

Análise Matemática C

Diego Marcon

20 de maio de 2023

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 em que me inspirei foram [1, 3, 7, 11, 14, 15].

O semestre de 2023-1 é o quarto em que estas notas são utilizadas na disciplina e muitos já contribuíram para sua melhoria, indicando erros e sugestões. Deixo meu agradecimento especial às seguintes pessoas que contribuíram de maneira significativa: Guilherme de Lima Feltes, Lucas da Rocha Schwengber, Luiz Eduardo Morais Reck, Marnes Augusto Hoff, Rafael Haag Petasny, Thiago Oliveira Ferreira e William Matheus Michel Braucks.

Conteúdo

1	Derivada primeira	4
1.1	Funções reais de uma variável real	6
1.2	Funções vetoriais de uma variável real – curvas em \mathbb{R}^m	6
1.3	Funções escalares de várias variáveis	8
1.4	Regra da Cadeia	10
2	Teoremas de Valor Médio	10
3	Derivada segunda e derivadas de ordem mais alta	11
4	Fórmula de Taylor	13
4.1	Caso especial $k = 2$; Hessiana	15
5	Problemas de otimização - máximos e mínimos	15
5.1	Comentário sobre pontos críticos não degenerados	18
6	Funções Convexas	18
6.1	Regularidade de primeira ordem	20
6.2	Regularidade de segunda ordem – o Teorema de Alexandrov	21
7	O Teorema da Função Inversa	22
8	O Teorema da Função Implícita	27
9	Formas canônicas locais	31
10	Mergulhos	35

11 Superfícies diferenciáveis em \mathbb{R}^m	37
11.1 Aplicações diferenciáveis entre superfícies	39
11.2 Derivada de aplicações entre superfícies	41
12 Valores regulares	42
13 Superfícies Orientáveis	45
13.1 Orientação de espaços vetoriais	45
13.2 Orientação de superfícies	46
14 Multiplicadores de Lagrange	50
15 Integral de Riemann em \mathbb{R}^n	53
15.1 Conjuntos de medida zero e o Teorema de Lebesgue	56
15.2 Conjuntos Jordan mensuráveis	60
15.3 Cálculo de integrais de modo iterativo	62
15.4 Teorema de Darboux – somas de Riemann mais gerais	64
15.5 Fórmula de mudança de variáveis	67
15.5.1 Coordenadas polares	71
15.5.2 Coordenadas esféricas	71
16 Formas diferenciais	73
17 Produto exterior – motivação	78
18 Produto tensorial	80
18.1 Produto tensorial de dois espaços vetoriais	81
18.2 Identificações naturais	84
18.2.1 Vetores são tensores: $V \simeq \mathbb{R} \otimes V$	84
18.2.2 Transformações lineares são tensores: $\mathcal{L}(W; V) \simeq V \otimes W^*$	85
18.2.3 Transformações bilineares são tensores: $\mathcal{B}(V \times W; \mathbb{R}) \simeq V^* \otimes W^*$	86
18.3 Produto tensorial de vários espaços vetoriais	86
18.4 Álgebra tensorial	87
19 Produto exterior	88
19.1 Álgebra exterior ou álgebra de Grassmann	91
19.2 Identificação I; Aplicações bilineares alternadas	92
19.3 Identificação II; Aplicações multilineares alternadas	93
20 Formas diferenciais de grau p e a derivada exterior	94
20.1 O <i>pullback</i> de formas diferenciais	97
21 Formas exatas e fechadas	100
22 Variedades diferenciáveis	102
22.1 Vetores tangentes e o espaço tangente	104
22.1.1 Vetores tangentes como velocidades de curvas	105
22.1.2 Vetores tangentes como derivações	106
22.1.3 A identificação entre as duas definições dadas	106
22.1.4 Em coordenadas	109
22.2 Exemplos	110
22.3 Diferencial	111
23 Partições da Unidade	111

24	Integrais de Superfície	115
24.1	Motivação e definições iniciais de integral	117
24.2	Definição definitiva via partições da unidade	118
25	Variedades com bordo	119
25.1	Espaço tangente em pontos de ∂M	122
25.2	Orientação induzida no bordo ∂M	122
26	Teorema de Stokes	123
26.1	O Teorema da Divergência de Gauss-Ostrogradsky	126
26.2	O Teorema do Rotacional de Kelvin-Stokes	126

1 Derivada primeira

Vamos considerar os espaços vetoriais \mathbb{R}^n e \mathbb{R}^m , onde $m, n \in \{1, 2, 3, \dots\}$. 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 $x_0 \in U$ quando existe uma transformação linear¹ que “aproxima” f perto de x_0 . Mais precisamente, f é **diferenciável** em $x_0 \in U$ quando existe uma transformação linear $T : \mathbb{R}^n \rightarrow \mathbb{R}^m$ tal que, para todo $x \in U$,

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

Nesse caso², denotamos $T = f'(x_0)$ ou $T = Df(x_0)$ e dizemos que T é a **derivada** de f em x_0 .

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 $x_0 \in U \subseteq \mathbb{R}^n$.

(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 $f : U \subseteq \mathbb{R}^n \rightarrow \mathbb{R}^m$ no ponto $x \in U$ e 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 de um vetor $v = e_j$ pertencente à base canônica de \mathbb{R}^n , são bastante comuns as notações:

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

Uma função $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$, cuja variável vetorial está em \mathbb{R}^n e cujos valores são vetores em \mathbb{R}^m , pode ser escrita em termos de suas componentes

$$\begin{aligned} f_1 : \mathbb{R}^n &\rightarrow \mathbb{R} \\ f_2 : \mathbb{R}^n &\rightarrow \mathbb{R} \\ &\vdots \\ f_m : \mathbb{R}^n &\rightarrow \mathbb{R} \end{aligned}$$

¹Na 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 .

²Rigorosamente, antes disso deveríamos mostrar que, no caso de f ser diferenciável, a transformação T que, pela definição, existe, é única. Este é um exercício da nossa primeira lista de exercícios.

que satisfazem,³ para todo $x \in \mathbb{R}^n$,

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

Assim, em termos das componentes da função f , podemos escrever

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

Lembramos de Álgebra Linear que a matriz canônica de uma transformação linear é aquela 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: $\det [f'(x)]$. Tanto a matriz como o seu determinante tem grande importância e podemos nos referir a ambos como o Jacobiano quando o contexto deixar claro o que se quer dizer. Quando o contexto não deixar claro, nos referimos à **matriz Jacobiana** e ao **determinante Jacobiano**.

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. Na realidade, os conceitos de continuidade e diferenciabilidade fazem sentido para funções definidas (e com valores) em espaços vetoriais normados (lista de exercícios), de modo que a identificação acima nem é de fato necessária.

³Vetores nos espaços \mathbb{R}^k são tradicionalmente denotados por vetores coluna e nós seguimos esta convenção. Em cursos de cálculo, é comum escrever um vetor $v \in \mathbb{R}^k$ na base canônica por $v = (v_1, v_2, \dots, v_k)$ ou mesmo $v = \langle v_1, v_2, \dots, v_k \rangle$. Eventualmente, usamos esta notação ao escrever, por exemplo, as variáveis de uma função; isto não deve se confundir, no entanto, com um vetor linha.

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 $x \in U$ e $v \in \mathbb{R}^n$, existem as derivadas direcionais $\frac{\partial f}{\partial v}(x)$ e, além disso, para cada $v \in \mathbb{R}^n$,

$$\frac{\partial f}{\partial v} : U \rightarrow \mathbb{R}^m$$

é uma função contínua.

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 $f'(a) : \mathbb{R} \rightarrow \mathbb{R}$ que satisfaz

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

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 1×1 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 $h \in \mathbb{R}$, não nulo, existe o inverso multiplicativo $h^{-1} \in \mathbb{R}$. Neste caso, podemos escrever, como de costume,

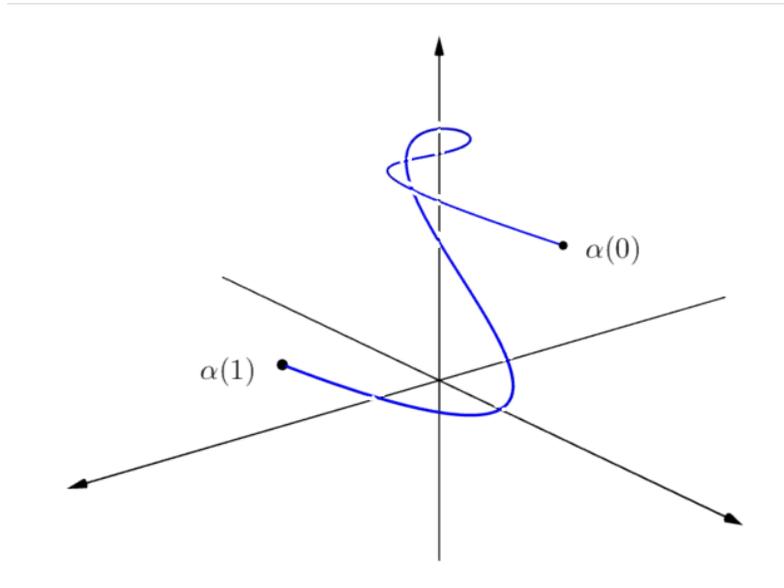
$$f'(a) = \lim_{h \rightarrow 0} \left(\frac{f(a+h) - f(a)}{h} - \frac{r(h)}{h} \right) = \lim_{h \rightarrow 0} \frac{f(a+h) - f(a)}{h}.$$

Em Cálculo 1, muitas vezes se estuda derivada como o coeficiente da reta tangente ao gráfico da função. É justamente isto que nossa definição diz, com as devidas identificações, ao escrever

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

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 α é uma **parametrização** da curva C . Ou ainda, que α é uma **curva parametrizada**. Isto é porque α , 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 α de curva.

Em coordenadas (usualmente usando a base canônica de \mathbb{R}^m), podemos escrever

$$\alpha(t) = \begin{bmatrix} \alpha_1(t) \\ \alpha_2(t) \\ \vdots \\ \alpha_m(t) \end{bmatrix} \in \mathbb{R}^m.$$

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

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

Acima escrevemos, naturalmente, as derivadas “parciais” como

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

Ao olharmos para a definição de diferenciabilidade em que aparece $h \in \mathbb{R}$, notamos (obviamente) que h é um número real. Assim, identificando a transformação linear $\alpha'(t)$ com a sua matriz canônica associada, é possível usar as representações acima para escrever, como de costume,

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

Geometricamente, após estas identificações, $\alpha'(t)$ é o **vetor tangente** à curva α 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 que a curva. A identificação acima é natural (ver exercício abaixo) e pode ser usada para considerar como equivalentes os objetos

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

Exercício 3. A identificação $\mathcal{L}(\mathbb{R}, \mathbb{R}^m) \simeq \mathbb{R}^m$ é um isomorfismo canônico, isto é, independe de escolha de bases para os espaços vetoriais 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 (Regra do produto). Para $\alpha : \mathbb{R} \longrightarrow \mathbb{R}^m$ e $\beta : \mathbb{R} \longrightarrow \mathbb{R}^m$ curvas diferenciáveis, mostre a regra do produto

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

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

$$\alpha'(t) \perp \alpha''(t) \quad \text{para todo } t \in \mathbb{R}.$$

Em outras palavras, se uma curva é tal que a magnitude do vetor velocidade permanece constante, então seu vetor velocidade é ortogonal ao seu vetor aceleração em todos os pontos da curva.

1.3 Funções escalares de várias variáveis

Este é o caso $m = 1$, de modo que funções da forma $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ são os nossos objetos de estudo. Supondo f diferenciável, a derivada de f em um ponto $a \in \mathbb{R}^n$ é uma transformação linear do tipo $f'(a) : \mathbb{R}^n \longrightarrow \mathbb{R}$. Observe que, neste caso, 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 podem ser calculadas como segue

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

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), uma vez fixado o sistema de coordenadas canônico, 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 **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 definido⁴, nos elementos de \mathcal{B} , por

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

⁴Notamos que, apesar de estarmos definindo os funcionais dx_i pela expressão (6), 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).

Desta forma,

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

Relembrando (3), segue 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

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

Esta fórmula torna rigorosa (como uma expressão no espaço dual) a representação do diferencial de f no ponto a em termos dos “diferenciais” dx_i , com $i \in \{1, 2, \dots, n\}$.

1.4 Regra da Cadeia

Terminamos esta seção lembrando como analisar a diferenciabilidade da composições de funções, que 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,

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

Observe que o que temos do lado direito da identidade acima é uma composição de transformações lineares que, em coordenadas, pode ser visto como um produto de matrizes. Olhando para as componentes deste produto, podemos escrever, para $i = 1, 2, \dots, p$ e $j = 1, 2, \dots, n$,

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

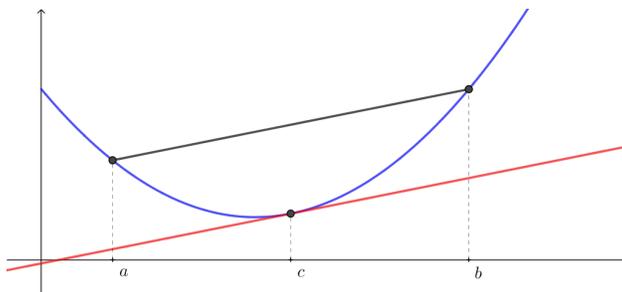
2 Teoremas de Valor Médio

Lembra que, em uma variável, vale o seguinte.

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.



Uma versão moderna deste resultado aparece já no trabalho de Cauchy em 1823, embora versões preliminares tenham aparecido anteriormente (ver wikipedia.org/meanvalue).

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

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 contido⁵ em U . Então, existe $c \in (x, y)$ tal que*

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

Já que c pertence ao segmento que une x e y , podemos escrever $c = (1-t)x + ty$ para algum $t \in (0, 1)$.

Demonstração. Inicialmente mostramos o caso especial $n = 1$ (ou seja, provamos que a desigualdade vale para curvas). Neste caso, podemos supor que $U = [a, b] \subseteq \mathbb{R}$. Para reduzir a prova ao caso unidimensional, definimos uma função auxiliar $\phi : U \rightarrow \mathbb{R}$ por

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

Geometricamente, esta expressão está relacionada ao ângulo que o vetor posição $f(t)$ faz com o vetor “deslocamento” $f(y) - f(x)$. Pelo Teorema do Valor Médio (Teorema 3 acima) e pela definição da norma de operadores, 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 $n = 1$ e $f(y) \neq f(x)$. Por outro lado, se $f(y) = f(x)$, (8) é trivial.

Para o caso geral, consideramos uma parametrização do segmento entre x e y e recaímos no caso que já mostramos restringindo a função f a este segmento: 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 ao definir $c = (1-t_0)x + t_0y$. □

Exercício 6. Prove o Teorema 4 diretamente por integração. É necessário adicionar alguma hipótese para formalizar a demonstração sugerida neste exercício?

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

3 Derivada segunda e derivadas de ordem mais alta

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)

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

⁵Supomos convexidade apenas para que o segmento $[x, y]$ esteja inteiramente contido em U . É possível enunciar versões sem convexidade do domínio.

que satisfaz (poderíamos enunciar de outras maneiras, análogas a (i)–(iv) do início destas notas)

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

A norma que aparece no numerador acima é a norma de operadores em $\mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$, isto é: para $T \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$,

$$\|T\| := \sup_{|x|=1} |T(x)|.$$

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

Exercício 8. Considerando

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

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, a segunda derivada $f''(x_0)$ de f em 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,

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

Em uma notação talvez seja mais comum ao leitor, podemos escrever

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

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, a quarta derivada como uma forma quadrilinear (ou 4-linear) e assim por diante. Além disso, existem identificações naturais análogas às do Exercício 8 para estes casos. Por exemplo,

$$\begin{aligned} \{\mathcal{T} : \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^n \longrightarrow \mathbb{R}^m; \mathcal{T} \text{ é trilinear}\} &\simeq \mathcal{L}(\mathbb{R}^n, \mathcal{L}(\mathbb{R}^n, \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m))) \\ &\simeq \mathcal{L}(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n; \mathbb{R}^m)) \end{aligned}$$

Fica de exercício ao leitor descrever a identificação de maneira precisa.

Nós dizemos ainda 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;
- \vdots
- f é de classe C^k quando existem as derivadas de f até ordem k e $f^{(k)}$ é contínua;
- \vdots
- f é de classe C^∞ quando existem as derivadas de todas as ordens de f .

4 Fórmula de Taylor

Uma das versões (com resto de Lagrange) da Fórmula de Taylor unidimensional nos diz que se $f : \mathbb{R} \rightarrow \mathbb{R}$ é uma função de classe C^{k+1} , então existe ξ no intervalo aberto entre a e x tal que

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

O objetivo desta seção é obter uma versão multidimensional deste resultado e representá-la de maneira eficiente.

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 ξ no segmento de a para x tal que*

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

Observamos como a notação do teorema acima é escolhida convenientemente de maneira a lembrar a fórmula unidimensional. Nós deveríamos escrever, mais precisamente,

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

A transformação $f^{(j)}(a)$ é uma aplicação j -linear e $f^{(j)}(a) \cdot (x-a)^j$ (para nós) significa aplicar o mesmo vetor $x-a$ nas j entradas de $f^{(j)}(a)$.

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

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

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 conveniente 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. Assim, ao escrever $D^\alpha f(a)$, o número natural α_i que está na i -ésima posição do multi-índice $\alpha \in \mathbb{N}^n$ indica que são tomadas α_i derivadas parciais com respeito à variável x_i . Por exemplo, para $f : \mathbb{R}^5 \rightarrow \mathbb{R}$, denotamos

$$\frac{\partial^6 f}{\partial x_2^3 \partial x_3^2 \partial x_4} (a) = D^\alpha f(a), \quad \text{onde } \alpha = (0, 3, 2, 1, 0).$$

Também é comum utilizar multi-índices para denotar monômios em várias variáveis: com o mesmo multi-índice α dado no exemplo acima, temos

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

De maneira geral, denotamos $\alpha = (\alpha_1, \alpha_2, \dots, \alpha_n) \in \mathbb{N}^n$ e $|\alpha| = \alpha_1 + \alpha_2 + \dots + \alpha_n$ e escrevemos

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

de modo que, nestes casos, $|\alpha|$ denota o grau do monômio e a ordem de diferenciação, respectivamente. Ainda, é também comum escrever

$$\alpha! = \alpha_1! \alpha_2! \dots \alpha_n! \quad \text{e} \quad \binom{n}{\alpha} = \frac{n!}{\alpha!} = \frac{n!}{\alpha_1! \alpha_2! \dots \alpha_n!},$$

o que simplifica a notação de expansões multinomiais como feito abaixo.

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

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

Na notação introduzida acima, podemos escrever, mais sucintamente,

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

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

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

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

Apesar de a notação do lado direito acima parecer pouco natural à primeira vista, é um dos objetivos (mais adiante no curso) justificar que, de fato, esta é uma maneira muito conveniente de se pensar na derivada de f na direção v . Em seguida, iterando e utilizando o Teorema Multinomial, obtemos⁶

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

Portanto, (10) é o mesmo que

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

⁶O Teorema Multinomial funciona sempre que pudermos distribuir os produtos e utilizar a comutatividade dos elementos. No caso das derivadas parciais, é possível comutar sempre que a função f for suficientemente regular, de modo que podemos aplicar a fórmula dada no Teorema Multinomial. Qual é a regularidade necessária? Este é o conteúdo do Lema de Schwarz.

4.1 Caso especial $k = 2$; Hessiana

Para a maior parte das aplicações, é suficiente escrever a Fórmula de Taylor até apenas ordem 2. 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 a matriz simétrica associada à Hessiana: utilizando a base canônica,

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

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, enunciamos 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 ξ 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. Da maneira que está enunciado, é realmente necessário supor $f \in C^3$?

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 primeira ordem (no contexto de otimização, a ordem se refere à ordem de diferenciação analisada na condição).

Proposição 7. *Suponhamos que $a \in \text{int } U$ é um ponto interior e de mínimo local para $f : U \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$. Suponhamos ainda 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, vale

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

Se w é um vetor unitário na mesma direção de v , podemos escrever $w = tv$. Logo, temos que, para qualquer t suficientemente pequeno,

$$df(a) \cdot w \geq -\frac{r(tv)}{t}.$$

Fazendo $t \rightarrow 0$, concluímos que $df(a) \cdot w \geq 0$. Além disso, este raciocínio vale para todo w unitário. Em particular, podemos trocar w por $-w$ e segue que $df(a) \cdot w = 0$ para todo w unitário. Isto implica que $df(a)$ é o funcional linear nulo, ou seja, $df(a) = 0$. \square

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

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

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

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

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

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

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

Como na proposição anterior, podemos escrever $h = tw$ com w unitário, dividir por t^2 e fazer $t \rightarrow 0$ para concluir 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 $|v|^2$. □

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:

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

Dividindo por $|h|^2$ podemos escrever

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

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

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

Agora, como ρ vai para zero quando $h \rightarrow 0$, existe $\delta > 0$ tal que

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

concluindo a demonstração. □

Exercício 13. Adaptar o raciocínio da Proposição 9 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**, que quer dizer que $\langle \nabla^2 f(a)v, v \rangle$ “muda de sinal”, isto é, assume tanto valores positivos quanto negativos ao variar v .

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!

Por outro lado, 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

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

Logo,

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

Em seguida, juntando os termos com x_2 :

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

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 Inversa⁷ g é invertível em uma vizinhança de a . Em particular, é injetiva e $g(x) \neq g(a) = 0$ nesta vizinhança se $x \neq a$. 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**^{8,9} quando f não é identicamente $+\infty$ e, para todo $x, y \in \mathbb{R}^n$ e para todo $0 < \lambda < 1$, tem-se

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

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

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

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

Podemos verificar que $\text{dom } f$ é um conjunto convexo (exercício). Além disso, a definição pode ser estendida para englobar mais de dois pontos:

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, o campo gradiente ∇f é um campo monótono em seus pontos de diferenciabilidade, ou seja, vale que

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

⁷Mais diretamente, o Passo 1 da prova do Teorema da Função Inversa que demonstramos mais adiante, mostra a injetividade de g em uma vizinhança de a , o que basta para a prova desta proposição.

⁸No 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].

⁹Permitir que a função assuma o valor $+\infty$ é interessante por vários motivos. Alguns deles seguem. Trabalhar com funções com mesmo domínio: todo espaço \mathbb{R}^n (para tanto, é possível estender funções convexas para além do seu domínio convexo de definição pelo valor $+\infty$, mantendo convexidade). O supremo de funções convexas é assim uma função convexa e podemos escrever (14) e (15) sem demasiada preocupação de o supremo ser finito. Além disso, o domínio efetivo de f , que pode ser pensado como o conjunto onde a função estaria usualmente definida, é convexo de modo que nosso conceito é mais geral. Ainda, outras propriedades relacionadas à Transformada de Legendre (e semicontinuidade) são endereçadas na lista de exercícios.

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. \square

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,

$$f(y) = \sup_{x \in \mathbb{R}^n} [f(x) + \langle \nabla f(x), y - x \rangle]. \quad (14)$$

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

Exercício 16. Mostre que o supremo pontual de uma família qualquer de funções convexas é também uma função convexa, isto é, dadas f_λ convexas, temos que

$$f(y) := \sup_{\lambda \in \Lambda} f_\lambda(y) \quad (15)$$

é convexa.

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

Proposição 15. *Seja $f : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$ 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. \Rightarrow 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.

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

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

e o resultado segue do anterior.

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:

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

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 particular¹⁰, f é diferenciável em quase todo ponto de $\text{dom } f$.*

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 $n = 1$, por ser um aberto convexo da reta, $\text{int dom } f$ deve ser um intervalo (a, b) . Em dimensão um, é possível caracterizar funções convexas como segue (faça um desenho e forneça uma interpretação geométrica dessa condição): dados $a < x < b$, temos

$$\frac{f(x) - f(a)}{x - a} \leq \frac{f(b) - f(a)}{b - a} \leq \frac{f(b) - f(x)}{b - x}. \quad (16)$$

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,

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

¹⁰O Teorema de Rademacher (ver, por exemplo, [5, Section 3.1.2]) afirma que funções Lipschitz são diferenciáveis em quase todo ponto, o que diferenciabilidade exceto, no máximo em um conjunto de medida nula (a definição de conjuntos de medida nula aparece mais adiante no curso e é semelhante à que se conhece da análise na reta).

Considerando $s \in (0, 1)$ tal que $x = sa + (1 - s)b$, provamos a segunda desigualdade em (16).
 Considere¹¹ $(c, d) \subset\subset (a, b)$. Iterando (16), 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 $x, y \in (c, d)$. □

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 têm comprimento $2r$ e são paralelas aos 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,

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

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 que¹²

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

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 $B_{r-\delta}(x_0)$. □

6.2 Regularidade de segunda ordem – o Teorema de Alexandrov

Em seguida, descrevemos superficialmente 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.

¹¹Lê-se a notação $B \subset\subset A$ como “ B compactamente contido em A ”, o que significa que $B \subseteq \bar{B} \subset A$ e que \bar{B} deve ser compacto. Assim, se A é um conjunto aberto, considerar um conjunto aberto $B \subset\subset A$ significa considerar um aberto $B \subset A$ “que está longe da fronteira de A ”, já que devemos ter $\text{dist}(B, \partial A) > 0$.

¹²A escolha de z pode ser mais facilmente descrita geometricamente: dados $x, y \in B_{r-\delta}(x_0)$, traçamos a reta que contém os pontos x e y e, a partir do ponto y , andamos uma distância δ na direção contrária à de x , definindo o ponto z . Dessa forma, precisamos considerar o raio $r - \delta$ para que o ponto z pertença a $B_r(x_0)$. Além disso, y está, por construção, no interior do intervalo de extremidades x e z e pode ser escrito como combinação convexa destes pontos.

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. □

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 Google¹³ mostrou que o Teorema da Função Inversa já aparece no trabalho de Ulisse Dini¹⁴, em “Lezioni di Analisi Infinitesimale”, de 1877.



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 suponhamos que, para algum $a \in E$, a transformação linear $f'(a) : \mathbb{R}^n \rightarrow \mathbb{R}^n$ 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) *a função f é injetiva em U ;*
- (ii) *$f(U) = V$, de modo que $f : U \rightarrow V$ é uma bijeção;*

¹³<https://mathoverflow.net/questions/94323/who-was-the-first-to-formulate-the-inverse-function-theorem>, visualizado em 30 de Janeiro de 2018

¹⁴Matemático 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 seqüências 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 em 03 de Janeiro de 2018. Segue um retrato de Dini (que também pode ser encontrado no link acima).

(iii) a função inversa f^{-1} é de classe $C^1(V; U)$ e vale que, para $y = f(x) \in V$

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

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 $f(a)$.

Exercício 17. Mostrar que, se $f \in C^k$ e $f'(a)$ é um isomorfismo, então f é um difeomorfismo de classe C^k 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

$$\|B - A\| < \frac{1}{\|A^{-1}\|}, \quad (17)$$

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 $M(n)$ ¹⁵.

(ii) A aplicação $\Phi : GL(n) \rightarrow GL(n)$ dada por $\Phi(A) = A^{-1}$, é um homeomorfismo¹⁶.

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

$$k|x| \leq |Bx|, \quad (18)$$

o que implica que B é injetiva. Sendo uma matriz quadrada, concluímos que é invertível. Para verificar (18), 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 (18) com

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

Para provar (ii), notamos que

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

¹⁵Poderíamos mostrar diretamente que $GL(n)$ é aberto pela continuidade da função determinante:

$$GL(n) = \{A \in M(n); \det A \neq 0\}.$$

Por outro lado, o item (i) contém uma informação mais precisa sobre o tamanho da bola centrada em A que contém apenas matrizes invertíveis. No caso de uma matriz A invertível que é também normal (isto é, que comuta com sua adjunta A^*), temos (verifique!)

$$\|A\| = \sup\{|\lambda|; \lambda \text{ é autovalor de } A\} \text{ enquanto } \|A^{-1}\|^{-1} = \min\{|\lambda|; \lambda \text{ é autovalor de } A\}.$$

Assim, $B_r(A) \subset GL(n)$ onde o raio r é o menor (em magnitude) autovalor de A . Isto está em concordância com o fato que uma matriz não é invertível se, e somente se, o valor zero é um de seus autovalores.

¹⁶Isto é, Φ é contínua, invertível (neste caso, $\Phi^{-1} = \Phi$) e com inversa contínua. Na realidade, Φ é um difeomorfismo: fica como exercício verificar que

$$\Phi'(A) \cdot H = -A^{-1}HA^{-1}.$$

Este resultado concorda com o caso (comutativo) unidimensional: $\phi(t) = 1/t \implies \phi'(t)h = -h/t^2$.

Além disso, fazendo $x = B^{-1}y$ em (18), 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

$$\|\Phi(B) - \Phi(A)\| = \|B^{-1} - A^{-1}\| \leq \|B^{-1}\| \|A - B\| \|A^{-1}\| \leq \frac{\|A^{-1}\|^2}{1 - \|A^{-1}\|\|B - A\|} \|B - A\|. \quad (19)$$

Observamos que, para $A \in GL(n)$ fixado, o lado direito de (19) vai para zero quando $\|B - A\| \rightarrow 0$. Logo, dado $\varepsilon > 0$, existe $\delta > 0$ tal que o lado direito de (19) fica menor do que ε 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 resultados independentes podem ser úteis em outras situações.

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 $\varepsilon > 0$,

$$|f(x+h) - f(x)| \geq \varepsilon|h|. \quad (20)$$

Sendo a desigualdade (20) 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 (18) que é a versão linear do argumento de injetividade.

Devemos, assim, mostrar (20). Note que, pela desigualdade triangular,

$$|f(x+h) - f(x)| \geq |f'(a) \cdot h| - |f(x+h) - f(x) - f'(a) \cdot h|. \quad (21)$$

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 $R > 0$ 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

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

Intuitivamente, nos restringimos à direção h e é como se estivéssemos tentando fazer a aproximação linear de f em x , mas utilizando $f'(a)$ ao invés de $f'(x)$. Por isso que a hipótese de que f' é contínua vai ser importante. Pela Desigualdade do Valor Médio, existe $t_0 \in (0, 1)$ que satisfaz

$$|f(x+h) - f(x) - f'(a) \cdot h| = |F(1) - F(0)| \leq |F'(t_0)|. \quad (22)$$

Temos ainda

$$|F'(t_0)| = |f'(x+t_0h) \cdot h - f'(a) \cdot h| \leq \|f'(x+t_0h) - f'(a)\| |h|. \quad (23)$$

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

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

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

Agora, utilizamos (22) e (23) 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

$$\varepsilon|h| = \varepsilon|f'(a)^{-1}f'(a) \cdot h| \leq \varepsilon\|f'(a)^{-1}\| |f'(a) \cdot h| < \frac{|f'(a) \cdot h|}{2}. \quad (24)$$

Assim, (21) implica

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

onde, na última desigualdade, utilizamos (24) novamente.

