sen} \phi \cos \theta \begin{vmatrix} \operatorname{sen} \phi \cos \theta & \rho \cos \phi \cos \theta \\ \cos \phi & -\rho \operatorname{sen} \phi \end{vmatrix} \right| \\
 &= \left| \rho \operatorname{sen} \phi \operatorname{sen} \theta \left( -\rho \operatorname{sen}^2 \phi \operatorname{sen} \theta - \rho \cos^2 \phi \operatorname{sen} \theta \right) + \rho \operatorname{sen} \phi \cos \theta \left( -\rho \operatorname{sen}^2 \phi \cos \theta - \rho \cos^2 \phi \cos \theta \right) \right| \\
 &= \left| \rho \operatorname{sen} \phi \operatorname{sen} \theta \left( -\rho \operatorname{sen} \theta \right) + \rho \operatorname{sen} \phi \cos \theta \left( -\rho \cos \theta \right) \right| = \left| -\rho^2 \operatorname{sen} \phi \left( \operatorname{sen}^2 \theta + \cos^2 \theta \right) \right| \\
 &= \left| -\rho^2 \operatorname{sen} \phi \right| = \rho^2 \operatorname{sen} \phi.
 \end{aligned}$$

Na última igualdade, utilizamos que  $\phi \in (0, \pi)$  implica  $\operatorname{sen} \phi > 0$ , de modo que o valor absoluto está considerado corretamente. Portanto, a fórmula de mudança de variáveis implica que, em coordenadas esféricas:

$$\iint_K f(x, y, z) \, d(x, y, z) = \iiint_{\phi^{-1}(K)} f(\rho \operatorname{sen} \phi \cos \theta, \rho \operatorname{sen} \phi \operatorname{sen} \theta, \rho \cos \phi) \rho^2 \operatorname{sen} \phi \, d(\rho, \theta, \phi),$$

como é de costume em cursos de cálculo.

## 16 Formas diferenciais

Em análise vetorial e geometria diferencial, formas diferenciais aparecem como uma maneira de estudar integrais em curvas e em superfícies de maneira unificada. Como veremos, é mais natural, ou até “inescapável”, estudar formas diferenciais em uma variedade diferenciável  $M$ . No entanto, vamos estudar inicialmente o caso especial  $M = \mathbb{R}^n$ , que é provavelmente mais familiar aos leitores dessas notas. A discussão que segue é baseada em diversas fontes; as principais são [1, 3, 11, 12]. A referência [1] não é um livro texto sobre o assunto e, na realidade, fala sobre objetos matemáticos mais sofisticados do que estes que tratamos nestas notas. No entanto, expõe de maneira particularmente clara os pontos essenciais da teoria e nossa discussão seguiu algumas de suas linhas.

Um caso especial de forma diferencial é o do diferencial de uma função, como se estuda em Cálculo [a bit of history, cartan, grassmann, etc...](#)



É importante identificar as classes de conjuntos em que estamos interessados. Por exemplo, é usual pensar em um campo vetorial em  $\mathbb{R}^n$  como uma função do tipo

$$F : \Omega \subseteq \mathbb{R}^n \rightarrow \mathbb{R}^n.$$

É também comum imaginar que, para  $x \in \Omega$  fixado, o vetor  $F(x)$  tem seu “ponto inicial” em  $x$ . Talvez fosse melhor escrever algo do tipo  $F(x) \in \mathbb{R}_x^n$  e, de fato, vamos fazer isto em algum momento do nosso curso. Dada uma função diferenciável  $f : \Omega \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$ , podemos considerar o campo gradiente dado por

$$F = \nabla f : \Omega \subseteq \mathbb{R}^n \rightarrow \mathbb{R}^n.$$

Lembramos que o gradiente é um operador que depende da estrutura de produto interno escolhida. O diferencial de uma função, por outro lado, depende apenas da “estrutura diferenciável” do espaço  $\mathbb{R}^n$  e é uma aplicação do tipo

$$df : \Omega \subseteq \mathbb{R}^n \rightarrow (\mathbb{R}^n)^*.$$

Mais geralmente, definimos uma **forma diferencial de grau um**, também conhecido como um **campo de formas lineares**, ou apenas **1-forma**, em  $\Omega \subseteq \mathbb{R}^n$  como uma aplicação diferenciável do tipo

$$\omega : \Omega \rightarrow (\mathbb{R}^n)^*.$$

Assim, para cada  $x \in \Omega$ , temos que  $\omega(x)$  é um funcional linear em  $\mathbb{R}^n$ . Em outras palavras, para cada  $x \in \mathbb{R}^n$ , tem-se

$$\omega(x) \cdot (av + bw) = a\omega(x) \cdot v + b\omega(x) \cdot w \quad \forall v, w \in \mathbb{R}^n \text{ e } \forall a, b \in \mathbb{R}.$$

São comuns as notações  $\omega(x) \cdot v = \omega(x)[v] = \omega(x)(v)$ . Em coordenadas, podemos escrever

$$\omega(x) = \sum_{i=1}^n a_i(x) dx^i,$$

onde  $\{dx^i\}$  é a base canônica do espaço dual, como em (6).

**Exercício 37.** Mostre que para uma 1-forma  $\omega : \Omega \rightarrow (\mathbb{R}^n)^*$ , as seguintes afirmações são equivalentes:

- (i) a forma  $\omega$  é diferenciável (ou classe  $C^k$ );
- (ii) cada uma das componentes  $a_i$  é diferenciável (ou classe  $C^k$ );
- (iii) para cada  $v \in \mathbb{R}^n$ , a função  $x \mapsto \omega(x)[v]$  são diferenciável (ou classe  $C^k$ )

**Exemplo 80.** Dada uma função  $f : \Omega \rightarrow \mathbb{R}$  de classe  $C^k$ , o diferencial  $df : \Omega \rightarrow (\mathbb{R}^n)^*$ , dado por

$$df(x) = \sum_{i=1}^n \frac{\partial f}{\partial x_i}(x) dx^i \in (\mathbb{R}^n)^*$$

é uma 1-forma de classe  $C^{k-1}$ .

Dada uma forma diferencial de ordem um  $\omega : \Omega \rightarrow (\mathbb{R}^n)^*$  e uma curva diferenciável  $\gamma : [a, b] \rightarrow \Omega$ , podemos definir a **integral de linha** de  $\omega$  ao longo de  $\gamma$  por

$$\int_{\gamma} \omega := \int_a^b \omega(\gamma(t)) \cdot \gamma'(t) dt. \quad (41)$$

Poderíamos definir, mais geralmente, a integral de uma 1-forma ao longo de caminhos mais gerais, não necessariamente diferenciáveis. No entanto, a definição seria um pouco repetitiva e entediante:

$$\int_{\gamma} \omega := \lim_{|\mathcal{P}| \rightarrow 0} \sum_{i=0}^{N+1} \omega(\gamma(t_i^*)) \cdot (\gamma(t_{i+1}) - \gamma(t_i)),$$

onde  $\mathcal{P} = \{a = t_0 < t_1 < \dots < t_N < t_{N+1} = b\}$ . Quando  $\gamma$  é diferenciável, o Teorema do Valor Médio pode ser utilizado para mostrar a equivalência das definições.

Mais imediato é estender a definição em (41) para caminhos  $C^1$  por partes. Dizemos que  $\gamma : [a, b] \rightarrow \Omega$  é um **caminhos  $C^1$  por partes** quando é contínuo e existem  $a < t_1 < \dots < t_N < b$  tais que as restrições  $\gamma_i := \gamma|_{[t_i, t_{i+1}]}$  são de classe  $C^1$ . A integral da forma  $\omega$  é então dada por

$$\int_{\gamma} \omega = \sum_{i=0}^N \int_{\gamma_i} \omega. \quad (42)$$

Isto vale pois o integrando em (41) pode estar definido a menos de um conjunto de medida nula.

Observamos que a integral de linha depende da curva e não da parametrização, apesar de nossa notação aparentemente indicar o contrário.

**Proposição 81.** Seja  $\omega : \Omega \rightarrow (\mathbb{R}^n)^*$  uma 1-forma contínua e  $\gamma : [a, b] \rightarrow \Omega$  um caminho  $C^1$  por partes. Se  $\phi : [c, d] \rightarrow [a, b]$  é de classe  $C^1$  com  $\phi(c) = a$  e  $\phi(d) = b$ , então

$$\int_{\gamma} \omega = \int_{\gamma \circ \phi} \omega.$$

*Demonstração.* Pelo teorema de mudança de variáveis para integrais unidimensionais:

$$\int_{\gamma} \omega = \int_a^b \omega(\gamma(t)) \cdot \gamma'(t) dt = \int_c^d \omega(\gamma(\phi(s))) \cdot \gamma'(\phi(s)) \phi'(s) ds = \int_{\gamma \circ \phi} \omega. \quad \square$$

Em seguida, provamos algumas propriedades da integral de linha definida acima.

**Proposição 82 (Propriedades da integral).** *A integração de formas diferenciais contínuas ao longo de curvas satisfaz as seguintes propriedades:*

(i) *Linearidade: para  $\alpha, \beta \in \mathbb{R}$  e para  $\omega_1, \omega_2 \in \Omega^1(\mathbb{R}^n)$*

$$\int_{\gamma} \alpha \omega_1 + \beta \omega_2 = \alpha \int_{\gamma} \omega_1 + \beta \int_{\gamma} \omega_2.$$

(ii) *Se  $\gamma^-$  tem a orientação inversa de  $\gamma$ , então*

$$\int_{\gamma^-} \omega = - \int_{\gamma} \omega.$$

(iii) *Concatenação de caminhos: se  $\gamma = \gamma_1 * \gamma_2$ , então*

$$\int_{\gamma} \omega = \int_{\gamma_1} \omega + \int_{\gamma_2} \omega.$$

(iv) *Estimativa básica da integral:*

$$\left| \int_{\gamma} \omega \right| \leq \sup_{\gamma} \|\omega\| \cdot \ell(\gamma).$$

*Em particular, se  $\|\omega(x)\| \leq M$ , para todo  $x \in \gamma$ , tem-se*

$$\left| \int_{\gamma} \omega \right| \leq M \cdot \ell(\gamma).$$

*Demonstração.* O item (i) é imediato da definição e é deixado como exercício. Para provar (ii), observamos que se  $\gamma(t), t \in [a, b]$ , então

$$\gamma^-(t) = z(a + b - t), t \in [a, b],$$

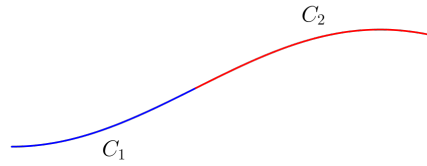
percorre a mesma curva no sentido contrário. Logo, (i) é uma simples mudança de coordenadas unidimensional.



Para mostrar (iii), a concatenação de caminhos é definida como na figura abaixo (encontre uma fórmula explícita). Note que não é necessário diferenciabilidade no ponto de encontro. Dessa forma, o item (iii) é consequência direta de (42)

Finalmente, passamos à prova de (iv), que segue da propriedade semelhante para integrais reais:

$$\left| \int_{\gamma} \omega \right| = \left| \int_a^b \omega(\gamma(t)) \cdot \gamma'(t) dt \right| \leq \int_a^b |\omega(\gamma(t)) \cdot \gamma'(t)| dt \leq \sup_{t \in [a, b]} \|\omega(\gamma(t))\| \int_a^b |\gamma'(t)| dt. \quad \square$$



**Observação 83.** Em cursos de Cálculo, define-se a integral de linha de um campo de vetores ao longo de uma curva. Notamos que, fixado o produto escalar usual em  $\mathbb{R}^n$ , o Teorema de Representação de Riesz nos dá uma correspondência biunívoca entre elementos de  $(\mathbb{R}^n)^*$  e elementos de  $\mathbb{R}^n$ . Assim,

$$\omega(x) = \sum_{i=1}^n a_i(x) dx^i \text{ é uma 1-forma } \longleftrightarrow F(x) = \sum_{i=1}^n a^i(x) e_i \text{ é um campo vetorial.}$$

Logo, por dualidade, podemos definir

$$\int_{\gamma} F := \int_a^b \langle F(\gamma(t)), \gamma'(t) \rangle dt$$

Esta é fisicamente interpretada, mais naturalmente (será mesmo?), como o trabalho que o campo de forças  $F$  realiza ao longo da curva  $\gamma$ .

Uma forma diferencial  $\omega$  é dita **exata** em  $\Omega$  quando existe uma função diferenciável  $f : \Omega \rightarrow \mathbb{R}$ , conhecida como um potencial escalar de  $\omega$ , tal que<sup>19</sup>

$$\omega = df.$$

Em vista da Observação 83, a seguinte proposição não deve surpreender.

**Proposição 84.** *Seja  $\Omega \subseteq \mathbb{R}^n$  aberto e conexo e  $\omega$  uma 1-forma. São equivalentes:*

- (i)  $\omega$  é exata em  $\Omega$ ;
- (ii) dados dois pontos  $x, y \in \Omega$ , a integral de  $\omega$  é independente do caminho  $C^1$  por partes  $\gamma : [a, b] \rightarrow \Omega$  que une  $x$  e  $y$ ;
- (iii) para toda curva  $C^1$  por partes e fechada em  $\Omega$ , a integral de  $\omega$  é zero.

*Demonstração.* Vejamos que (i)  $\implies$  (ii). Por simplicidade, vamos supor  $\gamma \in C^1([a, b])$ , o caso geral ficando como exercício. Sendo  $\omega$  exata em  $\Omega$ , existe um potencial escalar  $f : \Omega \rightarrow \mathbb{R}$  tal que  $\omega = df$ . Logo, pela Regra da Cadeia,

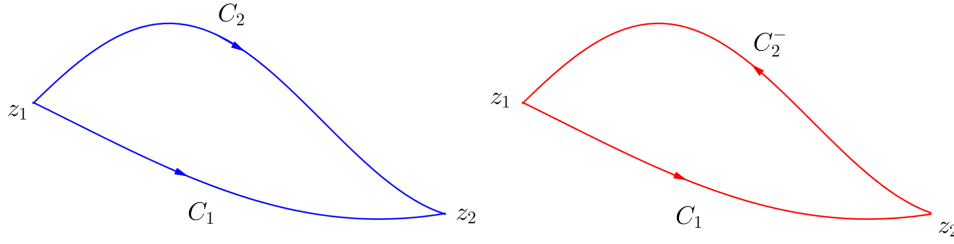
$$\int_{\gamma} \omega = \int_a^b df(\gamma(t)) \cdot \gamma'(t) dt = \int_a^b \frac{d}{dt} f(\gamma(t)) dt = f(\gamma(b)) - f(\gamma(a)),$$

que mostra o que queríamos.

A mesma conta acima mostra que (i)  $\implies$  (iii); de fato, para uma curva fechada, tem-se  $\gamma(a) = \gamma(b)$ .

Para mostrar que (iii)  $\implies$  (ii), sejam duas curvas  $\gamma_1$  e  $\gamma_2$ . A curva  $\gamma := \gamma_1 * \gamma_2^{-1}$ , como na figura, é fechada.

<sup>19</sup>Isto é para formas  $\omega$  que o campo gradiente é para campos vetoriais.



Pela hipótese (iii) e pela proposição anterior, temos

$$0 = \int_{\gamma} \omega = \int_{\gamma_1} \omega - \int_{\gamma_2} \omega,$$

isto é, a integral de  $\omega$  depende apenas das extremidades da curva.

Finalmente, mostramos que (ii)  $\implies$  (i). Fixamos  $x_0 \in \Omega$  qualquer. Sendo  $\Omega$  aberto e conexo, é conexo por caminhos. Dado  $x \in \Omega$ , considera  $\gamma_x$  um caminho que une  $x_0$  a  $x$  e definimos

$$f(x) := \int_{\gamma_x} \omega.$$

Por (ii),  $f$  está bem definida. Se  $v \in \mathbb{R}^n$  é pequeno suficiente, então o segmento de reta que une  $x$  até  $x + v$  está todo contido em  $\Omega$ . Parametrizando este segmento por  $\gamma(t) = x + tv$ ,  $t \in [0, 1]$ , podemos escrever

$$f(x+v) - f(x) - \omega(x) \cdot v = \int_{\gamma_{x+v}} \omega - \int_{\gamma_x} \omega - \omega(x) \cdot v = \int_0^1 [\omega(x+tv) - \omega(x)] \cdot v dt.$$

Agora, pela continuidade de  $\omega$ , dado  $\varepsilon > 0$ , existe  $\delta > 0$  tal que

$$|h| < \delta \implies \|\omega(x+h) - \omega(x)\| < \varepsilon.$$

Portanto,  $|v| < \delta$  implica, como queríamos, que

$$|f(x+v) - f(x) - \omega(x) \cdot v| \leq \varepsilon|v|. \quad \square$$

**Exemplo 85.** Considere em  $\mathbb{R}^2 \setminus \{(0,0)\}$  a 1-forma

$$\omega(x,y) := \frac{-y}{x^2+y^2} dx + \frac{x}{x^2+y^2} dy.$$

Se  $\gamma : [0, 2\pi] \rightarrow \mathbb{R}^2$  é o círculo de raio um centrado na origem  $\gamma(t) = (\cos t, \sin t)$ , então

$$\begin{aligned} \int_{\gamma} \omega &= \int_0^{2\pi} \omega(\gamma(t)) \cdot \gamma'(t) dt \\ &= \int_0^{2\pi} \frac{-\sin t}{\cos^2 t + \sin^2 t} (-\sin t) + \frac{\cos t}{\cos^2 t + \sin^2 t} \cos t dt \\ &= \int_0^{2\pi} dt = 2\pi. \end{aligned}$$

Utilizando a proposição anterior, item (iii), concluímos que  $\omega$  não é exata. Esta forma é conhecida como “*elemento de ângulo*”. O nome vem da propriedade que a integral da forma  $\omega$  como acima pode ser interpretada como  $2\pi$  vezes o número (líquido) de voltas que uma curva fechada  $\gamma$  faz em torno da origem no sentido anti-horário.

Mais precisamente, dada uma curva  $\gamma : [a, b] \rightarrow \mathbb{R}^2 \setminus \{(0,0)\}$  qualquer (não necessariamente fechada), escrevemos em coordenadas polares

$$\gamma(t) = (x(t), y(t)) = (r(t) \cos \theta(t), r(t) \sin \theta(t)).$$

Então, para calcular a intergal, calculamos

$$\begin{aligned}x' &= r' \cos \theta - r\theta' \sin \theta \\y' &= r' \sin \theta + r\theta' \cos \theta\end{aligned}$$

que implica em

$$-yx' + xy' = -rr' \cos \theta \sin \theta + r^2\theta' \sin^2 \theta + rr' \cos \theta \sin \theta + r^2\theta' \cos^2 \theta = r^2\theta'.$$

Logo,

$$\int_{\gamma} \omega = \int_a^b \frac{-y(t)x'(t) + x(t)y'(t)}{x(t)^2 + y(t)^2} dt = \int_a^b \theta'(t) dt = \theta(b) - \theta(a).$$

Observamos ainda que a variação do ângulo  $\theta$  é positiva quando o movimento é no sentido anti-horário. Além disso, no caso de a curva  $\gamma$  ser fechada, tem-se que  $\theta(b) - \theta(a)$  é um múltiplo de  $2\pi$  e

$$\frac{1}{2\pi} \int_{\gamma} \omega$$

é um número inteiro conhecido como “*número de voltas*” (do inglês, “winding number” poderia ser algo como “número de enrolamentos” ou “número de revoluções”).<sup>20</sup>

Além do interesse acima, a forma desse exemplo serve como contra-exemplo em várias situações, como veremos.  $\triangleleft$

## 17 Produto exterior – motivação

O produto vetorial de dois vetores  $u$  e  $v$  do espaço tridimensional  $\mathbb{R}^3$  pode ser definido como um vetor, geralmente denotado por  $u \times v$ , com as duas principais propriedades geométricas:

- (i)  $u \times v \in \mathbb{R}^3$  é ortogonal a ambos  $u$  e  $v$  e;
- (ii) a magnitude  $\|u \times v\|$  de  $u \times v$  é igual à área do paralelogramo gerado pelos vetores  $u$  e  $v$  no plano  $\text{Span}\{u, v\} \subset \mathbb{R}^3$ .

Esta definição não está muito bem formulada, pois existem dois vetores em  $\mathbb{R}^3$  com estas propriedades e, portanto, alguma escolha adicional, menos natural, deve ser feita. Esta escolha é a famosa “regra da mão direita”, que diz para considerarmos o sentido que torne  $\{u, v, u \times v\}$  uma base positiva de  $\mathbb{R}^3$  (isto é, com a mesma orientação da base canônica). Além do mais, esta definição é específica de dimensão 3.

Queremos definir um produto entre dois vetores de um espaço vetorial  $V$  de dimensão finita qualquer. Vamos pedir que este produto tenha propriedades “razoáveis” e mantenha a seguinte propriedade característica do produto vetorial:

$$u \times v = -v \times u.$$

Neste contexto, é mais comum denotar o produto com símbolo  $\wedge$ . Definimos a **álgebra exterior** como a álgebra  $\Lambda V$  gerada pelos elementos de  $V$  com o produto abstrato  $\wedge$  e a relação

$$u \wedge v = -v \wedge u.$$

<sup>20</sup>O leitor familiarizado com variável complexa pode ter reconhecido que a forma elemento de ângulo acima desempenha papel semelhante ao que a forma  $dz/z$  desempenha em  $\mathbb{C} \setminus \{0\}$ . Isto se deve ao seguinte:

$$\frac{dz}{z} = \frac{dx + i dy}{x + iy} = \frac{x dx + y dy}{x^2 + y^2} + i \frac{-y dx + x dy}{x^2 + y^2} = d \left[ \log \sqrt{x^2 + y^2} \right] + i \frac{-y dx + x dy}{x^2 + y^2}.$$

Como a parte real é uma forma exata, tem-se, pela proposição acima, que

$$\frac{1}{2\pi i} \int_{\gamma} \frac{dz}{z} = \frac{1}{2\pi} \int_{\gamma} \omega = \text{número de voltas que } \gamma \text{ percorre em torno da origem.}$$

um pouco mais precisamente,  $\Lambda V$  consiste de todas as constantes, todos os vetores de  $V$ , todos os produtos da forma  $u \wedge v$  e combinações lineares destes, todos os produtos da forma  $u \wedge v \wedge w$  e combinações lineares destes, e assim por diante.

Como um exemplo importante, tomamos  $(\mathbb{R}^3)^*$ , podemos formar o conjunto  $\Lambda(\mathbb{R}^3)^*$ , cujos elementos são:

- Todas as constantes de  $\mathbb{R} =: \Lambda^0(\mathbb{R}^3)^*$ ;
- Todos os vetores de  $(\mathbb{R}^3)^* =: \Lambda^1(\mathbb{R}^3)^*$ , que são gerados por  $dx, dy, dz$ ;
- Elementos de “grau dois” elementares ou 2-vetores elementares: utilizando a anti-simetria e (a propriedade que segue da anti-simetria)  $u \wedge u = 0$ , obtemos elementos de  $\Lambda(\mathbb{R}^3)^*$  da forma

$$\begin{aligned} v \wedge w &= (v_x dx + v_y dy + v_z dz) \wedge (w_x dx + w_y dy + w_z dz) \\ &= (v_y w_z - v_z w_y) dy \wedge dz + (v_z w_x - v_x w_z) dz \wedge dx + (v_x w_y - v_y w_x) dz \wedge dy \\ &= \det \begin{bmatrix} v_y & v_z \\ w_y & w_z \end{bmatrix} dy \wedge dz - \det \begin{bmatrix} v_x & v_z \\ w_x & w_z \end{bmatrix} dz \wedge dx + \det \begin{bmatrix} v_x & v_y \\ w_x & w_y \end{bmatrix} dy \wedge dz; \end{aligned} \quad (43)$$

- Quaisquer combinações lineares de elementos de grau 2 como acima, formando um espaço vetorial denotado por  $\Lambda^2(\mathbb{R}^3)^*$ ;
- Elementos de “grau três” elementares ou 3-vetores elementares e quaisquer combinações lineares deles:

$$\begin{aligned} u \wedge v \wedge w &= (u_x dx + u_y dy + u_z dz) \wedge (v \wedge w) \\ &= (v_y w_z - v_z w_y) dy \wedge dz + (v_z w_x - v_x w_z) dz \wedge dx + (v_x w_y - v_y w_x) dz \wedge dy \\ &= \left\{ u_x \det \begin{bmatrix} v_y & v_z \\ w_y & w_z \end{bmatrix} - u_y \det \begin{bmatrix} v_x & v_z \\ w_x & w_z \end{bmatrix} + u_z \det \begin{bmatrix} v_x & v_y \\ w_x & w_y \end{bmatrix} \right\} dx \wedge dy \wedge dz \\ &= \det \begin{bmatrix} u_x & u_y & u_z \\ v_x & v_y & v_z \\ w_x & w_y & w_z \end{bmatrix} dx \wedge dy \wedge dz; \end{aligned} \quad (44)$$

Estes elementos formam um espaço vetorial denotado por  $\Lambda^3(\mathbb{R}^3)^*$ ;

- Estes são todos, pois elementos de grau 4 ou mais num espaço tridimensional são nulos; de fato,  $u \wedge v \wedge w \wedge z$  em termos da base canônica  $dx, dy, dz$  sempre vai ter um elemento “repetido” do tipo  $dx \wedge dy \wedge dx \wedge dz = 0$ .<sup>21</sup> Logo, podemos escrever que  $\Lambda^k(\mathbb{R}^3)^* = \{0\}$ , para todo  $k > 3$ .

A álgebra exterior  $\Lambda(\mathbb{R}^3)^*$  é escrita então como a soma direta

$$\Lambda(\mathbb{R}^3)^* = \Lambda^0(\mathbb{R}^3)^* \oplus \Lambda^1(\mathbb{R}^3)^* \oplus \Lambda^2(\mathbb{R}^3)^* \oplus \Lambda^3(\mathbb{R}^3)^*$$

e um elemento arbitrário deste espaço é da forma

$$\begin{aligned} a + \sum_{i=1}^3 b_i dx^i + \sum_{i,j=1}^3 c_i dx^i \wedge dx^j + d dx \wedge dy \wedge dz \quad \text{ou, se preferir,} \\ \left( a, \sum_{i=1}^3 b_i dx^i, \sum_{i,j=1}^3 c_i dx^i \wedge dx^j, d dx \wedge dy \wedge dz \right). \end{aligned}$$

Por economia de notação, escrevemos acima  $dx^1 = dx$ ,  $dx^2 = dy$  e  $dx^3 = dz$ .

<sup>21</sup>Cuidado! Não é sempre verdade que  $v \wedge v = 0$ . Em 4 dimensões,  $v = dx \wedge dy + dz \wedge dw$  satisfaz  $v \wedge v = 2dx \wedge dy \wedge dz \wedge dw$ .



## 18 Produto tensorial

Apesar de nossas notas estarem em um nível (um pouco) mais elementar, o texto de Keith Conrad, disponível [online](#)<sup>22</sup>, foi bastante inspirador para a discussão que segue sobre o produto tensorial e o produto exterior. Outras referências, mais clássicas, incluem [15, 18]. A descrição destes tópicos em [1] também foi inspiradora.

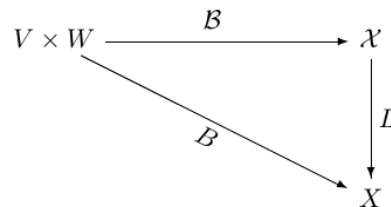
Dados dois espaços vetoriais  $V$  e  $W$  de dimensão finita, podemos pensar que o espaço tensorial  $V \otimes W$  é o espaço vetorial

$$V \otimes W = \text{Span}\{v_i \otimes w_j\}$$

onde  $\{v_i \otimes w_j\}$  é, por definição, um conjunto linearmente independente,  $\{v_i\}$  é uma base de  $V$  e  $\{w_j\}$  é uma base de  $W$ . Além disso,  $u \otimes v$  deve ser linear tanto em  $u$  quanto em  $v$ . Isto não é muito interessante nem no sentido matemático nem no sentido físico, pois precisamos fixar sistemas de coordenadas – bases para os espaços  $V$  e  $W$  – e utilizá-las explicitamente na definição. Logo, não é claro se o espaço é independente da escolha desses sistemas de coordenadas.

De forma mais abstrata, o espaço tensorial pode ser definido como um espaço com uma propriedade universal: “transformar aplicações bilineares em aplicações lineares”. Enunciamos o teorema que prova a existência do espaço tensorial e dá um sentido mais preciso para essa frase. Apresentamos em seguida uma prova que funciona mesmo em dimensão infinita.

**Teorema 86.** *Sejam  $V$  e  $W$  espaços vetoriais. Então, a menos de um isomorfismo, existe um único espaço vetorial  $\mathcal{X}$  e uma aplicação bilinear  $\mathcal{B} : V \times W \rightarrow \mathcal{X}$  que satisfazem: dada qualquer aplicação bilinear  $B : V \times W \rightarrow P$  (com valores em um espaço vetorial qualquer  $P$ ), existe (exatamente) uma aplicação linear  $L : \mathcal{X} \rightarrow P$  tal que  $B = L \circ \mathcal{B}$ .*



Antes da prova, passamos a alguns comentários sobre o teorema acima. Notamos inicialmente que não há referência alguma para bases de  $V$  e  $W$ , de modo que o produto tensorial é independente do sistema de coordenadas utilizado.

O par  $(\mathcal{X}, \mathcal{B})$  é o espaço tensorial que queríamos definir e sua propriedade universal é a propriedade enunciada no teorema acima. A notação utilizada é

$$\mathcal{X} = V \otimes W \quad \text{e} \quad \mathcal{B} = \otimes.$$

Os elementos de  $V \otimes W$  são chamados de *tensores*. Podemos “construir” tensores de  $V \otimes W$  a partir de elementos  $v \in V$  e  $w \in W$  olhando para a imagem pela aplicação bilinear  $\otimes$ . Assim,

$$\otimes(v, w) = v \otimes w \in V \otimes W,$$

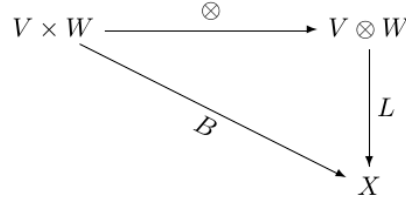
onde a igualdade acima representa apenas uma notação, que é a usual. Elementos da forma  $v \otimes w$  são chamados *tensores elementares*. Notamos, no entanto, que nem todo tensor é elementar, *e.g.*,  $e_1 \otimes e_1 + e_2 \otimes e_2 \in \mathbb{R}^2 \otimes \mathbb{R}^2$  não é um tensor elementar.

Na notação acima, a propriedade universal pode ser reescrita como: se  $B$  é uma aplicação bilinear com domínio  $V \times W$ , então existe uma única aplicação linear  $L$  com domínio  $V \otimes W$  tal que

$$B(v, w) = L(v \otimes w).$$

Isto é o que se quer dizer quando se fala: “o produto tensorial tem a propriedade de transformar aplicações bilineares em aplicações lineares”.

<sup>22</sup><http://www.math.uconn.edu/~kconrad/blurbs/linmultialg/tensorprod.pdf>



Propositalmente, não explicitamos o corpo dos escalares utilizados, pois toda a construção acima vale para  $\mathcal{K}$ -espaços vetoriais gerais. Na verdade, a construção acima funciona para módulos sobre um anel  $\mathcal{R}$ . A única diferença é que para módulos, podemos não ter base, mas vale a seguinte versão do Teorema 87 que aparecerá adiante: se  $V = \text{Span}\{v_i\}$  e  $W = \text{Span}\{w_j\}$ , então  $V \otimes W = \text{Span}\{v_i \otimes w_j\}$ .

Passamos finalmente à prova da existência do espaço tensorial.

*Demonstração do Teorema 86.* Separamos a prova do teorema em alguns passos.

*Passo 1. Construção de  $\mathcal{X}$  e de  $\mathcal{B}$ .* Começamos com o produto Cartesiano usual  $V \times W$  e construímos o espaço vetorial livre  $\mathcal{F}(V \times W)$ . Mais explicitamente,  $\mathcal{F}(V \times W)$  é o espaço de todas as combinações lineares formais finitas de elementos de  $V \times W$ . Os elementos de  $\mathcal{F}(V \times W)$  são gerados pelos elementos de  $V \times W$  desconsiderando sua estrutura de espaço vetorial; apenas como um conjunto de objetos. Para fixar ideias, poderíamos imaginar os elementos  $(v, w) \in \mathcal{F}(V \times W)$  como elementos abstratos da forma  $\delta_{(v,w)}$ . Assim, por definição, todo elemento de  $\mathcal{F}(V \times W)$  é uma combinação linear finita de elementos do tipo  $\delta_{(v,w)}$ . Além disso, tem-se independência linear

$$a_1 \delta_{(v_1, w_1)} + a_2 \delta_{(v_2, w_2)} + \cdots + a_k \delta_{(v_k, w_k)} = 0 \implies a_1 = a_2 = \cdots = a_k = 0.$$

Note que o espaço vetorial  $\mathcal{F}(V \times W)$  é gigantesco! Sua base tem a mesma cardinalidade de  $V \times W$ !

Agora, a ideia é definir  $\mathcal{X} = V \otimes W$  como o espaço mais geral possível que tenha relações de bilinearidade. Mais geral aqui quer dizer “com menos restrições”. Para este fim, consideramos o subespaço vetorial  $\mathfrak{R}$  de  $\mathcal{F}(V \times W)$  gerado pelos seguintes elementos:

- (i)  $(v_1 + v_2, w) - (v_1, w) - (v_2, w)$  para  $v_1, v_2 \in V$  e  $w \in W$ ;
- (ii)  $(cv, w) - c(v, w)$  para  $c \in \mathbb{R}$  (ou em um corpo qualquer ou em um anel),  $v \in V$  e  $w \in W$ ;
- (iii)  $(v, w_1 + w_2) - (v, w_1) - (v, w_2)$  para  $v \in V$  e  $w_1, w_2 \in W$ ;
- (iv)  $(v, cw) - c(v, w)$  para  $c \in \mathbb{R}$  (ou em um corpo qualquer ou em um anel),  $v \in V$  e  $w \in W$ .