Passo 2. Escrevemos $U := B_R(a)$ e $V := f(U)$ e temos, pelo Passo 1, que $f : U \rightarrow V$ é uma bijeção. Vamos verificar, nas hipóteses do teorema, que V é um aberto do \mathbb{R}^n ; na verdade, o argumento que utilizamos abaixo mostra que $f : U \rightarrow V$ 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 f^{-1} é uma aplicação contínua¹⁷.

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 $\overline{B_r(x_0)} \subset U$. Afirmamos 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

$$|y - f(x_0)| < \frac{\varepsilon r}{2}. \quad (25)$$

A função $\phi(x) := |y - f(x)|$ é contínua no conjunto compacto $\overline{B_r(x_0)}$. Logo, assume um valor mínimo em algum ponto $x^* \in \overline{B_r(x_0)}$. Primeiro, vamos verificar que x^* pertence ao interior da bola: se fosse $|x^* - x_0| = r$, teríamos que, usando (20) e (25),

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

Mas daí, utilizando (25) novamente,

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

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

Agora, notamos que x^* também minimiza a função $\psi(x) = |f(x) - y|^2$ em $\overline{B_r(x_0)}$. Logo, sendo x^* um ponto interior, temos que é um ponto crítico, isto é, $2(f(x^*) - y)^T f'(x^*) = 0$. Como $f'(x^*)$ é bijetiva, concluímos que deve ser $y = f(x^*)$. Lembrando que $x^* \in B_r(x_0)$, temos $y \in f(B_r(x_0))$, 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

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

Desta forma, temos as relações

$$k = f(x+h) - f(x) \quad \text{e} \quad h = f^{-1}(y+k) - f^{-1}(y). \quad (26)$$

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 (26), temos

$$f^{-1}(y+k) - f^{-1}(y) = h = f'(x)^{-1} \cdot k - f'(x)^{-1} \cdot r(h) =: f'(x)^{-1} \cdot k + \rho(k). \quad (27)$$

¹⁷Uma aplicação g é contínua em \mathbb{R}^n se, e somente se, $g^{-1}(A)$ é aberto para qualquer A aberto.

Por (20), 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

$$\frac{|\rho(k)|}{|k|} = \frac{|f'(x)^{-1} \cdot r(h)|}{|k|} \leq \frac{\|f'(x)^{-1}\| |r(h)|}{|k|} = \frac{\|f'(x)^{-1}\| |h|}{|f(x+h) - f(x)|} \cdot \frac{|r(h)|}{|h|} \stackrel{(20)}{\leq} \frac{\|f'(x)^{-1}\|}{\varepsilon} \cdot \frac{|r(h)|}{|h|}.$$

Portanto,

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

e (27) 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,

$$[f'(x, y)]_{\text{Can}} = \begin{bmatrix} e^x \cos y & e^x \sin y \\ -e^x \sin y & e^x \cos y \end{bmatrix} \implies \det f'(x, y) = e^{2x} > 0 \quad \text{para todo } (x, y) \in \mathbb{R}^2.$$

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, já que f não é injetiva: $f(x, y + 2\pi) = f(x, y)$.

Exercício 18. 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 19. 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 .

Os pontos críticos tais que $\nabla^2 f(a)$ é invertível são chamados de pontos críticos não degenerados. Neste caso, o exercício afirma que todos os pontos críticos não degenerados de uma função de classe C^2 são isolados.

Observação 21. O Teorema da Função Inversa pode ser interpretado como uma afirmação sobre existência e unicidade de soluções para sistemas (possivelmente) não lineares de equações. Se $E \subseteq \mathbb{R}^n$ é um aberto e $f : E \rightarrow \mathbb{R}^n$ é de classe C^1 , nós podemos escrever

$$y = f(x) = (f_1(x), f_2(x), \dots, f_n(x)).$$

Além disso, se $x = (x_1, x_2, \dots, x_n)$ e $y = (y_1, y_2, \dots, y_n)$, a igualdade $y = f(x)$ é equivalente a

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

Desta maneira, resolver o sistema equivale a encontrar funções $\xi_1, \xi_2, \dots, \xi_n$ de modo que

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

O Teorema da Função Inversa nos diz que, em uma vizinhança de um ponto $a = (a_1, a_2, \dots, a_n)$, isto é possível quando a matriz das derivadas parciais

$$\left[\frac{\partial f_i}{\partial x_j}(a) \right]_{n \times n}$$

é uma matriz invertível. No caso de sistemas lineares, a matriz que deve ser invertível (como condição suficiente para resolução do sistema) é a própria matriz que define o sistema linear, como era de se esperar.

8 O Teorema da Função Implícita

O Teorema da Função Implícita, historicamente, apareceu como uma condição suficiente 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 em termos das demais variáveis:

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

Esta formulação, em termos de sistemas lineares, sugere uma relação direta com o Teorema da Função Inversa.

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

Exemplo 22. Vamos resolver o sistema linear $Ax = 0$, onde a matriz A , de ordem 3×5 , 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. Neste caso, como indicam as contas abaixo, nós poderíamos considerar, por exemplo, $y_1 := x_4$ e $y_2 := x_5$, embora não seja necessário mudar a notação.

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

$$\dim \text{Nuc } A + \dim \text{Im } A = 5.$$

A dimensão do núcleo de A representa o número de variáveis livres do sistema homogêneo associado. Ao escolher convenientemente as variáveis livres, sabemos que as demais podem ser escritas em termos das livres.

Mais geralmente, temos o seguinte exercício de Álgebra Linear.

Exercício 20. Seja $T \in \mathcal{L}(\mathbb{R}^{n+k}, \mathbb{R}^n)$ tal que

$$T(x, 0) = 0 \implies x = 0. \quad (28)$$

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 enunciamos 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 20, então a conclusão também é válida, localmente, para a função f , que é possivelmente não linear. Aparentemente, na Itália, o Teorema da Função Implícita também é conhecido como Teorema de Dini.

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

Exercício 21. 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 23 (Teorema da Função Implícita). *Seja $f : E \subseteq \mathbb{R}^{n+m} \rightarrow \mathbb{R}^n$ uma função de classe C^1 . Suponhamos que $(a, b) \in E$ é tal que*

- $f(a, b) = 0$ e;
- $f'(a, b) : \mathbb{R}^{n+m} \rightarrow \mathbb{R}^n$ satisfaz algum dos (e, portanto, todos os) itens do Exercício 21.

Então,

- existem vizinhanças V de $a \in \mathbb{R}^n$ e W de $b \in \mathbb{R}^m$ com $V \times W \subseteq E$ e;
- existe única função $\xi : W \rightarrow V$ de classe C^1

tais que, para todo $(x, y) \in V \times W$,

$$f(x, y) = 0 \iff x = \xi(y).$$

Além disso, as vizinhanças V e W podem ser tomadas de modo que $\partial_x f(x, y)$ é invertível, para qualquer $(x, y) \in V \times W$, e temos

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

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

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

Dessa forma, diferenciando coordenada a coordenada, obtemos

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

Como $F'(a, b)$ está definida em espaços vetoriais de mesma dimensão, para verificar que é invertível, basta verificar que é injetiva. Sendo assim estudamos o núcleo de $F'(a, b)$:

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

Nossa hipótese garante então que $h = 0$ e, logo, $F'(a, b)$ é invertível. Pelo Teorema da Função Inversa, existem vizinhanças (abertas) U_1 de $(a, b) \in \mathbb{R}^{n+m}$ e U_2 de $(0, b) \in \mathbb{R}^{n+m}$ onde $F|_{U_1} : U_1 \rightarrow U_2$ admite uma inversa $G := (F|_{U_1})^{-1}$. Olhando para as “componentes” de G , podemos escrever

$$G(x, y) = (g(x, y), h(x, y)), \quad (x, y) \in U_2.$$

Como G é de classe C^1 , também suas componentes o são. Já que $F \circ G = \text{id}$, para todo $(x, y) \in U_2$, devemos ter

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

Em particular, para todo $(x, y) \in U_2$, temos

$$h(x, y) = y \quad \text{e} \quad f(g(x, y), y) = x.$$

Esta última identidade descreve os diferentes “conjuntos de nível” de f na vizinhança U_2 e sugere a definição $\xi(y) := g(0, y)$.

Concluimos a demonstração observando que é possível escolher as vizinhanças convenientemente como no enunciado: sendo U_1 aberto que contém o ponto (a, b) , consideramos uma vizinhança “retangular” da forma $V \times \tilde{W} \subseteq U_1$, onde V é uma vizinhança de $a \in \mathbb{R}^n$ e \tilde{W} é uma vizinhança de $b \in \mathbb{R}^m$. Em seguida, como $F|_{U_1}$ é um difeomorfismo, o conjunto $F(V \times \tilde{W}) \subseteq U_2$ é aberto. Daí escolhemos uma vizinhança $W \subseteq \tilde{W}$ de b tal que $\{0\} \times W \subseteq F(V \times \tilde{W})$. Assim, fica bem definida a função $\xi : W \rightarrow V$ por

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

A unicidade de ξ segue da injetividade de $F|_{U_1} : U_1 \rightarrow U_2$. □

Exercício 22. Pela Regra da Cadeia, escrever uma equação diferencial que a aplicação ξ 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 24. 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

$$\begin{bmatrix} 2x & 2y & 2z \end{bmatrix}.$$

Qualquer coluna dessa matriz forma um conjunto linearmente independente (com um vetor só) 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 25. *Seja $f : E \subseteq \mathbb{R}^{k+1} \rightarrow \mathbb{R}$ uma função de classe C^1 . Suponhamos também que o ponto $(a_1, a_2, \dots, a_k, a_{k+1}) \in E$ 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, existem vizinhanças W do ponto $(a_1, a_2, \dots, a_k) \in \mathbb{R}^k$ e V de $a_{k+1} \in \mathbb{R}$ e existe única função $\xi : W \rightarrow V$ de classe C^1 tais que $a_{k+1} = \xi(a_1, a_2, \dots, a_k)$ e, para todo $(x_1, x_2, \dots, x_k, y) \in W \times V$,

$$f(x_1, x_2, \dots, x_k, y) = 0 \iff y = \xi(a_1, a_2, \dots, a_k).$$

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 26 (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

$$\begin{bmatrix} 2 & -4 \\ -6 & 0 \end{bmatrix},$$

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^∞ tais que

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

Devemos ainda ter as condições (conferir com o Exercício 22)

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

ou ainda,

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

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 (são os difeomorfismos locais).

Exercício 23. 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.

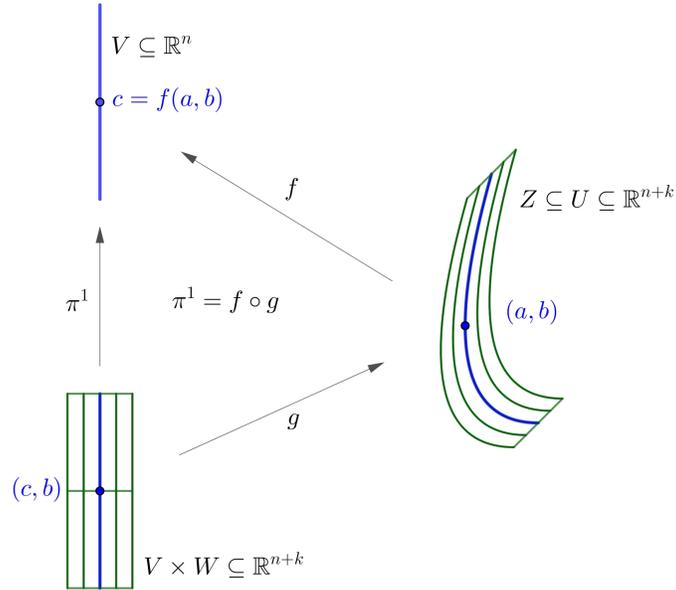
Teorema 27 (Forma Local de Submersões). *Seja $f : U \subseteq \mathbb{R}^{n+m} \rightarrow \mathbb{R}^n$ uma submersão de classe $C^k(U; \mathbb{R}^n)$ e $(a, b) \in U \subseteq \mathbb{R}^{n+m}$. Então, existem*

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

tais que, para todo $(u, v) \in V \times W$, 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+m} da forma mais intuitiva possível em \mathbb{R}^{n+m} e aplicar o Teorema da Função Inversa. Definimos $F : \mathbb{R}^{n+m} \rightarrow \mathbb{R}^{n+m}$ por

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

Dessa forma,

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

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 21:

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

Pelo Teorema da Função inversa, existe uma vizinhança Z de $(a, b) \in \mathbb{R}^{n+m}$ e uma vizinhança (que podemos tomar da forma) $V \times W$ de $(c, b) \in \mathbb{R}^{n+m}$ 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 28. A função $\pi^1 : \mathbb{R}^{n+m} \rightarrow \mathbb{R}^n$ 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 29 (Forma Local de Imersões). *Consideramos $f : U \subseteq \mathbb{R}^n \rightarrow \mathbb{R}^{n+m}$ uma imersão de classe $C^k(U; \mathbb{R}^{n+m})$ e $a \in U \subseteq \mathbb{R}^n$. Então, existem*

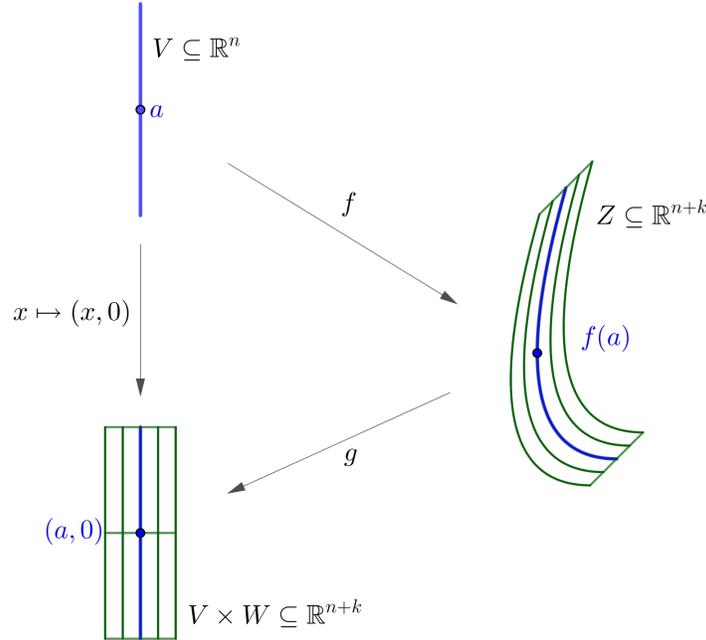
- (i) uma vizinhança W de $0 \in \mathbb{R}^m$;
- (ii) uma vizinhança $V \subseteq U$ de $a \in \mathbb{R}^n$;
- (iii) uma vizinhança Z de $f(a) \in \mathbb{R}^{n+m}$ e;

(iv) um difeomorfismo $g : Z \rightarrow V \times W$ de classe C^k

tais que, para todo $u \in V$, 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 . Logo,

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

Completamos este conjunto até uma base, isto é, consideramos

$$\left\{ \frac{\partial f}{\partial x_1}(a), \frac{\partial f}{\partial x_2}(a), \dots, \frac{\partial f}{\partial x_n}(a), v_1, v_2, \dots, v_m \right\} \text{ base de } \mathbb{R}^{n+m}. \quad (29)$$

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

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

A notação acima implica, em particular, que os números y^i são as componentes do vetor $y \in \mathbb{R}^m$ na base canônica de \mathbb{R}^m . Assim, vetor y é levado por F em um vetor que tem estas mesmas componentes na base $\{v_1, v_2, \dots, v_m\} \subset \mathbb{R}^m$. A expressão para a derivada de F no ponto $(a, 0)$ e na direção do vetor $(h, \ell) \in \mathbb{R}^{n+m}$ é

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

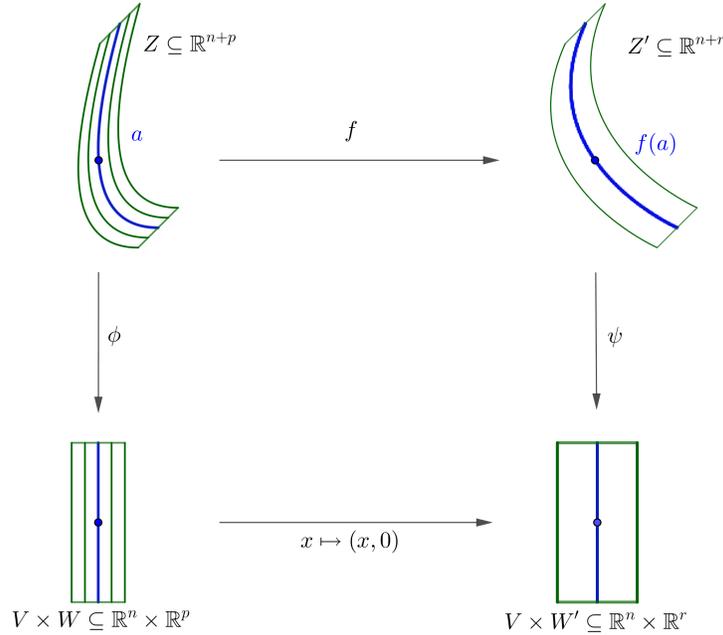
De maneira análoga, os números ℓ^i denotam as componentes do vetor $\ell \in \mathbb{R}^m$ na base canônica de \mathbb{R}^m . Denotando ainda por h^i as componentes de h na base canônica de \mathbb{R}^n , nós observamos que $F'(a, 0) \cdot (h, \ell) = 0$ pode ser reescrito como

$$h^1 \frac{\partial f}{\partial x_1}(a) + h^2 \frac{\partial f}{\partial x_2}(a) + \dots + h^n \frac{\partial f}{\partial x_n}(a) + \ell^1 v_1 + \ell^2 v_2 + \dots + \ell^m v_m = 0$$

Mas, por (29), estes vetores formam uma base de \mathbb{R}^{n+m} e então só pode ser $F'(a, 0) \cdot (h, \ell) = 0$ no caso em que $(h, \ell) = 0$ (definição de independência linear). Logo, $F'(a, 0)$ é um isomorfismo. Pelo Teorema da Aplicação Inversa, F é um difeomorfismo entre vizinhanças de $(a, 0)$ e de $F(a, 0) = f(a)$, respectivamente. Logo, existe uma vizinhança (que podemos tomar da forma) $V \times W$ de $(a, 0)$ 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, enunciemos um teorema que engloba todos os anteriores como casos particulares:



Teorema 30 (Teorema do Posto Constante). *Seja $f : U \subseteq \mathbb{R}^{n+p} \rightarrow \mathbb{R}^{n+r}$ uma aplicação de classe C^k 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 de 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. Sem perda de generalidade, vamos assumir que o posto de f é constante em todo o aberto U (caso contrário, bastaria restringirmos f a uma vizinhança aberta de a onde o posto é constante – esta existe por hipótese). Reordenando as coordenadas, se necessário, nós podemos assumir que a matriz

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

é invertível para todo $x \in U$ (lembra que o conjunto das matrizes invertíveis é aberto!). 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 ϕ é um difeomorfismo de uma vizinhança Z de $a \in U$ em uma vizinhança (que pode ser tomada da forma) $V \times W$ de $\phi(a) \in \mathbb{R}^{n+p}$. As componentes da função $\phi^{-1} : V \times W \rightarrow Z$ são da forma

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

e a definição de inversa nos dá, para todo $(x, y) \in V \times W$,

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

Logo, as componentes da função $f \circ \phi^{-1} : V \times W \rightarrow \mathbb{R}^{n+r}$ podem ser descritas como

$$\begin{aligned} (f \circ \phi^{-1})(x, y) &= ((f_1, f_2, \dots, f_n)(\varphi(x, y), y), (f_{n+1}, f_{n+2}, \dots, f_{n+r})(\varphi(x, y), y)) \\ &= (x_1, \dots, x_n, (f_{n+1} \circ \phi^{-1})(x, y), \dots, (f_{n+r} \circ \phi^{-1})(x, y)). \end{aligned}$$

Assim, na base canônica, a representação matricial da derivada da função $f \circ \phi^{-1}$ em um ponto (x, y) é

$$[(f \circ \phi^{-1})'(x, y)]_{\text{Can}} = \begin{bmatrix} I_n & 0 \\ * & \partial_{y_j}(f_i \circ \phi^{-1}) \end{bmatrix},$$

onde no canto inferior direito apenas são consideradas as componentes com $i > n$. Sendo o posto de $(f \circ \phi^{-1})'(x, y)$ exatamente igual a n , temos que necessariamente $\partial_{y_j}(f_i \circ \phi^{-1}) = 0$ para todo j e todo $i > n$; caso contrário, a dimensão da imagem seria maior do que n . Logo, as funções

$$(x, y) \mapsto f_i \circ \phi^{-1}(x, y), \quad \text{para } i > n, \quad (30)$$

são independentes de y (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).

Finalmente, em uma vizinhança \tilde{Z} de $f(a)$, definimos $\psi : \tilde{Z} \subseteq \mathbb{R}^{n+r} \rightarrow \mathbb{R}^{n+r}$ por

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

onde $w \in \mathbb{R}^n$. Temos que, por (30), ψ está bem definida. Além disso,

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

de modo que ψ é um difeomorfismo de uma vizinhança Z' de $f(a)$ em uma vizinhança (que pode ser tomada da forma) $V \times W'$. Ainda mais, por construção,

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

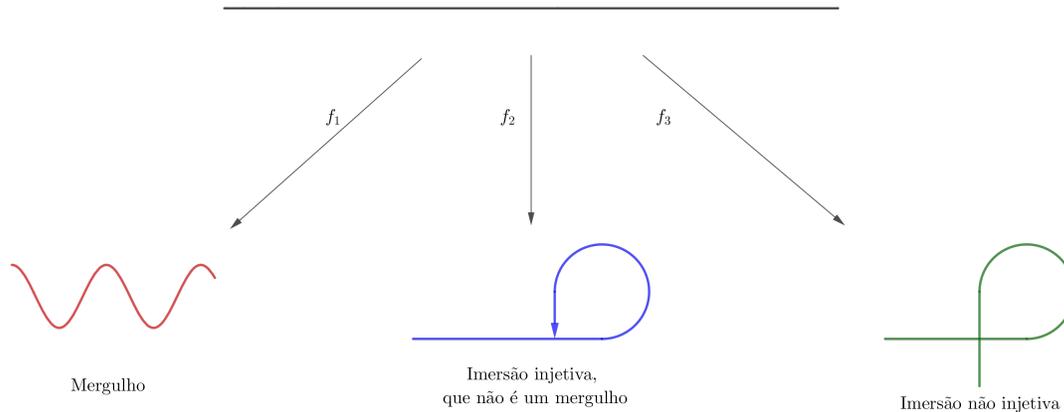
Exercício 24. Obter as formas canônicas locais como corolários do Teorema do Posto.

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, ϕ for um homeomorfismo sobre a sua imagem.

Nós utilizamos este conceito adiante na descrição de parametrizações de superfícies diferenciáveis. Para fazer um comparativo, deixamos os seguintes comentários:

- (i) uma imersão ϕ é, por definição, uma aplicação tal que $\phi'(x)$ é injetiva, para todo $x \in U$;
 - (ii) no entanto, uma imersão não necessariamente é uma aplicação injetiva. A sua imagem pode “conter autointerseções”;
 - (iii) e ainda, uma imersão pode ser injetiva sem ser um homeomorfismo sobre a sua imagem.
- Geometricamente, é instrutivo termos os seguintes exemplos em mente, 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 a aplicação f em sua forma canônica local, como nas seções anteriores.

Em dimensões maiores, nós podemos ter em mente os mesmos tipos de exemplos. Um toro pode ser mergulhado em \mathbb{R}^3 ; nós podemos ter imersões injetivas que não são mergulhos pensando em uma versão bidimensional da curva azul acima (imagina que estamos vendo uma superfície de perfil); exemplos de imersões não injetivas incluem a Garrafa de Klein e a Superfície de Boy.

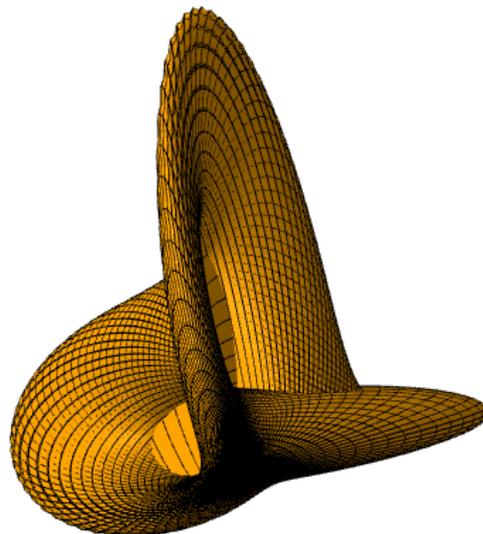
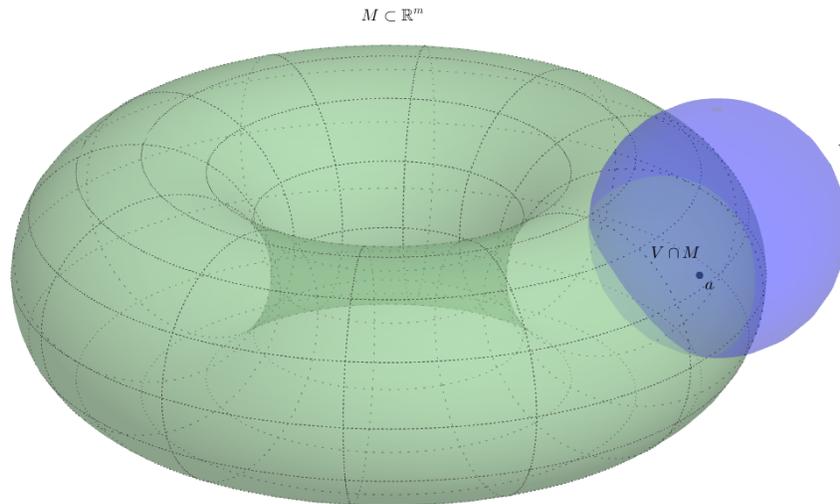


Figura 1: Superfície de Boy construída com Sage – imersão do plano projetivo real no espaço \mathbb{R}^3 , encontrada por Werner Boy em 1901. Observamos que o plano projetivo não pode ser mergulhado em \mathbb{R}^3 . Para descrição, construção, etc; ver [Wikipedia](#).

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. Fazemos de pelo menos duas formas diferentes ao longo do nosso curso. Segue a primeira neste capítulo; a outra no capítulo intitulado Variedades Diferenciáveis.

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 ϕ 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 classe C^k) de um aberto $U \subseteq \mathbb{R}^n$. Observamos que V é um aberto do espaço ambiente (na figura representado por uma bola em \mathbb{R}^3) enquanto M possivelmente tem “dimensão mais baixa” (no desenho, representado pelo toro não sólido).

Exemplo 31. 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 ϕ 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 ou mesmo da esfera $x^2 + y^2 + z^2 = 1$.

Proposição 32. 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 ϕ 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 π 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

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

Como ψ é de classe C^k , também f o é. □

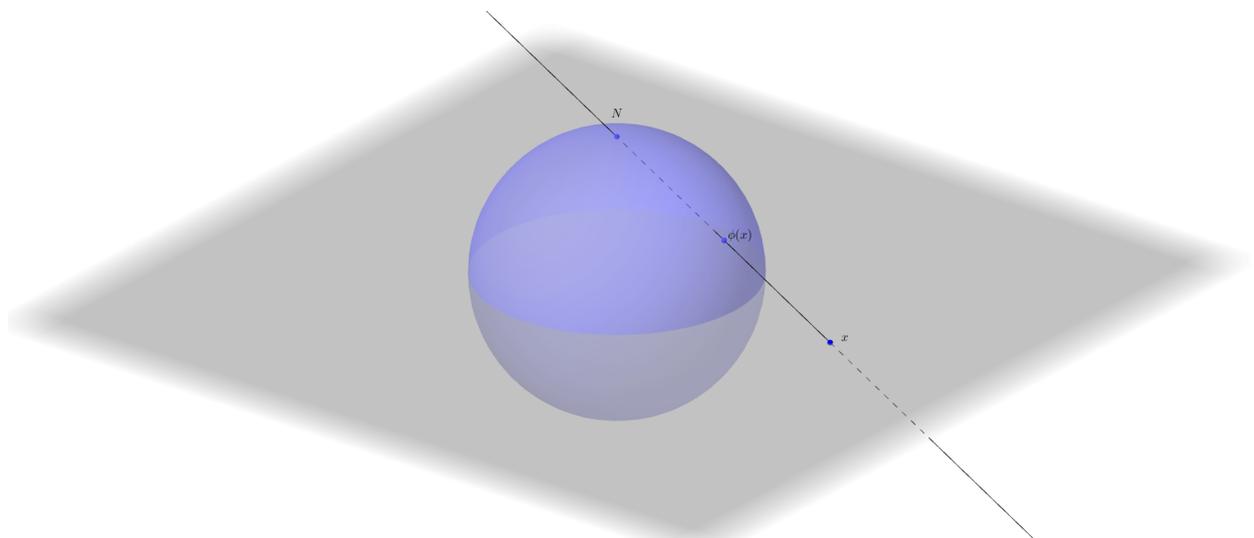
Exemplo 33. A esfera $\mathbb{S}^n = \{(x_1, x_2, \dots, x_{n+1}) \in \mathbb{R}^{n+1}; x_1^2 + x_2^2 + \dots + x_{n+1}^2 = 1\} \subseteq \mathbb{R}^{n+1}$ é uma superfície n -dimensional e de classe C^∞ de \mathbb{R}^{n+1} .

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 ϕ é um mergulho. Fazendo o análogo para o polo sul (sinal negativo antes da raiz quadrada 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 $\phi(x) \in \mathbb{S}^n \setminus \{N\}$ e daí:

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

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 ϕ e obter que ϕ é uma parametrização de classe C^∞ para o conjunto $\mathbb{S}^n \setminus \{N\}$.

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 25. Obter uma parametrização do toro (superfície que parece uma rosquinha - ver figura da página 37) em \mathbb{R}^3 .

Exercício 26 (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.

11.1 Aplicações diferenciáveis entre superfícies

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

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

Inicialmente, nós observamos que a definição usual de diferenciabilidade exige que o domínio da função seja um subconjunto aberto de \mathbb{R}^m . Por outro lado, quando a dimensão de M é menor do que a dimensão do espaço ambiente, a superfície não é um subconjunto aberto¹⁸ em \mathbb{R}^m , de modo que o conceito usual não se estende trivialmente e deve ser fundamentalmente diferente.