Nossa motivação é que os elementos do espaço quociente satisfaçam a bilinearidade desejada. O espaço  $\mathcal{X}$  é então definido como o quociente

$$\mathcal{X} := \mathcal{F}(V \times W) / \mathfrak{R},$$

que tem naturalmente uma estrutura de espaço vetorial (ver, por exemplo, [2, página 8]). Denotamos as classes de equivalência  $(v, w) + \mathfrak{R}$  por  $\mathcal{B}(v, w)$ . Observamos, no entanto, que estas não são todas as classes possíveis. A bilinearidade da transformação  $(v, w) \mapsto \mathcal{B}(v, w)$  é consequência direta das relações que definem as classes de equivalência. Por exemplo, a aditividade na primeira coordenada pode ser justificada como

$$\mathcal{B}(v_1 + v_2, w) = \mathcal{B}(v_1, w) + \mathcal{B}(v_2, w) \text{ pois } (v_1 + v_2, w) - (v_1, w) - (v_2, w) \in \mathfrak{R}.$$

*Passo 2. Propriedade universal.* Mostramos que vale a propriedade universal enunciada. Seja  $P$  um espaço vetorial arbitrário e  $B : V \times W \rightarrow P$  uma aplicação bilinear qualquer. Definimos uma aplicação linear

$$\ell : \mathcal{F}(V \times W) \rightarrow P$$

nos elementos da base de  $\mathcal{F}(V \times W)$  como

$$\ell[(v, w)] := B(v, w).$$

Para mostrar que  $\ell$  passa a uma aplicação (quociente) em  $\mathcal{X}$ , devemos mostrar que  $\ell$  se anula em  $\mathfrak{N}$ , isto é, que  $\mathfrak{N} \subseteq \ker \ell$ . Façamos uma das propriedades: considere o elemento

$$(v_1 + v_2, w) - (v_1, w) - (v_2, w) \in \mathfrak{N}.$$

Sendo  $\ell$  linear e  $B$  bilinear

$$\begin{aligned} \ell[(v_1 + v_2, w) - ((v_1, w) + (v_2, w))] &= \ell[(v_1 + v_2, w)] - \ell[(v_1, w)] - \ell[(v_2, w)] \\ &= B(v_1 + v_2, w) - B(v_1, w) - B(v_2, w) \\ &= 0. \end{aligned}$$

Portanto, a aplicação  $L : \mathcal{X} \rightarrow P$  dada por  $L((v, w) + \mathfrak{N}) := \ell[(v, w)]$  está bem definida e é linear. Modificando a notação, isto significa que

$$L(\mathcal{B}(v, w)) = \ell[(v, w)] = B(v, w),$$

como queríamos.

**Passo 3. Unicidade a menos de isomorfismo.** Suponhamos que existam também  $\mathcal{X}'$  e  $\mathcal{B}' : V \times W \rightarrow \mathcal{X}'$  satisfazendo as hipóteses do teorema. Aplicamos duas vezes a propriedade universal (uma vez para cada espaço):

1. Sendo  $\mathcal{B}' : V \times W \rightarrow \mathcal{X}'$  bilinear, existe  $L : \mathcal{X} \rightarrow \mathcal{X}'$  linear tal que  $L(\mathcal{B}(v, w)) = \mathcal{B}'(v, w)$ ;
2. Sendo  $\mathcal{B} : V \times W \rightarrow \mathcal{X}$  bilinear, existe  $L' : \mathcal{X}' \rightarrow \mathcal{X}$  linear tal que  $L'(\mathcal{B}'(v, w)) = \mathcal{B}(v, w)$ .

Os elementos da forma  $\mathcal{B}(v, w)$  geram  $\mathcal{X}$ . Logo, de

$$L'[L(\mathcal{B}(v, w))] = L'(\mathcal{B}'(v, w)) = \mathcal{B}(v, w)$$

concluimos que  $L' \circ L = \text{Id}_{\mathcal{X}}$ . Analogamente,  $L \circ L' = \text{Id}_{\mathcal{X}'}$ . Portanto,  $L$  é um isomorfismo de espaços vetoriais com  $L^{-1} = L'$ .  $\square$

Analisando a prova do Teorema 1, vemos que todo elemento de  $V \otimes W$  pode ser escrito como uma soma de tensores elementares da forma  $v \otimes w = \mathcal{B}(v, w)$ . Assim, o nosso segundo parágrafo pode ser justificado como segue:

**Teorema 87.** Se  $\{v_i\}$  e  $\{w_j\}$  são bases de  $V$  e  $W$ , respectivamente, então  $\{v_i \otimes w_j\}$  é uma base de  $V \otimes W$ . Em particular, se  $V$  e  $W$  são de dimensão finita,  $n$  e  $m$ , respectivamente, então  $V \otimes W$  tem dimensão  $n \cdot m$ .

*Demonstração.* Pelo que observamos logo acima, basta verificar que o conjunto  $\text{Span}\{v_i \otimes w_j\}$  gera os tensores elementares. De fato, por bilinearidade,

$$v \otimes w = \left( \sum_{i=1}^d x^i v_i \right) \otimes \left( \sum_{j=1}^m y^j w_j \right) = \sum_{i=1}^d \sum_{j=1}^m x^i y^j v_i \otimes w_j. \quad (45) \quad \square$$

**Observação 88.** A Equação (45) é utilizada quando se pensa no produto tensorial de dois vetores  $v = [x^i] \in \mathbb{R}^n$  e  $w = [y^j] \in \mathbb{R}^m$  em termos mais elementares e se escreve

$$v \otimes w = [x^i y^j]_{m \times n} = \begin{bmatrix} x^1 y^1 & x^1 y^2 & \cdots & x^1 y^n \\ x^2 y^1 & x^2 y^2 & \cdots & x^2 y^n \\ \vdots & \vdots & \ddots & \vdots \\ x^m y^1 & x^m y^2 & \cdots & x^m y^n \end{bmatrix}.$$

Esta descrição não é um isomorfismo canônico entre  $\mathbb{R}^n \otimes \mathbb{R}^m$  e matrizes de ordem  $m \times n$ . Olhar, no entanto, a identificação 2 abaixo,

A seguir, mostramos como vários espaços “já conhecidos” podem ser identificados como espaços tensoriais. Além disso, as identificações abaixo devem tornar o trabalho com tensores menos “abstrato” e mais “palpável” do que a descrição rigorosa acima.

Nos restringimos a identificações naturais, isto é, isomorfismos canônicos entre espaços vetoriais, isto é (de novo!), isomorfismos que não dependam de escolha de bases ou de qualquer outra estrutura dos espaços, como produtos internos.

Por simplicidade, vamos também considerar espaços vetoriais de dimensão finita.

**1.**  $\mathbb{R} \otimes V \simeq V$

Suponhamos que  $V$  é um espaço vetorial real. Intuitivamente, tem-se  $\mathbb{R} \otimes V = \text{Span}\{1 \otimes v_i\}$  e logo podemos escrever qualquer elemento  $x \in \mathbb{R} \otimes V$  como

$$x = \sum_{i=1}^d x^i (1 \otimes v_i) = \sum_{i=1}^d x^i \otimes v_i.$$

No entanto, sem recorrer a coordenadas, é fácil ver que  $r \otimes v \mapsto rv$  é um isomorfismo (canônico, pois não faz referência à bases).

**2.**  $V \otimes W^* \simeq \mathcal{L}(W; V)$

Dado um tensor elementar  $v \otimes \varphi \in V \otimes W^*$ , associamos uma transformação linear  $v \otimes \varphi : W \rightarrow V$ :

$$(v \otimes \varphi)(w) := \varphi(w)v.$$

Vamos mostrar que qualquer  $T \in \mathcal{L}(W; V)$  pode ser obtida por combinações destas elementares: de fato, sejam  $\{v_i\}$  e  $\{w_j\}$  bases de  $V$  e  $W$ , respectivamente. Se  $\{dw^j\}$  é base de  $W^*$ , dual a  $\{w_j\}$ , temos

$$T(x) = T\left(\sum_{j=1}^d x^j w_j\right) = \sum_{j=1}^d x^j T(w_j) = \sum_{j=1}^d \sum_{i=1}^m (x^j T_j^i) v_i = \sum_{i,j} T_j^i dw^j(x) v_i = \sum_{i,j} T_j^i (v_i \otimes dw^j)(x).$$

**Observação 89.** É claro que poderíamos ter escrito  $W^* \otimes V$  acima, já que sempre  $U \otimes V$  é canonicamente isomorfo a  $V \otimes U$ . Esta última afirmação segue diretamente do Teorema 86.

**3.**  $V^* \otimes W^* \simeq \mathcal{B}(V \times W; \mathbb{R})$

Acima, denotamos por  $\mathcal{B}(V \times W; \mathbb{R})$  o espaço das transformações bilineares  $B : V \times W \rightarrow \mathbb{R}$ . Definimos

$$\varphi \otimes \psi \in V^* \otimes W^* \xrightarrow{\Phi} (\varphi \otimes \psi)(v, w) := \varphi(v)\psi(w),$$

que se estende, por linearidade, a uma transformação linear  $\Phi : V^* \otimes W^* \rightarrow \mathcal{B}(V \times W; \mathbb{R})$ .

Agora, fixadas bases para  $V$  e  $W$  (e suas respectivas bases duais), dada  $B : V \times W \rightarrow \mathbb{R}$ , vamos ver que está na imagem de  $\Phi$ : denotando  $B_{ij} = B(v_i, w_j)$ ,

$$B(x, y) = \sum_{j=1}^d \sum_{i=1}^m x^i y^j B_{ij} = \sum_{i,j} B_{ij} dv^i(x) dw^j(y) = \left( \sum_{i,j} B_{ij} \Phi(dv^i \otimes dw^j) \right)(x, y),$$

de modo que

$$B = \Phi \left( \sum_{i,j} B_{ij} dv^i \otimes dw^j \right).$$

**Observação 90.** Este ponto de vista é interessante se quisermos pensar em tensores como aplicações multilineares: elementos de  $V^* \otimes W^*$  se identificam com aplicações bilineares da forma

$$B : V \times W \rightarrow \mathbb{R}.$$

Em outras palavras, tensores covariantes de segunda ordem podem ser pensadas como aplicações bilineares do tipo acima.

Todas as construções acima podem ser feitas para um número maior de espaços vetoriais  $V_1, V_2, \dots, V_r$ . Nossa propriedade universal passa a ser “transformar aplicações  $r$ -lineares em transformações lineares”. Além disso, temos um isomorfismo canônico

$$V_1^* \otimes V_2^* \otimes \dots \otimes V_r^* \simeq \mathcal{B}(V, V, \dots, V; \mathbb{R})$$

dado por

$$(f_1 \otimes f_2 \otimes \dots \otimes f_r)(v_1, v_2, \dots, v_r) = f_1(v_1)f_2(v_2) \dots f_r(v_r).$$

Desta forma, podemos pensar no tensor elementar  $f_1 \otimes f_2 \otimes \dots \otimes f_r$  como um objeto abstrato com relações de linearidade em cada componente ou como uma forma  $r$ -linear (multilinear).

No caso especial em que  $V_i = V$  para todo  $i$ , denotamos

$$\mathcal{T}^r V = \otimes^r V := V \otimes V \otimes \dots \otimes V.$$

Denotando o espaço vetorial das formas  $r$ -lineares por  $\mathcal{L}_r(V)$ , isto é,

$$\mathcal{L}_r(V) := \mathcal{B}(V, V, \dots, V; \mathbb{R}),$$

temos da discussão acima um isomorfismo canônico

$$\mathcal{T}^r(V^*) \simeq \mathcal{L}_r(V).$$

Lembramos as propriedades básicas das formas  $r$ -lineares e deixamos sua demonstração como exercício para o leitor.

**Exercício 38.** Seja  $\{e_1, e_2, \dots, e_n\}$  uma base de  $V$ . Então

- (i) Mostre que  $\dim \mathcal{L}_r(V) = n^r$  e exiba uma base de  $\mathcal{L}_r(V)$ .
- (ii) Mostre que uma aplicação  $r$ -linear é completamente determinada pelos seus valores nos elementos da forma  $(e_{i_1}, e_{i_2}, \dots, e_{i_r})$ , para  $i_j \in \{1, 2, \dots, n\}$ .

## 18.1 Álgebra tensorial

Lembramos alguma terminologia algébrica. Uma **álgebra sobre um corpo**  $\mathbb{K}$  é um conjunto  $X$  que é um  $\mathbb{K}$ -espaço vetorial e que, adicionalmente, possui uma operação (bilinear) de produto entre elementos de  $X$ . Além disso, quando as operações são associativas relativas também ao produto, dizemos que  $X$  é uma álgebra associativa. Mais explicitamente, uma **álgebra associativa** sobre um corpo  $\mathbb{K}$  é um conjunto  $X$  com operações  $+$  e  $\times$  de modo que

1)  $(X, +, \times)$  é um anel. Em outras palavras,

- (i)  $(X, +)$  é um grupo abeliano: para todo  $x, y, z \in X$ , vale
  - Associatividade:  $(x + y) + z = x + (y + z)$ ;
  - Identidade aditiva: existe  $0 \in X$  tal que  $x + 0 = 0 + x = x$ ;
  - Inverso: existe  $y \in X$  tal que  $x + y = 0$ ;
  - Comutatividade:  $x + y = y + x$ .

- (ii)  $(X, \times)$  é um semigrupo com identidade: para todo  $x, y, z \in X$ , vale
  - Associatividade:  $(x \times y) \times z = x \times (y \times z)$ ;
  - Identidade multiplicativa: existe  $1 \in X$  tal que  $x \times 1 = 1 \times x = x$ ;

(iii) Distributividade da multiplicação com respeito a adição:

$$x \times (y + z) = x \times y + x \times z \quad \text{e} \quad (x + y) \times z = x \times z + y \times z.$$

2)  $(X, +)$  é um  $\mathbb{K}$ -espaço vetorial. Em outras palavras,

- (i)  $(X, +)$  é um grupo abeliano, como acima;
- (ii) Propriedades da multiplicação por escalar:
  - Compatibilidade com produtos de escalares:  $a(bx) = (ab)x$ ;
  - Identidade escalar:  $1x = x$ ;
  - Distributividade da adição em  $X$ :  $a(x + y) = ax + ay$ ;
  - Distributividade da adição em  $\mathbb{K}$ :  $(a + b)x = ax + bx$ .

3) Associatividade do produto com o produto por escalar:

$$c(x \times y) = (cx) \times y = x \times (cy).$$

A soma direta dos espaços tensoriais definidos na seção anterior formam uma álgebra associativa. Denotamos

$$\mathcal{T}^0V := \mathbb{R} \quad \text{e} \quad \mathcal{T}^1V := V$$

e denotamos

$$\mathcal{T}V = \bigoplus_{k=1}^{+\infty} \mathcal{T}^kV := \mathbb{R} \oplus V \oplus \mathcal{T}^1V \oplus \mathcal{T}^2V \oplus \cdots \oplus \mathcal{T}^kV \oplus \cdots.$$

O produto de elementos de ordens diferentes  $x \in \mathcal{T}^rV$  e  $y \in \mathcal{T}^sV$  é dado por

$$x \otimes y \in \mathcal{T}^{r+s}V.$$

## 19 Produto exterior

Queremos definir o produto exterior  $\Lambda V$  como a álgebra associativa “mais geral” que satisfaça a propriedade de “anticomutatividade”. Nesta seção, vamos formalizar esta ideia.

Dado um espaço vetorial  $V$ , o produto exterior de  $k$  cópias de  $V$ , denotado por

$$\Lambda^kV = V \wedge V \wedge \cdots \wedge V,$$

é obtido como um quociente do produto tensorial  $\mathcal{T}^kV = V \otimes V \otimes \cdots \otimes V$  ao adicionar “relações de alternância”. Mais precisamente, o **produto exterior** de  $k$  cópias de  $V$  é definido por

$$\Lambda^kV := \mathcal{T}^kV / \mathcal{S}_k,$$

onde  $\mathcal{S}_k$  é o subespaço vetorial de  $\mathcal{T}^kV$  gerado pelos elementos da forma

$$v_1 \otimes v_2 \otimes \cdots \otimes v_k \quad \text{onde} \quad v_i = v_j \quad \text{para algum} \quad i \neq j.$$

A classe de equivalência de um tensor elementar é dita um **produto exterior elementar**, que é denotado por

$$v_1 \wedge v_2 \wedge \cdots \wedge v_k := v_1 \otimes v_2 \otimes \cdots \otimes v_k + \mathcal{S}_k.$$

A aplicação quociente  $\mathcal{T}^kV \rightarrow \Lambda^kV$  é sobrejetiva e todo tensor de ordem  $k$  é gerado como soma (finita) de tensores elementares. Daí segue que todo elemento de  $\Lambda^kV$  pode ser escrito como soma (finita) de produtos exteriores elementares: todo  $\omega \in \Lambda^kV$  pode ser escrito como

$$\omega = \sum_{i=1}^N v_1^i \wedge v_2^i \wedge \cdots \wedge v_k^i.$$

**Observação 91.** Notamos que a imagem de um tensor elementar  $v_1 \otimes v_2 \otimes \cdots \otimes v_k \in \mathcal{S}_k$  pela aplicação quociente é, trivialmente,

$$v_1 \wedge v_2 \wedge \cdots \wedge v_k = 0.$$

Observamos que, se  $k > n$ , então todos os tensores elementares pertencem a  $\mathcal{S}_k$  e, logo,  $\Lambda^kV = \{0\}$ .

Para  $V$  de dimensão finita igual a  $n$ , vamos construir uma base para  $\Lambda^k V$  a partir de uma base  $\{dx^1, dx^2, \dots, dx^n\}$  de  $V$ . Denotamos a base de  $V$  desta maneira porque, como já havíamos mencionado anteriormente, nosso maior interesse será quando  $V = (\mathbb{R}^n)^*$ ; de qualquer maneira, a discussão que segue vale para espaços de dimensão finita qualquer.

Observamos inicialmente que, tautologicamente,  $\Lambda^0 V = \mathbb{R}$  e  $\Lambda^1 V = V$ . Além disso, já vimos que  $\Lambda^k V = \{0\}$  quando  $k > n$ . Logo, basta estudarmos bases para os espaços  $\Lambda^k V$  cujo índice  $k \in \{2, 3, \dots, n\}$ .

Consideramos um produto exterior elementar da forma

$$v_1 \wedge v_2 \wedge \dots \wedge v_k.$$

Cada  $v_i \in V$  pode ser representado na base de  $V$  como

$$v_i = \sum_{j=1}^n a_{ij} dx^j.$$

Logo, por multilinearidade,

$$\begin{aligned} v_1 \wedge v_2 \wedge \dots \wedge v_k &= \left( \sum_{j_1=1}^n a_{1j_1} dx^{j_1} \right) \wedge \left( \sum_{j_2=1}^n a_{2j_2} dx^{j_2} \right) \wedge \dots \wedge \left( \sum_{j_k=1}^n a_{kj_k} dx^{j_k} \right) \\ &= \sum_{j_1=1}^n \sum_{j_2=1}^n \dots \sum_{j_k=1}^n a_{1j_1} a_{2j_2} \dots a_{kj_k} dx^{j_1} \wedge dx^{j_2} \wedge \dots \wedge dx^{j_k}. \end{aligned} \quad (46)$$

Vários dos produtos exteriores acima são “repetidos” ou iguais a “zero”, por causa da propriedade de anticomutatividade. De fato, quando dois índices  $j_i$  são iguais, obtemos produto exterior zero. Por outro lado, quando são todos diferentes, podemos fazer a seguinte análise: escrevemos

$$j_1 = \sigma(i_1), \quad j_2 = \sigma(i_2), \quad \dots, \quad j_k = \sigma(i_k),$$

onde  $\sigma : I \rightarrow I$  é a permutação do conjunto

$$I := \{i_1 < i_2 < \dots < i_k\} = \{j_1, j_2, \dots, j_k\}$$

que reordena os índices  $j$  em ordem crescente. Qualquer permutação  $\sigma$  pode ser escrita como produto (composição) de transposições

$$\sigma = \sigma_1 \sigma_2 \dots \sigma_{N(\sigma)}.$$

Lembramos que uma transposição é uma permutação que apenas troca dois elementos de lugar. Por exemplo,  $\sigma : \{1, 2, \dots, m\} \rightarrow \{1, 2, \dots, m\}$  dada por

$$\sigma(1, 2, \dots, i, \dots, j, \dots, m) = (1, 2, \dots, j, \dots, i, \dots, m)$$

é uma transposição. Qualquer permutação  $\sigma$  pode ser escrita como produto de transposições

$$\sigma = \sigma_1 \sigma_2 \dots \sigma_N.$$

É possível mostrar [2, Corolário 4.19 e Lema 1.15] que a “paridade de  $N$ ” é um invariante de  $\sigma$  (é a paridade que é invariante, não o número  $N$ , que nem único é). Em outras palavras, dadas duas representações de  $\sigma$  como composição de transposições, uma a partir de  $N_1$  transposições e outra a partir de  $N_2$ , tem-se

$$N_1 \equiv N_2 \pmod{2}.$$

Segue que está bem definido o “sinal” de  $\sigma$ :

$$\text{sgn}(\sigma) := (-1)^{N(\sigma)} = \begin{cases} 1, & \text{se } N(\sigma) \text{ é par} \\ -1, & \text{se } N(\sigma) \text{ é ímpar} \end{cases}$$

Com estas considerações em mente, para índices distintos fixados  $j_1, j_2, \dots, j_k$ , consideramos um reordenamento do tipo

$$I := \{j_1, j_2, \dots, j_k\} = \{i_1 < i_2 < \dots < i_k\}$$

e a permutação  $\sigma : I \rightarrow I$  dada por

$$\sigma(i_1) = j_1, \quad \sigma(i_2) = j_2, \quad \dots, \quad \sigma(i_k) = j_k.$$

Logo, podemos escrever

$$dx^{j_1} \wedge dx^{j_2} \wedge \dots \wedge dx^{j_k} = dx^{\sigma(i_1)} \wedge dx^{\sigma(i_2)} \wedge \dots \wedge dx^{\sigma(i_k)} = \text{sgn}(\sigma) dx^{i_1} \wedge dx^{i_2} \wedge \dots \wedge dx^{i_k}.$$

Portanto, denotando o conjunto de permutações de  $k$  elementos de  $I$  por  $S_k(I)$ , (46) pode ser reescrita como

$$\begin{aligned} v_1 \wedge v_2 \wedge \dots \wedge v_k &= \sum_{1 \leq i_1 < i_2 < \dots < i_k \leq n} \left( \sum_{\sigma \in S_k(I)} \text{sgn}(\sigma) a_{1\sigma(i_1)} a_{2\sigma(i_2)} \dots a_{k\sigma(i_k)} \right) dx^{i_1} \wedge dx^{i_2} \wedge \dots \wedge dx^{i_k} \\ &= \sum_{1 \leq i_1 < i_2 < \dots < i_k \leq n} \det \begin{bmatrix} a_{1i_1} & \dots & a_{1i_k} \\ \vdots & \ddots & \vdots \\ a_{ki_1} & \dots & a_{ki_k} \end{bmatrix} dx^{i_1} \wedge dx^{i_2} \wedge \dots \wedge dx^{i_k}. \end{aligned} \quad (47)$$

A Fórmula (47) está em pleno acordo com os cálculos informais que fizemos na Seção 17, ver Fórmulas (43) e (44). É também a essência da prova do seguinte resultado:

**Proposição 92.** *Seja  $V$  um espaço de dimensão finita e igual a  $n$  e  $\{dx^1, dx^2, \dots, dx^n\} \subset V$  uma base de  $V$ . Então,*

$$\mathfrak{B}_k := \{dx^{i_1} \wedge dx^{i_2} \wedge \dots \wedge dx^{i_k}; 1 \leq i_1 < i_2 < \dots < i_k \leq n\}$$

é uma base de  $\Lambda^k V$ . Em particular,

$$\dim \Lambda^k V = \binom{n}{k} = \frac{n!}{k!(n-k)!}.$$

*Demonstração.* Pela Fórmula (47), sabemos que o conjunto  $\mathfrak{B}_k$  definido acima gera  $\Lambda^k V$ . Ainda falta mostrar que  $\mathfrak{B}_k$  é linearmente independente.

Notamos que  $\mathfrak{B}_n$  consiste de apenas um elemento que é não nulo:  $dx^1 \wedge dx^2 \wedge \dots \wedge dx^n \neq 0$ , pois  $dx^1 \otimes dx^2 \otimes \dots \otimes dx^n \notin \mathcal{S}_n$ . Para  $2 \leq k < n$ , suponhamos que

$$\sum_{1 \leq i_1 < i_2 < \dots < i_k \leq n} c_{i_1 i_2 \dots i_k} dx^{i_1} \wedge dx^{i_2} \wedge \dots \wedge dx^{i_k} = 0. \quad (48)$$

Para mostrar que os coeficientes são todos nulos, vamos fazer o produto exterior “com os elementos que faltam” para chegar em  $dx^1 \wedge dx^2 \wedge \dots \wedge dx^n$ . Dados os índices  $I = \{1 \leq i_1 < i_2 < \dots < i_k \leq n\}$ , consideramos o conjunto de índices  $J = \{j_1, j_2, \dots, j_{n-k}\}$  tal que

$$I \cup J = \{1, 2, \dots, n\}.$$

Notamos ainda que todos os produtos exteriores elementares que aparecem na soma em (48) possuem algum índice que está em  $J$ , exceto aquele cujos índices são exatamente os de  $I$ . Logo,

$$0 = \left( \sum c_{i_1 i_2 \dots i_k} dx^{i_1} \wedge dx^{i_2} \wedge \dots \wedge dx^{i_k} \right) \wedge dx^{j_1} \wedge dx^{j_2} \wedge \dots \wedge dx^{j_{n-k}} = \pm c_{i_1 i_2 \dots i_k} dx^1 \wedge dx^2 \wedge \dots \wedge dx^n.$$

Segue, como queríamos, que  $c_{i_1 i_2 \dots i_k} = 0$  □



## 19.1 Álgebra exterior ou álgebra de Grassmann



falar sobre grassmann e motivação histórica

A **álgebra de Grassmann** ou **álgebra exterior** é definida como

$$\Lambda V := \mathbb{R} \oplus V \oplus \Lambda^2 V \oplus \Lambda^3 V \oplus \dots \oplus \Lambda^{n-1} V \oplus \Lambda^n V,$$

onde o produto  $\wedge : \Lambda^k V \times \Lambda^m V \rightarrow \Lambda^{k+m} V$  é uma operação bilinear, obtida dos produtos elementares por “concatenação”. Mais precisamente, temos

$$\omega = \sum_I a_I dx^I \in \Lambda^k V \quad \text{e} \quad \eta = \sum_J a_J dx^J \in \Lambda^m V \rightsquigarrow \omega \wedge \eta := \sum_{I,J} a_I b_J dx^I \wedge dx^J \in \Lambda^{k+m}.$$

A fórmula binomial usual

$$(x + y)^n = \sum_{k=0}^n \binom{n}{k} x^{n-k} y^k,$$

aplicada para  $x = y = 1$ , fornece a dimensão da álgebra de Grassmann:

$$\dim \Lambda V = \sum_{k=0}^n \binom{n}{k} = (1 + 1)^n = 2^n.$$

Observamos que o produto exterior de elementos de graus diferentes não é necessariamente anti-simétrico.

**Exercício 39.** Mostre que, se  $\omega \in \Lambda^k V$  e  $\eta \in \Lambda^m V$ , então

$$\omega \wedge \eta = (-1)^{km} \eta \wedge \omega.$$

**Exercício 40.** Descreva a álgebra de Grassmann  $\Lambda(\mathbb{R}^n)^*$ , no mesmo espírito da Seção 17 (agora tudo é rigoroso), para dimensões  $n$  igual a 1, 2, 3 e 4.

## 19.2 Identificação I; Aplicações bilineares alternadas

Uma aplicação  $\phi : V \times V \rightarrow \mathbb{R}$  é dita **bilinear alternada** quando

$$\phi(v, v) = 0 \quad \text{para todo} \quad v \in V. \tag{49}$$

Observamos que  $\phi$  é alternada se, e somente se, é **anti-simétrica**,<sup>23</sup>

$$\phi(v, w) = -\phi(w, v) \quad \text{para todo } v, w \in V. \quad (50)$$

Denotamos o espaço vetorial das formas bilineares alternadas, que é claramente um subespaço vetorial de  $\mathcal{B}(V \times V; \mathbb{R})$ , por  $\mathcal{A}_2(V)$ . Vamos mostrar, em analogia com a identificação  $V^* \otimes V^* \simeq \mathcal{B}(V \times V; \mathbb{R})$ , que existe um isomorfismo canônico

$$\Phi : \Lambda^2(V^*) \rightarrow \mathcal{A}_2(V).$$

Em um produto elementar  $f \wedge g \in \Lambda^2(V^*)$ , definimos uma aplicação bilinear alternada por

$$f \wedge g \mapsto (f \wedge g)(v, w) := \det \begin{bmatrix} f(v) & f(w) \\ g(v) & g(w) \end{bmatrix} \quad (51)$$

e estendemos, por linearidade, a uma aplicação linear  $\Phi : \Lambda^2(V^*) \rightarrow \mathcal{A}_2(V)$ . A definição acima é motivada, como veremos, pela fórmula (47) acima. Pelas propriedades básicas do determinante, a fórmula (51) realmente define uma forma bilinear alternada. Notamos que, fixada uma base de  $V$  e a base dual de  $V^*$ , temos

$$(\mathrm{d}x^{i_1} \wedge \mathrm{d}x^{i_2})(e_{j_1}, e_{j_2}) := \det \begin{bmatrix} \mathrm{d}x^{i_1}(e_{j_1}) & \mathrm{d}x^{i_1}(e_{j_2}) \\ \mathrm{d}x^{i_2}(e_{j_1}) & \mathrm{d}x^{i_2}(e_{j_2}) \end{bmatrix} = \begin{cases} 1 & \text{se } i_1 = j_1 \text{ e } i_2 = j_2 \\ 0 & \text{nos outros casos} \end{cases}$$

Logo, usando a bilinearidade, obtemos

$$(\mathrm{d}x^i \wedge \mathrm{d}x^j)(v, w) := \det \begin{bmatrix} \mathrm{d}x^i(v) & \mathrm{d}x^j(w) \\ \mathrm{d}x^i(w) & \mathrm{d}x^j(v) \end{bmatrix} = v^i w^j - v^j w^i.$$

Por outro lado, dada uma forma bilinear alternada  $\phi : V \times V \rightarrow \mathbb{R}$ , podemos escrever

$$\phi(v, w) = \sum_{i,j=1}^n v^i w^j \phi(e_i, e_j) = \sum_{i<j} \phi(e_i, e_j)(v^i w^j - v^j w^i) = \sum_{i<j} \phi(e_i, e_j)(\mathrm{d}x^i \wedge \mathrm{d}x^j)(v, w),$$

de modo que

$$\phi = \sum_{i<j} \phi(e_i, e_j) \mathrm{d}x^i \wedge \mathrm{d}x^j.$$

### 19.3 Identificação II; Aplicações multilineares alternadas

O que foi feito na seção anterior se generaliza facilmente para o espaço  $\Lambda^n(V^*)$ , fornecendo um isomorfismo

$$\Lambda^n(V^*) \simeq \mathcal{A}_n(V)$$

com as formas multilineares alternadas. A generalização é simples, mas entediante, pois precisamos trabalhar com mais índices.

Dizemos que uma forma multilinear  $\phi : V \times \dots \times V \rightarrow \mathbb{R}$  é

– **alternada** quando

$$\phi(v_1, v_2, \dots, v_r) = 0 \quad \text{sempre que } v_i = v_j \quad \text{para algum } i \neq j.$$

– **anti-simétrica** ou **skew-simétrica** quando muda de sinal ao trocarmos dois vetores de lugar:

$$\phi(v_1, v_2, \dots, v_i, \dots, v_j, \dots, v_r) = -\phi(v_1, v_2, \dots, v_j, \dots, v_i, \dots, v_r).$$

<sup>23</sup>As condições (49) e (50) não são equivalentes quando consideramos espaços vetoriais sobre corpos de característica dois. Por exemplo, em  $\mathbb{Z}_2$  tem-se  $[2] = [0]$ . No nosso caso, estaremos sempre trabalhando sobre  $\mathbb{R}$ , onde as equivalências valem.

O conjunto das formas multilineares alternadas (com  $n$  cópias de  $V$ ) é denotado por  $\mathcal{A}_n(V)$  e claramente é um subespaço vetorial do espaço das formas multilineares.

**Exercício 41.** Seja  $\phi : V \times \cdots \times V \rightarrow \mathbb{R}$  uma forma multilinear. Mostre que são equivalentes:

- (i)  $\phi$  é alternada;
- (ii)  $\phi$  é anti-simétrica;
- (iii) Para qualquer permutação de índices  $\sigma : \{1, 2, \dots, n\} \rightarrow \{1, 2, \dots, n\}$ , tem-se

$$\phi(v_{\sigma(1)}, v_{\sigma(2)}, \dots, v_{\sigma(n)}) = \text{sgn}(\sigma) \phi(v_1, v_2, \dots, v_n).$$

A identificação que procuramos é definida em produtos exteriores elementares como segue: dado  $f_1 \wedge f_2 \wedge \cdots \wedge f_k \in \Lambda^k(V^*)$ , definimos um elemento de  $\mathcal{A}_k(V)$  por

$$(f_1 \wedge f_2 \wedge \cdots \wedge f_k)(v_1, v_2, \dots, v_k) := \det [f_i(v_j)] = \det \begin{bmatrix} f_1(v_1) & f_1(v_2) & \cdots & f_1(v_k) \\ f_2(v_1) & f_2(v_2) & \cdots & f_2(v_k) \\ \vdots & \vdots & \ddots & \vdots \\ f_k(v_1) & f_k(v_2) & \cdots & f_k(v_k) \end{bmatrix}. \quad (52)$$

Por linearidade, obtemos uma aplicação  $\Phi : \Lambda^k(V^*) \rightarrow \mathcal{A}_k(V)$  que é um isomorfismo de espaços vetoriais.