Agora, ao lidar com superfícies diferenciáveis de dimensão n em \mathbb{R}^m , temos cartas à nossa disposição que nos permitem “trabalhar” localmente como se estivéssemos em um aberto do \mathbb{R}^n . A ideia é então dizer que uma função é diferenciável se o for quando escrita em termos de qualquer carta local. Passamos a descrever esta construção de maneira mais rigorosa.

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, no entanto, que a noção de diferenciabilidade deve ser uma propriedade intrínseca de M , ou seja, não pode depender do sistema de coordenadas particular utilizado. Se eu utilizo o sistema de coordenadas local ϕ e o leitor outro sistema ψ , devemos ser capazes de justificar que

$$f \circ \phi \text{ é diferenciável} \iff f \circ \psi \text{ é diferenciável.}$$

Se este não fosse o caso, a função poderia ser diferenciável segundo uma carta e não diferenciável segundo outra carta, o que indicaria que o conceito está mal definido.

Neste sentido que o lema abaixo é de importância fundamental.

¹⁸Na realidade, nesta situação em que $\dim M < m$, a superfície M tem interior vazio em \mathbb{R}^m , de modo que nenhum de seus pontos é interior.

Lema 34. Seja $M \subseteq \mathbb{R}^m$ uma superfície de dimensão n e de classe C^k e seja $\phi : U \rightarrow \phi(U) \subseteq M$ 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.

Em particular, dada outra carta $\psi : V \rightarrow \phi(U) \subseteq M$ para a mesma vizinhança coordenada, temos

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

Demonstração. Começamos por provar (i). Note inicialmente que este item não segue da regra da cadeia, pois a inversa ϕ^{-1} não está definida em um subconjunto aberto de \mathbb{R}^m . Mas sabemos que ϕ 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 ϕ imersão, existe um sistema de coordenadas onde ϕ 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 π 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, ϕ^{-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

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

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 ξ é um difeomorfismo. □

Nós enfatizamos que, apenas equipados com o lema acima, pudemos definir diferenciabilidade para funções definidas em superfícies de maneira consistente. Observamos ainda que, embora a diferenciabilidade esteja bem definida, não é óbvio que tipo de objeto deveria ser a derivada de uma aplicação entre superfícies. Como motivação, nós vamos pensar novamente no caso Euclidiano de uma função $f : U \rightarrow \mathbb{R}^d$, definida em um aberto $U \subseteq \mathbb{R}^m$. A derivada de f associa a um vetor $v \in \mathbb{R}^m$, a derivada direcional de f na direção v . Fixado $v \in \mathbb{R}^m$, seja $\alpha : (-\delta, \delta) \rightarrow U$ uma curva diferenciável satisfazendo

$$\alpha(0) = x \quad \text{e} \quad \alpha'(0) = v.$$

Por exemplo, α poderia ser um segmento de reta da forma $t \mapsto x + tv$, que está bem-definido em U para $\delta > 0$ suficientemente pequeno, já que U é aberto. Pela Regra da Cadeia, temos

$$\left. \frac{d}{dt} f(\alpha(t)) \right|_{t=0} = df(x) \cdot v.$$

Em outras palavras, a aplicação $df(x) : \mathbb{R}^m \rightarrow \mathbb{R}^d$ é uma aplicação que associa ao vetor $v \in \mathbb{R}^m$, tangente à curva α em x , o vetor $df(x) \cdot v \in \mathbb{R}^d$, tangente à curva $f \circ \alpha$ em $f(x)$. Esta é a ideia a ser explorada. Inicialmente, contudo, esclarecemos como o chamado espaço tangente $T_x M$ a uma superfície M está relacionado com “velocidades de curvas” para então importar uma noção de derivada para aplicações entre superfícies.

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 ϕ 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 35. *Temos a seguinte caracterização para o espaço tangente:*

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

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 ϕ uma parametrização de uma vizinhança de x que contém toda a curva γ (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 $T_x M$ e 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 27. Prove esta última afirmação.

11.2 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** quando¹⁹, 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 pelo Lema 34.

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

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

¹⁹Para 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.

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

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

Utilizando a Proposição 35, note que $f'(x) \cdot v \in T_{f(x)} N$ pois $f \circ \gamma$ é uma curva diferenciável 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 **aplicação induzida** por f ou de **push-forward** de f .

12 Valores regulares

Seja $f : U \subset \mathbb{R}^d \rightarrow \mathbb{R}^m$ 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 $f'(x) : \mathbb{R}^d \rightarrow \mathbb{R}^m$ sobrejetiva. Em particular, deve-se ter $d \geq m$. No teorema abaixo, escrevemos $d = m + n$.

Teorema 36 (Superfície de nível em um valor regular). *Seja $f : U \subset \mathbb{R}^{m+n} \rightarrow \mathbb{R}^m$ uma função de classe C^k . Se $c \in \mathbb{R}^m$ é um valor regular de f , então $M = f^{-1}(c)$ é uma superfície de dimensão n e de classe C^k em \mathbb{R}^{m+n} . Além disso,*

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

Demonstração. Este teorema pode ser visto como um “refraseamento” do Teorema da Função Implícita. Sendo $f'(w)$ sobrejetiva, sua matriz canônica possui m colunas linearmente independentes. Reordenando as coordenadas, se necessário, podemos supor que são as m primeiras. Escrevemos $w = (x, y) \in \mathbb{R}^{m+n}$. As primeiras colunas linearmente independentes implicam que

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

Logo, localmente, $x = \xi(y)$ onde ξ é 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, $T_x M \subseteq \text{Nuc } f'(x)$. Sendo $\dim \text{Im } f'(x) = m$, devemos ter $\dim \text{Nuc } f'(x) = n$. Portanto, só pode ser que $T_x M = \text{Nuc } f'(x)$. \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 hiperfície²⁰) em \mathbb{R}^{n+1} e o espaço tangente a $M = f^{-1}(c)$ em $x \in M$ é

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

Em outras palavras, o vetor gradiente é ortogonal à superfície de nível de um valor regular de f e o plano tangente à superfície de nível $f^{-1}(c)$ no ponto x é o subespaço ortogonal a $\nabla f(x)$.

²⁰Uma **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 37. A esfera $\mathbb{S}^n \subset \mathbb{R}^{n+1}$ pode ser vista como a imagem inversa de $c = 1$ pela função

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

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 hipersuperfície de classe C^∞ cujo espaço tangente é

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

Exemplo 38. Seja $A \in M(n)$ uma matriz real simétrica e consideramos a forma quadrática associada $f : \mathbb{R}^n \rightarrow \mathbb{R}$ definida 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 hipersuperfície de classe C^∞ em \mathbb{R}^n . Seu espaço tangente em um ponto $x \in M$ pode ser caracterizado por

$$T_x M = \{Ax\}^\perp.$$

A esfera de raio $r > 0$ é o caso especial em que $A = I$ e $c = r^2$.

Exemplo 39. Consideramos o grupo especial linear, definido como²¹

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

e a aplicação $f : M(n) \rightarrow \mathbb{R}$ 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^∞ em $M(n)$.

Vamos analisar o espaço tangente $T_I(SL(\mathbb{R}^n))$.²² 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,

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

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

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

²¹A abreviação vem do inglês, em que este grupo é conhecido como *Special Linear Group*.

²² $SL(\mathbb{R}^n)$ 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. Considere o grupo ortogonal e o conjunto das matrizes simétricas (ou auto-adjuntas reais):

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

Seja $f : M(n) \longrightarrow \mathcal{S}(n)$ a aplicação dada por

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

Assim, $\mathcal{O}(n) = f^{-1}(I)$ e vamos verificar que I é um valor regular de f . Temos (exercício)

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

Queremos mostrar que, para $A \in \mathcal{O}(n)$, a transformação $f'(A)$ é sobrejetiva: dada uma matriz simétrica $S \in \mathcal{S}(n)$, observamos que, sendo $AA^T = I$ e $S = S^T$,

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

Assim, $S \in \text{Im } f'(A)$ e segue que $\mathcal{O}(n)$ é uma superfície de classe C^∞ em $M(n)$ e de dimensão

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

O espaço tangente a $\mathcal{O}(n)$ em I pode ser caracterizado como

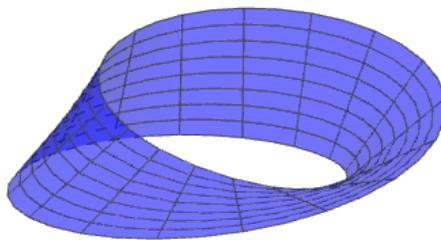
$$T_I(\mathcal{O}(n)) = \{H \in M(n); H + H^T = 0\},$$

que é o conjunto das matrizes anti-simétricas (ou anti-auto-adjuntas reais) e que, sendo um espaço tangente a $\mathcal{O}(n)$, deve ter a mesma dimensão de $\mathcal{O}(n)$ (embora obter esta dimensão seja muito mais elementar com métodos puramente algébricos).

Exemplo 41. Nem toda superfície é imagem inversa de valor regular. Se uma hipersuperfície da forma $M = f^{-1}(c)$ para c valor regular e $f \in C^1$, então

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

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. Na linguagem da seção seguinte, nós acabamos de verificar que toda superfície de nível, obtida como imagem inversa de um valor regular, é orientável. Desta, maneira, qualquer superfície não-orientável não pode ser obtida como imagem inversa de valor regular. Como mostramos na próxima seção, um exemplo é a faixa de Möbius do desenho abaixo.



Proposição 42. 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 $f : U \times \mathbb{R}^{m-n} \rightarrow \mathbb{R}^{m-n}$ por

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

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

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

é certamente sobrejetiva. □

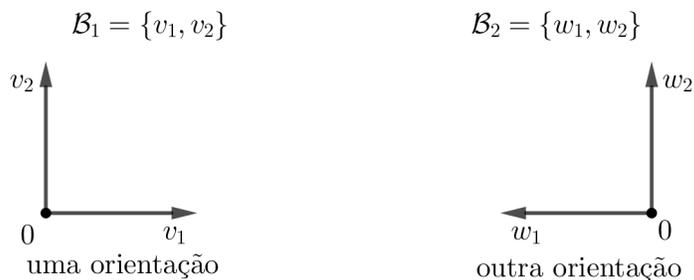
Observação 43. Note que nem toda superfície que é imagem inversa de valor regular é o gráfico de uma função. Considere, por exemplo, a esfera. Mais geralmente, o gráfico de uma função é homeomorfo ao seu domínio e é, por consequência, aberto. Dessa maneira, nenhuma superfície compacta pode ser o gráfico 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 tem a ver com uma 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 todo $x \in U$, é obtida a partir da parametrização ϕ , 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 ϕ e ψ pode coincidir ou não (dãã!). É possível decidir qual é o caso olhando para a matriz de mudança de bases. Pela regra da cadeia,

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

Escrevendo²³ $y = (\psi^{-1} \circ \phi)(x)$, 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 dizemos que $\phi : U \rightarrow \phi(U) \subset M$ e $\psi : V \rightarrow \psi(V) \subset M$ são duas cartas **coerentes** ou **compatíveis** quando

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

²³É 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 $y_i(x) = (\psi^{-1} \circ \phi)_i(x)$.

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 paratrizações 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 maximal²⁴. 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 44. Qualquer superfície que possa ser coberta por uma vizinhança coordenada só, como gráficos de funções, são orientáveis.

Exemplo 45. Subconjuntos abertos de superfícies orientáveis são também orientáveis; basta considerar a restrição das cartas do atlas.

Exemplo 46. 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 ϕ e ψ , 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

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

Exemplo 47. Devido ao exemplo anterior, \mathbb{S}^n é orientável, para todo $n > 1$; ver projeção estereográfica da página 38. Por que excluimos o caso $n = 1$?

Exercício 28. Mostre que \mathbb{S}^1 é orientável.

Exercício 29. 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 “estiver restrito” a permanecer sobre a superfície M . 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 $v \in T_x M$.

Vejamos agora que, no caso de hipercifícies, há uma caracterização de orientabilidade em termos de campos normais a M .

Teorema 48. *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.*

²⁴Um atlas \mathcal{A} é dito maximal se, para ϕ 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).

Demonstração. $\boxed{\Rightarrow}$ 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 satisfaz²⁵

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

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

$$\begin{bmatrix} \frac{\partial \phi}{\partial x_1}(x_0) & \cdots & \frac{\partial \phi}{\partial x_n}(x_0) & u(x) \end{bmatrix}^T = \begin{bmatrix} \left(\frac{\partial y_i}{\partial x_j}(x_0) \right)_{n \times n} & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \frac{\partial \psi}{\partial x_1}(y_0) & \cdots & \frac{\partial \psi}{\partial x_n}(y_0) & u(x) \end{bmatrix}^T. \quad (31)$$

Vejamus 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 $T_x M = \text{Nuc } g'(x)$. 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 .

$\boxed{\Leftarrow}$ 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 ψ 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 ψ 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 (31). \square

²⁵Ao completar a base de $T_x M$ a uma base de \mathbb{R}^{n+1} com um vetor unitário, existem apenas duas possibilidades; escolhemos àquela que resulta em uma base positiva de \mathbb{R}^{n+1} .

Exemplo 49. Vamos construir parametrizações para a **Faixa de Möbius**²⁶ e em seguida mostrar que é uma superfície não orientável. A faixa de Möbius (ver figura da página 44) 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° , retornando ao segmento original, mas “invertido”.

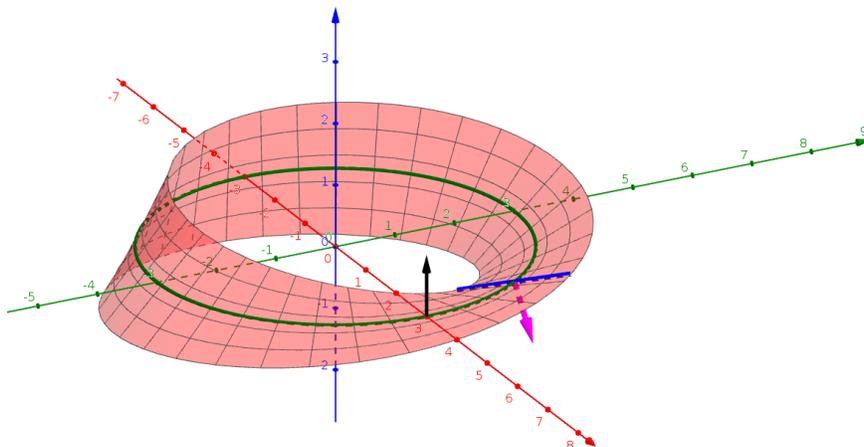
Fixado um parâmetro $t \in (0, 2\pi)$, o centro do segmento se desloca para o ponto $x_0 \cdot u(t)$, onde $u(t) = (\cos t, \sin t, 0)$. Consideramos o plano em \mathbb{R}^3 gerado pelos vetores unitários $u(t)$ e e_3 ; o semi-círculo de raio s centrado no ponto $x_0 \cdot u(t)$ deste plano pode ser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



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 parametrizações²⁷.

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, caso fosse possível estender este campo, deveríamos ter $v(0) = -v(2\pi)$, o que é uma contradição. Isto mostra que a faixa de Möbius não é orientável. As contas ficam como exercício.

Vejamos que uma das implicações do Teorema 48 segue válida para superfícies de codimensão maior. A prova é bastante semelhante.

²⁶A 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”.

²⁷Isto não contradiz o exemplo 46, pois a união das vizinhanças parametrizadas não é conexa.

Teorema 50. *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

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

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 ψ satisfizer

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

em seu domínio conexo V , podemos “trocar o sinal” de ψ 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 equivalente a que exibimos em (31), 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 51. *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, $T_x M = \text{Nuc } f'(x)$ 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 52 (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 $c \in \mathbb{R}$. Sejam $U \subset \mathbb{R}^{n+1}$ 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*

$$\nabla f(x) = \lambda \nabla g(x). \quad (32)$$

Em particular, se a restrição $f|_M$ tem um mínimo ou máximo em x_0 , então vale (32).

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

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

Se, além disso, $x \in M$ é um ponto crítico de $f|_M$ 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 $x \in M$ é um ponto crítico da restrição $f|_M$ se, e somente se, $\nabla g(x)$ 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 53 (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*

$$\nabla f(x) = \lambda^T g'(x). \quad (33)$$

Mais explicitamente,

$$\nabla f(x) = \lambda_1 \nabla g_1(x) + \lambda_2 \nabla g_2(x) + \cdots + \lambda_k \nabla g_k(x). \quad (34)$$

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 $\nabla f(x) \in (\text{Nuc } A)^\perp$, onde $A = g'(x)$. Isto porque

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

Exemplo 54. 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, devemos procurar por um mínimo dentre os pontos $x_0 \in M$ que satisfazem $\nabla f(x_0) \perp T_{x_0} M$

Exemplo 55. 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

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

Matrizes simétricas possuem autovalores reais. Se λ_{\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 para operadores simétricos da Álgebra Linear.

Exercício 30. 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, λ_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 λ_0 é o multiplicador de Lagrange associado, então, para todo $v \in \text{Nuc } g'(x_0)$, 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 esta definição se estende 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

$$\text{vol}(R) = (b_1 - a_1) \cdot (b_2 - a_2) \cdot \cdots \cdot (b_n - a_n). \quad (35)$$

A definição de um retângulo aberto ou de um retângulo fechado é clara e o volume é definido pela mesma fórmula (35) acima.

Seja $f : R \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$ uma função limitada. Consideramos uma partição do 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 (denotada por) \mathcal{P} do retângulo R . Com a notação

$$m_\lambda := \inf_{x \in R_\lambda} f(x) \quad \text{e} \quad M_\lambda := \sup_{x \in R_\lambda} f(x), \quad (36)$$

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 se faz em cursos mais introdutórios. 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 56. 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 57. 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 58. Uma função $f : R \subset \mathbb{R}^n \rightarrow \mathbb{R}$ 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 59. Se $f : R \subset \mathbb{R}^n \rightarrow \mathbb{R}$ 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 60. Uma função $f : R \subset \mathbb{R}^n \rightarrow \mathbb{R}$ 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 61 (Propriedades da integral). As propriedades abaixo são satisfeitas pela integral de Riemann:

(i) Linearidade: se $f, g : R \subset \mathbb{R}^n \rightarrow \mathbb{R}$ 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 $f : R \subset \mathbb{R}^n \rightarrow \mathbb{R}$ 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 $f : R \subset \mathbb{R}^n \rightarrow \mathbb{R}$ é integrável, então $|f| : R \subset \mathbb{R}^n \rightarrow \mathbb{R}$ é integrável e tem-se

$$\left| \int_R f(x) dx \right| \leq \int_R |f(x)| dx. \quad (37)$$

(iv) Valor médio: se $f : R \subset \mathbb{R}^n \rightarrow \mathbb{R}$ 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 como 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 (37).

(iv) Sendo f limitada, podemos escrever $f(x) \in [m, M]$ para todo $x \in R$. Pelo item (ii) e pelo Exemplo 56, 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 $\{R_i\}_{i \in \mathbb{N}}$ cuja soma dos volumes é menor do que ε , isto é,

$$A \subseteq \bigcup_{i=1}^{+\infty} R_i \quad \text{e} \quad \sum_{i=1}^{+\infty} \text{vol}(R_i) < \varepsilon.$$

Exercício 31. 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 62. 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 62 acima pode ser usada para verificar o seguinte exercício.

Exercício 32. 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 33. 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 63. Retângulos não degenerados (isto é, um produto cartesiano de intervalos não vazios) **não** são conjuntos de medida zero. De fato, seja R um retângulo não degenerado. 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 χ_{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 56, e a propriedade (ii) da Proposição 61 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 ε . \triangleleft

Exemplo 64. Uma consequência imediata do exemplo anterior é que conjuntos cujo interior é não vazio não podem ter medida nula.

Teorema 65. 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 ℓ_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 66. 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 67. 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 68. 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**.

O próximo teorema é o mais importante desta seção.

Teorema 69 (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 $\text{osc}(f) : R \subset \mathbb{R}^n \rightarrow \mathbb{R}$ 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 como se comporta a oscilação de f em bolas cada vez “mais próximas” do ponto $x \in R$. Ou ainda, mede quão ruim é a descontinuidade de f em x . Notamos que (prove como exercício)

$$f \text{ é contínua no ponto } x \iff \text{osc}(f)[x] = 0.$$

\Rightarrow 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. Pelo 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_ε 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 η 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_ε por retângulos abertos cuja soma dos volumes é menor do que ε , 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 $\chi_A : R \subset \mathbb{R}^n \rightarrow \mathbb{R}$, 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** de A por

$$\text{vol } A := \int_R \chi_A(x) \, dx.$$

Uma consequência imediata do Teorema de Lebesgue é o seguinte.

Proposição 70. *Um conjunto limitado $A \subset \mathbb{R}^n$ é Jordan mensurável se, e somente se, ∂A tem medida (de Lebesgue) nula.*

Demonstração. Suponhamos, por simplicidade que $A \subset \text{int } R$; caso contrário, bastaria considerar um retângulo maior. Basta então observar que o conjunto dos pontos de descontinuidade de χ_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 χ_A é constante em $\text{int}(A)$ e em $\text{int}(\mathbb{R}^n \setminus A)$. \square

Exemplo 71. 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 72. Mais geralmente (ver Corolário 68), se a fronteira de um conjunto limitado for um conjunto n -contavelmente retificável, então este conjunto é Jordan mensurável.

Exercício 34. 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 35. Seja K um conjunto compacto. Mostre que são equivalentes:

- (i) $\text{vol}(K) = 0$;
- (ii) K tem medida zero.

Exercício 36. Discuta a equivalência das seguintes afirmações:

- (i) A é Jordan mensurável;
- (ii) ∂A tem medida nula;
- (iii) ∂A tem volume zero.

Observamos que ser A de medida nula (segundo Lebesgue) não implica em Jordan mensurabilidade, ver Exemplo 57. Na linguagem desta seção, mostramos no Exemplo 57 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 73. *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

$$\int_{\underline{R}} \chi_A(x) \, dx = 0.$$

Se A é Jordan mensurável, isto é equivalente a

$$\text{vol}(A) = \int_{\underline{R}} \chi_A(x) \, dx = \int_{\underline{R}} \chi_A(x) \, dx = 0. \quad \square$$

Passamos à definição de integral de uma função limitada $f : A \subset \mathbb{R}^n \rightarrow \mathbb{R}$ definida em um conjunto Jordan mensurável A . Consideramos R um retângulo que contém A e, por abuso de notação²⁸, consideramos também a extensão de f como zero fora de A : definimos a função $f\chi_A : R \subset \mathbb{R}^n \rightarrow \mathbb{R}$ 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 69 – 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 61 – que vimos em domínios retangulares também se estendem trivialmente para integrais em domínios Jordan mensuráveis.

Exercício 37. Verificar estas últimas afirmações!

Proposição 74. *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. Além disso, se $\text{int}(A) \cap \text{int}(B) = \emptyset$, temos*

$$\int_{A \cup B} f(x) \, dx = \int_A f(x) \, dx + \int_B f(x) \, dx.$$

²⁸As funções $f : A \rightarrow \mathbb{R}$ e $\chi_A : R \subset \mathbb{R}^n \rightarrow \mathbb{R}$ 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, ∂A e ∂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 75. *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

$$\int_{R \times Q} f(x, y) \, d(x, y) = \int_R \left[\int_{\underline{Q}} f(x, y) \, dy \right] dx = \int_R \left[\int_{\overline{Q}} f(x, y) \, dy \right] dx. \quad (38)$$

Enunciado semelhante vale trocando os papéis de x e de y .

Demonstração. Vamos provar apenas a primeira igualdade, pois a outra pode ser obtida de maneira análoga. Toda partição de $R \times Q$ por retângulos cujas arestas são paralelas aos eixos coordenados 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 (36)). 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 76. *Considere o conjunto*

$$A = \{(x, y) \in [a, b] \times \mathbb{R}; \varphi(x) \leq y \leq \psi(x)\},$$

onde φ e ψ 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 φ e ψ 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 escolher 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 77 (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 e 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, é um conceito geométrico que serve para os nossos propósitos.

Lema 78. *Seja $B \subset A$ com $\text{vol}(B) = 0$. Temos que, dado $\varepsilon > 0$, existe $\delta > 0$ tal que*

$$D \text{ decomposição de } A \text{ com } |\mathcal{D}| < \delta \implies \sum_{d(X_i, B) < \delta} \text{vol}(X_i) < \varepsilon.$$

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 ε . 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 2δ 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

$$D \text{ decomposição de } A \text{ com } |\mathcal{D}| < \delta \implies \sum_{d(X_i, B) < \delta} \text{vol}(X_i) \leq \sum_{i=1}^N \text{vol}(R_i^{2\delta}) < \varepsilon \quad \square$$

Demonstração do Teorema 77. 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 \hat{A} um retângulo que contém A e $\hat{f} := f \chi_{\hat{A}}$. 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 34, 35 e 36. Logo, o Lema 78 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

$$\overline{\int_A} f(x) dx \leq S(f; \mathcal{D}). \quad (39)$$

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 78, 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

$$\begin{aligned} S(f; \mathcal{D}) &= \sum_{X_i \subset \text{int } R} M_{X_i} \cdot \text{vol}(X_i) + \sum_{X_i \cap B \neq \emptyset} M_{X_i} \cdot \text{vol}(X_i) \\ &\leq \sum_{R \in \mathcal{P}} M_R \cdot \text{vol}(R) + (\max_A f) \sum_{X_i \cap B \neq \emptyset} \text{vol}(X_i) \\ &< \overline{\int_A} f(x) dx + (1 + \max_A f) \cdot \varepsilon. \end{aligned} \quad (40)$$

Juntando (39) e (40), 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 Sanduíche, 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 conhecida como método de integração por substituição): se $I \subseteq \mathbb{R}$ é um intervalo, $\phi : [a, b] \rightarrow I$ 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 $x = \phi(t) \rightsquigarrow dx = \phi'(t) dt$. 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 79 (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, vemos algumas aplicações clássicas deste resultado, a saber, 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 $n = 1$, com $\phi : \mathbb{R} \rightarrow \mathbb{R}$ função 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 é

$$S(f; \mathcal{P}) = \sum_{i=1}^N M_{[t_i, t_{i+1}]}(f) \cdot (t_i - t_{i-1})$$

Observe, sendo ϕ invertível, que $t_i = \phi(s_i) = cs_i + d$ implica que $\mathcal{Q} = \{s_i\}$ é uma partição de K e temos que

$$S(f \circ \phi; \mathcal{Q}) = \sum_{i=1}^N M_{[s_i, s_{i+1}]}(f \circ \phi) \cdot |s_i - s_{i-1}| = \frac{1}{|c|} \sum_{i=1}^N M_{[t_i, t_{i+1}]}(f) \cdot (t_i - t_{i-1}) = \frac{S(f; \mathcal{P})}{|c|}.$$

Nós escrevemos as diferenças na partição \mathcal{Q} em valor absoluto, pois ϕ não é necessariamente crescente. Como toda partição de K pode ser obtida desta maneira, concluímos que

$$\int_{\phi(K)} f(x) dx = |c| \int_K f(\phi(y)) dy = \int_K f(\phi(y)) |\det \phi'(y)| dy. \quad (41)$$

A igualdade entre as integrais inferiores é obtida analogamente, basta notar que também vale

$$M_{[s_i, s_{i+1}]}(f \circ \phi) = M_{[t_i, t_{i+1}]}(f) \quad \text{e logo} \quad I(f \circ \phi; \mathcal{Q}) = \frac{I(f; \mathcal{P})}{|c|}.$$

Passo 2. Vale para $\phi : \mathbb{R}^n \rightarrow \mathbb{R}^n$ linear invertível.

Na linguagem da Álgebra Linear, pode-se dizer que, sendo ϕ invertível, “ ϕ 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 ϕ , podemos chegar à matriz identidade fazendo uma sequência de dois tipos de operações elementares (que são descritas em termos de transformações invertíveis):

- Trocar duas linhas de lugar, o que corresponde a multiplicação pela transformação linear

$$T(x) = (x_1, x_2, \dots, x_{i-1}, x_j, x_{i+1}, \dots, x_{j-1}, x_i, x_{j+1}, \dots, x_{n-1}, x_n),$$

cuja matriz associada é (trocar as colunas i e j de lugar na matriz identidade)

$$[T]_{\text{Can}} = \begin{bmatrix} 1 & 0 & \cdots & 0 & \cdots & 0 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & \cdots & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & & \vdots & & \vdots & & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & \cdots & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & & \vdots & & \vdots & & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & \cdots & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & & \vdots & & \vdots & & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & \cdots & 0 & \cdots & 1 & 0 \\ 0 & 0 & \cdots & 0 & \cdots & 0 & \cdots & 0 & 1 \end{bmatrix}.$$

- Adicionar um múltiplo de uma linha a um múltiplo de outra linha, o que corresponde a multiplicar ϕ pela transformação linear

$$S(x) = (x_1, x_2, \dots, ax_i + bx_j, \dots, x_{n-1}, x_n),$$

cuja matriz canônica associada é da forma

$$[S]_{\text{Can}} = \begin{bmatrix} 1 & 0 & \cdots & 0 & \cdots & 0 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & \cdots & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & & \vdots & & \vdots & & \vdots & \vdots \\ 0 & 0 & \cdots & a & \cdots & b & \cdots & 0 & 0 \\ \vdots & \vdots & & \vdots & & \vdots & & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & \cdots & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & & \vdots & & \vdots & & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & \cdots & 0 & \cdots & 1 & 0 \\ 0 & 0 & \cdots & 0 & \cdots & 0 & \cdots & 0 & 1 \end{bmatrix}.$$

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 destas transformações são também transformações elementares e do mesmo respectivo tipo, podemos escrever ϕ como um produto de transformações destes dois tipos.

Em seguida, 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 $\psi = T_1 T_2$ (lembre que ψ linear implica que $\psi'(y) = \psi$ para todo y):

$$\int_{\psi(K)} f = \int_{T_2(K)} (f \circ T_1) |\det T_1| = \int_K f(T_1 T_2(y)) |\det T_1| |\det T_2| dy = \int_K f(\psi(y)) |\det \psi'(y)| dy.$$

Resta então mostrar que a fórmula vale para as aplicações 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,

$$\int_{T(R)} f \chi_K = \int_R (f \circ T) \chi_K \quad \text{e também} \quad \int_{T(R)} f \chi_K = \int_R \overline{(f \circ T) \chi_K},$$

como queríamos.

Agora provamos que 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: escrevemos as projeções $\pi^1 : \mathbb{R}^n \rightarrow \mathbb{R}$ e $\pi^2 : \mathbb{R}^n \rightarrow \mathbb{R}^{n-1}$, dadas por

$$\pi^1(x_1, y) = x_1 \quad \text{e} \quad \pi^2(x_1, y) = y.$$

Temos que $\pi^1 \circ S$ é uma aplicação afim como no Passo 1 enquanto $\pi^2 \circ S = \pi^2$. Logo,

$$\begin{aligned} \int_{S(K)} f(x) dx &= \int_{\pi^2 S(K)} \left(\int_{\pi^1 S(K)} f(x_1, y) dx_1 \right) dy \\ &= \int_{\pi^2(K)} \left(|a| \int_{\pi^1(K)} f(ax_1 + bx_2, y) dx_1 \right) dy \\ &= \int_K f(S(x)) |\det S| dx. \end{aligned}$$

Passo 3. Vale para $\phi : \mathbb{R}^n \rightarrow \mathbb{R}^n$ geral, como no enunciado do teorema.