**Exercício 42.** Mostre que  $\Phi$  é, de fato, um isomorfismo de espaços vetoriais.

## 20 Formas diferenciais de grau $p$ e a derivada exterior

Uma **forma diferencial de grau  $p$** , também conhecida como  **$p$ -forma** ou **forma exterior de grau  $p$** , é uma aplicação da forma

$$\omega : U \subseteq \mathbb{R}^n \longrightarrow \Lambda^p(\mathbb{R}^n)^* \simeq \mathcal{A}_p(\mathbb{R}^n).$$

Em outras palavras, uma forma diferencial de grau  $p$  é um campo de formas exteriores de grau  $p$ . Vamos denotar o conjunto das formas diferenciais de grau  $p$  por  $\Omega^p(U)$ . Para cada  $x \in U$  fixado, temos um elemento  $\omega(x) \in \Lambda^p(\mathbb{R}^n)^*$ , que pode ser escrito como

$$\omega(x) = \sum_I a_I(x) dx^I, \quad (53)$$

onde a soma é tomada sobre partições “crescentes”  $I = \{1 \leq i_1 < i_2 < \cdots < i_p \leq n\}$ . Lembre que estamos utilizando a notação

$$a_I(x) := a_{i_1 i_2 \dots i_p}(x) \quad \text{e} \quad dx^I := dx^{i_1} \wedge dx^{i_2} \wedge \cdots \wedge dx^{i_p}.$$

**Exercício 43.** Mostre que para uma  $p$ -forma  $\omega : U \rightarrow \Lambda^p(\mathbb{R}^n)^*$ , as afirmações são equivalentes:

- (i) a forma  $\omega$  é diferenciável (ou classe  $C^k$ );
- (ii) cada uma das componentes  $a_I$  é diferenciável (ou classe  $C^k$ );
- (iii) para  $v_1, v_2, \dots, v_p \in \mathbb{R}^n$ , a função  $x \mapsto \omega(x)[v_1, v_2, \dots, v_p]$  é diferenciável (ou classe  $C^k$ ).

Convencionamos que  $\Omega^0(U)$  é o conjunto das funções diferenciáveis (na maioria dos contextos, se supõe suaves) em  $U$ .

Podemos ainda considerar a álgebra

$$\Omega(U) := \Omega^0(U) \oplus \Omega^1(U) \oplus \Omega^2(U) \oplus \cdots \oplus \Omega^n(U),$$

com a noção de produto exterior que provém da álgebra exterior da subseção anterior.

A ideia agora é definir uma noção de diferenciabilidade de formas diferenciais. Dada uma forma diferencial  $\omega$  de grau  $p$ , como em (53), definimos a **derivada exterior** de  $\omega$  como  $d\omega \in \Omega^{p+1}(U)$  dada por:

$$\omega(x) = \sum_I a_I(x) dx^I \rightsquigarrow d\omega(x) := \sum_I da_I(x) \wedge dx^I = \sum_I \sum_{j=1}^n \frac{\partial a_I}{\partial x^j}(x) dx^j \wedge dx^I. \quad (54)$$

Vamos mostrar que esta noção de diferencial de formas satisfaz as propriedades desejadas.

**Teorema 93.** *A derivada exterior é o único conjunto de aplicações da forma*

$$d : \Omega^p(U) \longrightarrow \Omega^{p+1}(U)$$

que satisfaz as seguintes propriedades

- 1) para  $p = 0$ , a aplicação  $d : \Omega^0(U) \rightarrow \Omega^1(U)$  é o diferencial usual de uma função.
- 2) para todo  $\omega, \mu \in \Omega(U)$  e para todo  $c_1, c_2 \in \mathbb{R}$ , temos linearidade:

$$d(c_1\omega + c_2\mu) = c_1d\omega + c_2d\mu.$$

- 3) para  $\omega \in \Omega^p(U)$  e  $\mu \in \Omega(U)$ , temos

$$d(\omega \wedge \mu) = (d\omega) \wedge \mu + (-1)^p \omega \wedge (d\mu).$$

- 4) para todo  $\omega \in \Omega(U)$ , tem-se  $d(d\omega) = 0$ ; em outras palavras,  $d^2 := d \circ d = 0$ .

*Demonstração.* Inicialmente, mostramos que a definição que aparece em (54) satisfaz as propriedades acima. A primeira é a definição do diferencial. A segunda segue diretamente de o diferencial de funções ser linear. Utilizando a linearidade, é suficiente verificar o terceiro item para as formas “elementares”, do tipo:

$$\omega = a_I(x) dx^I \quad \text{e} \quad \mu = b_J dx^J.$$

Abaixo, utilizamos, respectivamente, a propriedade do produto exterior, a definição da derivada exterior, a regra do produto para derivadas parciais, distributividade e o reordenamento dos índices (trocamos  $dx^k$  de lugar sucessivamente com os elementos  $dx^{i_1}, dx^{i_2}, \dots, dx^{i_p}$  e por isso temos  $p$  trocas de sinal):

$$\begin{aligned} d(\omega \wedge \mu) &= d\left[(a_I dx^I) \wedge (b_J dx^J)\right] \\ &= d(a_I b_J dx^I \wedge dx^J) \\ &= d(a_I b_J) \wedge dx^I \wedge dx^J \\ &= \sum_{k=1}^n \left( \frac{\partial a_I}{\partial x^k} b_J + a_I \frac{\partial b_J}{\partial x^k} \right) dx^k \wedge dx^I \wedge dx^J \\ &= \sum_{k=1}^n \frac{\partial a_I}{\partial x^k} b_J dx^k \wedge dx^I \wedge dx^J + \sum_{k=1}^n a_I \frac{\partial b_J}{\partial x^k} dx^k \wedge dx^I \wedge dx^J \\ &= \left( \sum_{k=1}^n \frac{\partial a_I}{\partial x^k} dx^k \wedge dx^I \right) \wedge (b_J dx^J) + (-1)^p (a_I dx^I) \wedge \left( \sum_{k=1}^n \frac{\partial b_J}{\partial x^k} dx^k \wedge dx^J \right) \\ &= (d\omega) \wedge \mu + (-1)^p \omega \wedge (d\mu). \end{aligned}$$

Agora, provamos o último item: para formas de classe  $C^2$ , temos comutatividade das derivadas parciais de segunda ordem; logo, vale

$$\begin{aligned}
 d^2\omega(x) &= d\left(\sum_I \sum_{j=1}^n \frac{\partial a_I}{\partial x^j}(x) dx^j \wedge dx^I\right) \\
 &= \left(\sum_I \sum_{j,k=1}^n \frac{\partial^2 a_I}{\partial x^k \partial x^j}(x) dx^k \wedge dx^j \wedge dx^I\right) \\
 &= \sum_I \sum_{j < k} \left(\frac{\partial^2 a_I}{\partial x^k \partial x^j}(x) - \frac{\partial^2 a_I}{\partial x^j \partial x^k}(x)\right) dx^k \wedge dx^j \wedge dx^I \\
 &= 0.
 \end{aligned}$$

Finalmente, mostramos a unicidade. Em outras palavras, as propriedades provadas acima determina completamente a derivada exterior. Mais precisamente, vamos mostrar que uma aplicação  $d$  com as propriedades acima deve ser dada pela fórmula em (54). Suponhamos dada uma função  $d$  que satisfaça as propriedades 1-4. Então

$$\begin{aligned}
 d\omega &= d\left(\sum_I a_I(x) dx^I\right) \\
 &\stackrel{2)}{=} \sum_I d(a_I(x) dx^I) \\
 &\stackrel{3) \text{ e } 1)}{=} \sum_I [da_I(x) \wedge dx^I + (-1)^0 a_I(x) \wedge d^2x^I] \\
 &\stackrel{4)}{=} \sum_I da_I(x) \wedge dx^I \quad \square
 \end{aligned}$$

É um ótimo exercício, analisar as definições e propriedades acima explicitamente em dimensão “baixa”. Vamos fazer as contas para dimensão  $n = 3$ .

**Exercício 44.** Escreva todas as possíveis formas e suas derivadas exteriores para cada uma das dimensões  $n = 1, 2$  e  $4$ .

Em  $\mathbb{R}^3$ , vamos denotar como de costume

$$dx^1 = dx, \quad dx^2 = dy \quad \text{e} \quad dx^3 = dz.$$

As formas de grau zero, elementos de  $\Omega^0(U)$ , são as funções e sua derivada exterior é o diferencial usual. Considere uma forma de grau 1 do tipo

$$\omega = \omega_x dx + \omega_y dy + \omega_z dz.$$

Temos

$$\begin{aligned}
 d\omega &= d\omega_x \wedge dx + d\omega_y \wedge dy + d\omega_z \wedge dz \\
 &= \left(\frac{\partial \omega_x}{\partial x} dx + \frac{\partial \omega_x}{\partial y} dy + \frac{\partial \omega_x}{\partial z} dz\right) \wedge dx + \left(\frac{\partial \omega_y}{\partial x} dx + \frac{\partial \omega_y}{\partial y} dy + \frac{\partial \omega_y}{\partial z} dz\right) \wedge dy \\
 &\quad + \left(\frac{\partial \omega_z}{\partial x} dx + \frac{\partial \omega_z}{\partial y} dy + \frac{\partial \omega_z}{\partial z} dz\right) \wedge dz \\
 &= \left(\frac{\partial \omega_z}{\partial y} - \frac{\partial \omega_y}{\partial z}\right) dy \wedge dz - \left(\frac{\partial \omega_x}{\partial z} - \frac{\partial \omega_z}{\partial x}\right) dz \wedge dx + \left(\frac{\partial \omega_y}{\partial x} - \frac{\partial \omega_x}{\partial y}\right) dx \wedge dy.
 \end{aligned}$$

Observe que não escrevemos, propositalmente, os índices em ordem “crescente”. Isto foi para “escancarar” que a derivada exterior tem as coordenadas do mesmo tipo que o “rotacional” de um

campo vetorial. Segue que a aplicação  $d : \Omega^1(U) \rightarrow \Omega^2(U)$  que é a derivada de uma 1-forma em  $\mathbb{R}^3$  pode ser vista como o rotacional. No entanto, nossa derivada é independente da “regra da mão direita”. O espaço  $\mathbb{R}^3$  é especial porque só em dimensão 3, vale que

$$\dim \Lambda^1(\mathbb{R}^3)^* = \binom{3}{1} = \binom{3}{2} = \dim \Lambda^2(\mathbb{R}^3)^*.$$

A regra da mão direita aparece se tentarmos, de maneira não natural, identificar uma 1-forma com uma 2-forma, através do **operador estrela de Hodge**  $\star : \Lambda^2(\mathbb{R}^3)^* \rightarrow \Lambda^1(\mathbb{R}^3)^*$ , definido nos elementos da base de  $\Lambda^2(\mathbb{R}^3)^*$  como:

$$\star(dy \wedge dz) = dx, \quad \star(dz \wedge dx) = dy \quad \text{e} \quad \star(dx \wedge dy) = dz.$$

Este é um dos motivos de o campo rotacional ser as vezes chamado de um “pseudo” campo vetorial. Comparar com a discussão da Seção 17.

Vejam agora a derivada exterior de uma 2-forma em  $\mathbb{R}^3$ : se escrevemos

$$\omega = \omega_{xy} dx \wedge dy + \omega_{yz} dy \wedge dz + \omega_{xz} dz \wedge dx,$$

então (já deixando de escrever os termos nulos)

$$\begin{aligned} d\omega &= d\omega_{xy} \wedge dx \wedge dy + d\omega_{xz} \wedge dx \wedge dz + d\omega_{yz} \wedge dy \wedge dz \\ &= \frac{\partial \omega_{xy}}{\partial z} dz \wedge dx \wedge dy + \frac{\partial \omega_{yz}}{\partial x} dx \wedge dy \wedge dz + \frac{\partial \omega_{xz}}{\partial y} dy \wedge dz \wedge dx \\ &= \left( \frac{\partial \omega_{yz}}{\partial x} + \frac{\partial \omega_{xz}}{\partial y} + \frac{\partial \omega_{xy}}{\partial z} \right) dx \wedge dy \wedge dz, \end{aligned}$$

que é o equivalente ao divergente do campo vetorial de coordenadas  $(\omega_{yz}, \omega_{xz}, \omega_{xy})$  (comparar com a identificação dada pelo operador estrela de Hodge).

Finalmente, é fácil de ver que se  $\omega$  é uma forma diferencial de grau 3 em  $\mathbb{R}^3$ , então  $d\omega = 0$ .

A conclusão é que a derivada exterior para formas em  $\mathbb{R}^3$  unifica os diferentes “tipos de derivada” do cálculo vetorial em um conceito só. Além disso, temos uma generalização para dimensão qualquer

**Exercício 45.** Seja  $\omega \in \Omega^{n-1}(U)$  uma forma de grau  $(n-1)$  em  $\mathbb{R}^n$ . Calcule  $d\omega$  e compare a sua resposta com o divergente de um campo vetorial em  $\mathbb{R}^n$ .

## 20.1 O pullback de formas diferenciais

Formas diferenciais são objetos “contravariantes” no seguinte sentido: dados  $M, N \subseteq \mathbb{R}^n$  e uma função diferenciável

$$f : M \rightarrow N$$

é possível induzir, como vamos descrever, uma aplicação entre formas diferenciais (que inverte a ordem dos espaços e, por isto, contravariante)

$$f^* : \Omega(N) \rightarrow \Omega(M).$$

É, de certa maneira, mais “natural” pensar em  $M, N$  superfícies diferenciáveis ou variedades diferenciáveis e, de fato, vamos fazer isto em breve. Por enquanto, nossa discussão é para conjuntos  $M, N \subseteq \mathbb{R}^n$  abertos, apesar de estarmos fazendo distinção entre pontos e vetores tangentes.

Lembramos da Álgebra Linear (ver, por exemplo, [2, Seção 3.7]) que, dados dois espaços vetoriais  $V$  e  $W$  e uma transformação linear  $T : V \rightarrow W$ , podemos associar a  $T$ , por dualidade, a **transposta** de  $T$ :

$$T^* : W^* \rightarrow V^* \quad \text{dada por} \quad T^*(\varphi)[v] := \varphi(T(v))$$

que também é uma transformação linear.

Para cada  $x \in M$ , a derivada de  $f$  em  $x$  é uma transformação linear

$$df(x) : T_x M \rightarrow T_{f(x)} N$$

que leva vetores tangentes a  $M$  em vetores tangentes a  $N$ . Isto foi discutido na Subseção 11.1, onde concluímos que o diferencial pode ser visto como uma operação que *push-forward* (“empurra para frente”) vetores tangentes a  $M$  para vetores tangentes a  $N$ <sup>24</sup>. É também comum encontrarmos a notação

$$f_* := df(x) : T_x M \rightarrow T_{f(x)} N$$

Por dualidade, como acima, está associada a transposta de  $f_*$  (que é usualmente denotada por):

$$f^* := df(x)^* : T_{f(x)}^* N \rightarrow T_x^* M.$$

Mais explicitamente, para  $\phi \in T_{f(x)}^* N$ , o funcional  $f^* \phi \in T_x^* M$  é definido como

$$f^*(\phi)[v] := \phi(f_*(v)).$$

Uma forma diferencial de grau 0 em  $N$  é simplesmente uma função diferenciável  $\varphi : N \rightarrow \mathbb{R}$ . O **pullback** de  $\varphi$  por  $f$  é definido simplesmente como a composição

$$f^* \varphi := \varphi \circ f : M \rightarrow \mathbb{R}.$$

Observe como não é naturalmente possível fazer um “push-forward” de  $\varphi$  por  $f$ .

Uma forma diferencial de grau 1 em  $N$  pode ser pensada como uma aplicação que, para cada  $y \in N$ , associa um “covetor” tangente

$$\omega(y) \in T_y^* N \simeq (\mathbb{R}^n)^*.$$

Assim, dada uma 1-forma  $\omega$  definida em  $N$ , definimos o **pullback** de  $\omega$  por  $f$  como a 1-forma diferencial  $f^* \omega$  em  $M$  definida como

$$(f^* \omega)(x)[v] := \omega(f(x))[f_*(v)].$$

Na notação que utilizávamos para a derivada de  $f$  até então, deveríamos escrever:

$$(f^* \omega)(x)[v] = \omega(f(x))[df(x) \cdot v].$$

Afirmamos que a derivada exterior é **natural**. Isto significa o diagrama comutativo abaixo:

$$\begin{array}{ccc} \Omega^0(N) & \xrightarrow{f^*} & \Omega^0(M) \\ \downarrow d & & \downarrow d \\ \Omega^1(N) & \xrightarrow{f^*} & \Omega^1(M) \end{array}$$

Em outras palavras, para toda função diferenciável  $\varphi : N \rightarrow \mathbb{R}$ , temos

$$f^*(d\varphi) = d(f^* \varphi). \tag{55}$$

Se pensarmos em  $f$  como uma mudança de variáveis, por exemplo, um difeomorfismo, a fórmula (55) acima afirma que a derivada exterior é independente do sistema de coordenadas escolhido, no sentido que podemos ou calcular o diferencial de uma função e depois mudar coordenadas, ou então primeiro mudar coordenadas e depois diferenciar, e vamos obter o mesmo resultado.