Note que $D_{f \circ \phi} = \phi^{-1}(D_f)$. Sendo ϕ e ϕ^{-1} ambas localmente de Lipschitz, temos que (ver Teorema 65)

$$f \text{ é integrável em } \phi(K) \iff f \circ \phi \text{ é integrável em } K.$$

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 $\phi(K)$ e $y_i \in Y_i$ é qualquer. Sendo ϕ 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.

Como a aplicação $\phi'(x_i)$ é linear invertível, então, pelo Passo 2, podemos escrever

$$\text{vol}(\phi(X_i)) = \text{vol} \left[\phi'(x_i) \phi'(x_i)^{-1}(\phi(X_i)) \right] = |\det \phi'(x_i)| \cdot \text{vol} \left[\phi'(x_i)^{-1}(\phi(X_i)) \right].$$

Afirmação: $\text{vol} \left[\phi'(x_i)^{-1}(\phi(X_i)) \right] \leq M_i^n \cdot \text{vol}(X_i)$, 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 (de aresta $\ell > 0$ e centro x_0) 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 $g(X) \subseteq \{y \in \mathbb{R}^n; |y - g(x_0)|_\infty \leq M\ell/2\}$, isto é, o conjunto imagem $g(X)$ está contido no cubo de aresta $M\ell$ e de centro $g(x_0)$. Segue que o volume de $g(X)$ é controlado por cima pelo volume do cubo:

$$\text{vol}(g(X)) \leq M^n \ell^n = M^n \text{vol}(X).$$

O caso geral em que X não é necessariamente um cubo segue deste, como passamos a justificar, considerando uma cobertura de X por cubos e utilizando a continuidade de 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

$$\begin{aligned} \int_{\phi(K)} f(x) dx &= \lim_{|\mathcal{D}| \rightarrow 0} \sum_{X_i \in \mathcal{D}} f(h(x_i)) \cdot \text{vol}(h(X_i)) \\ &\leq \lim_{|\mathcal{D}| \rightarrow 0} \sum_{X_i \in \mathcal{D}} f(h(x_i)) |\det \phi'(x_i)| M_i^n \cdot \text{vol}(X_i) \\ &= \lim_{|\mathcal{D}| \rightarrow 0} \sum_{X_i \in \mathcal{D}} f(h(x_i)) |\det \phi'(x_i)| (1 + (M_i - 1))^n \cdot \text{vol}(X_i) \\ &\stackrel{**}{=} \int_K f(\phi(y)) |\det \phi'(y)| dy. \end{aligned} \tag{42}$$

A última igualdade seria imediata, não fosse o termo $(1 + (M_i - 1))^n$. No entanto, pelo Lema 80 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 $|\mathcal{D}| \rightarrow 0$. 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 somatórios abaixo e veja que fica menor do que ε)

$$\lim_{|\mathcal{D}| \rightarrow 0} \sum_{X_i \in \mathcal{D}} f(h(x_i)) |\det \phi'(x_i)| (1 + (M_i - 1))^n \text{vol}(X_i) = \lim_{|\mathcal{D}| \rightarrow 0} \sum_{X_i \in \mathcal{D}} f(h(x_i)) |\det \phi'(x_i)| \text{vol}(X_i).$$

Finalmente, a desigualdade em (42) pode ser aplicada para ϕ^{-1} na última integral de (42), 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 esta subseção com a prova do lema que foi utilizado na demonstração acima.

Lema 80. *Seja X um conjunto compacto e Jordan mensurável e $\varphi : X \times X \rightarrow \mathbb{R}$ uma função contínua tal que $\varphi(x, x) = 1$. Então, dado $\varepsilon > 0$, existe $\delta > 0$ tal que*

$$|x - y| < \delta \implies |\varphi(x, y) - 1| < \varepsilon.$$

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 . Logo, como φ é contínua, isto contradiz ser $\varphi(a, a) = 1$. \square

15.5.1 Coordenadas polares

A função ϕ 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. A função φ que muda de coordenadas esféricas para Cartesianas pode ser escrita como

$$\varphi(\rho, \theta, \phi) = (\rho \sin \phi \cos \theta, \rho \sin \phi \sin \theta, \rho \cos \phi), \quad \text{onde } (\rho, \theta, \phi) \in U := (0, +\infty) \times (0, 2\pi) \times (0, \pi).$$

Assim, a matriz das derivadas parciais é

$$\begin{bmatrix} \frac{\partial \varphi_1}{\partial \rho} & \frac{\partial \varphi_1}{\partial \theta} & \frac{\partial \varphi_1}{\partial \phi} \\ \frac{\partial \varphi_2}{\partial \rho} & \frac{\partial \varphi_2}{\partial \theta} & \frac{\partial \varphi_2}{\partial \phi} \\ \frac{\partial \varphi_3}{\partial \rho} & \frac{\partial \varphi_3}{\partial \theta} & \frac{\partial \varphi_3}{\partial \phi} \end{bmatrix} = \begin{bmatrix} \sin \phi \cos \theta & -\rho \sin \phi \sin \theta & \rho \cos \phi \cos \theta \\ \sin \phi \sin \theta & \rho \sin \phi \cos \theta & \rho \cos \phi \sin \theta \\ \cos \phi & 0 & -\rho \sin \phi \end{bmatrix}.$$

cujo Jacobiano é (usando, por exemplo a segunda coluna para expandir em cofatores)

$$\begin{aligned} J\varphi &= \left| \rho \sin \phi \sin \theta \cdot \det \begin{bmatrix} \sin \phi \sin \theta & \rho \cos \phi \sin \theta \\ \cos \phi & -\rho \sin \phi \end{bmatrix} + \rho \sin \phi \cos \theta \cdot \det \begin{bmatrix} \sin \phi \cos \theta & \rho \cos \phi \cos \theta \\ \cos \phi & -\rho \sin \phi \end{bmatrix} \right| \\ &= \left| \rho \sin \phi \sin \theta \left(-\rho \sin^2 \phi \sin \theta - \rho \cos^2 \phi \sin \theta \right) + \rho \sin \phi \cos \theta \left(-\rho \sin^2 \phi \cos \theta - \rho \cos^2 \phi \cos \theta \right) \right| \\ &= \left| \rho \sin \phi \sin \theta \left(-\rho \sin \theta \right) + \rho \sin \phi \cos \theta \left(-\rho \cos \theta \right) \right| = \left| -\rho^2 \sin \phi \left(\sin^2 \theta + \cos^2 \theta \right) \right| \\ &= \left| -\rho^2 \sin \phi \right| = \rho^2 \sin \phi. \end{aligned}$$

Na última igualdade, utilizamos que $\phi \in (0, \pi)$ implica $\sin \phi > 0$, de modo que o valor absoluto está considerado corretamente. Portanto, a fórmula de mudança de variáveis implica que, em coordenadas esféricas,

$$\iint_K f(x, y, z) \, d(x, y, z) = \iiint_{\varphi^{-1}(K)} f(\rho \sin \phi \cos \theta, \rho \sin \phi \sin \theta, \rho \cos \phi) \rho^2 \sin \phi \, d(\rho, \theta, \phi),$$

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. E ainda, é mais natural, ou até mesmo “inescapável”, estudar formas diferenciais em uma superfície diferenciável ou 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 pontos essenciais da teoria e nossa discussão segue 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

$$F : \Omega \subseteq \mathbb{R}^n \longrightarrow \mathbb{R}^n.$$

É também comum “visualizar o campo F ” imaginando que, para $x \in \Omega$ fixado, o vetor $F(x)$ tem seu “ponto inicial” em x . Dessa forma, talvez devêssemos já escrever $F(x) \in T_x\mathbb{R}^n$ e, de fato, é o que vamos fazer em breve; no entanto, para tornar a notação menos carregada e as ideias mais elementares, adiamos esta formalidade. Além do mais, temos a identificação natural $T_x\mathbb{R}^n \simeq \mathbb{R}^n$ para todo $x \in \Omega$. Dada uma função diferenciável $f : \Omega \subseteq \mathbb{R}^n \longrightarrow \mathbb{R}$, podemos considerar o seu “campo gradiente” dado por

$$F = \nabla f : \Omega \subseteq \mathbb{R}^n \longrightarrow \mathbb{R}^n.$$

Lembramos que o gradiente é um operador que depende da estrutura de produto interno escolhida para o espaço \mathbb{R}^n (ver Observação 2 na Subseção 1.3). O diferencial de uma função, por outro lado, depende apenas da “estrutura diferenciável” do espaço \mathbb{R}^n e é uma aplicação do tipo

$$df : \Omega \subseteq \mathbb{R}^n \longrightarrow (\mathbb{R}^n)^*.$$

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

$$\omega : \Omega \longrightarrow (\mathbb{R}^n)^*.$$

Assim, para cada $x \in \Omega$, temos que $\omega(x) \in (\mathbb{R}^n)^*$ é um funcional linear em \mathbb{R}^n . Em outras palavras, para cada $x \in \Omega$, temos que $\omega(x) : \mathbb{R}^n \longrightarrow \mathbb{R}$ satisfaz

$$\omega(x) \cdot (av + bw) = a\omega(x) \cdot v + b\omega(x) \cdot w \quad \forall v, w \in \mathbb{R}^n \text{ e } \forall a, b \in \mathbb{R}.$$

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 38. Mostre que, para uma 1-forma $\omega : \Omega \longrightarrow (\mathbb{R}^n)^*$, as seguintes afirmações são equivalentes:

- (i) a forma ω é diferenciável (ou de classe C^k);
- (ii) cada uma das componentes a_i é diferenciável (ou de classe C^k);
- (iii) para cada $v \in \mathbb{R}^n$, a função $x \mapsto \omega(x)[v]$ é diferenciável (ou de classe C^k)

Exemplo 81. Dada uma função $f : \Omega \longrightarrow \mathbb{R}$ de classe C^k , o diferencial $df : \Omega \longrightarrow (\mathbb{R}^n)^*$, 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 $\omega : \Omega \longrightarrow (\mathbb{R}^n)^*$ e uma curva diferenciável $\gamma : [a, b] \rightarrow \Omega$, podemos definir a **integral de linha** de ω ao longo de γ por

$$\int_{\gamma} \omega := \int_a^b \omega(\gamma(t)) \cdot \gamma'(t) dt. \quad (43)$$

Poderíamos definir, mais geralmente, a integral de uma 1-forma ao longo de caminhos mais gerais, não necessariamente diferenciáveis. No entanto, a definição deveria “repetir” muito do que já foi feito anteriormente com somas de Riemann, deixaria nossos argumentos menos intuitivos e não viria associada a grandes vantagens teóricas. Logo, nós vamos nos restringir a caminhos diferenciáveis por partes. Para o leitor interessado, a integral de linha a partir de Somas de Riemann pode ser escrita como:

$$\int_{\gamma} \omega := \lim_{|\mathcal{P}| \rightarrow 0} \sum_{i=0}^{N+1} \omega(\gamma(t_i^*)) \cdot (\gamma(t_{i+1}) - \gamma(t_i)),$$

onde $\mathcal{P} = \{a = t_0 < t_1 < \dots < t_N < t_{N+1} = b\}$ é uma partição de $[a, b]$. Quando γ é diferenciável, o Teorema do Valor Médio pode ser utilizado para mostrar a equivalência das definições.

Mais imediato é estender a definição em (43) para caminhos C^1 por partes. Dizemos que $\gamma : [a, b] \longrightarrow \Omega$ 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 ω é então dada por

$$\int_{\gamma} \omega = \sum_{i=0}^N \int_{\gamma_i} \omega. \quad (44)$$

Isto vale pois o integrando em (43) pode estar definido a menos de um conjunto de medida nula.

Observamos que a integral de linha depende da curva e de sua orientação, mas não da parametrização, apesar de nossa notação aparentemente indicar o contrário.

Proposição 82. 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 83 (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 γ^- tem a orientação inversa de γ , 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

$$\gamma^-(t) = z(a + b - t), t \in [a, b],$$

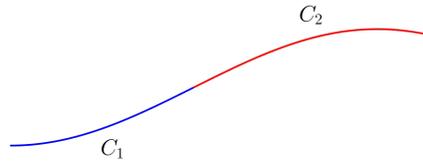
percorre a mesma curva no sentido contrário. Logo, (i) é uma simples mudança de coordenadas unidimensional.



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 (44)

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 84. 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,

$$\omega(x) = \sum_{i=1}^n a_i(x) dx^i \text{ é uma 1-forma} \iff F(x) = \sum_{i=1}^n a^i(x) e_i \text{ é um campo vetorial.}$$

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 em uma partícula que se movimenta ao longo da curva γ .

Uma forma diferencial ω é dita **exata** em Ω quando existe uma função diferenciável $f : \Omega \rightarrow \mathbb{R}$, conhecida como um potencial escalar de ω , tal que²⁹

$$\omega = df.$$

Em vista da Observação 84, a seguinte proposição não deve ser tão surpreendente (pensar na analogia entre formas exatas e campos conservativos).

Proposição 85. *Seja $\Omega \subseteq \mathbb{R}^n$ aberto e conexo e ω uma 1-forma. São equivalentes:*

- (i) ω é exata em Ω ;
- (ii) dados dois pontos $x, y \in \Omega$, a integral de linha de ω é independente do caminho C^1 por partes $\gamma : [a, b] \rightarrow \Omega$ que une os pontos x e y ;
- (iii) para toda curva C^1 por partes e fechada em Ω , a integral de ω é zero.

Demonstração. Vejamos que (i) \implies (ii). Por simplicidade, vamos supor $\gamma \in C^1([a, b]; \Omega)$, o caso geral ficando como exercício. Sendo ω exata em Ω , existe um potencial escalar $f : \Omega \rightarrow \mathbb{R}$ tal que $\omega = df$. Logo, pela Regra da Cadeia,

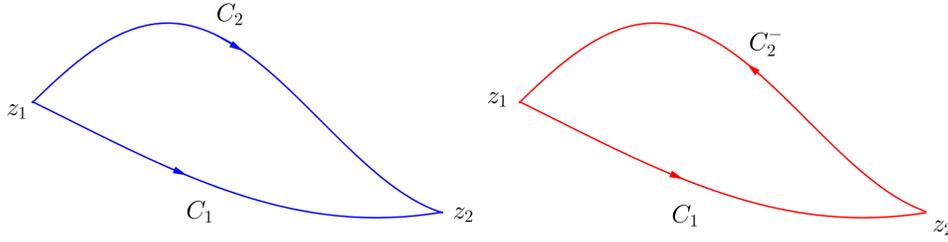
$$\int_{\gamma} \omega = \int_a^b df(\gamma(t)) \cdot \gamma'(t) dt = \int_a^b \frac{d}{dt} f(\gamma(t)) dt = f(\gamma(b)) - f(\gamma(a)) = f(y) - f(x),$$

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 γ_1 e γ_2 . A curva $\gamma := \gamma_1 * \gamma_2^{-}$, como na figura, é fechada.

²⁹Uma possível analogia é pensar que formas exatas estão para formas diferenciais assim como campos gradiente estão 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 ω depende apenas das extremidades de cada curva.

Finalmente, mostramos que (ii) \implies (i). Fixamos $x_0 \in \Omega$ qualquer. Sendo Ω aberto e conexo, é conexo por caminhos. Dado $x \in \Omega$, consideramos γ_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 suficientemente pequeno, então o segmento de reta que une x até $x+v$ está todo contido em Ω . Parametrizando este segmento por $\gamma(t) = x + tv$, $t \in [0, 1]$, podemos escrever

$$f(x+v) - f(x) - \omega(x) \cdot v = \int_{\gamma_x * \gamma} \omega - \int_{\gamma_x} \omega - \omega(x) \cdot v = \int_0^1 [\omega(x+tv) - \omega(x)] \cdot v dt.$$

Agora, pela continuidade de ω , dado $\varepsilon > 0$, existe $\delta > 0$ tal que

$$|h| < \delta \implies \|\omega(x+h) - \omega(x)\| < \varepsilon.$$

Portanto, $|v| < \delta$ implica, como queríamos, que

$$|f(x+v) - f(x) - \omega(x) \cdot v| \leq \varepsilon|v|,$$

isto é, f é diferenciável em x com $df(x) = \omega(x)$. □

Exemplo 86. 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 ω não é exata. Esta forma é conhecida como “*elemento de ângulo*”. O nome vem da seguinte propriedade (que passamos a descrever abaixo) de que a integral da forma elemento de ângulo pode ser interpretada como 2π vezes o número líquido de voltas que uma curva fechada γ faz em torno da origem no sentido anti-horário.

A palavra “líquido” se refere ao fato de que devemos contar o número de voltas no sentido anti-horário e subtrair o número de voltas no sentido horário.

Mais precisamente, dada uma curva $\gamma : [a, b] \rightarrow \mathbb{R}^2 \setminus \{(0, 0)\}$ qualquer (não necessariamente fechada), a escrevemos em coordenadas polares

$$\gamma(t) = (x(t), y(t)) = (r(t) \cos \theta(t), r(t) \sin \theta(t)).$$

Então, para calcular a integral de linha de ω ao longo de γ , precisamos das derivadas com respeito a t das componentes desta curva γ :

$$\begin{aligned} x' &= r' \cos \theta - r\theta' \sin \theta \\ y' &= r' \sin \theta + r\theta' \cos \theta \end{aligned}$$

o 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, como afirmado,

$$\int_{\gamma} \omega = \int_a^b \frac{-y(t)x'(t) + x(t)y'(t)}{x(t)^2 + y(t)^2} dt = \int_a^b \theta'(t) dt = \theta(b) - \theta(a).$$

Observamos ainda que a variação do ângulo θ é positiva quando o movimento é no sentido anti-horário. Além disso, no caso de a curva γ ser fechada, tem-se que $\theta(b) - \theta(a)$ é um múltiplo de 2π , de maneira que

$$\frac{1}{2\pi} \int_{\gamma} \omega$$

é um número inteiro conhecido como “*número de voltas*” (do inglês, “winding number” poderia ser algo como “número de enrolamentos” ou “número de revoluções”).³⁰

Além do interesse intrínseco acima, a forma elemento de ângulo 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 $\text{Ger}\{u, v\} \subset \mathbb{R}^3$.

Esta definição, no entanto, não está 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

³⁰O 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 sua integral ao longo de um caminho fechado é zero. Logo,

$$\frac{1}{2\pi i} \int_{\gamma} \frac{dz}{z} = \frac{1}{2\pi} \int_{\gamma} \omega = \text{número de voltas (líquido) que } \gamma \text{ percorre em torno da origem, no sentido anti-horário.}$$

definição é especial da dimensão 3: em dimensão 4, por exemplo, dados dois vetores quaisquer, existiria uma infinidade de vetores com as duas propriedades acima.

O nosso objetivo inicial passa a ser definir um produto entre dois vetores de um espaço vetorial V de dimensão finita qualquer que tenha propriedades semelhantes às do produto vetorial. 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 o símbolo \wedge e chamá-lo de **produto exterior** ou do inglês de **wedge product** (em referência ao símbolo do produto). Definimos a **álgebra exterior** como a álgebra ΛV gerada pelos elementos de V com o produto abstrato \wedge e a relação

$$u \wedge v = -v \wedge u.$$

Um pouco mais precisamente³¹, ΛV consiste de todas as constantes, todos os vetores de V , todos os produtos formais da forma $u \wedge v$ e combinações lineares destes, todos os produtos formais da forma $u \wedge v \wedge w$ e combinações lineares destes, e assim por diante.

Em termos práticos, o que estamos fazendo aqui é analisar o que aconteceria se definíssemos um produto de vetores de V que é abstrato e cuja única propriedade conhecida é a anti-simetria. Daí nós “juntamos” as constantes, os vetores, os produtos de dois vetores e de três vetores em um conjunto “grande” que será a álgebra exterior.

Como um exemplo importante, consideramos $V = (\mathbb{R}^3)^*$ e vamos descrever explicitamente o conjunto $\Lambda(\mathbb{R}^3)^*$, cujos elementos são:

- Todas as constantes (reais) em \mathbb{R} que, por simplicidade de notação, será também denotado por $\mathbb{R} := \Lambda^0(\mathbb{R}^3)^*$;
- Todos os vetores do espaço dual $(\mathbb{R}^3)^* =: \Lambda^1(\mathbb{R}^3)^*$:

$$v = v_x dx + v_y dy + v_z dz.$$

Acima, v_x, v_y e $v_z \in \mathbb{R}$ denotam as componentes do funcional linear v na base canônica $\{dx, dy, dz\}$ do espaço dual;

- 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

$$\begin{aligned} v \wedge w &= (v_x dx + v_y dy + v_z dz) \wedge (w_x dx + w_y dy + w_z dz) \\ &= (v_y w_z - v_z w_y) dy \wedge dz + (v_z w_x - v_x w_z) dz \wedge dx + (v_x w_y - v_y w_x) dx \wedge dy \\ &= \det \begin{bmatrix} v_y & v_z \\ w_y & w_z \end{bmatrix} dy \wedge dz - \det \begin{bmatrix} v_x & v_z \\ w_x & w_z \end{bmatrix} dz \wedge dx + \det \begin{bmatrix} v_x & v_y \\ w_x & w_y \end{bmatrix} dx \wedge dy; \end{aligned} \quad (45)$$

A observação crucial é que as componentes do vetor $v \wedge w$ acima são essencialmente as mesmas componentes do produto vetorial $v \times w$, exceto que não aparece qualquer tipo de orientação artificial como na regra da mão direita. Por um lado pode parecer um ponto negativo que $v \wedge w$ não é exatamente um “vetor comum” e sim um “2-vetor”. Por outro, este novo objeto possui as mesmas propriedades (i) e (ii) que pedíamos acima e independe de uma regra convencionada.

- Quaisquer combinações lineares de elementos de grau 2 como acima, formando um espaço vetorial denotado por $\Lambda^2(\mathbb{R}^3)^*$:

$$\Lambda^2(\mathbb{R}^3)^* := \left\{ v_{yz} dy \wedge dz + v_{zx} dz \wedge dx + v_{xy} dx \wedge dy; v_{yz}, v_{zx}, v_{xy} \in \mathbb{R} \right\};$$

³¹Isto será estudado com maior rigor em breve. O objetivo desta seção é o de construir a “álgebra exterior” no espaço tridimensional \mathbb{R}^3 de maneira mais informal e intuitiva.

- Elementos de “grau três” elementares ou 3-vetores elementares e quaisquer combinações lineares deles:

$$\begin{aligned}
u \wedge v \wedge w &= (u_x dx + u_y dy + u_z dz) \wedge (v \wedge w) \\
&= u \wedge \left[(v_y w_z - v_z w_y) dy \wedge dz + (v_z w_x - v_x w_z) dz \wedge dx + (v_x w_y - v_y w_x) dx \wedge dy \right] \\
&= \left\{ u_x \det \begin{bmatrix} v_y & v_z \\ w_y & w_z \end{bmatrix} - u_y \det \begin{bmatrix} v_x & v_z \\ w_x & w_z \end{bmatrix} + u_z \det \begin{bmatrix} v_x & v_y \\ w_x & w_y \end{bmatrix} \right\} dx \wedge dy \wedge dz \\
&= \det \begin{bmatrix} u_x & u_y & u_z \\ v_x & v_y & v_z \\ w_x & w_y & w_z \end{bmatrix} dx \wedge dy \wedge dz.
\end{aligned} \tag{46}$$

Estes elementos (e combinações lineares deles) formam um espaço vetorial denotado por $\Lambda^3(\mathbb{R}^3)^*$ que neste caso é unidimensional. Observamos ainda que (quase que gratuitamente) o produto exterior de três vetores captura ainda o chamado “produto misto” da geometria analítica, já que

$$\langle u \times v, w \rangle = \det \begin{bmatrix} u_x & u_y & u_z \\ v_x & v_y & v_z \\ w_x & w_y & w_z \end{bmatrix},$$

fazendo que os conceitos estejam unificados em um só. O produto exterior de três vetores em \mathbb{R}^3 captura, portanto, o volume com sinal do paralelepípedo gerado pelos três vetores $u, v, w \in \mathbb{R}^3$.

- Estes são todos os elementos da álgebra, 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 $dx \wedge dy \wedge dx \wedge dz = 0$.³² 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

$$\begin{aligned}
a + \sum_{i=1}^3 b_i dx^i + \sum_{i,j=1;i<j}^3 c_{ij} dx^i \wedge dx^j + d dx \wedge dy \wedge dz \quad \text{ou, se preferir,} \\
\left(a, \sum_{i=1}^3 b_i dx^i, \sum_{i,j=1;i<j}^3 c_{ij} dx^i \wedge dx^j, d dx \wedge dy \wedge dz \right).
\end{aligned}$$

Por economia de notação, escrevemos acima $dx^1 = dx$, $dx^2 = dy$ e $dx^3 = dz$.

Antes de proceder à construção formal do produto exterior e da álgebra exterior, nós vamos formalizar a noção de produto tensorial; contudo, seria possível construir diretamente a álgebra exterior, ver Exercício ??.

18 Produto tensorial

Apesar de nossas notas estarem em um nível (um pouco) mais elementar, o texto de Keith Conrad, disponível [online](#)³³, 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.

Existem pelo menos três abordagens (aparentemente) diferentes para falar de tensores:

³²Cuidado! Não é sempre verdade que $v \wedge v = 0$. Em 4 dimensões, $v = dx \wedge dy + dz \wedge dw$ satisfaz $v \wedge v = 2dx \wedge dy \wedge dz \wedge dw$.

³³<http://www.math.uconn.edu/~kconrad/blurbs/linmultialg/tensorprod.pdf>

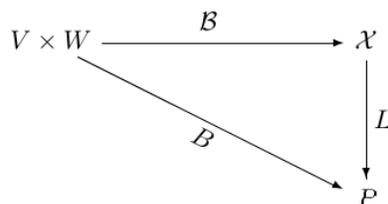
- “algébrica”: encontrada na maioria dos livros de álgebra, caracteriza tensores como objetos de um espaço vetorial que tem uma “propriedade universal”.
- “multilinear”: define tensores como aplicações multilineares entre espaços vetoriais.
- “em componentes”: apresentada em alguns cursos de física (mas não apenas), por exemplo no curso de Leonard Susskind³⁴, descreve tensores como um conjunto de componentes com um ou vários índices que se transformam de maneira particular quando realizamos mudanças de coordenadas.

É de nosso entendimento ser útil conhecer todas as três descrições acima e como relacioná-las entre si. Iniciamos com a descrição algébrica.

18.1 Produto tensorial de dois espaços vetoriais

De forma 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 87. *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: dada qualquer aplicação bilinear $B : V \times W \rightarrow P$, existe exatamente uma aplicação linear $L : \mathcal{X} \rightarrow P$ tal que $B = L \circ \mathcal{B}$. A notação utilizada é

$$\mathcal{X} = V \otimes W \quad \text{e} \quad \mathcal{B} = \otimes.$$

Os elementos do espaço vetorial $V \otimes W$ 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,

$$\otimes(v, w) = v \otimes w \in V \otimes W,$$

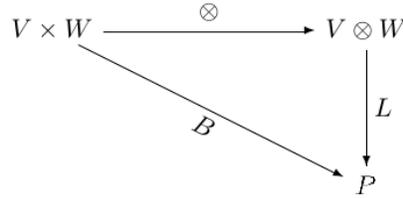
onde a igualdade acima representa apenas uma notação, que é a mais usual. Elementos da forma $v \otimes w$ são chamados **tensores elementares**. Notamos, no entanto, que nem todo tensor é elementar, *e.g.*, $e_1 \otimes e_1 + e_2 \otimes e_2 \in \mathbb{R}^2 \otimes \mathbb{R}^2$ não é um tensor elementar; isto pode ser visto como uma consequência do Teorema 88 abaixo.

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

$$B(v, w) = L(v \otimes w).$$

³⁴“Einstein’s General Relativity”, Lecture 4, a partir do minuto 00:02:00, disponível no YouTube no link [AQUI](#). O curso desse professor de Stanford é muito bom, btw :)

Isto é o que se quer dizer quando se fala: “o produto tensorial tem a propriedade de transformar aplicações bilineares em aplicações lineares”.



Propositalmente, não explicitamos o corpo \mathcal{K} dos escalares utilizado, 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 88 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 87. 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 menor número de restrições possível”. 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, a definição de espaço quociente em [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 ℓ passa³⁵ a uma aplicação (quociente) em \mathcal{X} , devemos mostrar que ℓ se anula em \mathfrak{N} , isto é, que $\mathfrak{N} \subseteq \ker \ell$. Fazemos uma das propriedades: considere o elemento

$$(v_1 + v_2, w) - (v_1, w) - (v_2, w) \in \mathfrak{N}.$$

Sendo ℓ 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, temos o seguinte:

Teorema 88. 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, para mostrar que o conjunto $\text{Span}\{v_i \otimes w_j\}$ gera $V \otimes W$ basta verificar que $\text{Span}\{v_i \otimes w_j\}$ gera qualquer tensor elementar. De fato, por bilinearidade,

$$v \otimes w = \left(\sum_{i=1}^n x^i v_i \right) \otimes \left(\sum_{j=1}^m y^j w_j \right) = \sum_{i=1}^n \sum_{j=1}^m x^i y^j v_i \otimes w_j. \quad (47)$$

Em seguida, verificamos que $\{v_i \otimes w_j; i = 1, 2, \dots, n \text{ e } j = 1, 2, \dots, m\}$ é linearmente independente. Para isto, consideramos constantes $c_{ij} \in \mathbb{R}$ tais que

$$\sum_{i=1}^n \sum_{j=1}^m c_{ij} v_i \otimes w_j = 0. \quad (48)$$

³⁵Isso significa garantir que a aplicação L abaixo fica bem definida.

Devemos verificar que $c_{ij} = 0$ para quaisquer índices i, j . Dados $k \in \{1, 2, \dots, n\}$ e $\ell \in \{1, 2, \dots, m\}$, definimos

$$B : V \times W \longrightarrow \mathbb{R} \\ (v, w) \longmapsto x^k y^\ell, \quad \text{onde} \quad v = \sum_{i=1}^n x^i v_i \quad \text{e} \quad w = \sum_{j=1}^m y^j w_j,$$

que é uma transformação bilinear. Pela propriedade universal de $V \otimes W$, existe uma transformação linear $L : V \otimes W \longrightarrow \mathbb{R}$ tal que $L(v \otimes w) = B(v, w) = x^k y^\ell$. Observamos que³⁶

$$L(e_i \otimes e_j) = \begin{cases} 1, & \text{se } i = k \text{ e } j = \ell \\ 0, & \text{caso contrário} \end{cases}.$$

Aplicando L em ambos os lados de (48) obtemos, como queríamos, que $c_{k\ell} = 0$. □

Observação 89. A Equação (47) 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 (deve-se pensar em coordenadas nas bases $v_i \otimes w_j$) e se escreve

$$v \otimes w = [x^i y^j]_{n \times m} = \begin{bmatrix} x^1 y^1 & x^1 y^2 & \dots & x^1 y^m \\ x^2 y^1 & x^2 y^2 & \dots & x^2 y^m \\ \vdots & \vdots & \ddots & \vdots \\ x^n y^1 & x^n y^2 & \dots & x^n y^m \end{bmatrix}.$$

Esta descrição não fornece um isomorfismo canônico entre $\mathbb{R}^n \otimes \mathbb{R}^m$ e matrizes de ordem $m \times n$ porque requer que escrevamos os vetores em bases fixadas. Mais natural é a identificação da Subseção 18.2.2 abaixo: $\mathcal{L}(\mathbb{R}^n; \mathbb{R}^m) \simeq \mathbb{R}^m \otimes (\mathbb{R}^n)^*$.

18.2 Identificações naturais

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 da subseção anterior.

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.

18.2.1 Vetores são tensores: $V \simeq \mathbb{R} \otimes V$

Suponhamos que V é um espaço vetorial real de dimensão d . Intuitivamente, fixada uma base $\{v_i\}$ de V , 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 de verificar que

$$\Phi : r \otimes v \longmapsto rv \tag{49}$$

se estende (por linearidade) a um isomorfismo canônico entre $\mathbb{R} \otimes V$ e V (isomorfismo que não faz referência a nenhuma base nem produto interno).

Neste momento, cabe a seguinte observação: em (49), nós definimos a imagem de um tensor elementar pela aplicação Φ . Pode-se verificar diretamente, pela bilinearidade da aplicação

³⁶A ideia toda é construir uma aplicação linear que satisfaz esta propriedade. Olhando para a propriedade universal, deve se tornar mais natural a definição da transformação bilinear B .

$(r, v) \mapsto r \otimes v$, que a extensão a todo $\mathbb{R} \otimes V$ de fato fica bem definida. Por outro lado, a construção de aplicações lineares a partir de relações de bilinearidade é justamente o que a propriedade universal do espaço tensorial permite fazer. Mais elegante, portanto, é o que segue. Definimos

$$\begin{aligned} B : \mathbb{R} \times V &\longrightarrow V \\ (r, v) &\longmapsto rv \end{aligned}$$

que é uma aplicação bilinear. Pela propriedade universal de $\mathbb{R} \otimes V$, existe única aplicação linear $\Phi : \mathbb{R} \otimes V \longrightarrow V$ tal que

$$\Phi(r \otimes v) = B(r, v) = rv.$$

Esta é a aplicação em (49). Fica como exercício a verificação de que é um isomorfismo.

18.2.2 Transformações lineares são tensores: $\mathcal{L}(W; V) \simeq V \otimes W^*$

A um tensor elementar $v \otimes \varphi \in V \otimes W^*$, associamos uma transformação linear $v \otimes \varphi : W \longrightarrow V$ da seguinte maneira:

$$(v \otimes \varphi)(w) := \varphi(w)v. \quad (50)$$

Estas são transformações lineares de posto um. Logo, não deve ser surpresa que aplicações lineares entre espaços de dimensão finita podem ser escritas como soma de aplicações deste tipo, como passamos a descrever.

Para definir uma aplicação linear entre $V \otimes W^*$ e $\mathcal{L}(W; V)$, nós utilizamos, como na subseção anterior, a propriedade universal. Definimos

$$\begin{aligned} B : V \times W^* &\longrightarrow \mathcal{L}(W; V) \\ (v, \varphi) &\longmapsto T_{v, \varphi} \end{aligned}$$

onde denotamos, temporariamente, por $T_{v, \varphi}$ a aplicação linear que aparece em (50), isto é, para cada $w \in W$, tem-se $T_{v, \varphi}(w) := \varphi(w)v$. Observando que B é uma aplicação bilinear, temos que existe única aplicação linear $\Phi : V \otimes W^* \longrightarrow \mathcal{L}(W; V)$ que satisfaz

$$\Phi(v \otimes \varphi) = T_{v, \varphi}.$$

A notação usual, já utilizada em (50), identifica $T_{v, \varphi} = \Phi(v \otimes \varphi) \simeq v \otimes \varphi$.

Para provar que Φ é um isomorfismo, vamos mostrar que é sobrejetiva, isto é, que qualquer transformação linear $T \in \mathcal{L}(W; V)$ pode ser obtida como combinações lineares de aplicações de posto um (que equivalem a tensores elementares). Isto é suficiente já que Φ é uma aplicação linear entre espaços vetoriais de mesma dimensão. Sejam $\{v_i\}$ e $\{w_j\}$ bases de V e W , respectivamente. Se $\{dw^j\}$ é base de W^* , dual a $\{w_j\}$, temos, para todo $x \in W$,

$$T(x) = T\left(\sum_{j=1}^d x^j w_j\right) = \sum_{j=1}^d x^j T(w_j) = \sum_{j=1}^d \sum_{i=1}^m (x^j T_j^i) v_i = \sum_{i,j} T_j^i dw^j(x) v_i = \sum_{i,j} T_j^i \Phi(v_i \otimes dw^j)(x).$$

Logo, como queríamos,

$$T = \sum_{i,j} T_j^i \Phi(v_i \otimes dw^j) = \Phi\left(\sum_{i,j} T_j^i v_i \otimes dw^j\right).$$

Como já mencionado, é usual omitir o isomorfismo Φ e escrever

$$T = \sum_{i,j} T_j^i v_i \otimes dw^j.$$

Observação 90. 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 87.

18.2.3 Transformações bilineares são tensores: $\mathcal{B}(V \times W; \mathbb{R}) \simeq V^* \otimes W^*$

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 estende³⁷ 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 Φ : 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 91. 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 de segunda ordem podem ser pensados como aplicações bilineares do tipo acima.

Observação 92. Por dualidade, também é possível exibir um isomorfismo natural

$$\mathcal{B}(V^* \times W^*; \mathbb{R}) \simeq V \otimes W.$$

Neste caso, a identificação é dada por

$$v \otimes w \in V \otimes W \xrightarrow{\Phi} (v, w)(\varphi, \psi) := \varphi(v)\psi(w).$$

Isto é coerente com a identificação natural entre qualquer espaço vetorial V , de dimensão finita, e seu bidual V^{**} .

18.3 Produto tensorial de vários espaços vetoriais

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”. O Teorema 87 pode ser enunciado com completa analogia e sua prova é também análoga – basta adicionar “mais relações de linearidade” na definição do subespaço \mathfrak{X} , uma para cada uma das componentes da r -upla (v_1, v_2, \dots, v_r) .

Além disso, assim como na subseção anterior, aplicações multilineares são naturalmente tensores através do isomorfismo canônico

$$\mathcal{B}(V_1, V_2, \dots, V_r; \mathbb{R}) \simeq V_1^* \otimes V_2^* \otimes \dots \otimes V_r^*$$

que é 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).

³⁷Como anteriormente, poderíamos tanto mostrar que a extensão linear está bem definida, o que é um pouco mais “braçal”, quanto utilizar a propriedade universal de $V^* \otimes W^*$. Deixamos os detalhes ao leitor.

No caso especial em que $V_i = V$ para todo i , utilizamos as notações especiais

$$\mathcal{T}^r V = \otimes^r V := V \otimes V \otimes \cdots \otimes V. \quad (51)$$

Denotando o espaço vetorial das formas r -lineares em V por $\mathcal{L}_r(V)$, isto é,

$$\mathcal{L}_r(V) := \mathcal{B}(V, V, \dots, V; \mathbb{R}),$$

tiramos da discussão acima que há um isomorfismo canônico

$$\otimes^r V^* \simeq \mathcal{L}_r(V).$$

Lembramos as propriedades básicas das formas r -lineares e deixamos sua demonstração como exercício para o leitor.

Exercício 39. 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 $i_j \in \{1, 2, \dots, n\}$.

Exercício 40. Um produto interno no espaço vetorial real \mathbb{R}^n é uma aplicação bilinear, simétrica e positiva definida, usualmente denotada como

$$(v, w) \mapsto \langle v, w \rangle.$$

Dada uma matriz A simétrica e positiva definida, podemos associar um produto interno pela fórmula

$$\Phi : A \mapsto \langle \cdot, \cdot \rangle_A := Av \cdot w,$$

onde o ponto acima é o produto escalar usual de \mathbb{R}^n .

- (i) Mostre que todo produto interno é da forma acima, isto é, mostre que Φ é um isomorfismo.
- (ii) Considerando as bases canônicas de \mathbb{R}^n e de $(\mathbb{R}^n)^*$, explique a identidade

$$\langle \cdot, \cdot \rangle_A = \sum_{i,j=1}^n a_{ij} dx^i \otimes dx^j.$$

O item anterior nos diz que todo produto interno de \mathbb{R}^n é um tensor deste tipo.

Campos de tensores deste tipo em uma variedade M são as métricas Riemannianas, como veremos (ou não!).

18.4 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 \boxtimes de modo que

1) $(X, +, \boxtimes)$ 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, \boxtimes) é um semigrupo com identidade: para todo $x, y, z \in X$, vale
 - Associatividade: $(x \boxtimes y) \boxtimes z = x \boxtimes (y \boxtimes z)$;
 - Identidade multiplicativa: existe $1 \in X$ tal que $x \boxtimes 1 = 1 \boxtimes x = x$;
- (iii) Distributividade da multiplicação com respeito à adição:

$$x \boxtimes (y + z) = x \boxtimes y + x \boxtimes z \quad \text{e} \quad (x + y) \boxtimes z = x \boxtimes z + y \boxtimes z.$$

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: $a(bx) = (ab)x$;
 - Unidade escalar: $1x = x$;
 - Distributividade da multiplicação por escalar com respeito à adição em X :

$$a(x + y) = ax + ay;$$

- Distributividade da multiplicação por escalar com respeito à adição em \mathbb{K} :

$$(a + b)x = ax + bx.$$

3) Associatividade do produto com o produto por escalar:

$$c(x \boxtimes y) = (cx) \boxtimes y = x \boxtimes (cy).$$

A soma direta dos espaços tensoriais definidos na seção anterior forma uma álgebra associativa com a operação de produto dada pelo produto tensorial de vetores. Mais precisamente, nós lembramos da notação em (51) e denotamos

$$\mathcal{T}^0 V := \mathbb{R} \quad \text{e} \quad \mathcal{T}^1 V := V.$$

A **álgebra tensorial** é definida como sendo o conjunto

$$\mathcal{T}V = \bigoplus_{k=0}^{+\infty} \mathcal{T}^k V := \mathbb{R} \oplus V \oplus \mathcal{T}^2 V \oplus \dots \oplus \mathcal{T}^k V \oplus \dots$$

e o produto de elementos de ordens diferentes $x \in \mathcal{T}^r V$ e $y \in \mathcal{T}^s V$ é dado por concatenação:

$$x \otimes y \in \mathcal{T}^{r+s} V.$$

19 Produto exterior

Queremos definir o produto exterior ΛV como a álgebra associativa “mais geral” que satisfaça a propriedade de “anticomutatividade” do produto. Nesta seção, vamos formalizar esta ideia.

Dado um espaço vetorial V , o produto exterior de k cópias de V , denotado por

$$\Lambda^k V = V \wedge V \wedge \dots \wedge V,$$

é obtido como um quociente do produto tensorial $\mathcal{T}^k V = V \otimes V \otimes \dots \otimes V$ ao adicionar “relações de alternância”. Mais precisamente, o **produto exterior** de k cópias de V é definido por

$$\Lambda^k V := \mathcal{T}^k V / \mathcal{S}_k,$$

onde \mathcal{S}_k é o subespaço vetorial de $\mathcal{T}^k V$ gerado pelos elementos da forma

$$v_1 \otimes v_2 \otimes \cdots \otimes v_k \quad \text{onde} \quad v_i = v_j \quad \text{para algum par de índices } i, j \text{ com } i \neq j.$$

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 $\mathcal{T}^k V \rightarrow \Lambda^k V$ é sobrejetiva e todo tensor de ordem k é gerado como soma (finita) de tensores elementares. Daí segue que todo elemento de $\Lambda^k V$ pode ser escrito como soma (finita) de produtos exteriores elementares: todo $\omega \in \Lambda^k V$ pode ser escrito como

$$\omega = \sum_{i=1}^N v_1^i \wedge v_2^i \wedge \cdots \wedge v_k^i.$$

Observação 93. 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, pela própria definição do espaço quociente,

$$v_1 \wedge v_2 \wedge \cdots \wedge v_k = 0.$$

Assim, se $k > n$, todos os tensores elementares de ordem k pertencem a \mathcal{S}_k e, logo, $\Lambda^k V = \{0\}$.

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 quaisquer.

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 descrevermos bases para os espaços $\Lambda^k V$ onde $k \in \{2, 3, \dots, n\}$.

Consideramos um produto exterior elementar da forma

$$v_1 \wedge v_2 \wedge \cdots \wedge v_k.$$

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,

$$\begin{aligned} v_1 \wedge v_2 \wedge \cdots \wedge v_k &= \left(\sum_{j_1=1}^n a_{1j_1} dx^{j_1} \right) \wedge \left(\sum_{j_2=1}^n a_{2j_2} dx^{j_2} \right) \wedge \cdots \wedge \left(\sum_{j_k=1}^n a_{kj_k} dx^{j_k} \right) \\ &= \sum_{j_1=1}^n \sum_{j_2=1}^n \cdots \sum_{j_k=1}^n a_{1j_1} a_{2j_2} \cdots a_{kj_k} dx^{j_1} \wedge dx^{j_2} \wedge \cdots \wedge dx^{j_k}. \end{aligned} \quad (52)$$

Vários dos produtos exteriores elementares acima são “repetidos” ou iguais a “zero”, por causa da propriedade de anticomutatividade. De fato, quando dois índices 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

$$I := \{i_1 < i_2 < \cdots < i_k\} = \{j_1, j_2, \dots, j_k\}$$

que reordena os índices j em ordem crescente. Qualquer permutação σ pode ser escrita como produto (composição) de transposições

$$\sigma = \sigma_1 \sigma_2 \cdots \sigma_{N(\sigma)}.$$

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 σ pode ser escrita como produto de transposições

$$\sigma = \sigma_1 \sigma_2 \cdots \sigma_N.$$

É possível mostrar [2, Corolário 4.19 e Lema 1.15] que a “paridade de N ” é um invariante de σ (é a paridade que é invariante, não o número N , que nem único é). Em outras palavras, dadas duas representações de σ 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 σ :

$$\text{sgn}(\sigma) := (-1)^{N(\sigma)} = \begin{cases} 1, & \text{se } N(\sigma) \text{ é par} \\ -1, & \text{se } N(\sigma) \text{ é ímpar} \end{cases}$$

Com estas considerações em mente, para índices distintos fixados j_1, j_2, \dots, j_k , consideramos um reordenamento do tipo

$$I := \{j_1, j_2, \dots, j_k\} = \{i_1 < i_2 < \cdots < i_k\}$$

e a permutação $\sigma : I \rightarrow I$ dada por

$$\sigma(i_1) = j_1, \quad \sigma(i_2) = j_2, \quad \dots, \quad \sigma(i_k) = j_k.$$

Vamos justificar que podemos escrever

$$dx^{j_1} \wedge dx^{j_2} \wedge \cdots \wedge dx^{j_k} = dx^{\sigma(i_1)} \wedge dx^{\sigma(i_2)} \wedge \cdots \wedge dx^{\sigma(i_k)} = \text{sgn}(\sigma) dx^{i_1} \wedge dx^{i_2} \wedge \cdots \wedge dx^{i_k}.$$

De fato, a cada transposição feita, deve-se efetuar uma troca de sinal donde o sinal da permutação surge naturalmente. Portanto, denotando o conjunto de permutações de k elementos de I por $S_k(I)$, (52) pode ser reescrita como

$$\begin{aligned} v_1 \wedge v_2 \wedge \cdots \wedge v_k &= \sum_{1 \leq i_1 < i_2 < \cdots < i_k \leq n} \left(\sum_{\sigma \in S_k(I)} \text{sgn}(\sigma) a_{1\sigma(i_1)} a_{2\sigma(i_2)} \cdots a_{k\sigma(i_k)} \right) dx^{i_1} \wedge dx^{i_2} \wedge \cdots \wedge dx^{i_k} \\ &= \sum_{1 \leq i_1 < i_2 < \cdots < i_k \leq n} \det \begin{bmatrix} a_{1i_1} & \cdots & a_{1i_k} \\ \vdots & \ddots & \vdots \\ a_{ki_1} & \cdots & a_{ki_k} \end{bmatrix} dx^{i_1} \wedge dx^{i_2} \wedge \cdots \wedge dx^{i_k}. \end{aligned} \quad (53)$$

A Fórmula (53) está em pleno acordo com os cálculos informais que fizemos na Seção 17, ver Fórmulas (45) e (46). É também a essência da prova do seguinte resultado:

Proposição 94. *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,*

$$\mathfrak{B}_k := \{dx^{i_1} \wedge dx^{i_2} \wedge \cdots \wedge dx^{i_k}; 1 \leq i_1 < i_2 < \cdots < i_k \leq n\}$$

é uma base de $\Lambda^k V$. Em particular,

$$\dim \Lambda^k V = \binom{n}{k} = \frac{n!}{k!(n-k)!}.$$

Demonstração. Pela Fórmula (53), 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: $dx^1 \wedge dx^2 \wedge \cdots \wedge dx^n \neq 0$, pois $dx^1 \otimes dx^2 \otimes \cdots \otimes dx^n \notin \mathcal{S}_n$. Segue que \mathfrak{B}_n é linearmente independente. Para $2 \leq k < n$, suponhamos que

$$\sum_{1 \leq i_1 < i_2 < \cdots < i_k \leq n} c_{i_1 i_2 \dots i_k} dx^{i_1} \wedge dx^{i_2} \wedge \cdots \wedge dx^{i_k} = 0. \quad (54)$$

Afirmamos que, neste caso, todos os coeficientes $c_{i_1 i_2 \dots i_k}$ são nulos. Isto pode ser verificado como segue: vamos fazer o produto exterior “com os elementos que faltam” para chegar em $dx^1 \wedge dx^2 \wedge \cdots \wedge dx^n$. Dado o conjunto de índices $I = \{1 \leq i_1 < i_2 < \cdots < i_k \leq n\}$, consideramos o conjunto $J = \{j_1, j_2, \dots, j_{n-k}\}$ que contém os índices complementares aos índices de I , isto é,

$$I \cup J = \{1, 2, \dots, n\}.$$

Analisando a cardinalidade dos conjuntos, inferimos que eles devem ser disjuntos. Notamos ainda que todos os produtos exteriores elementares que aparecem na soma em (54) possuem algum índice que está em J , exceto aquele cujos índices são exatamente os de I . Logo,

$$0 = \left(\sum c_{i_1 i_2 \dots i_k} dx^{i_1} \wedge dx^{i_2} \wedge \cdots \wedge dx^{i_k} \right) \wedge dx^{j_1} \wedge dx^{j_2} \wedge \cdots \wedge dx^{j_{n-k}} = \pm c_{i_1 i_2 \dots i_k} dx^1 \wedge dx^2 \wedge \cdots \wedge dx^n.$$

Segue, como queríamos, que $c_{i_1 i_2 \dots i_k} = 0$.

A afirmação sobre a dimensão de $\Lambda^k V$ fica como um exercício (de análise combinatória). \square

19.1 Álgebra exterior ou álgebra de Grassmann



falar sobre grassmann e motivação histórica

A **álgebra de Grassmann** ou **álgebra exterior** é definida como a soma direta

$$\Lambda V := \mathbb{R} \oplus V \oplus \Lambda^2 V \oplus \Lambda^3 V \oplus \cdots \oplus \Lambda^{n-1} V \oplus \Lambda^n V,$$

onde o produto entre elementos de ordens distintas é determinado pela aplicação

$$\wedge : \Lambda^k V \times \Lambda^m V \longrightarrow \Lambda^{k+m} V,$$

que define uma operação bilinear, obtida a partir de produtos elementares por “concatenação”. Mais explicitamente, para elementos

$$\omega = \sum_I a_I dx^I \in \Lambda^k V \quad \text{e} \quad \eta = \sum_J a_J dx^J \in \Lambda^m V$$

definimos

$$\omega \wedge \eta := \sum_{I,J} a_I b_J dx^I \wedge dx^J \in \Lambda^{k+m}.$$

A fórmula binomial usual

$$(x+y)^n = \sum_{k=0}^n \binom{n}{k} x^{n-k} y^k,$$

aplicada em $x = y = 1$ 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, mas vale o seguinte.

Exercício 41. Mostre que, se $\omega \in \Lambda^k V$ e $\eta \in \Lambda^m V$, então

$$\omega \wedge \eta = (-1)^{km} \eta \wedge \omega.$$

Exercício 42. 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

$$\phi(v, v) = 0 \quad \text{para todo } v \in V. \quad (55)$$

Observamos que ϕ é alternada se, e somente se, ϕ é **anti-simétrica**³⁸, isto é,

$$\phi(v, w) = -\phi(w, v) \quad \text{para todo } v, w \in V. \quad (56)$$

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

$$f \wedge g \mapsto (f \wedge g)(v, w) := \det \begin{bmatrix} f(v) & f(w) \\ g(v) & g(w) \end{bmatrix} \quad (57)$$

e estendemos, por linearidade, a uma aplicação linear $\Phi : \Lambda^2(V^*) \rightarrow \mathcal{A}_2(V)$. A definição acima é motivada, como explicamos abaixo, pela fórmula (53) acima. Pelas propriedades básicas do determinante, a fórmula (57) realmente define uma forma bilinear alternada. Notamos que, fixada uma base de V e a base dual de V^* , temos

$$(dx^{i_1} \wedge dx^{i_2})(e_{j_1}, e_{j_2}) := \det \begin{bmatrix} dx^{i_1}(e_{j_1}) & dx^{i_1}(e_{j_2}) \\ dx^{i_2}(e_{j_1}) & dx^{i_2}(e_{j_2}) \end{bmatrix} = \begin{cases} 1 & \text{se } i_1 = j_1 \text{ e } i_2 = j_2 \\ 0 & \text{nos outros casos} \end{cases}$$

Logo, usando a bilinearidade, obtemos

$$(dx^i \wedge dx^j)(v, w) := \det \begin{bmatrix} dx^i(v) & dx^j(w) \\ dx^i(w) & dx^j(v) \end{bmatrix} = v^i w^j - v^j w^i.$$

³⁸As condições (55) e (56) 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 a equivalência vale.

Por outro lado, dada uma forma bilinear alternada $\phi : V \times V \rightarrow \mathbb{R}$, podemos escrever

$$\phi(v, w) = \sum_{i,j=1}^n v^i w^j \phi(e_i, e_j) = \sum_{i < j} \phi(e_i, e_j) (v^i w^j - v^j w^i) = \sum_{i < j} \phi(e_i, e_j) (dx^i \wedge dx^j)(v, w),$$

de modo que

$$\phi = \sum_{i < j} \phi(e_i, e_j) dx^i \wedge dx^j.$$

19.3 Identificação II; Aplicações multilineares alternadas

O que foi feito na seção anterior pode ser generalizado para o produto exterior de “mais cópias” de V^* e nós obtemos o espaço $\Lambda^n(V^*)$ e um isomorfismo

$$\Lambda^n(V^*) \simeq \mathcal{A}_n(V)$$

com o espaço das formas multilineares alternadas. A generalização é simples, mas tediosa, pois precisamos trabalhar com ainda mais índices.

Dizemos que uma forma multilinear $\phi : V \times \cdots \times V \rightarrow \mathbb{R}$ é

– **alternada** quando

$$\phi(v_1, v_2, \dots, v_r) = 0 \quad \text{sempre que} \quad v_i = v_j \quad \text{para algum par} \quad i \neq j.$$

– **anti-simétrica** ou **skew-simétrica** quando muda de sinal ao trocarmos dois vetores de lugar:

$$\phi(v_1, v_2, \dots, v_i, \dots, v_j, \dots, v_r) = -\phi(v_1, v_2, \dots, v_j, \dots, v_i, \dots, v_r).$$

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 43. Seja $\phi : V \times \cdots \times V \rightarrow \mathbb{R}$ uma forma multilinear. Mostre que são equivalentes:

- (i) ϕ é alternada;
- (ii) ϕ 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

$$(f_1 \wedge f_2 \wedge \cdots \wedge f_k)(v_1, v_2, \dots, v_k) := \det [f_i(v_j)] = \det \begin{bmatrix} f_1(v_1) & f_1(v_2) & \cdots & f_1(v_k) \\ f_2(v_1) & f_2(v_2) & \cdots & f_2(v_k) \\ \vdots & \vdots & \ddots & \vdots \\ f_k(v_1) & f_k(v_2) & \cdots & f_k(v_k) \end{bmatrix}. \quad (58)$$

Por linearidade, obtemos uma aplicação $\Phi : \Lambda^k(V^*) \rightarrow \mathcal{A}_k(V)$ que é um isomorfismo de espaços vetoriais.

Exercício 44. Mostre que Φ é, de fato, um isomorfismo de espaços vetoriais.

20 Formas diferenciais de grau p e a derivada exterior

Uma **forma diferencial de ordem 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 alternadas de ordem 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

$$\omega(x) = \sum_I a_I(x) dx^I, \quad (59)$$

onde a soma é tomada sobre partições “crescentes” $I = \{1 \leq i_1 < i_2 < \dots < i_p \leq n\}$. A notação abreviada acima é a seguinte:

$$a_I(x) := a_{i_1 i_2 \dots i_p}(x) \quad \text{e} \quad dx^I := dx^{i_1} \wedge dx^{i_2} \wedge \dots \wedge dx^{i_p}.$$

Exercício 45. Mostre que para uma p -forma $\omega : U \longrightarrow \Lambda^p(\mathbb{R}^n)^*$, as seguintes afirmações são equivalentes:

- (i) a forma ω é diferenciável (ou classe C^k) como aplicação com valores no espaço vetorial $\Lambda^p(\mathbb{R}^n)^*$;
- (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 $x \longmapsto \omega(x)[v_1, v_2, \dots, v_p]$ é diferenciável (ou classe C^k).

Convencionamos que $\Omega^0(U)$ é o conjunto das funções diferenciáveis (na maioria dos contextos, se supõe suaves, ou seja, de classe C^∞) em U .

Podemos ainda considerar a álgebra de todas as formas diferenciais (de todas as ordens)

$$\Omega(U) := \Omega^0(U) \oplus \Omega^1(U) \oplus \Omega^2(U) \oplus \dots \oplus \Omega^n(U),$$

cujas noção de produto é o produto exterior que provém da subseção anterior.

A ideia agora é definir uma outra noção de diferenciabilidade de formas diferenciais, diferente do conceito de derivada como aplicação com valores em um espaço vetorial que poderia sugerir o Exercício 45. Esta nova noção, a de derivada exterior, nos permite, por exemplo, estender de uma vez só e em apenas um conceito, as noções fundamentais do cálculo vetorial de gradiente, divergente e rotacional (em \mathbb{R}^3). Dada uma forma diferencial ω de grau p , como em (59), definimos a **derivada exterior** de ω como sendo a forma $d\omega \in \Omega^{p+1}(U)$, de ordem $p+1$, dada por:

$$d\omega(x) := \sum_I da_I(x) \wedge dx^I = \sum_I \sum_{j=1}^n \frac{\partial a_I}{\partial x^j}(x) dx^j \wedge dx^I. \quad (60)$$

Passamos a mostrar que esta noção de diferencial de formas satisfaz as seguintes propriedades.

Teorema 95. A derivada exterior é o único conjunto de aplicações³⁹

$$d : \Omega^p(U) \longrightarrow \Omega^{p+1}(U)$$

que satisfaz as seguintes propriedades:

- 1) para $p = 0$, a aplicação $d : \Omega^0(U) \longrightarrow \Omega^1(U)$ é o diferencial usual de uma função.

³⁹Aqui, “conjunto de aplicações” se refere ao fato de termos uma aplicação para cada valor de p , o que muda o domínio e o contradomínio da diferenciação exterior. Tecnicamente, nós deveríamos explicitar esta dependência escrevendo d_p ao invés de d . Contudo, optamos por deixar a notação menos poluída, já que a ordem da forma diferencial é usualmente clara pelo contexto.

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

$$d(\omega \wedge \mu) = (d\omega) \wedge \mu + (-1)^p \omega \wedge (d\mu).$$

Observe que o grau de μ não interfere na fórmula.

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 (60) satisfaz as propriedades acima. A primeira é a definição do diferencial. A segunda segue diretamente de o diferencial de funções ser um objeto 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 determinam completamente a derivada exterior. Mais precisamente, vamos mostrar que uma aplicação

d com as propriedades enunciadas no teorema só pode ser dada pela fórmula em (60). Suponhamos dada uma função d que satisfaça as propriedades 1–4. Então

$$\begin{aligned} d\omega &= d\left(\sum_I a_I(x) dx^I\right) \\ &\stackrel{2)}{=} \sum_I d(a_I(x) dx^I) \\ &\stackrel{3) \text{ e } 1)}{=} \sum_I [da_I(x) \wedge dx^I + (-1)^0 a_I(x) \wedge d^2x^I] \\ &\stackrel{4)}{=} \sum_I da_I(x) \wedge dx^I \end{aligned} \quad \square$$

É um ótimo exercício, analisar as definições e propriedades acima explicitamente em dimensão “baixa”. Vamos fazer as contas para dimensão $n = 3$.

Exercício 46. Escreva todas as possíveis formas e suas derivadas exteriores para cada uma das dimensões $n = 1, 2$ e 4 .