<sup>24</sup>Na ocasião, por causa desta propriedade, chamamos os vetores tangentes de objetos “covariantes”.

Para provar (55), basta utilizar as definições cuidadosamente e a Regra da Cadeia. De fato, temos que  $d\varphi(f(x)) \in T_{f(x)}^*N$  e  $f^*(d\varphi)$  é uma 1-forma em  $M$ ; logo, para  $x \in M$  e  $v \in T_xM$ , calculamos

$$\begin{aligned} f^*(d\varphi)(x)[v] &= d\varphi(f(x)) [df(x) \cdot v] \\ &= d(\varphi \circ f)(x)[v] \\ &= d(f^*\varphi)(x)[v]. \end{aligned}$$

Para refletir sobre a definição e seu significado como mudança de variáveis, recomendamos o seguinte exercício.

**Exercício 46.** Considere  $\omega : \mathbb{R}^2 \setminus \{(0, 0)\} \rightarrow (\mathbb{R}^2)^*$  a forma elemento de ângulo

$$\omega(x) = \frac{-y}{x^2 + y^2} dx + \frac{x}{x^2 + y^2} dy$$

e  $f : (0, +\infty) \times (0, 2\pi) \rightarrow \mathbb{R}^2 \setminus \{(0, 0)\}$  dada por

$$f(r, \theta) := (r \cos \theta, r \sin \theta).$$

Calcule  $f^*\omega$  e interprete a sua resposta.

Em seguida, estendemos nossa definição anterior para o pullback de formas diferenciais de grau maior.

**Proposição 94.** O pullback de formas diferenciais

$$f^* : \Omega(N) \rightarrow \Omega(M)$$

fica completamente determinado pelas propriedades:

- 1) O pullback de 0-formas (funções) e de 1-formas coincide com as definições dadas acima.
- 2) *Linearidade:* para  $\omega, \mu \in \Omega(N)$  e  $a_1, a_2 \in \mathbb{R}$ , temos

$$f^*(a_1\omega + a_2\mu) = a_1f^*(\omega) + a_2f^*(\mu).$$

- 3) *Concordância com produto exterior:* para  $\omega, \mu \in \Omega(N)$ , temos

$$f^*(\omega \wedge \mu) = f^*(\omega) \wedge f^*(\mu).$$

*Demonstração.* Começamos observando que se  $g$  é uma 0-forma e  $\omega$  é uma 1-forma, a propriedade 1) e as definições anteriores implicam

$$\begin{aligned} f^*(g\omega)(x)[v] &= (g\omega)(f(x)) [df(x) \cdot v] \\ &= g(f(x)) \omega(f(x)) [df(x) \cdot v] \\ &= (f^*g)(x) (f^*\omega)(x)[v]. \end{aligned}$$

Logo,  $f^*(g\omega) = (f^*g)(f^*\omega)$ .

Por linearidade, basta agora determinar  $f^*$  em formas de grau  $p$  elementares. Seja então

$$\omega(x) = a(x) dx^{i_1} \wedge dx^{i_2} \wedge \dots \wedge dx^{i_p} \in \Omega^p(N).$$

Pela propriedade que mostramos acima e a propriedade 3) do enunciado, devemos ter

$$f^*\omega = (f^*a) (f^*dx^{i_1}) \wedge (f^*dx^{i_2}) \wedge \dots \wedge (f^*dx^{i_p}). \quad (56)$$

A fórmula (56) mostra a unicidade de  $f^*$  a partir das propriedades 1) a 3). Resta mostrar que  $f^*$  definida por (56) satisfaz estas propriedades. Deixamos a cargo do leitor, como exercício.  $\square$



Em seguida, mostramos que a derivada exterior é natural na álgebra das formas diferenciais. Em outras palavras, vamos mostrar que o diagrama abaixo é comutativo.

$$\begin{array}{ccc} \Omega^p(N) & \xrightarrow{f^*} & \Omega^p(M) \\ \downarrow d & & \downarrow d \\ \Omega^{p+1}(N) & \xrightarrow{f^*} & \Omega^{p+1}(M) \end{array}$$

**Proposição 95.** Para  $f : M \rightarrow N$  diferenciável e para  $\omega$  uma  $p$ -forma diferencial, temos

$$f^*(d\omega) = d(f^*\omega).$$

*Demonstração.* A prova é uma consequência direta das propriedades que definem  $f^*$ , da definição da derivada exterior (na passagem da linha 3 para a linha 4) e da naturalidade que já sabemos que vale para funções:

$$\begin{aligned} f^*(d\omega) &= f^*\left(\sum_I da_I \wedge dx^{i_1} \wedge dx^{i_2} \wedge \cdots \wedge dx^{i_p}\right) \\ &= \sum_I (f^*da_I) \wedge (f^*dx^{i_1}) \wedge (f^*dx^{i_2}) \wedge \cdots \wedge (f^*dx^{i_p}) \\ &= \sum_I d(f^*a_I) \wedge d(f^*x^{i_1}) \wedge d(f^*x^{i_2}) \wedge \cdots \wedge d(f^*x^{i_p}) \\ &= d\left(\sum_I (f^*a_I) d(f^*x^{i_1}) \wedge d(f^*x^{i_2}) \wedge \cdots \wedge d(f^*x^{i_p})\right) \\ &= d\left(\sum_I (f^*a_I) (f^*dx^{i_1}) \wedge (f^*dx^{i_2}) \wedge \cdots \wedge (f^*dx^{i_p})\right) \\ &= d\left(f^*\left(\sum_I a_I dx^{i_1} \wedge dx^{i_2} \wedge \cdots \wedge dx^{i_p}\right)\right) \\ &= d(f^*\omega). \end{aligned} \quad \square$$

Encerramos esta seção com um exercício que indica como descrever o *pullback* de  $p$ -formas pensadas como campos de formas  $p$ -lineares alternadas.

**Exercício 47.** Utilize o isomorfismo canônico  $\Lambda^p(\mathbb{R}^n)^* \simeq \mathcal{A}_p(\mathbb{R}^n)$  para mostrar que o *pullback* de uma  $p$ -forma é dado por

$$(f^*\omega)(x)(v_1, v_2, \dots, v_p) = \omega(f(x))[df(x) \cdot v_1, df(x) \cdot v_2, \dots, df(x) \cdot v_p].$$

## 21 Formas exatas e fechadas

De acordo com o Teorema 93, item 4), se uma 1-forma for o diferencial de uma função, digamos  $\mu = d\phi$ , então temos

$$d\mu = d(d\phi) = 0.$$

No caso de uma forma diferencial ter derivada exterior nula, nós dizemos que esta forma diferencial é uma **forma fechada**. No caso de uma forma diferencial ser o diferencial de alguma função, nós dizemos que a forma diferencial é uma **forma exata**. Segue do Teorema 93 citado acima que toda forma exata é fechada. Vamos analisar uma possível recíproca deste resultado.

A forma elemento de ângulo estudada do Exemplo 85 não é exata, como consequência da Proposição 84, mas é uma forma fechada. Isto mostra que somente podemos esperar por uma

recíproca parcial. Analisando mais de perto a prova da Proposição 84, o que pode nos prevenir de definir uma função potencial no caso geral é que a integral de linha pode ser dependente do caminho escolhido para integrar a forma diferencial.

Mais explicitamente, suponhamos que  $\omega : U \subseteq \mathbb{R}^n \rightarrow (\mathbb{R}^n)^*$  é uma forma diferencial de grau um definida em um aberto conexo  $U$ . Fixamos um ponto  $x_0 \in U$  arbitrário. Sendo  $U$  conexo por caminhos, podemos encontrar um caminho  $\gamma_x : [a, b] \rightarrow U$  com  $\gamma_x(a) = x_0$  e  $\gamma_x(b) = x$ . Vamos analisar a possibilidade de definir

$$\phi(x) := \int_{\gamma_x} \omega. \quad (57)$$

Como já explicamos,  $\phi$  pode estar “mal definida”, pois escolhas diferentes de caminhos poderiam resultar em integrais com valores diferentes. De qualquer maneira, vamos estudar como varia o valor da integral

$$\int_{\gamma} \omega$$

quando fizermos variações diferenciáveis na curva  $\gamma$ . Mais precisamente, suponhamos que existe uma família de curvas diferenciáveis

$$\gamma : [0, 1] \times [a, b] \rightarrow \mathbb{R}^n$$

tal que

$$\gamma(s, a) = x_0 \quad \text{e} \quad \gamma(s, b) = x \quad \text{para todo } s \in [0, 1].$$

Suponhamos que  $\gamma$  é diferenciável em ambas as variáveis  $s$  e  $t$ . Intuitivamente, temos, para cada  $s \in [0, 1]$ , uma curva  $\gamma_s := \gamma(s, \cdot) : [a, b] \rightarrow U$  com as extremidades fixas:  $\gamma_s(a) = x_0$  e  $\gamma_s(b) = x$ . Estamos, portanto, supondo que existe uma interpolação diferenciável entre as duas curvas  $\gamma_0$  e  $\gamma_1$ . Denotamos

$$I_s := \int_{\gamma_s} \omega = \int_a^b \omega(\gamma(s, t)) \left( \frac{\partial \gamma}{\partial t}(s, t) \right) dt.$$

Se

$$\omega(x) = \sum_{i=1}^n a_i(x) dx^i,$$

sabemos que

$$d\omega(x) = \sum_{i,j=1}^n \frac{\partial a_i}{\partial x^j}(x) dx^i \wedge dx^j.$$

Além disso, lembrando da identificação de 2-vetores com formas bilineares alternadas, temos

$$d\omega(x)[v, w] = \sum_{i,j=1}^n \frac{\partial a_i}{\partial x^j}(x) (dx^i \wedge dx^j)[v, w] = \sum_{i,j=1}^n \frac{\partial a_i}{\partial x^j}(x) (v^i w^j - v^j w^i).$$

O integrando pode ser escrito como

$$\omega(\gamma(s, t)) \left( \frac{\partial \gamma}{\partial t}(s, t) \right) = \sum_{i=1}^n a_i(\gamma(s, t)) \frac{\partial \gamma^i}{\partial t}(s, t).$$

Vamos identificar a dependência de  $I_s$  na variável  $s$ :

$$\frac{\partial I_s}{\partial s} = \sum_{i=1}^n \frac{\partial}{\partial s} \int_a^b a_i(\gamma(s, t)) \frac{\partial \gamma^i}{\partial t}(s, t) dt.$$

Passando a derivada para dentro da integral (pode?) e aplicando a regra do produto, obtemos

$$\frac{\partial I_s}{\partial s} = \sum_{i=1}^n \int_a^b \left( a_i(\gamma(s, t)) \frac{\partial^2 \gamma^i}{\partial s \partial t}(s, t) + \sum_{j=1}^n \frac{\partial a_i}{\partial x^j}(\gamma(s, t)) \frac{\partial \gamma^j}{\partial s}(s, t) \frac{\partial \gamma^i}{\partial t}(s, t) \right) dt.$$

Logo, por integração por partes,

$$\begin{aligned}
\frac{\partial I_s}{\partial s} &= \sum_{i,j=1}^n \int_a^b \left( -\frac{\partial a_i}{\partial x^j}(\gamma(s,t)) \frac{\partial \gamma^j}{\partial t}(s,t) \frac{\partial \gamma^i}{\partial s}(s,t) + \frac{\partial a_i}{\partial x^j}(\gamma(s,t)) \frac{\partial \gamma^j}{\partial s}(s,t) \frac{\partial \gamma^i}{\partial t}(s,t) \right) dt \\
&= \int_a^b \sum_{i,j=1}^n \frac{\partial a_i}{\partial x^j}(\gamma(s,t)) \left( \frac{\partial \gamma^j}{\partial s}(s,t) \frac{\partial \gamma^i}{\partial t}(s,t) - \frac{\partial \gamma^i}{\partial s}(s,t) \frac{\partial \gamma^j}{\partial t}(s,t) \right) dt \\
&= \int_a^b \sum_{i,j=1}^n \frac{\partial a_i}{\partial x^j}(\gamma(s,t)) (dx^i \wedge dx^j) \left( \frac{\partial \gamma}{\partial s}(s,t), \frac{\partial \gamma}{\partial t}(s,t) \right) dt \\
&= \int_a^b d\omega(\gamma(s,t)) (\partial_t \gamma(s,t), \partial_s \gamma(s,t)) dt.
\end{aligned}$$

Assim, se for  $\omega$  fechada, segue que a integral não depende dos caminhos  $\gamma_s$ . Se quaisquer duas curvas puderem ser conectadas por uma família de caminhos como estes que utilizamos acima, então é possível definir a função potencial pela Fórmula (57).

Dois caminhos  $\gamma_0 : [a, b] \rightarrow U$  e  $\gamma_1 : [a, b] \rightarrow U$  que tem as mesmas extremidades, isto é, tais que  $\gamma_0(a) = \gamma_1(a)$  e  $\gamma_0(b) = \gamma_1(b)$ , são ditos **homotópicos** em  $U$  quando existe uma função *contínua*

$$\gamma : [0, 1] \times [a, b] \rightarrow U$$

tal que

$$\gamma(0, t) = \gamma_0(t) \quad \text{e} \quad \gamma(1, t) = \gamma_1(t) \quad \text{para todo } t \in [a, b]$$

e também

$$\gamma(s, a) = \gamma(a) \quad \text{e} \quad \gamma(s, b) = \gamma(b) \quad \text{para todo } s \in [0, 1].$$

Também dizemos que a função  $\gamma$  é uma **homotopia** entre os caminhos  $\gamma_0$  e  $\gamma_1$ .

A nossa discussão acima (quase) implica na seguinte proposição.

**Proposição 96.** Se  $\gamma_0$  e  $\gamma_1$  são caminhos suaves por partes e homotópicos em um conjunto  $U$  e  $\omega$  é uma 1-forma fechada em  $U$ , então

$$\int_{\gamma_0} \omega = \int_{\gamma_1} \omega.$$

*Demonstração.* Para completar a prova da proposição, resta mostrar que, dados dois caminhos suaves por partes *continuamente* homotópicos, é possível encontrar uma homotopia *diferenciável*, para justificar as contas acima. **fazer** □

Se um conjunto  $U$  é tal que quaisquer dois caminhos suaves por partes são homotópicos, então nós dizemos que  $U$  é **simplesmente conexo**. Intuitivamente,  $U$  deve ser um conjunto “sem buracos”, já que dadas duas curvas podemos “transformar uma na outra continuamente sem sair de  $U$ ”. O seguinte resultado deve agora soar muito natural (e a prova é consequência direta da discussão acima e da prova da Proposição 84).

**Teorema 97.** Seja  $\omega \in \Omega^1(U)$  uma 1-forma diferencial em um conjunto simplesmente conexo  $U$ . Então,  $\omega$  é uma forma fechada se, e somente se, é uma forma exata.

## 22 Variedades diferenciáveis

Nesta seção, vamos introduzir um conceito fundamental de geometria diferencial, o de variedades diferenciáveis. Várias referências foram consultadas; entre elas, [1, 4, 8, 9, 17].

Uma variedade diferenciável é um espaço topológico que localmente se parece com o espaço Euclidiano de maneira a permitir diferenciabilidade e, por consequência, o uso de técnicas do Cálculo Diferencial. De forma mais precisa, uma **variedade diferenciável** (de classe  $C^\infty$ ) de dimensão  $n$  é um espaço topológico  $M$  junto com um conjunto de funções injetivas, chamadas

**cartas**,  $\phi_\alpha : U_\alpha \rightarrow \mathbb{R}^n$ , onde  $U_\alpha$  são abertos em  $M$  e cobrem  $M$ , e as funções de transição  $\phi_\alpha \circ \phi_\beta^{-1}$  são de classe  $C^\infty$  nos conjuntos onde estão definidas. Ao conjunto das cartas  $\{\phi_\alpha\}$  chamamos **atlas**.

A ideia é utilizar a conclusão fundamental do Teorema 33, item (ii), para o caso de conjuntos mais gerais do que subconjuntos do  $\mathbb{R}^n$ . Esta é a propriedade que nos permitiu definir diferenciabilidade de maneira coerente.

**Exemplo 98.** Qualquer dos espaços  $\mathbb{R}^n$ . Uma carta que cobre todo o espaço é a aplicação identidade.

**Exemplo 99 (Superfícies diferenciáveis).** Qualquer superfície diferenciável é uma variedade diferenciável de um espaço  $\mathbb{R}^n$ , pois, como vimos, as mudanças de parâmetros são difeomorfismos.

**Exemplo 100 (Espaço projetivo real).** Consideramos

$$\mathbb{P}^2(\mathbb{R}) := \{\text{retas em } \mathbb{R}^3 \text{ que passam pela origem}\}.$$

Este é um exemplo não trivial, no sentido de ser essencialmente diferente de uma superfície diferenciável. Observamos ainda, sem demonstração<sup>25</sup>, que  $\mathbb{P}^2(\mathbb{R})$  é uma variedade de dimensão 2 que não pode ser mergulhada em  $\mathbb{R}^3$ .