Para a base do espaço dual ao \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 agora uma forma de grau 1 do tipo

$$\omega = \omega_x dx + \omega_y dy + \omega_z dz.$$

Aqui, as funções ω_x, ω_y e ω_z denotam as componentes da forma ω (e não derivadas parciais). 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. A aplicação $d : \Omega^1(U) \rightarrow \Omega^2(U)$, que é a derivada de uma 1-forma em \mathbb{R}^3 , pode ser, em um sentido a ser tornado mais preciso, identificada com o rotacional de um campo vetorial em \mathbb{R}^3 . No entanto, a derivada exterior aparece como uma operação mais natural para formas e é independente da famosa “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 em \mathbb{R}^3 , 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 também é um dos motivos de o campo rotacional ser às 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 : escrevemos

$$\omega = \omega_{xy} dx \wedge dy + \omega_{yz} dy \wedge dz + \omega_{xz} dz \wedge dx,$$

onde os índices novamente indicam as componentes da forma (e não derivadas parciais). Então, já deixando de escrever os termos que são nulos,

$$\begin{aligned} d\omega &= d\omega_{xy} \wedge dx \wedge dy + d\omega_{xz} \wedge dz \wedge dx + d\omega_{yz} \wedge dy \wedge dz \\ &= \frac{\partial \omega_{xy}}{\partial z} dz \wedge dx \wedge dy + \frac{\partial \omega_{yz}}{\partial x} dx \wedge dy \wedge dz + \frac{\partial \omega_{xz}}{\partial y} dy \wedge dz \wedge dx \\ &= \left(\frac{\partial \omega_{yz}}{\partial x} + \frac{\partial \omega_{xz}}{\partial y} + \frac{\partial \omega_{xy}}{\partial z} \right) dx \wedge dy \wedge dz, \end{aligned}$$

que é o equivalente ao divergente do campo vetorial de coordenadas $(\omega_{yz}, \omega_{xz}, \omega_{xy})$ (como seria a identificação dada pelo operador estrela de Hodge neste caso?).

Finalmente, é fácil de ver que se ω 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 47. Seja $\omega \in \Omega^{n-1}(U)$ uma forma de ordem $(n-1)$ 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

$$f : M \longrightarrow N$$

é possível induzir, como passamos a descrever, uma aplicação entre formas diferenciais (que inverte a ordem dos espaços e, por isto, contravariante)

$$f^* : \Omega(N) \longrightarrow \Omega(M).$$

Nesta discussão é, de certa maneira, mais “natural” pensar em M, N superfícies diferenciáveis ou variedades diferenciáveis para explicitar os espaços em que os objetos “vivem”. Pode-se pensar que $M, N \subseteq \mathbb{R}^n$ são conjuntos abertos, apesar de estarmos fazendo distinção entre pontos e vetores tangentes, como no caso de superfícies e variedades diferenciáveis.

Lembramos da Álgebra Linear (ver, por exemplo, [2, Seção 3.7]) que, dados dois espaços vetoriais V e W e uma transformação linear $T : V \longrightarrow W$, podemos associar a T , por dualidade, a chamada **transposta** de T :

$$T^* : W^* \longrightarrow V^* \quad \text{dada por} \quad T^*(\varphi)[v] := \varphi(T(v)) \quad (61)$$

que também é uma transformação linear.

Para cada $x \in M$, a derivada de f em x é uma transformação linear

$$df(x) : T_x M \longrightarrow T_{f(x)} N$$

que leva vetores tangentes a M em vetores tangentes a N . Isto foi discutido na Subseção 11.2, 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 ⁴⁰. É também comum encontrarmos a notação $f_* := df : TM \longrightarrow TN$, ou ainda,

$$f_*|_x := df(x) : T_x M \longrightarrow T_{f(x)} N$$

⁴⁰Na ocasião, por causa desta propriedade, chamamos os vetores tangentes de objetos “covariantes”.

Por dualidade, como acima, está associada a transposta de f_* (que é usualmente denotada por):

$$f^*|_x := df(x)^* : T_{f(x)}^*N \longrightarrow T_x^*M.$$

Mais explicitamente, para $\phi \in T_{f(x)}^*N$, o funcional $f^*\phi \in T_x^*M$ é definido como (ver (61))

$$f^*(\phi)[v] := \phi(f_*(v)).$$

Uma forma diferencial de grau 0 em N é, por definição, simplesmente uma função diferenciável $\varphi : N \longrightarrow \mathbb{R}$. O **pullback**⁴¹ de φ por f é então definido como a composição

$$f^*\varphi := \varphi \circ f : M \longrightarrow \mathbb{R}.$$

Observe como não é naturalmente possível fazer um “push-forward” de φ por f (tente!).

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 ω definida em N , definimos o **pullback** de ω por f como a 1-forma diferencial $f^*\omega$ em M dada por

$$(f^*\omega)(x)[v] := \omega(f(x))[f_*(v)].$$

Na notação que utilizávamos para a derivada de f até então, deveríamos escrever:

$$(f^*\omega)(x)[v] = \omega(f(x))[df(x) \cdot v].$$

Dada a definição de *pullback* de 1-formas, vamos mostrar que a derivada exterior é um operador **natural**, no sentido de ser comutativo o diagrama 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 $\varphi : N \longrightarrow \mathbb{R}$, temos

$$f^*(d\varphi) = d(f^*\varphi). \tag{62}$$

Se pensarmos em f como uma mudança de variáveis (por exemplo, um difeomorfismo)⁴², a fórmula (62) 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.

Para provar (62), 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.

⁴¹O termo em inglês significa, literalmente, puxar para trás. Isto ilustra o fato de que nossa definição pega uma função definida no contradomínio de f e “puxa” para uma função definida no domínio.

⁴²Além do mais, a importância do pullback se dará, como veremos, justamente ao considerarmos mudanças de coordenadas!

Exercício 48. Considere $\omega : \mathbb{R}^2 \setminus \{(0, 0)\} \rightarrow (\mathbb{R}^2)^*$ a forma elemento de ângulo vista no Exemplo 86 da Seção 16:

$$\omega(x) = \frac{-y}{x^2 + y^2} dx + \frac{x}{x^2 + y^2} dy$$

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 96. 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 ω 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

$$\omega(x) = a(x) dx^{i_1} \wedge dx^{i_2} \wedge \cdots \wedge dx^{i_p} \in \Omega^p(N).$$

Pela propriedade que mostramos acima e a propriedade 3) do enunciado, devemos ter

$$f^*\omega = (f^*a) (f^*dx^{i_1}) \wedge (f^*dx^{i_2}) \wedge \cdots \wedge (f^*dx^{i_p}). \quad (63)$$

A fórmula (63) mostra a unicidade de f^* a partir das propriedades 1) a 3). Resta mostrar que f^* definida por (63) 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 97. Para $f : M \rightarrow N$ diferenciável e para ω 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:

$$\begin{aligned} f^*(d\omega) &= f^*\left(\sum_I da_I \wedge dx^{i_1} \wedge dx^{i_2} \wedge \cdots \wedge dx^{i_p}\right) \\ &= \sum_I (f^*da_I) \wedge (f^*dx^{i_1}) \wedge (f^*dx^{i_2}) \wedge \cdots \wedge (f^*dx^{i_p}) \\ &= \sum_I d(f^*a_I) \wedge d(f^*x^{i_1}) \wedge d(f^*x^{i_2}) \wedge \cdots \wedge d(f^*x^{i_p}) \\ &= d\left(\sum_I (f^*a_I) d(f^*x^{i_1}) \wedge d(f^*x^{i_2}) \wedge \cdots \wedge d(f^*x^{i_p})\right) \\ &= d\left(\sum_I (f^*a_I) (f^*dx^{i_1}) \wedge (f^*dx^{i_2}) \wedge \cdots \wedge (f^*dx^{i_p})\right) \\ &= d\left(f^*\left(\sum_I a_I dx^{i_1} \wedge dx^{i_2} \wedge \cdots \wedge dx^{i_p}\right)\right) \\ &= d(f^*\omega). \end{aligned}$$

□

Encerramos esta seção com um exercício que indica como descrever o *pullback* de p -formas pensadas como campos de formas p -lineares alternadas.

Exercício 49. Mostre que se a aplicação $f : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ é diferenciável e se $\omega = dy^1 \wedge dy^2 \wedge dy^3$, então

$$(f^*\omega)(x) = \det f'(x) dx^1 \wedge dx^2 \wedge dx^3.$$

Generalize para \mathbb{R}^n .

Exercício 50. 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 95, 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 95 citado acima que toda forma exata é fechada. Vamos analisar uma possível recíproca deste resultado.

A forma elemento de ângulo estudada do Exemplo 86 não é exata, como consequência da Proposição 85, mas é uma forma fechada. Isto mostra que somente podemos esperar por uma recíproca parcial. Analisando mais de perto a prova da Proposição 85, 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

$$\phi(x) := \int_{\gamma_x} \omega. \quad (64)$$

Como já explicamos, ϕ 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 γ . 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 γ 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 γ_0 e γ_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 ω fechada, segue que a integral não depende dos caminhos γ_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 (64).

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 γ é uma **homotopia** entre os caminhos γ_0 e γ_1 .

A nossa discussão acima (quase) implica na seguinte proposição.

Proposição 98. Se γ_0 e γ_1 são caminhos suaves por partes e homotópicos em um conjunto U e ω 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 85).

Teorema 99. Seja $\omega \in \Omega^1(U)$ uma 1-forma diferencial em um conjunto simplesmente conexo U . Então, ω é uma forma fechada se, e somente se, é uma forma exata.

22 Variedades diferenciáveis

Nesta seção, introduzimos um conceito fundamental de geometria diferencial, a saber, o de variedade diferenciável. Várias referências foram consultadas; entre elas, [1, 4, 8, 9, 17].

Intuitivamente, uma variedade diferenciável é um espaço topológico que localmente se parece com um aberto de um espaço vetorial, de tal maneira a permitir o conceito de diferenciabilidade e, como consequência, o uso de técnicas usuais do cálculo diferencial.

Uma **variedade diferenciável**⁴³ de dimensão n é um espaço topológico⁴⁴ M junto com um conjunto de funções injetivas, chamadas **cartas**, **mapas** ou **parametrizações**, $\phi_\alpha : U_\alpha \rightarrow \mathbb{R}^n$, onde os domínios $U_\alpha \subseteq M$ são abertos em M que cobrem todo o conjunto M e as funções de transição $\phi_\alpha \circ \phi_\beta^{-1}$ são diferenciáveis⁴⁵ nos conjuntos onde estão definidas. O conjunto de todas as cartas $\{\phi_\alpha\}$ de uma variedade diferenciável M é chamado um **atlas** de M .

A ideia desta definição é de utilizar a conclusão fundamental do Teorema 34, item (ii), para **definir** uma noção de estrutura diferenciável em conjuntos mais gerais do que subconjuntos de \mathbb{R}^n . Esta é a propriedade que nos permite definir diferenciabilidade de maneira coerente.

Exemplo 100. O espaço \mathbb{R}^n é uma variedade diferenciável de dimensão n . Para ver isto, basta considerar (apenas) uma carta que cobre todo o espaço, que é a aplicação identidade. De maneira similar, todo espaço vetorial de dimensão n é uma variedade diferenciável de dimensão n . Além disso, qualquer subconjunto aberto de um espaço de dimensão n é uma variedade de dimensão n .

Exemplo 101 (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 102 (Espaço projetivo real). Consideramos

$$\mathbb{P}^2(\mathbb{R}) := \{\text{retas em } \mathbb{R}^3 \text{ que passam pela origem}\}.$$

Vamos mostrar que $\mathbb{P}^2(\mathbb{R})$ possui uma estrutura de variedade diferenciável de dimensão dois. Note que este é um exemplo não trivial, no sentido de ser essencialmente diferente de uma superfície diferenciável (não é um subconjunto de algum espaço Euclidiano). Observamos ainda, sem demonstração⁴⁶, que $\mathbb{P}^2(\mathbb{R})$ é uma variedade de dimensão dois que não pode ser “visualizada” (mais precisamente, 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

$$\phi_2([x_1, x_2, x_3]) := \left(\frac{x_1}{x_2}, \frac{x_3}{x_2} \right) \quad \text{e} \quad \phi_3([x_1, x_2, x_3]) := \left(\frac{x_1}{x_3}, \frac{x_2}{x_3} \right).$$

⁴³Nestas notas, como de costume na maioria dos textos de geometria, nossas variedades são de classe C^∞ ; para evitar a fadiga, não escrevemos “variedade diferenciável de classe C^∞ ”.

⁴⁴A noção de “espaço topológico” aparece aqui apenas para que se possa fazer referência a “conjuntos abertos” U_α de M ; não são necessários conhecimentos de topologia para acompanhar o texto.

⁴⁵Nesta parte das notas e muitas vezes neste contexto, a palavra diferenciável é um sinônimo de ser “de classe C^∞ ”. Fazemos isso nessa seção, a menos que haja menção contrária explícita.

⁴⁶No 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.

A função ϕ_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^∞ . De fato, se considerarmos, por exemplo, ϕ_1 e ϕ_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^∞ em $\phi_1(U_1 \cap U_2)$.

Em uma variedade diferenciável, é possível definir diferenciabilidade de funções, assim como o fizemos para superfícies diferenciáveis, “olhando para a função em coordenadas”, isto é, através de imagens por cartas. Assim, uma função $f : M \rightarrow \mathbb{R}$ é dita **diferenciável** em M quando $f \circ \phi_\alpha^{-1} : A_\alpha \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$ é diferenciável, para toda parametrização $\phi_\alpha : U_\alpha \subseteq M \rightarrow A_\alpha$. 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 fato, se duas cartas ϕ e ψ parametrizam a mesma região U de M e $f \circ \phi^{-1}$ é diferenciável, então também $f \circ \psi^{-1}$ é diferenciável, pois

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

De maneira análoga, f é de classe C^k ou de classe C^∞ quando $f \circ \phi_\alpha^{-1}$ é de classe C^k ou de classe C^∞ , respectivamente.

22.1 Vetores tangentes e o espaço tangente

Quando a variedade está “dentro” de um espaço euclidiano, poderíamos definir vetores tangentes assim como se faz para superfícies diferenciáveis, a saber, como elementos da imagem do diferencial de uma parametrização qualquer. No caso mais abstrato em que nos encontramos, é necessário obter alguma caracterização diferente. Por enquanto, sequer sabemos definir o diferencial de uma função definida em uma variedade e, na realidade o conceito de diferencial só será possível uma vez definidos os vetores tangentes.

Intuitivamente, nos espaços do tipo \mathbb{R}^n , campos de vetores podem ser pensados geometricamente como um conjunto de setas, uma sobre cada ponto do espaço. Usualmente, campos vetoriais podem representar um campo de força ou de velocidade de fluidos. Cada um desses vetores ou setas representa a direção e o sentido que uma partícula andaria sob o efeito do campo vetorial. Ou então, cada uma das setas indica como calcular derivadas direcionais de funções escalares. Estas duas maneiras de pensar indicam os caminhos mais comuns de se definir o espaço tangente a uma variedade diferenciável.

Inicialmente, vamos pensar de maneira heurística no caso $M = \mathbb{R}^n$. Dada uma função diferenciável $f : \mathbb{R}^n \rightarrow \mathbb{R}$, a derivada direcional de f no ponto x e na direção de um vetor $v \in \mathbb{R}^n$ pode ser calculada considerando uma curva α com $\alpha(0) = x$ e $\alpha'(0) = v$; tem-se

$$\frac{\partial f}{\partial v}(x) = \frac{d}{dt} f(\alpha(t)) \Big|_{t=0} = \sum_{i=1}^n \alpha'_i(0) \frac{\partial f}{\partial x_i}(x) = \left(\sum_{i=1}^n \alpha'_i(0) \frac{\partial}{\partial x_i} \Big|_x \right) f, \quad (65)$$

onde $\alpha(t) = (\alpha_1(t), \alpha_2(t), \dots, \alpha_n(t))$ e, na última igualdade, denotamos por

$$\frac{\partial}{\partial x_i} \Big|_x : C^\infty(\mathbb{R}^n) \longrightarrow \mathbb{R}$$

o funcional linear definido por

$$\frac{\partial}{\partial x_i} \Big|_x (f) := \frac{\partial f}{\partial x_i}(x). \quad (66)$$

Pensando em um vetor tangente como um objeto que “age” em funções diferenciáveis, a fórmula (65) indica então que deve ser

$$v = \sum_{i=1}^n v_i \frac{\partial}{\partial x_i} \Big|_x,$$

onde $v_i = \alpha'_i(0)$ são as coordenadas de v . Além disso, temos que os funcionais lineares definidos em (66) são candidatos naturais para formar uma base do espaço vetorial de todos os vetores tangentes.

As duas ideias que pretendemos explorar estão contidas acima. Uma ideia é pensar que vetores tangentes são vetores velocidade de curvas. Outra é escrever v como um funcional linear da forma $v : C^\infty(M) \rightarrow \mathbb{R}$ com propriedades do tipo de uma “derivada” e pensar em $v(f)$ como a derivada de f em x e na direção v .

22.1.1 Vetores tangentes como velocidades de curvas

Um dos pontos de vista é pensar em vetores tangentes à variedades como velocidades de curvas que passam por x com a “mesma direção”. Para isto, identificamos curvas cujos vetores velocidade em cartas coincidem. Mais precisamente, fixado um ponto $x \in M$ e dadas duas curvas diferenciáveis α e β com $\alpha(0) = x = \beta(0)$, definimos uma relação de equivalência por⁴⁷

$$\alpha \sim \beta \iff \frac{d}{dt} \phi(\alpha(t)) \Big|_{t=0} = \frac{d}{dt} \phi(\beta(t)) \Big|_{t=0} \quad \text{para alguma carta } \phi.$$

Fica por conta do leitor a verificação de que isto de fato define uma relação de equivalência no conjunto das curvas diferenciáveis que passam pelo ponto x em $t = 0$. Definimos um **vetor tangente** a M em x como sendo uma classe de equivalência $v = [\gamma]$ para alguma curva diferenciável γ com $\gamma(0) = x$. O conjunto de todos os vetores tangentes a M no ponto $x \in M$ é chamado de **espaço tangente** a M em x e é denotado por $T_x M$. Ou seja, tem-se

$$T_x M := \{v; v = [\gamma] \text{ para alguma curva diferenciável } \gamma \text{ tal que } \gamma(0) = x\}.$$

Não parece óbvio que isto define um espaço linear.

Proposição 103. *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 ϕ que parametriza uma vizinhança de x em M , definimos⁴⁸ $d\phi(x) : T_x M \rightarrow \mathbb{R}^n$ por

$$d\phi(x)(v) := \frac{d}{dt} \phi(\gamma(t)) \Big|_{t=0},$$

onde γ 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 definição da relação de equivalência, vemos que $d\phi(x)$ é injetiva. Para mostrar que é também sobrejetiva, consideramos $w \in \mathbb{R}^n$ e definimos

$$\alpha(t) := \phi^{-1}(\phi(x) + tw) \quad \text{de modo que} \quad \frac{d}{dt} \phi(\alpha(t)) \Big|_{t=0} = w.$$

⁴⁷Pela diferenciabilidade das funções de transição, a igualdade vale, na realidade, para toda carta que contém as curvas α e β . Logo, o espaço tangente é independente da parametrização escolhida.

⁴⁸Neste ponto, $d\phi(x)$ é apenas uma notação para esta função que está sendo definida no início da prova da Proposição 103 e não deve se deve pensar que estamos “diferenciando” ϕ . Por outro lado, a notação é consistente com o diferencial de uma função que definimos mais adiante.

A curva α deve ser pensada como um segmento de reta em \mathbb{R}^n que passa por $\phi(x)$ com velocidade w que é “levado” do espaço \mathbb{R}^n para a variedade M através de ϕ^{-1} .

Sendo $d\phi(x)$ uma bijeção, podemos equipar $T_x M$ com a estrutura de espaço vetorial induzida por $d\phi(x)$, o que torna o conjunto $T_x M$ um espaço vetorial de dimensão n e a aplicação $d\phi(x)$ um isomorfismo. 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

$$(\psi \circ \phi^{-1})'(\phi(x)) : \mathbb{R}^n \longrightarrow \mathbb{R}^n$$

é um isomorfismo. Portanto, 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 γ é qualquer representante da classe de equivalência $[\gamma]$. Assim, continuamos com a fórmula usual

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

embora esta seja agora apenas a definição da aplicação $d\phi(x)$ em conjunto com uma notação conveniente para os vetores tangentes (que são classes de equivalência).

22.1.2 Vetores tangentes como derivações

Um outro ponto de vista, talvez 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 um funcional linear da forma $v_x : C^\infty(M) \longrightarrow \mathbb{R}$ 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 motivada pela regra da diferenciação de um produto de funções. Leia-se que “a derivada de fg na direção v é a derivada de f na direção v vezes g mais o produto de f pela derivada de g na direção v ”. 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 nossas duas definições são equivalentes).

Nesta definição, ao contrário da anterior, o espaço tangente $\tilde{T}_x M$ possui naturalmente uma estrutura de espaço vetorial pois é um subespaço vetorial do espaço dual $C^\infty(M)^*$ (lembre que esta estrutura é dada pela soma de funcionais lineares e a multiplicação de um funcional linear por um escalar). No entanto, não é claro que a dimensão de $\tilde{T}_x M$ seja a mesma da variedade M . Em verdade, nem é claro que $\tilde{T}_x M$ deva ter dimensão finita!

Apesar desta descrição ser um pouco mais abstrata, é a mais útil e é a que utilizamos com mais frequência. Na subseção seguinte, mostramos que ambas as descrições que utilizamos são equivalentes e como passar de uma para a outra.

22.1.3 A identificação entre as duas definições dadas

Dado $[\gamma] \in T_x M$, associamos uma derivação $v_\gamma \in \tilde{T}_x M$ pela identidade

$$v_\gamma(f) := \left. \frac{d}{dt} f(\gamma(t)) \right|_{t=0} \quad \text{para } f \in C^\infty(M). \quad (67)$$

Note como esta definição indica que, se quisermos pensar em vetores tangentes como velocidades de curvas, ainda assim somos “permitidos” calcular derivadas direcionais de funções. Mas precisamos ainda verificar algumas compatibilidades:

- v_γ está bem definido porque independe do elemento da classe $[\gamma]$ escolhido: dados $\gamma_1, \gamma_2 \in [\gamma]$ e uma parametrização ϕ , a definição de $[\gamma]$ implica

$$\left. \frac{d}{dt} f(\gamma_1(t)) \right|_{t=0} = d(f \circ \phi^{-1})(\phi(x)) \left. \frac{d}{dt} \phi(\gamma_1(t)) \right|_{t=0} = d(f \circ \phi^{-1})(\phi(x)) \left. \frac{d}{dt} \phi(\gamma_2(t)) \right|_{t=0} = \left. \frac{d}{dt} f(\gamma_2(t)) \right|_{t=0}.$$

Na primeira (e analogamente na última) igualdade, nós escrevemos $f \circ \gamma_1 = (f \circ \phi^{-1}) \circ (\phi \circ \gamma_1)$ e utilizamos a regra da cadeia usual já que as funções em parêntesis estão definidas em espaços Euclidianos.

- v_γ é linear: dados $f, g \in C^\infty(M)$ e $a, b \in \mathbb{R}$, temos

$$v_\gamma(af + bg) := \left. \frac{d}{dt} (af + bg)(\gamma(t)) \right|_{t=0} = \left. \frac{d}{dt} (af(\gamma(t)) + bg(\gamma(t))) \right|_{t=0} = av_\gamma(f) + bv_\gamma(g).$$

A última igualdade segue da linearidade da derivada de funções reais.

- Vale a regra de Leibniz:

$$v_\gamma(fg) = \left. \frac{d}{dt} [f(\gamma(t))g(\gamma(t))] \right|_{t=0} = \left. \frac{d}{dt} f(\gamma(t)) \right|_{t=0} g(\gamma(0)) + f(\gamma(0)) \left. \frac{d}{dt} g(\gamma(t)) \right|_{t=0}.$$

Assim, de fato (67) define um elemento de $\tilde{T}_x M$.

Proposição 104. A aplicação

$$\begin{aligned} \Phi : T_x M &\longrightarrow \tilde{T}_x M \\ [\gamma] &\longmapsto v_\gamma \end{aligned}$$

é um isomorfismo de espaços vetoriais.

Demonstração. Inicialmente, vamos mostrar que Φ é uma aplicação linear. É evidente que devemos utilizar a estrutura linear de $T_x M$ para provar linearidade. Lembre que esta estrutura foi definida na demonstração da Proposição 103. Dados $a, b \in \mathbb{R}$ e $[\alpha], [\beta] \in T_x M$, vamos verificar que

$$v_{a[\alpha]+b[\beta]} = av_{[\alpha]} + bv_{[\beta]}.$$

Em outras palavras, devemos provar que, para toda função $f \in C^\infty(M)$, tem-se

$$\left. \frac{d}{dt} f(a\alpha(t) + b\beta(t)) \right|_{t=0} = a \left. \frac{d}{dt} f(\alpha(t)) \right|_{t=0} + b \left. \frac{d}{dt} f(\beta(t)) \right|_{t=0}.$$

Dada uma parametrização $\phi : U \subset M \longrightarrow \mathbb{R}^n$ em torno do ponto $x \in M$, calculamos (observe como é utilizada, nas três igualdades centrais, a estrutura linear de $T_x M$ a partir da definição de $d\phi(x)$, como deveria ser)

$$\begin{aligned} \left. \frac{d}{dt} f(a\alpha(t) + b\beta(t)) \right|_{t=0} &= d(f \circ \phi^{-1})(\phi(x)) \left. \frac{d}{dt} \phi(a\alpha(t) + b\beta(t)) \right|_{t=0} \\ &= d(f \circ \phi^{-1})(\phi(x)) d\phi(x) (a[\alpha] + b[\beta]) \\ &= a d(f \circ \phi^{-1})(\phi(x)) \cdot (d\phi(x) \cdot [\alpha]) + b d(f \circ \phi^{-1})(\phi(x)) \cdot (d\phi(x) \cdot [\beta]) \\ &= a d(f \circ \phi^{-1})(\phi(x)) \left. \frac{d}{dt} \phi(\alpha(t)) \right|_{t=0} + b d(f \circ \phi^{-1})(\phi(x)) \left. \frac{d}{dt} \phi(\beta(t)) \right|_{t=0} \\ &= a \left. \frac{d}{dt} f(\alpha(t)) \right|_{t=0} + b \left. \frac{d}{dt} f(\beta(t)) \right|_{t=0}. \end{aligned}$$

Em seguida, mostramos que Φ é injetiva: dados $v, w \in T_x M$ satisfazendo $\Phi(v) = \Phi(w)$, escolhemos representantes α e β de v e w , respectivamente. Devemos ter

$$\left. \frac{d}{dt} f(\alpha(t)) \right|_{t=0} = \left. \frac{d}{dt} f(\beta(t)) \right|_{t=0} \quad \text{para toda } f \in C^\infty(M).$$

Dada uma parametrização $\phi : U \subset M \rightarrow \mathbb{R}^n$, vamos considerar $f = \phi^j : U \rightarrow \mathbb{R}$, onde $\phi^j := \pi^j \circ \phi$ é a projeção na j -ésima coordenada de ϕ . Olhando coordenada a coordenada, temos⁴⁹

$$\frac{d}{dt}(\phi \circ \alpha)(t) \Big|_{t=0} = \frac{d}{dt}(\phi \circ \beta)(t) \Big|_{t=0}.$$

Desta forma, $\alpha \sim \beta$ e logo $v = [\alpha] = [\beta] = w$.

Finalmente, mostramos a sobrejetividade de Φ . Devemos mostrar que, dada uma derivação qualquer $v \in \tilde{T}_x M$, podemos encontrar uma curva γ tal que $v = v_\gamma$ como em (67). 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 ϕ . 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 e pela definição do vetor a , temos

$$v_\gamma(f) = \frac{d}{dt}f(\gamma(t)) \Big|_{t=0} = d(f \circ \phi^{-1})(\phi(x_0)) \cdot (\phi \circ \gamma)'(0) = \sum_{i=1}^n \frac{\partial(f \circ \phi^{-1})}{\partial x^i}(\phi(x_0)) v(\phi^i).$$

A última igualdade é como se calcula a derivada direcional de funções de \mathbb{R}^n . 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, obtemos⁵⁰

$$\begin{aligned} v(f) &= \sum_{i=1}^n v(\phi^i - \phi^i(x)) (g_i \circ \phi)(x) + (\phi^i(x) - \phi^i(x)) v(g_i \circ \phi) \\ &= \sum_{i=1}^n a^i g_i(x_0) = \sum_{i=1}^n \frac{\partial(f \circ \phi^{-1})}{\partial x^i}(\phi(x_0)) v(\phi^i). \end{aligned}$$

Portanto, $v(f) = v_\gamma(f)$, como queríamos demonstrar. \square

⁴⁹Estritamente falando, deveríamos ter uma função ϕ 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 restrições topológicas que devem ser impostas sobre M .

⁵⁰Deveríamos, 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 segunda propriedade que utilizamos: se $f \equiv c$ é uma função constante, então $v(c) = 0$. De fato, a regra de Leibniz implica que

$$v(1) = v(1 \cdot 1) = 1 \cdot v(1) + v(1) \cdot 1 = 2 \cdot v(1) \implies v(1) = 0 \implies v(c) = c \cdot v(1) = 0.$$

22.1.4 Em coordenadas

Em seguida, vamos escrever vetores tangentes em coordenadas locais, o que é importante tanto para obter bases locais para espaços tangentes como para fazer contas localmente. Levando em consideração a identificação da subseção anterior, temos que, dado $v \in T_x M$, existe uma curva $\gamma : I \rightarrow M$ com $\gamma(0) = x$ e tal que $v = v_\gamma$. Podemos escrever

$$vf = v(f) = \left. \frac{d}{dt} f(\gamma(t)) \right|_{t=0}.$$

Consideramos uma carta $\phi : U \rightarrow \mathbb{R}^n$ em uma vizinhança $U \subseteq M$ de x . 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 com superscritos⁵¹, um ponto pertencente a $U \subseteq M$ pode ser escrito da forma $\phi^{-1}(x^1, x^2, \dots, x^n)$ e a curva γ na carta ϕ como

$$\phi(\gamma(t)) = (x^1(t), x^2(t), \dots, x^n(t)).$$

Logo,

$$v(f) = \sum_{i=1}^n \frac{dx^i}{dt}(0) \frac{\partial(f \circ \phi^{-1})}{\partial x^i}(\phi(x)).$$

Esta fórmula está rigorosamente justificada; comparar com a fórmula heurística (65). Na verdade, é possível “chegar mais perto de (65)” definindo

$$\left. \frac{\partial}{\partial x^i} \right|_x : C^\infty(M) \rightarrow \mathbb{R} \quad \text{por} \quad \left. \frac{\partial}{\partial x^i} \right|_x (f) := \frac{\partial(f \circ \phi^{-1})}{\partial x^i}(\phi(x)), \quad (68)$$

que define uma derivação linear, ou seja, um elemento de $T_x M$. Pelas contas acima, temos

$$v = \sum_{i=1}^n (x^i)'(0) \left. \frac{\partial}{\partial x^i} \right|_x =: \sum_{i=1}^n v^i \left. \frac{\partial}{\partial x^i} \right|_x.$$

Acima, definimos $v^i = (x^i)'(0)$. Denotamos ainda

$$\left. \frac{\partial f}{\partial x^i} \right|_x := \frac{\partial(f \circ \phi^{-1})}{\partial x^i} \circ \phi \quad \text{ou} \quad \left. \frac{\partial f}{\partial x^i} \right|_x := \frac{\partial(f \circ \phi^{-1})}{\partial x^i}(\phi(x))$$

para escrever a fórmula acima como

$$vf = \sum_{i=1}^n v^i \left. \frac{\partial f}{\partial x^i} \right|_x, \quad (69)$$

que é exatamente a fórmula (65), mas em sua versão para a variedade M . Pelo visto na seção anterior, toda derivação linear é dessa forma, o que mostra que os elementos definidos em (68) formam uma base de $T_x M$ para cada ponto x da vizinhança coordenada U : são linearmente independentes (por análise dimensional) e

$$T_x M = \text{Span} \left\{ \left. \frac{\partial}{\partial x^1} \right|_x, \left. \frac{\partial}{\partial x^2} \right|_x, \dots, \left. \frac{\partial}{\partial x^n} \right|_x \right\}.$$

Na definição (68), descrevemos uma base para o espaço tangente “mais pensado” como um espaço de derivações. Uma análise da prova da Proposição 104 nos permite descrever uma base de $T_x M$ em termos de classes de equivalências de curvas:

$$\left. \frac{\partial}{\partial x^i} \right|_x = [\gamma_i]$$

⁵¹Essa convenção é conveniente para utilização da notação de Einstein, que explicamos em outra seção.

onde $\gamma_i : (-\delta, \delta) \rightarrow M$ é a curva dada em coordenadas por

$$\phi(\gamma_i(t)) = \phi(x) + te_i.$$

Explicitamente, escolhamos $\delta > 0$ pequeno o suficiente para que $\phi(x) + te_i \in \phi(U)$ sempre que $|t| < \delta$ e definimos

$$\gamma_i(t) := \phi^{-1}(\phi(x) + te_i).$$

Concluimos que uma base para o espaço tangente formado por classes de equivalências de curvas é dada da seguinte maneira: consideramos segmentos em \mathbb{R}^n que passam por $\phi(x)$ com velocidades dadas pela base canônica de \mathbb{R}^n e depois “levamos” estes segmentos para cima de M pela aplicação ϕ^{-1} .