Podemos, no entanto, identificar  $\mathbb{P}^2(\mathbb{R})$  com o espaço quociente

$$\mathbb{P}^2(\mathbb{R}) \simeq (\mathbb{R}^3 \setminus \{0\}) / \sim$$

onde  $\sim$  é a relação de equivalência que identifica todos os pontos que estão sobre uma mesma reta em  $\mathbb{R}^3$ :

$$x \sim y \iff x = \lambda y \text{ para algum } \lambda \neq 0.$$

Agora, consideramos as vizinhanças

$$U_1 := \{[x]; x_1 \neq 0\}, \quad U_2 := \{[x]; x_2 \neq 0\} \quad \text{e} \quad U_3 := \{[x]; x_3 \neq 0\},$$

que certamente cobrem  $\mathbb{P}^2(\mathbb{R})$ . A função  $\phi_1 : U_1 \rightarrow \mathbb{R}^2$  é definida por

$$\phi_1([x_1, x_2, x_3]) := \left( \frac{x_2}{x_1}, \frac{x_3}{x_1} \right)$$

Analogamente, as outras duas funções  $\phi_2 : U_2 \rightarrow \mathbb{R}^2$  e  $\phi_3 : U_3 \rightarrow \mathbb{R}^2$  são definidas por

$$\phi_2([x_1, x_2, x_3]) := \left( \frac{x_1}{x_2}, \frac{x_3}{x_2} \right) \quad \text{e} \quad \phi_3([x_1, x_2, x_3]) := \left( \frac{x_1}{x_3}, \frac{x_2}{x_3} \right).$$

A função  $\phi_1$  está bem definida (e as outras analogamente), pois  $[x] = [y]$  implica  $x = \lambda y$  para algum  $\lambda \neq 0$ ; logo,

$$\frac{y_2}{y_1} = \frac{\lambda x_2}{\lambda x_1} = \frac{x_2}{x_1} \quad \text{e} \quad \frac{y_3}{y_1} = \frac{\lambda x_3}{\lambda x_1} = \frac{x_3}{x_1}.$$

São também bijetivas e com inversas iguais a

$$\phi_1^{-1}(u, v) = [1, u, v], \quad \phi_2^{-1}(u, v) = [u, 1, v] \quad \text{e} \quad \phi_3^{-1}(u, v) = [u, v, 1]$$

Resta mostrar que as mudanças de parâmetros são de classe  $C^\infty$ . De fato, se considerarmos, por exemplo,  $\phi_1$  e  $\phi_2$ , temos

$$\phi_2(\phi_1^{-1}(u, v)) = \phi_2([1, u, v]) = \left( \frac{1}{u}, \frac{v}{u} \right)$$

que é de classe  $C^\infty$  em  $\phi_1(U_1 \cap U_2)$ .

<sup>25</sup>No entanto, a ideia da prova é a seguinte: mostra-se que  $\mathbb{P}^2(\mathbb{R})$  é compacta e não orientável. Em seguida, se pudesse ser mergulhada em  $\mathbb{R}^3$ , limitaria uma região compacta em  $\mathbb{R}^3$  (este é o conteúdo do Teorema de Separação de Jordan–Brouwer), cuja fronteira é  $\mathbb{P}^2(\mathbb{R})$ . No entanto, toda hipersuperfície compacta de  $\mathbb{R}^3$  (ou de  $\mathbb{R}^n$ ) é orientável, o que gera uma contradição.

Em uma variedade diferenciável, é possível falar de diferenciabilidade de funções “olhando em coordenadas”, isto é, através das imagens por cartas. Uma função  $f : M \rightarrow \mathbb{R}$  é dita **diferenciável** em  $M$  quando  $f \circ \phi_\alpha^{-1} : A \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$  é diferenciável, para todas as parametrizações  $\phi_\alpha$ . Isto equivale a dizer que, quando vista “em coordenadas”,  $f$  é uma função diferenciável. A condição de as funções de transição serem suaves é fundamental para que este conceito esteja bem definido.

De forma análoga,  $f$  é de classe  $C^k$  ou de classe  $C^\infty$  quando  $f \circ \phi_\alpha^{-1}$  é de classe  $C^k$  ou de classe  $C^\infty$ , respectivamente.

## 22.1 Espaço tangente

Quando a variedade está “dentro” de um espaço euclidiano, poderíamos definir vetores tangentes assim como fazemos para superfícies diferenciáveis, a saber, como elementos da imagem do diferencial de uma parametrização qualquer. No caso mais abstrato que nos encontramos, é necessário obter alguma caracterização diferente.

Intuitivamente, nos espaços do tipo  $\mathbb{R}^n$ , campos de vetores aparecem como um conjunto de setas, uma sobre cada ponto do espaço. Usualmente, campos vetoriais podem representar campos de força ou de velocidade de fluidos. Cada uma dessas setas representa a direção e o sentido que uma partícula andaria sob o efeito do tal campo vetorial. Ou então, cada uma das setas poderia indicar a derivada direcional a ser calculada de funções escalares. Estas duas maneiras de pensar indicam os caminhos mais comuns de se definir o espaço tangente a uma variedade diferenciável.

Vamos inicialmente pensar em  $f : \mathbb{R}^n \rightarrow \mathbb{R}$ . Ainda não estamos pensando rigorosamente. A derivada direcional de  $f$  no ponto  $x$  e na direção de um vetor  $v$  pode ser calculada ao considerar uma curva  $\alpha$  com  $\alpha(0) = x$  e  $\alpha'(0) = v$ ; tem-se

$$\frac{\partial f}{\partial v}(x) = \left. \frac{d}{dt} f(\alpha(t)) \right|_{t=0} = \sum_{i=1}^n \alpha'_i(0) \frac{\partial f}{\partial x_i}(x) = \left( \sum_{i=1}^n \alpha'_i(0) \frac{\partial}{\partial x_i} \Big|_{t=0} \right) f, \quad (58)$$

onde  $\alpha(t) = (\alpha_1(t), \alpha_2(t), \dots, \alpha_n(t))$ . Em particular,  $\alpha'_i(0)$  são as coordenadas de  $v \in \mathbb{R}^n$ . A ideia é escrever  $v$  como um funcional linear da forma  $v : C^\infty(\mathbb{R}^n) \rightarrow \mathbb{R}$  e pensar em  $v(f)$  como a derivada de  $f$  em  $x$  e na direção  $v$ . Para isso,  $v$  deve ainda satisfazer propriedades do tipo de uma “derivação”.

Mais rigorosamente, um dos pontos de vista é pensar nos vetores tangentes à variedades como velocidades de curvas que passam por  $x$  com a “mesma direção”. Mais precisamente, dadas duas curvas diferenciáveis  $\alpha$  e  $\beta$  com  $\alpha(0) = x = \beta(0)$ , definimos uma relação de equivalência por<sup>26</sup>

$$\alpha \sim \beta \iff \left. \frac{d}{dt} \phi(\alpha(t)) \right|_{t=0} = \left. \frac{d}{dt} \phi(\beta(t)) \right|_{t=0} \quad \text{para alguma carta } \phi.$$

Um **vetor tangente** é uma classe de equivalência  $v = [\gamma]$  para alguma curva diferenciável  $\gamma$  com  $\gamma(0) = x$  e o conjunto dos vetores tangentes a  $M$  no ponto  $x \in M$  é chamado de **espaço tangente** e denotado por  $T_x M$ .

**Proposição 101.** *Se  $M$  é uma variedade diferenciável de dimensão  $n$ , então  $T_x M$  é um espaço vetorial de dimensão  $n$ .*

*Demonstração.* Dado um ponto  $x \in M$  e uma carta  $\phi$  que parametriza uma vizinhança de  $x$  em  $M$ , definimos  $d\phi_x : T_x M \rightarrow \mathbb{R}^n$  por

$$d\phi_x(v) := \left. \frac{d}{dt} \phi(\gamma(t)) \right|_{t=0},$$

onde  $\gamma$  é qualquer representante da classe de equivalência  $v = [\gamma]$ . Esta aplicação está bem definida, como pode ser visto diretamente da definição da relação de equivalência. Também pela

<sup>26</sup>Pela diferenciabilidade das funções de transição, a igualdade vale, na realidade, para toda carta que contém as curvas  $\alpha$  e  $\beta$ . Logo, o espaço tangente é independente da parametrização escolhida.

definição da relação de equivalência, vemos que  $d\phi_x$  é injetiva. Para mostrar que é também sobrejetiva, dado  $w \in \mathbb{R}^n$ , definimos

$$\alpha(t) = \phi^{-1}(\phi(x) + tw) \quad \text{de modo que} \quad \left. \frac{d}{dt} \phi(\alpha(t)) \right|_{t=0} = w.$$

Sendo  $d\phi_x$  uma bijeção, podemos equipar  $T_x M$  com a estrutura de espaço vetorial induzida pela aplicação inversa  $d\phi_x^{-1}$ , tornando  $T_x M$  um espaço vetorial de dimensão  $n$ . De maneira mais explícita, as operações em  $T_x M$  são definidas como

$$v + w := d\phi_x^{-1}(d\phi_x(v) + d\phi_x(w)) \quad \text{e} \quad cv := d\phi_x^{-1}(c d\phi_x(v)).$$

Sendo as funções de transição difeomorfismos, temos que

$$(\psi \circ \phi^{-1})'(\phi(x)) : \mathbb{R}^n \rightarrow \mathbb{R}^n$$

é um isomorfismo. Portanto, as duas parametrizações quaisquer de vizinhanças de  $x$  geram em  $T_x M$  a mesma estrutura de espaço vetorial.  $\square$

É também comum utilizar a notação  $\gamma'(0) = [\gamma]$ , onde  $\gamma$  é qualquer representante da classe de equivalência  $[\gamma]$ . Assim, continuamos com a fórmula usual

$$d\phi_{\alpha(0)}(\gamma'(0)) = \left. \frac{d}{dt} \phi(\gamma(t)) \right|_{t=0},$$

embora esta seja agora apenas a definição da aplicação  $d\phi_x$ .

Um outro ponto de vista, menos geométrico mas bastante conveniente, é olhar para vetores tangentes como objetos que agem sobre funções diferenciáveis da mesma maneira que “tomar derivadas direcionais”. Em outras palavras, um vetor tangente é visto como o que nos permite fazer a derivada direcional de funções suaves. Nós dizemos que  $v_x : C^\infty(M) \rightarrow \mathbb{R}$  é um **vetor tangente** a  $M$  no ponto  $x \in M$  quando é uma **derivação** em  $x$ , isto é, quando é um funcional linear que satisfaz a regra de Leibniz no ponto  $x$ :

$$v_x(fg) = v_x(f) \cdot g(x) + f(x) \cdot v_x(g).$$

Esta propriedade é claramente motivada pela regra da diferenciação de um produto de funções. Ao espaço das derivações como acima chamamos de **espaço tangente** a  $M$  em  $x$  e denotamos por  $\tilde{T}_x M$  (o til será deixado de lado quando mostrarmos que as duas definições são equivalentes). Naturalmente,  $\tilde{T}_x M$  tem uma estrutura de espaço vetorial (dada pela soma de funcionais lineares e a multiplicação de um funcional linear por um escalar).

Observamos que, dado  $[\gamma] \in T_x M$ , podemos associar uma derivação em  $\tilde{T}_x M$  pela identidade

$$v_\gamma(f) := \left. \frac{d}{dt} f(\gamma(t)) \right|_{t=0} \quad \text{para } f \in C^\infty(M). \quad (59)$$

Para verificar que de fato (59) define um elemento de  $\tilde{T}_x M$ , consideramos uma carta  $\phi : U \rightarrow \mathbb{R}^n$  em uma vizinhança  $U$  de  $x$ . Em coordenadas, temos

$$v_\gamma(f) = \left. \frac{d}{dt} (f \circ \phi^{-1} \circ \phi \circ \gamma)(t) \right|_{t=0} = d(f \circ \phi^{-1})(\phi(x)) \cdot \left. \frac{d}{dt} (\phi \circ \gamma)(t) \right|_{t=0}.$$

Escrevendo as variáveis de  $\mathbb{R}^n$  da inversa da parametrização como  $\phi^{-1}(x^1, x^2, \dots, x^n)$  e a curva  $\gamma$  na carta  $\phi$  por

$$\phi(\gamma(t)) = (x^1(t), x^2(t), \dots, x^n(t)),$$

temos

$$v_\gamma(f) = \sum_{i=1}^n \frac{dx^i}{dt}(0) \frac{\partial (f \circ \phi^{-1})}{\partial x^i}(\phi(x)).$$

Comparar esta fórmula com (58). A linearidade de  $v_\gamma$  segue daí:

$$\begin{aligned} v_\gamma(af + bg) &= \sum_{i=1}^n \frac{dx^i}{dt}(0) \frac{\partial((af + bg) \circ \phi^{-1})}{\partial x^i}(\phi(x)) \\ &= a \sum_{i=1}^n \frac{dx^i}{dt}(0) \frac{\partial(f \circ \phi^{-1})}{\partial x^i}(\phi(x)) + b \sum_{i=1}^n \frac{dx^i}{dt}(0) \frac{\partial(g \circ \phi^{-1})}{\partial x^i}(\phi(x)) \\ &= av_\gamma(f) + bv_\gamma(g). \end{aligned}$$

A regra de Leibniz é ainda mais direta:

$$v_\gamma(fg) = \frac{d}{dt} [f(\gamma(t))g(\gamma(t))] \Big|_{t=0} = \frac{d}{dt} f(\gamma(t)) \Big|_{t=0} g(\gamma(0)) + f(\gamma(0)) \frac{d}{dt} g(\gamma(t)) \Big|_{t=0}.$$

Logo,  $[\gamma] \in T_x M \implies v_\gamma \in \tilde{T}_x M$ .

Observamos ainda que, definindo

$$\frac{\partial}{\partial x^i} \Big|_x : C^\infty(M) \rightarrow \mathbb{R} \quad \text{por} \quad \frac{\partial}{\partial x^i} \Big|_x (f) := \frac{\partial(f \circ \phi^{-1})}{\partial x^i}(\phi(x)), \quad (60)$$

obtemos uma derivação linear, ou seja, um elemento de  $\tilde{T}_x M$ . Pelas contas acima, temos

$$v_\gamma = \sum_{i=1}^n (x^i)'(0) \frac{\partial}{\partial x^i} \Big|_x,$$

que é exatamente a fórmula (58), mas em sua versão para a variedade  $M$ . Em seguida, vamos mostrar que toda derivação linear é dessa forma, o que mostra que as duas noções de espaço tangente que definimos são equivalentes. Além disso, a equação (60) define uma base de  $T_x M$  para cada ponto  $x$  da vizinhança coordenada  $U$ :

$$T_x M = \text{Span} \left\{ \frac{\partial}{\partial x^1} \Big|_x, \frac{\partial}{\partial x^2} \Big|_x, \dots, \frac{\partial}{\partial x^n} \Big|_x \right\}.$$

Tudo isto é o conteúdo da proposição seguinte.

**Proposição 102.** *A aplicação*

$$\begin{aligned} \Phi : T_x M &\rightarrow \tilde{T}_x M \\ [\gamma] &\mapsto v_\gamma \end{aligned}$$

*é um isomorfismo de espaços vetoriais.*

*Demonstração.* Inicialmente, vamos mostrar que  $\Phi$  é injetiva: dados  $v, w \in T_x M$  satisfazendo  $\Phi(v) = \Phi(w)$ , escolhemos respectivos representantes  $\alpha$  e  $\beta$ . Assim, devemos ter

$$\frac{d}{dt} f(\alpha(t)) \Big|_{t=0} = \frac{d}{dt} f(\beta(t)) \Big|_{t=0} \quad \text{para toda } f \in C^\infty(M).$$

Dada uma parametrização  $\phi : U \subset M \rightarrow \mathbb{R}^n$ , vamos considerar  $f = \phi^j : U \rightarrow \mathbb{R}$  acima, onde  $\phi^j := \pi^j \circ \phi$  é a projeção na  $j$ -ésima coordenada de  $\phi$ . Olhando coordenada a coordenada, temos<sup>27</sup>

$$\frac{d}{dt} (\phi \circ \alpha)(t) \Big|_{t=0} = \frac{d}{dt} (\phi \circ \beta)(t) \Big|_{t=0}.$$

<sup>27</sup> Estritamente falando, deveríamos ter uma função  $\phi$  definida em  $M$  e não apenas em um subconjunto  $U \subset M$ . Este detalhe técnico está relacionado com a existência de funções de corte e, conseqüentemente, com as restrições topológicas que impomos sobre  $M$ . Mostraremos a existência de funções de corte na Seção 23.

Desta forma,  $\alpha \sim \beta$  e logo  $v = [\alpha] = [\beta] = w$ .