22.2 Exemplos

Muitos dos exemplos que temos em mente são superfícies diferenciáveis em um espaço Euclidiano. E por isso é importante saber como identificar nossa noção abstrata de espaço tangente com a noção usual. Nessa seção vamos mostrar alguns exemplos de espaços tangentes nesse sentido.

Dada uma superfície diferenciável $S \subset \mathbb{R}^m$ de dimensão n , para cada ponto $x \in S$, temos uma vizinhança $U \subset \mathbb{R}^m$ de x e uma parametrização⁵² $\phi^{-1} : A \rightarrow U \cap S$ tal que

$$d(\phi^{-1})(y) : \mathbb{R}^n \rightarrow \mathbb{R}^m$$

é injetiva para todo $y \in A = \phi(U)$. No nosso curso, assim como em cursos de geometria diferencial ou de análise vetorial, o espaço tangente a S no ponto x é definido como a imagem de $d(\phi^{-1})(\phi(x))$ e assim, por definição, temos um isomorfismo

$$d(\phi^{-1})(\phi(x)) : \mathbb{R}^n \rightarrow T_x S$$

Este isomorfismo é justamente o inverso do que foi definido na demonstração da Proposição 103. Portanto, o espaço tangente que definimos na seção anterior, no caso de superfícies diferenciáveis de um espaço Euclidiano, é isomorfo ao que já estávamos acostumados.

Exemplo 105. Considere $M = \mathbb{S}^2 \subseteq \mathbb{R}^3$ a esfera unitária de \mathbb{R}^3 . Levando em consideração o que discutimos acima, podemos considerar, como de costume,

$$T_x \mathbb{S}^2 = \{v \in \mathbb{R}^3; x \cdot v = 0\}.$$

Exemplo 106. Para o espaço \mathbb{R}^n , escolhendo por carta a aplicação identidade, temos

$$d(\text{Id})(x) \cdot v := \left. \frac{d}{dt}(x + tv) \right|_{t=0} \quad \text{para } v \in T_x \mathbb{R}^n.$$

Acima, escolhamos a curva $\gamma(t) = x + tv$ que satisfaz o requisitos $\gamma(0) = x$ e $\gamma \in v$. Nós podemos identificar naturalmente “todos os espaços tangentes” ao fazer “translações para a origem”. Para x fixado, a translação

$$T(y) = y + x$$

serve para identificar os espaços tangentes. Assim, podemos escrever todos os espaços tangentes através de curvas que passam pela origem. Deste modo,

$$T_x \mathbb{R}^n \simeq T_0 \mathbb{R}^n = \mathbb{R}^n.$$

Exemplos de outros espaços com esta propriedade, de os espaços tangentes se identificarem com o espaço tangente na origem, incluem todos os chamados Grupos de Lie.

⁵²Utilizamos a inversa para que nossa notação fique coerente com a definição de variedade diferenciável que utilizamos, onde parametrizações vão da superfície para um aberto do espaço Euclidiano \mathbb{R}^n .

22.3 Diferencial

Definimos o conceito de diferenciabilidade de uma função real $f : M \rightarrow \mathbb{R}$ na primeira seção deste capítulo. No entanto, para falar na *derivada* ou no *diferencial* de f , é necessário o conceito de espaço tangente.

Em termos de classes de equivalências de curvas, definimos a aplicação

$$\begin{aligned} df(x) : T_x M &\longrightarrow \mathbb{R} \\ [\alpha] &\longmapsto \left. \frac{d}{dt}(f \circ \alpha)(t) \right|_{t=0} \end{aligned}$$

que é chamada de a **derivada** ou o **diferencial** de f em x . É fácil ver, como já fizemos, que isto está bem definido independentemente do representante $\alpha \in [\alpha]$ escolhido.

A **derivada** $df(x) : T_x M \rightarrow \mathbb{R}$ também pode ser descrita em termos de derivações: para a derivação $v \in T_x M$, escrevemos

$$df(x) \cdot v := vf.$$

Por esta identidade, estamos definido o diferencial como a “derivada direcional” de f na “direção” v . Pela identificação apresentada da seção anterior, deve ser claro que estas duas descrições estão coerentes.

O diferencial de uma função f é uma transformação linear definida em $T_x M$, isto é, um elemento do espaço dual ao espaço tangente:

$$df(x) \in (T_x M)^*.$$

Também é comum a notação $T_x^* M$ para este dual e dizer que $df(x)$ é um vetor cotangente ou um covetor. Mais ainda, df é um caso particular de 1-forma definida em M . Mais precisamente, uma 1-forma em M pode ser definida como

$$\omega : M \rightarrow T^* M \quad \text{com} \quad \omega(x) \in T_x^* M, \quad \forall x \in M.$$

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 coleção⁵³ $\mathcal{X} = \{U_\alpha\}_{\alpha \in A}$ de abertos de M , uma **partição da unidade subordinada a \mathcal{X}** é uma família de funções diferenciáveis $\{f_\alpha\} \subseteq C^\infty(M)$ com as seguintes propriedades:

- (i) Tem-se $0 \leq f_\alpha(x) \leq 1$ para todo $x \in M$ e para todo índice α ;
- (ii) No complementar de U_α , temos $f_\alpha \equiv 0$. Uma outra maneira de dizer isto é o seguinte: definimos o **suporte** de f_α 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_α 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 uma família **localmente finita**.

⁵³Aqui, A simplesmente denota um conjunto de índices qualquer.

(iv) Para todo $x \in M$, temos

$$\sum_{\alpha} f_{\alpha}(x) = 1. \quad (70)$$

O item (iii) é importante para que a soma que aparece neste item (iv) é sempre uma *soma finita*, não sendo necessárias preocupações 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 (70), podemos (obviamente!) escrever

$$\varphi(x) = \sum_{\alpha} f_{\alpha}(x)\varphi(x),$$

ou ainda,

$$\varphi = \sum_{\alpha} f_{\alpha}\varphi. \quad (71)$$

As funções $f_{\alpha}\varphi$ são definidas em M , mas apenas tem valores não nulos em U_{α} . Assim, o estudo local (em cada U_{α}) de φ deve implicar em um estudo global de φ por conta da identidade (71). O nosso interesse, particularmente, é o de ser possível definir localmente a integração de formas diferenciais em variedades a partir de um conjunto de cartas. Com uma partição da unidade em mãos, é possível estender este conceito e definir a integral de uma forma diferencial globalmente em M .

Passamos a provar a existência de partições da unidade.

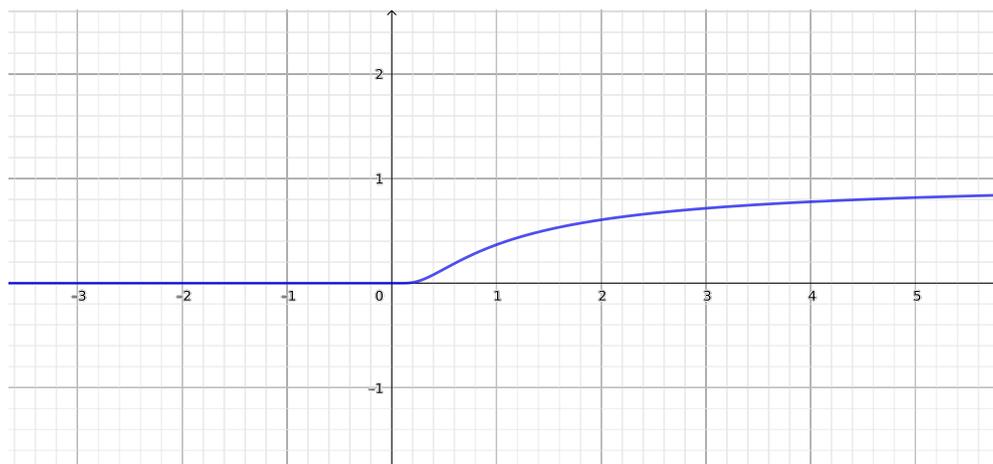
Teorema 107. *Seja M um variedade diferenciável e $\mathcal{X} = \{U_{\alpha}\}_{\alpha \in A}$ uma cobertura aberta de M . Então existe uma partição diferenciável da unidade subordinada a \mathcal{X} .*

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.

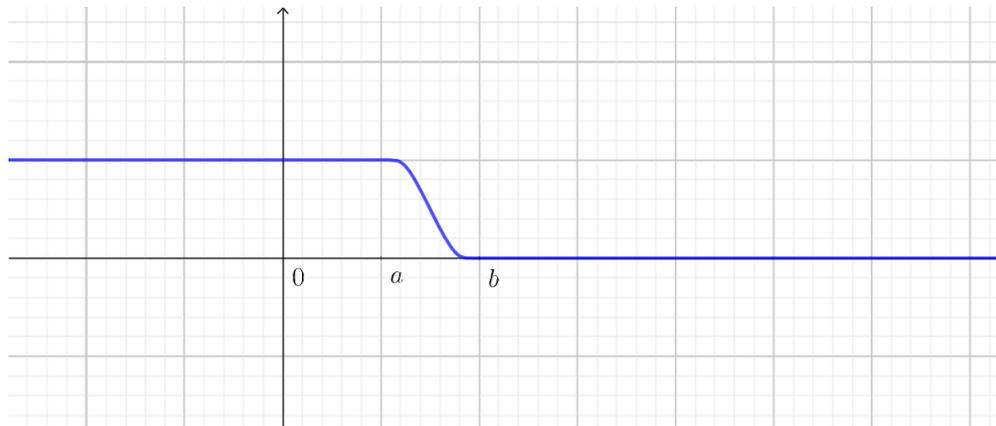


Agora, dados dois números reais satisfazendo $0 < a < b$, definimos $h : \mathbb{R} \rightarrow \mathbb{R}$ por

$$h(x) := \frac{g(b-x)}{g(b-x) + g(x-a)},$$

de modo que h é de classe C^∞ e tem-se

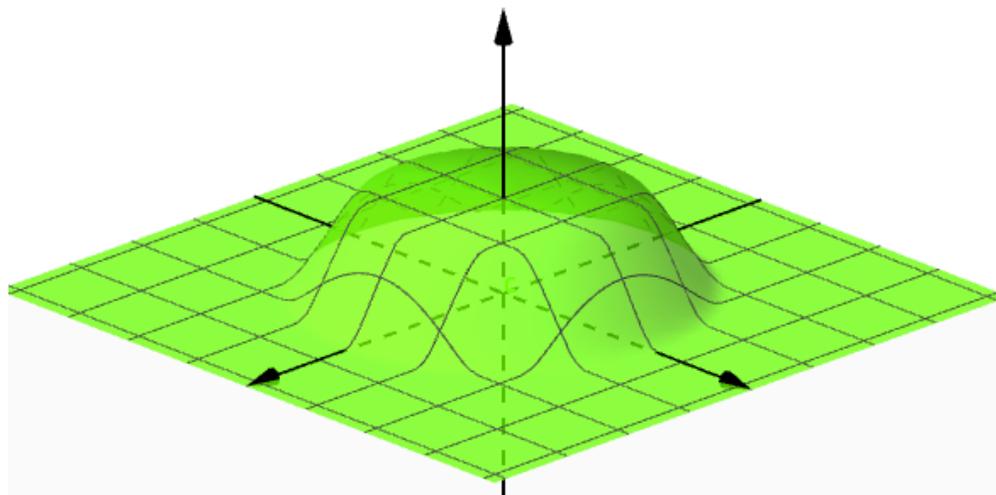
$$h(x) = \begin{cases} 1 & \text{se } x \leq a \\ \in (0, 1) & \text{se } a < x < b \\ 0 & \text{se } x \geq b \end{cases}$$



Desta maneira, podemos ainda construir uma função $H \in C^\infty(\mathbb{R}^n; \mathbb{R})$ satisfazendo

$$H \equiv 1 \text{ em } B_a(0), \quad 0 < H < 1 \text{ em } B_b(0) \setminus B_a(0), \quad \text{e} \quad H \equiv 0 \text{ em } \mathbb{R}^n \setminus B_b(0).$$

De fato, basta definir radialmente $H(x) := h(|x|)$.



A ideia da construção de uma partição da unidade para M é de considerar vizinhanças coordenadas convenientes de M que formam uma família localmente finita e depois construir funções, definidas em M , análogas à função H .

Prova do Teorema 107. Nós consideramos inicialmente o caso em que a variedade M pode ser coberta por apenas uma vizinhança coordenada $\phi : M \rightarrow \mathbb{R}^n$. Este caso é mais elementar por M ser homeomorfa a \mathbb{R}^n e, conseqüentemente, ser a “estrutura de conjuntos abertos” (topologia) de M equivalente à de \mathbb{R}^n . Nós definimos

$$\mathcal{B} := \left\{ B_r(x); \begin{array}{l} B_r(x) \subset \mathbb{R}^n \text{ é uma bola aberta com raio } r \in \mathbb{Q} \text{ cujo centro } x \text{ possui} \\ \text{coordenadas racionais e tal que } B_{r'}(x) \subset \phi(M) \text{ para algum } r' > r \end{array} \right\}$$

Desta maneira, \mathcal{B} é um conjunto enumerável de vizinhanças abertas em \mathbb{R}^n que satisfaz a seguinte propriedade (cuja demonstração segue da densidade dos pontos de coordenadas racional e é deixada como exercício): dado um conjunto aberto $A \subseteq \phi(M)$ e $y \in A$, existe uma bola $B \in \mathcal{B}$ tal que $y \in B \subset A$. Em linguagem topológica, diz-se que \mathcal{B} é uma base para os abertos de $\phi(M)$, no sentido que todo aberto de $\phi(M)$ pode ser escrito como união de abertos de \mathcal{B} . A nossa base \mathcal{B} tem propriedades adicionais: é um conjunto enumerável de bolas cujo fecho (que é compacto) está inteiramente contido no aberto $\phi(M)$.

Como ϕ é um homeomorfismo (e, em particular, contínua), temos que imagens inversas de abertos são abertos. Logo,

$$\mathfrak{B} := \{\phi^{-1}(B); B \in \mathcal{B}\}$$

é uma base enumerável de abertos para M , isto é, dados um aberto $A \subset M$ e um ponto $p \in A$, existe $U \in \mathfrak{B}$ tal que $p \in U \subset A$. Os elementos $\phi^{-1}(B) \in \mathfrak{B}$ são chamados de bolas coordenadas de M , já que são pré-imagens pela carta ϕ de uma bola usual do \mathbb{R}^n . Além disso, pelas nossas definições, estas bolas coordenadas são abertos com fecho compacto em M .

Em seguida, afirmamos que, dada uma cobertura aberta $\mathcal{X} = \{U_\alpha\}_{\alpha \in A}$ da variedade M , existe uma subcobertura enumerável e localmente finita que consiste de elementos de \mathfrak{B} . De fato, seja $\{K_i\}_{i=1}^{+\infty}$ uma exaustão de M por compactos, isto é, uma coleção de compactos $K_i \subset M$ tais que⁵⁴

$$K_i \subset \text{int } K_{i+1} \subset M \quad \text{e} \quad M = \bigcup_{i=1}^{+\infty} K_i.$$

Daí definimos “anéis” compactos $V_j := K_{j+1} \setminus \text{int } K_j$ e “anéis” abertos $W_j := \text{int } K_{j+2} \setminus K_{j-1}$ e notamos que $V_j \subset W_j$. Para cada $x \in V_j$, existe $U_{\alpha(x)} \in \mathcal{X}$ tal que $x \in U_{\alpha(x)}$. Logo, existe $B_x \in \mathfrak{B}$ tal que $x \in B_x \subset W_j \cap U_{\alpha(x)}$, pois $W_j \cap U_{\alpha(x)}$ é aberto e \mathfrak{B} é uma base de abertos de M . Como V_j é compacto, a cobertura

$$V_j \subset \bigcup_{x \in V_j} B_x$$

admite uma subcobertura finita. A união, para todo $j \in \mathbb{N}$, destas coberturas finitas nos dá uma subcobertura enumerável de \mathcal{X} que consiste de elementos de \mathfrak{B} . Além disso, como $V_j \subset W_j$ e $W_j \cap W_k \neq \emptyset$ somente quando $k \in \{j-2, j-1, j, j+1, j+2\}$, segue que a subcobertura construída também é, como afirmado, localmente finita.

Denotamos por $\{B_i\} \subset \mathfrak{B}$ a subcobertura de \mathcal{X} obtida acima. Pelas nossas definições, podemos escrever $B_i = \phi^{-1}(B_{r_i}(x_i))$, para algum $r_i > 0$. Além disso, para algum $r'_i > r_i$, conforme definição de \mathcal{B} , denotamos $B'_i = \phi^{-1}(B_{r'_i}(x_i))$. Temos

$$\phi(B_i) = B_{r_i}(0), \quad \phi(\overline{B_i}) = \overline{B_{r_i}(0)} \quad \text{e} \quad \phi(B'_i) = B_{r'_i}(0).$$

Definimos $f : M \rightarrow \mathbb{R}$ a função de classe $C^\infty(M)$ dada por

$$f_i(x) := \begin{cases} H_i(\phi(x)), & \text{se } x \in B'_i \\ 0, & \text{se } x \in M \setminus \overline{B_i} \end{cases}$$

onde $H_i \in C^\infty(\mathbb{R}^n; \mathbb{R})$ é uma função satisfazendo

$$H_i \equiv 1 \text{ em } B_{r_i}(0), \quad 0 < H_i < 1 \text{ em } B_{r'_i}(0) \setminus B_{r_i}(0), \quad \text{e} \quad H_i \equiv 0 \text{ em } \mathbb{R}^n \setminus B_{r'_i}(0).$$

⁵⁴A existência de uma exaustão por compactos não é tão imediata e requer atenção. Se M é uma variedade diferenciável de Hausdorff e com base enumerável, é possível mostrar que existe tal exaustão. Nós não entraremos em detalhes topológicos neste texto. No entanto, a construção segue as seguintes linhas: enumera-se a base de abertos (cujos fechos são compactos) por $\{B_i\}_{i=1}^{+\infty}$. Definimos $K_1 = \overline{B_1}$ e, para argumentar por indução, assumimos definidos K_1, K_2, \dots, K_k tais que $B_j \subset K_j$ para todo $j \in \{1, 2, \dots, k\}$ e $K_j \subset \text{int } K_{j+1}$ para todo $j \in \{1, 2, \dots, k-1\}$. Em seguida, definimos

$$K_{k+1} = \overline{B_1} \cup \overline{B_2} \cup \dots \cup \overline{B_{m_k}},$$

onde $m_k \geq k+1$ é tomado grande o suficiente de modo a termos

$$K_k \subset \text{int } K_{k+1} \quad \text{e} \quad B_{k+1} \subset K_{k+1}.$$

Em particular, temos $\text{supp } f_i = \overline{B_i}$.

Agora, considere a função $f : M \rightarrow \mathbb{R}$ definida por

$$f(x) := \sum_{i=1}^{+\infty} f_i(x),$$

cuja soma é sempre de um número finito de termos, já que $\{B_i\} \subset \mathfrak{B}$ é uma subcobertura localmente finita de \mathcal{X} . Além disso, $f(x) > 0$ para todo $x \in M$, já que todo ponto de M está em alguma bola coordenada B_i . Logo, definindo $g : M \rightarrow \mathbb{R}$ por

$$g(x) := \frac{f_i(x)}{f(x)}, \quad \text{temos} \quad 0 \leq g_i \leq 1 \quad \text{e} \quad \sum_{i=1}^{+\infty} g_i(x) = 1.$$

Resta mostrar ainda que é possível reindexar nossas funções de modo a manter os mesmos índices $\alpha \in A$ da cobertura original. Note que, sendo uma subcobertura, podemos escolher, para cada i , algum índice $a(i) \in A$ tal que $B_i \subset U_{a(i)}$. Assim, fixado $\alpha \in A$, definimos $f_\alpha : M \rightarrow \mathbb{R}$ por

$$f_\alpha(x) := \sum_{i; a(i)=\alpha} g_i(x).$$

Temos

$$\text{supp } f_\alpha = \overline{\bigcup_{i; a(i)=\alpha} B_i} = \bigcup_{i; a(i)=\alpha} \overline{B_i} \subseteq U_\alpha \quad \text{e} \quad \sum_{i=1}^{+\infty} f_\alpha(x) = \sum_{i=1}^{+\infty} g_i(x) = 1.$$

Isto prova a existência de uma partição da unidade subordinada a \mathcal{X} no caso de uma variedade que pode ser coberta por apenas uma carta.

No caso geral, nós podemos considerar uma família de vizinhanças coordenadas que cobrem toda a variedade M . Cada vizinhança coordenada é, por si só, uma variedade diferenciável coberta por apenas uma parametrização e possui, como anteriormente, uma base formada por bolas coordenadas. A reunião de todas estas bases obtidas desta maneira resulta em uma base \mathfrak{B} para os conjuntos abertos de M . Como anteriormente, podemos extrair uma subcobertura enumerável e localmente finita de \mathcal{X} formada por elementos de \mathfrak{B} . Deve-se tomar o cuidado que agora cada bola coordenada B_i é uma vizinhança parametrizada por uma carta ϕ_i . A menos de escrever estes novos índices, ou seja, a menos de escrever ϕ_i no lugar de ϕ , o restante da prova segue *verbatim*. \square

24 Integrais de Superfície

Nesta seção, nós introduzimos o importante conceito de integral de superfície. Como já mencionado anteriormente, no nosso curso, a ideia é entender formas diferenciais como objetos que são mais naturais dentro de um símbolo de integral. Ou ainda, formas diferenciais são “integrandos naturais”. A notação de integral, desde os cursos iniciais de cálculo, vem acompanhada do símbolo dx e neste curso nós revisitamos este conceito para significar um “covetor”, isto é, um funcional linear. Isto permitiu que expressões clássicas do cálculo como

$$u = f(x) \implies du = f'(x) dx$$

ficassem rigorosamente justificadas, mas é questionário se esta não seria apenas uma maneira artificial de se justificar a notação clássica. Deixamos esta discussão mais filosófica um pouco de lado, espero que felizes com a beleza da unificação que a abordagem de formas diferenciais traz ao assunto. Além disso, o formalismo abriu caminhos para muitos avanços e é hoje indispensável para muito além de “enunciar diversos teoremas do cálculo vetorial clássico em um único”. Nós comentamos sobre algumas dessas direções para futuras difereções de estudo ao final do texto.

Passemos às definições. Devemos manter em mente que formas diferenciais de grau n serão os objetos em que vamos definir integrais em variedades n -dimensionais. As superfícies diferenciáveis em \mathbb{R}^m são os exemplos mais importantes para fixar as ideias.

Seja M uma variedade diferenciável. Uma forma diferencial de grau p é uma aplicação que associa a cada $x \in M$ uma forma exterior de grau p em $(T_x M)^*$, isto é

$$\omega : x \in M \mapsto \omega(x) \in \Lambda^p(T_x M)^*.$$

Equivalentemente, utilizando a identificação entre formas exteriores e formas multilineares alternadas, podemos escrever que uma forma diferencial de grau p em M é uma aplicação do tipo

$$\omega : x \in M \mapsto \omega(x) \in \mathcal{A}_p(T_x M).$$

Para entender melhor este objeto, que está definido globalmente na variedade M , é instrutivo escrevê-lo em coordenadas locais. Seja $\phi : U \subset M \rightarrow \phi(U) \subset \mathbb{R}^n$ uma parametrização de M tal que $x \in U$. A parametrização ϕ determina uma base para os espaços tangentes a pontos pertencentes a $U \subset M$. Mais explicitamente, ver Subseção 22.1.4,

$$\left\{ \frac{\partial}{\partial x^1} \Big|_x, \frac{\partial}{\partial x^2} \Big|_x, \dots, \frac{\partial}{\partial x^n} \Big|_x \right\} \subset T_x M$$

é uma base de $T_x M$, cuja base dual associada é denotada por

$$\left\{ dx^1 \Big|_x, dx^2 \Big|_x, \dots, dx^n \Big|_x \right\} \subset T_x^* M,$$

isto é, os funcionais lineares $dx^i \Big|_x : T_x M \rightarrow \mathbb{R}$ satisfazem

$$dx^i \Big|_x \cdot \frac{\partial}{\partial x^j} \Big|_x = \delta_j^i = \begin{cases} 1, & \text{se } i = j; \\ 0, & \text{se } i \neq j. \end{cases}$$

Com o intuito de “descarregar a notação”, o ponto x é frequentemente omitido e o ponto do espaço tangente (ou cotangente) considerado deve ser entendido pelo contexto. Nós observamos que, exceto pelo fato de que as bases são locais em M (todas as expressões acima dependem de ϕ e de U), os funcionais e a notação foram determinados de tal maneira que trabalhar com os objetos em M fique muito semelhante com o que já fazíamos em espaços Euclidianos do tipo \mathbb{R}^m .

Com a notação acima, $\omega(x) \in \Lambda^p T_x^* M$ pode ser escrita como

$$\omega(x) = \sum_{1 \leq i_1 < i_2 < \dots < i_p \leq n} a_{i_1 i_2 \dots i_p}(x) dx^{i_1} \Big|_x \wedge dx^{i_2} \Big|_x \wedge \dots \wedge dx^{i_p} \Big|_x =: \sum_I a_I(x) dx^I \Big|_x,$$

onde a soma é tomada, como em seções anteriores, sobre partições “crescentes” $I = \{1 \leq i_1 < i_2 < \dots < i_p \leq n\}$ em conjunto com a notação

$$a_I(x) := a_{i_1 i_2 \dots i_p}(x) \quad \text{e} \quad dx^I := dx^{i_1} \wedge dx^{i_2} \wedge \dots \wedge dx^{i_p}.$$

Comparar estas expressões com a representação (59) de uma p -forma em \mathbb{R}^n . A diferença é que, em uma variedade diferenciável, as bases variam diferenciavelmente com $x \in U$ por conta da carta, enquanto que, em um espaço Euclidiano, todos os espaços tangentes podem ser identificados canonicamente.

Observação 108. Observamos que o contradomínio de ω não está explicitado na definição. Para escrever um contradomínio rigorosamente, devemos recorrer a um fibrado de formas exteriores do tipo

$$\wedge^p T^* M := \bigcup_{x \in M} \wedge^p T_x^* M.$$

Podemos ainda definir $\pi : \wedge^p T^*M \rightarrow M$ como segue: se $\eta \in \wedge^p T_x^*M$, definimos $\pi(\eta) = x$. Com esta aplicação, denominada projeção, $\wedge^p T^*M$ possui uma estrutura de fibrado vetorial⁵⁵ e uma forma diferencial de ordem p em M pode ser descrita como uma aplicação

$$\omega : M \rightarrow \wedge^p T^*M \quad \text{tal que} \quad \pi \circ \omega = \text{Id}_M.$$

Diz-se também que as formas diferenciais de grau p são as “seções do fibrado” $\wedge^p T^*M$.

24.1 Motivação e definições iniciais de integral

Seja ω uma forma diferencial de grau n em \mathbb{R}^n e observe que o grau da forma é igual à dimensão do espaço Euclidiano. Como a dimensão do espaço vetorial $\wedge^n(\mathbb{R}^n)^*$ é igual a um, podemos escrever, para todo $x \in \mathbb{R}^n$,

$$\omega(x) = f(x) dx^1 \wedge dx^2 \wedge \cdots \wedge dx^n.$$

Assumindo que o conjunto

$$\text{supp } \omega := \{x \in \mathbb{R}^n; \omega(x) \neq 0\},$$

denominado **suporte de ω** , é compacto e que ω é uma forma diferencial contínua, segue que f possui suporte compacto, é contínua e, em particular, é integrável. Nós definimos a integral de ω em \mathbb{R}^n (na realidade, a integral se dá em um conjunto limitado, o suporte de ω , como permitido na seção de integração de Riemann) pela fórmula

$$\int_{\mathbb{R}^n} \omega := \int_R f(x) dx,$$

onde R é um retângulo que contém $\text{supp } \omega = \text{supp } f$.

Enfatizamos a essência tautológica desta definição: ao denotar

$$x = (x^1, x^2, \dots, x^n) \quad \text{e} \quad dx = dx^1 dx^2 \cdots dx^n,$$

nossa definição se reescreve como

$$\int_{\mathbb{R}^n} f(x) dx^1 \wedge dx^2 \wedge \cdots \wedge dx^n := \int_{\mathbb{R}^n} f(x) dx^1 dx^2 \cdots dx^n,$$

onde o lado direito denota a integral de Riemann que estudamos anteriormente.

Observação 109. Para que esta definição não gere contradições, nós devemos fixar uma orientação (utilizamos a orientação em que a base canônica é positiva) para \mathbb{R}^n . Caso contrário, poderíamos mudar o sinal do lado esquerdo sem mudar o sinal do lado direito; por exemplo, utilizando que $dx^1 \wedge dx^2 = -dx^2 \wedge dx^1$. Dessa maneira, caso a forma esteja representada em uma base negativa de \mathbb{R}^n , nós devemos trocar o sinal na definição.