Finalmente, mostramos a sobrejetividade de  $\Phi$ . Devemos mostrar que, dada uma derivação qualquer  $v \in \tilde{T}_x M$ , podemos encontrar uma curva  $\gamma$  tal que  $v = v_\gamma$  como em (59). Seja  $\phi : U \rightarrow \mathbb{R}^n$  uma carta em uma vizinhança de  $x = \phi^{-1}(x_0)$  e seja  $\phi^i := \pi^i \circ \phi : U \rightarrow \mathbb{R}$  a  $i$ -ésima coordenada de  $\phi$ . Definimos

$$a^i := v(\phi^i) \in \mathbb{R} \quad \text{e} \quad a = (a^1, a^2, \dots, a^n) \in \mathbb{R}^n.$$

Para  $t$  suficientemente pequeno, está bem definida a curva

$$\gamma(t) := \phi^{-1}(x_0 + ta) \in U \subset M.$$

Desta forma, pela regra da cadeia usual do espaço euclidiano, segue que

$$v_\gamma(f) = \left. \frac{d}{dt} f(\gamma(t)) \right|_{t=0} = d(f \circ \phi^{-1})(\phi(x)) \cdot (\phi \circ \gamma)'(0) = \sum_{i=1}^n \frac{\partial(f \circ \phi^{-1})}{\partial x^i}(\phi(x)) v(\phi^i).$$

Por outro lado, o Teorema Fundamental do Cálculo aplicado para a função  $f \circ \phi^{-1} : \phi(U) \subset \mathbb{R}^n \rightarrow \mathbb{R}$  implica que (observe que  $\phi(U)$  é um aberto de  $\mathbb{R}^n$  e estamos aplicando o teorema usual da análise na reta)

$$\begin{aligned} (f \circ \phi^{-1})(y) - (f \circ \phi^{-1})(x_0) &= \int_0^1 \frac{d}{dt} (f \circ \phi^{-1})(x_0 + t(y - x_0)) dt \\ &= \sum_{i=1}^n (y^i - x_0^i) \int_0^1 \frac{\partial(f \circ \phi^{-1})}{\partial x^i}(x_0 + t(y - x_0)) dt \\ &=: \sum_{i=1}^n (y^i - x_0^i) g_i(y), \end{aligned}$$

onde, conforme indicado, definimos

$$g_i(y) = \int_0^1 \frac{\partial(f \circ \phi^{-1})}{\partial x^i}(x_0 + t(y - x_0)) dt.$$

Escrevendo  $w = \phi^{-1}(y) \in M$ , obtemos

$$f(w) - f(x) = \sum_{i=1}^n (\phi^i(w) - \phi^i(x)) (g_i \circ \phi)(w).$$

Aplicando  $v$  em ambos os lados, utilizando que  $v$  é linear e a regra de Leibniz, obtemos<sup>28</sup>

$$\begin{aligned} v(f) &= \sum_{i=1}^n v(\phi^i - \phi^i(x)) (g_i \circ \phi)(x) + (\phi^i(x) - \phi^i(x)) v(g_i \circ \phi) \\ &= \sum_{i=1}^n a^i g_i(x_0) = \sum_{i=1}^n \frac{\partial(f \circ \phi^{-1})}{\partial x^i}(\phi(x)) v(\phi^i). \end{aligned}$$

Portanto,  $v(f) = v_\gamma(f)$ , como queríamos demonstrar. □

<sup>28</sup>Deveríamos, mais precisamente, observar que  $v$  depende apenas dos valores de  $f$  em uma vizinhança do ponto  $x$ . Isto, novamente, depende da existência de funções de corte. Mais elementar é a observação que, se  $f \equiv c$  é uma função constante, então  $v(c) = 0$ : de fato, a regra de Leibniz implica que

$$v(1) = v(1 \cdot 1) = 1 \cdot v(1) + v(1) \cdot 1 = 2 \cdot v(1) \implies v(1) = 0 \implies v(c) = c \cdot v(1) = 0.$$



Na definição (60), descrevemos uma base para o espaço tangente “pensado” como derivações. Uma análise da prova da Proposição 102 nos permite descrever uma base de  $T_x M$  em termos de classes de equivalências de curvas:

$$\frac{\partial}{\partial x^i} \Big|_x = [\gamma]$$

onde  $\gamma : (-\delta, \delta) \rightarrow M$  é a curva dada em coordenadas por

$$\phi(\gamma(t)) = x_0 + te_i.$$

## 22.2 Diferencial

Sejam  $M$  e  $N$  variedades diferenciáveis e  $f : M \rightarrow N$  uma aplicação diferenciável. Nós já definimos diferenciabilidade  $f$ , mas ainda não falamos sobre a *derivada* ou o *diferencial* de  $f$ . Intuitivamente, uma curva em  $M$  é levada por  $f$  em uma curva em  $N$  e, logo, a derivada “deve” levar um vetor (tangente) de  $M$  em um vetor (tangente) de  $N$ . Nesta seção, vamos descrever esta construção.

Em termos de classes de equivalências de curvas, definimos a aplicação

$$\begin{aligned} df(x) : T_x M &\rightarrow T_{f(x)} N \\ [\alpha] &\rightarrow [f \circ \alpha] \end{aligned}$$

que é chamada de a **derivada** ou o **diferencial** de  $f$  em  $x$ . Nossa definição está bem, pois se  $\alpha \sim \beta$  são dois representantes da mesma classe, a Regra da Cadeia usual do  $\mathbb{R}^n$  implica que

$$\begin{aligned} \frac{d}{dt} [\psi((f \circ \alpha)(t))] \Big|_{t=0} &= d(\psi \circ f \circ \phi^{-1})(\phi(x)) \cdot \frac{d}{dt} (\phi \circ \alpha)(t) \Big|_{t=0} \\ &= d(\psi \circ f \circ \phi^{-1})(\phi(x)) \cdot \frac{d}{dt} (\phi \circ \beta)(t) \Big|_{t=0} \\ &= \frac{d}{dt} [\psi((f \circ \beta)(t))] \Big|_{t=0}, \end{aligned}$$

de modo que  $[f \circ \alpha] = [f \circ \beta]$ .

**Exercício 48.** Mostre que  $df(x)$  é uma transformação linear (como é a estrutura linear de  $T_x M$  como classe de equivalência de curvas?).

**Exercício 49.** Mostre que vale a Regra da Cadeia: se  $f : M \rightarrow N$  e  $g : N \rightarrow P$  são aplicações diferenciáveis entre variedades, então temos a seguinte identidade entre transformações lineares:

$$d(g \circ f)(x) = dg(f(x)) \cdot df(x).$$

O **diferencial**  $df(x) : T_x M \rightarrow T_{f(x)} N$  também pode ser descrito em termos de derivações: para a derivação  $v \in T_x M$ , a derivação  $df(x) \cdot v$  é a que, para a função  $g \in C^\infty(N)$ , associa

$$(df(x) \cdot v)(g) := v(g \circ f).$$

Esta fórmula de fato define uma derivação no ponto  $f(x)$ , pois

$$v[(gh) \circ f] = v[(g \circ f)(h \circ f)] = v(g \circ f) h(f(x)) + g(f(x)) v(h \circ f).$$

Claramente, é linear em  $g$ . Além disso, a aplicação  $df(x)$  definida desta maneira é claramente linear (em  $v$ ).

**Exercício 50.** Prove a Regra da Cadeia como no Exercício 49, mas pensando com “derivações”.

Podemos utilizar identificação natural entre  $T_{f(x)} \mathbb{R}$  e  $\mathbb{R}$  para definir o diferencial de funções e de curvas. Deixaremos indicado e como exercício (**fazer isto**).

## 23 Partições da Unidade

Uma ferramenta importantíssima na passagem do "local para o global" em análise geométrica é uma partição (diferenciável) da unidade. Seja  $M$  uma variedade diferenciável. Dada uma coleção  $\mathcal{X} = \{U_\alpha\}$  de abertos de  $M$ , uma **partição da unidade subordinada a  $\mathcal{X}$**  é uma família de funções diferenciáveis  $\{f_\alpha\} \subseteq C^\infty(M)$  tal que

- (i) Tem-se  $0 \leq f_\alpha(x) \leq 1$  para todo  $x \in M$  e para todo índice  $\alpha$ ;
- (ii) No complementar de  $U_\alpha$ , temos  $f_\alpha \equiv 0$ . Uma outra maneira de dizer isto é o seguinte: definimos o **suporte** de  $f_\alpha$  como sendo o conjunto (fechado)

$$\text{supp } f_\alpha := \overline{\{x \in M; f_\alpha(x) \neq 0\}}.$$

Assim, este item (ii) pode ser reescrito como  $\text{supp } f_\alpha \subseteq U_\alpha$ ;

- (iii) Para cada ponto  $x \in M$ , existe um aberto de  $M$  contendo  $x$  onde apenas um número finito de funções  $f_\alpha$  é não nulo. Outra maneira de enunciar a mesma propriedade é: dado um ponto  $x \in M$ , existe um aberto de  $M$  que contém  $x$  e que intersecciona apenas um número finito de conjuntos da família  $\{\text{supp } f_\alpha\}$ . Quando esta propriedade (iii) vale, nós dizemos que a família  $\{\text{supp } f_\alpha\}$  é **localmente finita**.

- (iv) Para todo  $x \in M$ , temos

$$\sum_{\alpha} f_\alpha(x) = 1. \quad (61)$$

Por conta do item (iii), a soma que aparece neste item (iv) é sempre uma *soma finita*, não sendo necessária a preocupação com convergência. Este é o item que motiva o nome "partição da unidade", pois escrevemos o número um como soma finita de elementos da forma  $f_\alpha(x)$ .

A importância de partições da unidade pode ser entendida (intuitivamente) como segue. Se  $\varphi : M \rightarrow \mathbb{R}$  é uma função diferenciável em  $M$ , a partir de (61), podemos (obviamente!) escrever

$$\varphi(x) = \sum_{\alpha} f_\alpha(x)\varphi(x),$$

ou ainda,

$$\varphi = \sum_{\alpha} f_\alpha \varphi. \quad (62)$$

As funções  $f_\alpha \varphi$  são definidas em  $M$ , mas apenas tem valores não nulos em  $U_\alpha$ . Assim, o estudo local (em  $U_\alpha$ ) de  $\varphi$  deve implicar em um estudo global de  $\varphi$  por conta da identidade (62).

Passamos a provar a existência de partições da unidade.

**Teorema 103.** *Seja  $M$  um variedade diferenciável e  $\mathcal{X} = \{U_\alpha\}_{\alpha \in A}$  uma cobertura aberta de  $M$ . Então existe uma partição diferenciável da unidade subordinada a  $\mathcal{X}$ .*

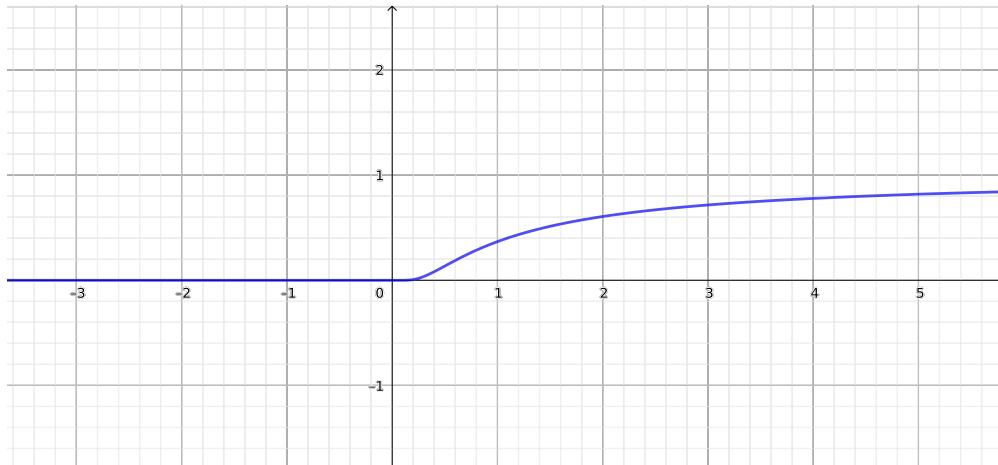
A prova do teorema que apresentamos segue as linhas de [9, 11]. Iniciamos mostrando a existência de funções "de corte".

**Exercício 51.** A função  $g : \mathbb{R} \rightarrow \mathbb{R}$  dada por

$$g(x) = \begin{cases} e^{-1/x} & \text{se } x > 0 \\ 0 & \text{se } x \leq 0 \end{cases}$$

é suave, isto é,  $g \in C^\infty(\mathbb{R})$ .

O gráfico da função  $g$  é como abaixo.

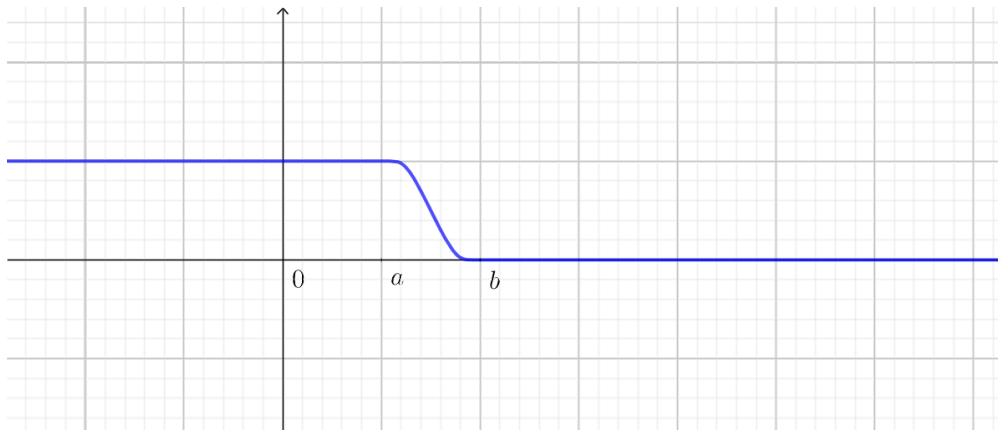


Agora, dados dois números reais satisfazendo  $0 < a < b$ , definimos  $h : \mathbb{R} \rightarrow \mathbb{R}$  por

$$h(x) := \frac{f(b-x)}{f(b-x) + f(x-a)},$$

de modo que

$$h(x) = \begin{cases} 1 & \text{se } x \leq a \\ \in (0, 1) & \text{se } a < x < b \\ 0 & \text{se } x \geq b \end{cases}$$



Desta maneira, podemos ainda construir uma função  $H \in C^\infty(\mathbb{R}^n)$  satisfazendo

$$H \equiv 1 \text{ em } B_a(0), \quad 0 < H < 1 \text{ em } B_b(0) \setminus B_a(0), \quad \text{e } H \equiv 0 \text{ em } \mathbb{R}^n \setminus B_b(0).$$

Basta definir radialmente  $H(x) := h(|x|)$ .

## Referências

- [1] John Baez and Javier P. Muniain. *Gauge fields, knots and gravity*, volume 4 of *Series on Knots and Everything*. World Scientific Publishing Co., Inc., River Edge, NJ, 1994.
- [2] Hamilton Prado Bueno. *Álgebra Linear - um segundo curso*, volume 06 of *Textos Universitários*. SBM, Rio de Janeiro, 2006.
- [3] Manfredo Perdigão do Carmo. *Formas Diferenciais e Aplicações*, volume 03 of *Coleção Fronteiras da Matemática*. SBM, Rio de Janeiro, 2015.
- [4] Manfredo Perdigão do Carmo. *Geometria Riemanniana*, volume 10 of *Coleção Projeto Euclides*. SBM, Rio de Janeiro, 2015.
- [5] L. Evans and R. Gariepy. *Measure theory and fine properties of functions*. CRC Press, 1992.
- [6] R.L. Fernandes. *Lições de Geometria Diferencial*. Lecture Notes, 2003.
- [7] Theodore Frankel. *The geometry of physics*. Cambridge University Press, Cambridge, third edition, 2012. An introduction.
- [8] Sylvestre Gallot, Dominique Hulin, and Jacques Lafontaine. *Riemannian geometry*. Universitext. Springer-Verlag, Berlin, third edition, 2004.
- [9] John M. Lee. *Introduction to smooth manifolds*, volume 218 of *Graduate Texts in Mathematics*. Springer, New York, second edition, 2013.
- [10] Elon Lages Lima. *Análise Real, Volume 2*, volume 13 of *Coleção Matemática Universitária*. IMPA, Rio de Janeiro, 2013.
- [11] Elon Lages Lima. *Curso de Análise Vol. 2*, volume 13 of *Coleção Projeto Euclides*. IMPA, Rio de Janeiro, 2015.
- [12] Elon Lages Lima. *Análise Real, Volume 3*, volume 16 of *Coleção Matemática Universitária*. IMPA, Rio de Janeiro, 2016.
- [13] R. Tyrrell Rockafellar. *Convex analysis*. Princeton Landmarks in Mathematics. Princeton University Press, Princeton, NJ, 1997. Reprint of the 1970 original, Princeton Paperbacks.
- [14] Walter Rudin. *Principles of mathematical analysis*. Second edition. McGraw-Hill Book Co., New York, 1964.
- [15] Michael Spivak. *Calculus on manifolds. A modern approach to classical theorems of advanced calculus*. W. A. Benjamin, Inc., New York-Amsterdam, 1965.
- [16] C. Villani. *Topics in optimal transportation*, volume 58 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2003.
- [17] C. Villani. *Optimal transport*, volume 338 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, 2009. Old and new.
- [18] Frank W. Warner. *Foundations of differentiable manifolds and Lie groups*, volume 94 of *Graduate Texts in Mathematics*. Springer-Verlag, New York-Berlin, 1983. Corrected reprint of the 1971 edition.