Em seguida, consideramos um difeomorfismo $g : U \rightarrow V$, onde $\text{supp } \omega \subset V = g(U)$. O pullback de ω pelo difeomorfismo g é a forma g^* cujo suporte $\text{supp } g^*\omega$ está contido em U . Pelo Exercício 49, temos que

$$(g^*\omega)(x) = \det g'(x) dx^1 \wedge dx^2 \wedge \cdots \wedge dx^n.$$

Caso o difeomorfismo g preserve orientação, o determinante acima é positivo e a fórmula de mudança de variáveis do Teorema 79 afirma que

$$\int_{g(U)} \omega = \int_U g^*\omega. \tag{72}$$

Caso g inverta orientação, temos que $\det g'(x) = -|\det g'(x)|$, mas o sinal na definição de integral de ω em $g(U)$ também deve ser invertido, de modo que a fórmula de mudança de variáveis nos dá a mesma expressão (72) neste caso também.

⁵⁵O que, exatamente, isto significa não é importante para o nosso texto.

Motivados por este fato e pensando numa carta de uma variedade diferenciável M como “uma mudança de variáveis”, a ideia é utilizar a fórmula (72) para *definir* integração de n -formas em variedades n -dimensionais. Entretanto, a dificuldade é que pode não existir uma carta global para a variedade M e o processo de “colar” vizinhanças coordenadas adjacentes é não trivial.

Se considerarmos inicialmente uma n -forma ω , contínua e de suporte compacto, definida em uma variedade diferenciável orientável de dimensão n que possa ser coberta por apenas uma parametrização (global) $\phi : M \rightarrow \phi(M) \subseteq \mathbb{R}^n$, é possível definir integração conforme motivamos acima. Definimos

$$\int_M \omega := \int_{\phi(M)} (\phi^{-1})^* \omega. \quad (73)$$

Temos que verificar que esta definição é independente da parametrização escolhida. Para isto, seja $\psi : M \rightarrow \psi(M) \subseteq \mathbb{R}^n$ uma parametrização que define a mesma orientação para M , isto é, tal que o difeomorfismo $\psi \circ \phi^{-1} : \phi(M) \rightarrow \psi(M)$ satisfaz

$$\det(\psi \circ \phi^{-1})'(y) > 0 \quad \text{para todo } y \in \phi(M).$$

Temos⁵⁶

$$\begin{aligned} \int_{\phi(M)} (\phi^{-1})^* \omega &= \int_{\phi(M)} (\phi^{-1})^* \omega \\ &= \int_{\phi(M)} (\psi^{-1} \circ \psi \circ \phi^{-1})^* \omega \\ &= \int_{\phi(M)} (\psi \circ \phi^{-1})^* (\psi^{-1})^* \omega \\ &= \int_{\psi(M)} (\psi^{-1})^* \omega, \end{aligned}$$

onde a última identidade é a fórmula de mudança de variáveis usual nesta nova notação, conforme no início desta subseção acima. Segue que a definição, como deveria, independe da parametrização escolhida.

A definição dada até agora é importante e funciona em vários casos interessantes:

- Superfícies parametrizadas, que por definição são cobertas por apenas uma carta;
- Formas diferenciais cujo suporte está contido em uma vizinhança parametrizada de M ;
- Para uma superfície diferenciável $M \subseteq \mathbb{R}^m$, quando uma carta parametriza um subconjunto de M de “medida total”, também podemos utilizar esta definição. Mais precisamente, a integral pode ser calculada com uma carta só se $\phi : U \subset M \rightarrow \phi(U) \subset \mathbb{R}^n$ é tal que $M \setminus U$ possui medida nula⁵⁷ em M .

24.2 Definição definitiva via partições da unidade

Seja M uma variedade diferenciável orientável (e orientada) M de dimensão n e $\omega \in \Omega^n(M)$ uma n -forma. Seja $\{U_i\}$ uma cobertura enumerável de M por vizinhanças coordenadas⁵⁸

$$M = \bigcup_{i \in \mathbb{N}} U_i \quad \text{com cartas} \quad \phi_i : U_i \rightarrow \phi_i(U_i) \subseteq \mathbb{R}^n.$$

⁵⁶Da segunda para a terceira linha, usamos que $(f \circ g)^* = g^* \circ f^*$, cuja verificação fica de exercício.

⁵⁷Aqui, o conceito de medida nula em M está sendo tomado de uma maneira informal e heurística. Deve ser pensado como medida nula “do ponto de vista da superfície”. Por exemplo, na esfera $\mathbb{S}^2 \subset \mathbb{R}^3$, as coordenadas esféricas apenas deixam de cobrir um semicírculo de um círculo máximo. Isto é uma curva unidimensional dentro de \mathbb{S}^2 que, do ponto de vista bidimensional da esfera, deve possuir medida nula. Não nos esforçaremos em formalizar esta ideia que exigiria algum esforço. Para estimular a curiosidade, dizemos que deve ser $\mathcal{H}^2(M \setminus U) = 0$, onde \mathcal{H}^2 é a medida de Hausdorff bidimensional em \mathbb{R}^3 .

⁵⁸Esta cobertura existe pelas restrições topológicas que usualmente são impostas em M , a saber, existência de base enumerável para a topologia. Esta hipótese também é conhecida como o segundo axioma de enumerabilidade. Isto já foi utilizado para provar existência de partições da unidade.

Pelo Teorema 107, existe uma partição da unidade subordinada a cobertura $\{U_i\}$, de modo que

$$f_i : M \longrightarrow [0, 1] \quad \text{satisfaz} \quad \text{supp } f_i \subset U_i \quad \text{e} \quad \sum_{i=1}^{+\infty} f_i(x) = 1.$$

Segue que a forma $f_i\omega \in \Omega^n(M)$ definida por

$$(f_i\omega)(x) := f_i(x)\omega(x)$$

possui suporte compacto contido em $\text{supp } f_i \subset U_i$ e seus valores, portanto, somente são não nulos na vizinhança coordenada U_i . Além disso,

$$\sum_{i=1}^{+\infty} (f_i\omega)(x) = 1.$$

Nós definimos

$$\int_M \omega := \sum_{i=1}^{+\infty} \int_M f_i\omega = \sum_{i=1}^{+\infty} \int_{\phi(U_i)} (\phi_i^{-1})^*(f_i\omega) \quad (74)$$

e observamos que a soma do lado direito é sempre finita (e não uma série). Observamos ainda que as integrais do lado direito independem das parametrizações escolhidas, mas devemos verificar que nosso conceito também é independente da partição da unidade considerada.

Lema 110. *A definição em (74) é independente da partição da unidade escolhida.*

Demonstração. Fixada uma cobertura por vizinhanças coordenadas, consideramos $\{f_i\}$ e $\{g_j\}$ partições da unidade subordinadas. Assim,

$$\begin{aligned} \sum_{i=1}^{+\infty} \int_M f_i\omega &= \sum_{j=1}^{+\infty} g_j \sum_{i=1}^{+\infty} \int_M f_i\omega \\ &= \sum_{j=1}^{+\infty} \sum_{i=1}^{+\infty} \int_M g_j f_i\omega \\ &= \sum_{i=1}^{+\infty} f_i \sum_{j=1}^{+\infty} \int_M g_j\omega \\ &= \sum_{j=1}^{+\infty} \int_M g_j\omega. \end{aligned}$$

Observamos que a família $\{\text{supp } f_i g_j\}_{i,j \in \mathbb{N}}$ continua sendo localmente finita, já o produto $f_i g_j$ se anula fora da interseção dos suportes de cada uma das funções. Logo, as somas em i, j são, na realidade, finitas e não há problemas ao comutar os somatórios acima. Provamos assim que nossa definição não depende da partição da unidade escolhida. \square

25 Variedades com bordo

Até o momento, consideramos superfícies e variedades diferenciáveis sem bordo⁵⁹. Os exemplos que temos em mente, sem bordo, são (abertos de) espaços Euclidianos, superfícies esféricas, toros, faixas de Möbius (sem círculo do bordo), etc. Nesta seção, introduzimos o conceito de

⁵⁹O uso da palavra bordo é intuitivo e, neste contexto, muitas vezes se confunde com fronteira. No entanto, a fronteira topológica de uma superfície de \mathbb{R}^m nem sempre coincide com o seu bordo. Por exemplo, um disco aberto em \mathbb{R}^2 é uma superfície sem bordo, isto é, seu bordo é o conjunto vazio; no entanto, sua fronteira topológica (em \mathbb{R}^2) é um círculo.

variedade com bordo que inclui conjuntos como bolas fechadas em \mathbb{R}^n , toros sólidos, além de todos os exemplos de variedades “sem bordo”.

A idéia da definição é a seguinte: uma variedade com bordo é um conjunto (mais precisamente, um espaço topológico) M tal que cada $x \in M$ possui uma vizinhança U que é homeomorfa ou a uma vizinhança aberta de \mathbb{R}^n ou a uma vizinhança aberta do semi-espaço

$$\mathbb{H}^n := \{(x^1, x^2, \dots, x^n) \in \mathbb{R}^n; x^n \geq 0\}$$

e, além disso, as mudanças de parâmetros são difeomorfismos. Intuitivamente, pontos de uma vizinhança coordenada (ϕ, U) são pontos interiores de M quando a n -ésima componente de $\phi(x)$ é positiva e são pontos do bordo de M quando a n -ésima componente é zero. Um detalhe sutil que deve ser averiguado é que, *a priori*, poderia acontecer de um ponto $x \in M$ estar ao mesmo tempo na imagem de duas cartas e numa delas ter sua última componente nula, enquanto na outra positiva. Além disso, outros pontos desta “definição” ainda precisam ser esclarecidos “mais urgentemente”, como, por exemplo, a diferenciabilidade das funções de mudança de parâmetros, definidas em abertos de \mathbb{H}^n que podem ser fechados de \mathbb{R}^n . Lembre que (no sentido usual) apenas sabemos diferenciar funções em abertos de espaços Euclidianos.

Dado um subconjunto $A \subseteq \mathbb{H}^n$, aberto em \mathbb{H}^n , e uma função $f : A \rightarrow B \subset \mathbb{R}^m$, nós dizemos que f é **diferenciável** em A quando existem vizinhanças (em \mathbb{R}^n) abertas $U \subseteq \mathbb{R}^n$ e $V \subset \mathbb{R}^m$ de A e de B , respectivamente, e uma função diferenciável (no sentido usual) $\tilde{f} : U \rightarrow V$ tal que

$$\tilde{f}(x) = f(x) \quad \text{para todo } x \in A.$$

Em outras palavras, uma função definida em um aberto de \mathbb{H}^n é uma função que possui uma extensão diferenciável a um aberto de \mathbb{R}^n . De maneira análoga, definimos funções de classe C^k , difeomorfismos, etc, em abertos de \mathbb{H}^n . Para pontos $x \in A \cap \text{int } \mathbb{H}^n$, esta nova noção de diferenciabilidade em x coincide com a usual. No entanto, a nova definição se aplica também para pontos $x \in A \cap \partial \mathbb{H}^n$.

Embora a noção de diferenciabilidade acima se aplique *verbatim* a subconjuntos arbitrários de \mathbb{R}^n , a noção de derivada $f'(x)$, para $x \in A$, pode não estar definida, no sentido de extensões diferentes possuírem derivadas distintas. No entanto, verificamos que, para o caso especial de abertos de \mathbb{H}^n , a noção de derivada está bem definida.

Proposição 111. *Sejam $A \subseteq \mathbb{H}^n$ um aberto em \mathbb{H}^n e $f : A \rightarrow \mathbb{R}^m$ uma função diferenciável em A . Então, dado $x \in A$, toda extensão diferenciável de f possui a mesma derivada em x .*

Demonstração. Seja \tilde{f} uma extensão diferenciável de f . Se $x \in \text{int } A \cap \mathbb{H}^n$, então f é diferenciável no sentido usual e não há o que provar. Suponhamos então que $x \in A \cap \partial \mathbb{H}^n$, isto é, que $x^n = 0$. Sendo \tilde{f} diferenciável (no sentido usual), é possível tomar limites, na definição de derivada, por qualquer direção possível. Se $v \in \mathbb{R}^n$ é tal que $v^n \geq 0$, então $x + tv \in \mathbb{H}^n$ para todo $t > 0$. Logo, sendo \tilde{f} uma extensão de f , tem-se

$$d\tilde{f}(x) \cdot v = \lim_{t \rightarrow 0^+} \tilde{f}(x + tv) = \lim_{t \rightarrow 0^+} f(x + tv).$$

Esta última expressão mostra que $d\tilde{f}(x) \cdot v$ é independente dos valores da extensão, pois depende apenas dos valores de f em $x + tv$ com $t > 0$. Analogamente, se $v^n < 0$, basta considerarmos o limite quando $t \rightarrow 0^-$, pois neste caso tem-se $x + tv \in \mathbb{H}^n$ para $t < 0$. Portanto, qualquer que seja $v \in \mathbb{R}^n$, a derivada de \tilde{f} em x , para qualquer extensão \tilde{f} , depende apenas dos valores de f em $A \subseteq \mathbb{H}^n$ como acima. \square

Podemos agora introduzir o conceito rigorosamente.

Uma **variedade com bordo** de dimensão n é um espaço topológico M junto com um conjunto de homeomorfismos, chamados cartas, mapas ou parametrizações, $\phi_\alpha : U_\alpha \rightarrow \phi_\alpha(U_\alpha) \subseteq \mathbb{H}^n$, onde os domínios $U_\alpha \subseteq M$ são abertos em M que cobrem todo o conjunto M e as funções de transição $\phi_\alpha \circ \phi_\beta^{-1}$ são diferenciáveis nos conjuntos onde estão definidas. O conjunto de todas as cartas $\{\phi_\alpha\}$ de uma variedade com bordo M é chamado um atlas de M .

Note como a diferença desta definição para a de variedade diferenciável (sem bordo) é apenas que as parametrizações são dadas no semi-espaço \mathbb{H}^n e que a diferenciabilidade das mudanças de parâmetro devem ser dadas de acordo com a definição discutida no início desta seção.

Passamos a discutir a noção de bordo ∂M de M . Heuristicamente, temos o seguinte:

- Se $\phi : U \subseteq M \rightarrow \phi(U) \subset \mathbb{H}^n$ é uma vizinhança coordenada de um ponto $x \in U$ tal que

$$\phi(U) \subseteq \mathbb{H}_+^n := \{(x^1, x^2, \dots, x^n) \in \mathbb{R}^n; x^n > 0\},$$

então $\phi(U)$ é também um conjunto aberto em \mathbb{R}^n e esta carta é “do mesmo tipo” das que vínhamos considerando até aqui. Neste caso, o ponto $x \in M$ é dito um ponto interior de M e o conjunto dos pontos interiores é denotado por $\text{int } M$.

- Por outro lado, denotando

$$\partial\mathbb{H}^n := \{(x^1, x^2, \dots, x^{n-1}, 0) \in \mathbb{R}^n; x^1, x_2, \dots, x^{n-1} \in \mathbb{R}\},$$

a vizinhança $\phi(U) \subseteq \mathbb{H}^n$ é aberta em \mathbb{H}^n e logo pode ser $\phi(U) \cap \partial\mathbb{H}^n \neq \emptyset$. Um ponto $x \in U$ tal que $\phi(x) \in \partial\mathbb{H}^n$ é chamado um ponto do bordo de M e o conjunto dos pontos do bordo é denotado por ∂M .

- Intuitivamente, temos que $\text{int } M$ é uma variedade diferenciável (sem bordo) de dimensão n , enquanto ∂M é uma variedade diferenciável de dimensão $n - 1$, também sem bordo. Nós vamos mostrar que, sendo M orientável, também ∂M é orientável e possui uma orientação naturalmente induzida pela orientação de M .

As definições acima apenas não são rigorosas porque, *a priori*, não é claro $\text{int } M \cap \partial M = \emptyset$, isto é, como não é imediato que o bordo é disjunto do conjunto dos pontos interiores, as definições podem não estar bem postas.

O próximo resultado é o que implica que o bordo de M está bem definido. Conhecido Teorema da Invariância (diferenciável) do Bordo.

Teorema 112. *Seja M uma variedade n -dimensional com bordo. Seja ainda $x \in M$ tal que existe uma carta $\phi : U \subset M \rightarrow \phi(U) \subset \mathbb{H}^n$ para a qual $\phi(x) \in \partial\mathbb{H}^n$. Então, $\psi(x) \in \partial\mathbb{H}^n$ para qualquer carta $\psi : V \subset M \rightarrow \psi(U) \subset \mathbb{H}^n$ com $x \in V$.*

Demonstração. Por contradição, suponhamos que existe uma carta ψ , como no enunciado, mas tal que $\psi(x) \in \text{int } \mathbb{H}^n$. Pela definição de variedade com bordo, tanto o homeomorfismo

$$\xi := \phi \circ \psi^{-1} : \psi(U \cap V) \rightarrow \phi(U \cap V)$$

quanto o homeomorfismo inverso ξ^{-1} são diferenciáveis no sentido de possuírem uma extensões diferenciáveis $\bar{\xi}$ e $\bar{\xi}^{-1}$, respectivamente, a abertos de \mathbb{R}^n . Em particular, existe um aberto W contendo $\phi(x)$ de maneira que $\bar{\xi}^{-1} : W \rightarrow \mathbb{R}^n$ é diferenciável e $\bar{\xi}^{-1} = \xi^{-1}$ em $W \cap \phi(U \cap V)$.

Agora, da hipótese de que $\psi(x) \in \text{int } \mathbb{H}^n$, segue que existe um aberto (em \mathbb{R}^n) Z de $\psi(x)$ onde ξ é diferenciável no sentido usual. Sem perda de generalidade, podemos assumir que Z está contido no aberto $\bar{\xi}^{-1}(W)$. Logo,

$$\bar{\xi}^{-1} \circ \xi|_Z = \xi^{-1} \circ \xi|_Z = \text{Id}|_Z \quad \text{donde} \quad d(\bar{\xi}^{-1})(\xi(y)) \cdot d\xi(y) = \text{Id}, \quad \forall y \in Z.$$

Em particular, $d\xi(y)$ é invertível e, por corolário do Teorema da Aplicação Inversa (ver na lista de exercícios), ξ é uma aplicação aberta. Portanto, $\xi(Z) \subset \mathbb{H}^n$ é também um aberto em \mathbb{R}^n que contém $\phi(x)$, contradizendo que $\phi(x) \in \partial\mathbb{H}^n$. \square

25.1 Espaço tangente em pontos de ∂M

Para uma superfície diferenciável n -dimensional $S \subseteq \mathbb{R}^m$, o fato de sem incoerências extensão diferenciável de uma carta ϕ definir $d\phi(x)$ coerentemente implica que o espaço tangente pode ser definido como $\text{Im } d\bar{\phi}(\mathbb{R}^n)$ para qualquer extensão $\bar{\phi}$ de ϕ . Esta imagem pode ser “visualizada” da mesma maneira que fizemos na Seção 13, mas agora no bordo de S . Ainda, o conjunto $\text{Im } d\bar{\phi}(\mathbb{H}^n)$ é um subespaço de $T_x S$ que é tangente à superfície $(n-1)$ -dimensional ∂S , isto é, $T_x(\partial S) = \text{Im } d\bar{\phi}(\mathbb{H}^n) \subset T_x S$.

No caso geral de uma variedade com bordo M , uma possibilidade é pensar nos vetores tangentes como derivações e fazer as restrições apropriadas para definir $T_x(\partial M)$. Denotamos

$$C_{\partial M}(M) := \{f : M \rightarrow \mathbb{R}; f \text{ é diferenciável em } M \text{ com bordo}\} \quad \text{e}$$

$$C^\infty(\partial M) := \{f : \partial M \rightarrow \mathbb{R}; f \text{ possui extensão diferenciável em vizinhança de } \partial M \text{ em } M\}.$$

Daí, definimos

$$T_x M := \{v : C_{\partial M}(M) \rightarrow \mathbb{R}; v \text{ é uma derivação linear em } x\} \quad \text{e}$$

$$T_x(\partial M) := \{v : C^\infty(\partial M) \rightarrow \mathbb{R}; v \text{ é uma derivação linear em } x\}.$$

Dessa maneira, dada uma função $f \in C^\infty(M)$, a sua restrição ao subconjunto ∂M , que por si só é uma variedade diferenciável, pertence a $C^\infty(\partial M)$, isto é, $f \in C^\infty(M) \implies f|_{\partial M} \in C^\infty(\partial M)$. Segue disto que

$$T_x(\partial M) \subset T_x M$$

como subespaço vetorial.

25.2 Orientação induzida no bordo ∂M

Em uma variedade com bordo M que é orientável com uma orientação (chamada de) positiva fixada, temos que, assim como no caso de variedades sem bordo, um atlas coerente induz orientações nos espaços tangentes $T_x M$. Mesmo quando $x \in \partial M$, a partir de uma carta positiva $\phi : U \subset M \rightarrow \mathbb{H}^n$, temos uma orientação positiva para $T_x M$ dada pela base

$$\mathcal{B} := \left\{ \frac{\partial}{\partial x^1} \Big|_x, \frac{\partial}{\partial x^2} \Big|_x, \dots, \frac{\partial}{\partial x^n} \Big|_x \right\},$$

onde (relembre)

$$\frac{\partial}{\partial x^i} \Big|_x (f) = \frac{\partial (f \circ \phi^{-1})}{\partial x^i} (\phi(x)).$$

Como $\phi(x) \in \partial \mathbb{H}^n$, temos que a derivada acima deve ser tomada em (qualquer) extensão diferenciável de $f \circ \phi^{-1}$.

Toda parametrização de uma vizinhança $U_0 \subseteq \partial M$ pode ser considerada como restrição de uma parametrização de uma vizinhança coordenada $U \subset M$ tal que $U \cap \partial M = U_0$. Seja \mathcal{A} um atlas coerente de M . A **orientação induzida** por \mathcal{A} em ∂M pode ser descrita da seguinte forma: nós definimos um conjunto de cartas \mathcal{A}_0 de ∂M por funções

$$\phi_0 : U_0 \subseteq \partial M \rightarrow \partial \mathbb{H}^n$$

cujas extensões $\phi : U \subseteq M \rightarrow \mathbb{H}^n$ pertencem a \mathcal{A} . Fica como exercício verificar que \mathcal{A} é um atlas coerente de ∂M . Em particular, isto mostra que se M é orientável, então também ∂M o é.

Uma descrição mais geométrica da orientação induzida em ∂M é dada em termos de vetores que “apontam para fora” de M . Digamos que um vetor tangente $v \in T_x M$ não pertence ao subespaço $T_x(\partial M)$. Sendo um hiperplano de $T_x M$, o subespaço $T_x(\partial M)$ divide $T_x M$ em dois semi-espaços disjuntos. Nós dizemos que $v \in T_x M \setminus T_x(\partial M)$ **aponta para fora** de M quando

$$v(\phi^n) < 0,$$

isto é, quando a n -ésima componente de v em termos da base \mathcal{B} de $T_x M$ for negativa. Note que $v(\phi^n) = 0$ significa que $v \in T_x(\partial M)$. Assim, uma base de $\{v_1, v_2, \dots, v_{n-1}\} \subset T_x(\partial M)$ é positiva quando, dado um vetor tangente $v \in T_x M$ que aponta para fora de M , temos que $\{v_1, v_2, \dots, v_{n-1}, v\}$ possui a mesma orientação de \mathcal{B} .

Exercício 52. Considere $\bar{B} \subset \mathbb{R}^3$ a bola fechada unitária e centrada na origem, que é uma variedade com bordo em \mathbb{R}^3 . Dada a orientação canônica de \bar{B} , encontre a orientação induzida em $\partial \bar{B} = \mathbb{S}^2$.

26 Teorema de Stokes

Nesta seção final, provamos o Teorema de Stokes que era um dos objetivos principais do nosso curso, em conjunto com o Teorema da Função Inversa, o Teorema da Função Implícita, as noções de superfície e de variedade diferenciáveis, Fórmula de Mudança de Variáveis.

O Teorema de Stokes geral que provamos abaixo traz toda essa novidade das formas diferenciais, produtos exteriores, diferenciação exterior, que transformou a maneira de se pensar em matemática e física no fim do século XIX e início do século XX. Em eletromagnetismo, por exemplo, se descobriu ser mais natural pensar em um campo elétrico como uma 1-forma enquanto o campo magnético é uma 2-forma. Já se sabia que magnético “pseudo-vetor”, mas o produto exterior captura todas as propriedades em um conceito só.

Teorema 113 (Teorema de Stokes). *Seja M uma variedade com bordo de dimensão n e $\omega \in \Omega^{n-1}(M)$ uma forma de suporte compacto. Então*

$$\int_M d\omega = \int_{\partial M} \omega.$$

Demonstração. Seja $\{f_i\}_{i \in \mathbb{N}}$ uma partição da unidade para M subordinada a uma cobertura $\{U_i\}_{i \in \mathbb{N}}$; em particular, $\text{supp } f_i \subset U_i$. Temos

$$\omega = \sum_{i=1}^{+\infty} (f_i \omega)$$

e ainda, pela linearidade da diferenciação exterior e da integral, temos

$$\int_M d\omega = \sum_{i=1}^{+\infty} \int_M d(f_i \omega).$$

Vejamos agora que basta mostrar o teorema localmente em U_i , isto é, basta mostrar que

$$\int_{U_i} d(f_i \omega) = \int_{U_i \cap \partial M} f_i \omega;$$

de fato, estando o suporte de f_i em U_i , podemos escrever

$$\int_M d\omega = \sum_{i=1}^{+\infty} \int_M d(f_i \omega) = \sum_{i=1}^{+\infty} \int_{U_i} d(f_i \omega) = \sum_{i=1}^{+\infty} \int_{U_i \cap \partial M} f_i \omega = \sum_{i=1}^{+\infty} \int_{\partial M} f_i \omega = \int_{\partial M} d\omega.$$

Este raciocínio prova ainda que basta verificar o teorema localmente em uma vizinhança coordenada. É isto que passamos a fazer, escrevendo ω ao invés de $f_i \omega$ abaixo.

A ideia agora é a seguinte. Se uma vizinhança coordenada não tem pontos do bordo, então mostramos que ambas integrais são nulas. Por outro lado, vale a fórmula também quando a vizinhança coordenada intersecciona o bordo.

Seja $\phi : U \subseteq M \rightarrow V \subseteq \mathbb{H}^n$ uma carta local, onde escrevemos $V := \phi(U)$. Na vizinhança coordenada U , podemos escrever ω em coordenadas como

$$\omega(x) = \sum_{i=1}^n (-1)^{i+1} a_i(\phi(x)) dx^1 \wedge dx^2 \wedge \cdots \wedge \widehat{dx^i} \wedge \cdots \wedge dx^n,$$

onde a notação $\widehat{dx^i}$ significa que o termo dx^i é omitido no produto exterior (que fica com $n-1$ fatores). A menos de redefinir o sinal dos coeficientes, não há perda em generalidade na introdução do fator $(-1)^{i+1}$; a sua motivação é o cálculo abaixo e a similaridade do coeficiente de $d\omega$ com o divergente de um campo de vetores. O diferencial exterior de ω pode ser calculado como

$$\begin{aligned} d\omega(x) &= \sum_{i=1}^n (-1)^{i+1} d(a_i \circ \phi)(x) \wedge dx^1 \wedge dx^2 \wedge \cdots \wedge \widehat{dx^i} \wedge \cdots \wedge dx^n \\ &= \sum_{i,j=1}^n (-1)^{i+1} \frac{\partial a_i}{\partial x^j}(x) dx^j \wedge dx^1 \wedge dx^2 \wedge \cdots \wedge \widehat{dx^i} \wedge \cdots \wedge dx^n \\ &= \sum_{i=1}^n (-1)^{i+1} \frac{\partial a_i}{\partial x^i}(x) dx^i \wedge dx^1 \wedge dx^2 \wedge \cdots \wedge \widehat{dx^i} \wedge \cdots \wedge dx^n \\ &= \sum_{i=1}^n \frac{\partial a_i}{\partial x^i}(x) dx^1 \wedge dx^2 \wedge \cdots \wedge dx^n \\ &= \left(\sum_{i=1}^n \frac{\partial a_i}{\partial x^i}(x) \right) dx^1 \wedge dx^2 \wedge \cdots \wedge dx^n. \end{aligned}$$

Por hipótese, o suporte de ω é compacto e, logo, podemos assumir que, para todo i , o suporte de $a_i \circ \phi : V \rightarrow \mathbb{R}$ está contido em um retângulo da forma $R = [c_1, d_1] \times [c_2, d_2] \times \cdots \times [c_n, d_n] \subset \mathbb{R}^n$. Os pontos de U que estão no bordo ∂M são os pontos $x \in U$ para os quais a última componente de $\phi(x)$ é nula, ou seja, $\phi(x) = (x^1, x^2, \dots, x^{n-1}, 0)$. Denotamos a carta induzida no bordo por $\psi : U \cap \partial M \rightarrow V_0 \subset \mathbb{R}^{n-1}$, que para $x \in U \cap \partial M$ satisfaz

$$\psi(x) = (x^1, x^2, \dots, x^{n-1}).$$

Acima, denotamos a imagem de ψ por V_0 . Além disso, o suporte da “restrição ao bordo” $i^*\omega$, onde $i : \partial M \rightarrow M$ é a inclusão, deve estar contido em $R_0 = [c_1, d_1] \times [c_2, d_2] \times \cdots \times [c_{n-1}, d_{n-1}] \subset \mathbb{R}^{n-1}$.

Assim, mostrar que

$$\int_U d\omega = \int_{U \cap \partial M} \omega$$

é equivalente a mostrar que

$$\int_R (\phi^{-1})^* d\omega = \int_{R_0} (\psi^{-1})^* \omega.$$

Ou ainda, é equivalente a mostrar que

$$\sum_{i=1}^n \int_{c_1}^{d_1} \int_{c_2}^{d_2} \cdots \int_{c_n}^{d_n} \frac{\partial a_i}{\partial x^i}(x) dx^1 dx^2 \cdots dx^n = \sum_{i=1}^n \int_{c_1}^{d_1} \cdots \int_{c_i}^{\widehat{d_i}} \cdots \int_{c_n}^{d_n} \zeta_i dx^1 \cdots \widehat{dx^i} \cdots dx^n, \quad (75)$$

onde novamente os “chapéus” indicam que o respectivo item deve ser “removido” e onde

$$\zeta_i(x^1, \dots, \widehat{x^i}, \dots, x^n) := a_i(x^1, \dots, d_i, \dots, x^n) - a_i(x^1, \dots, c_i, \dots, x^n).$$

Neste momento, observamos que, com exceção do uso da partição da unidade, a demonstração ainda não começou. Tudo o que fizemos até aqui foi introduzir notação e utilizar as nossas definições para representar os objetos localmente. E esta é o principal desafio do material: entender os

conceitos e o formalismo básico para que só então a teoria fique apresentada de maneira simples. Walter Rudin, no clássico “Principles of Mathematical Analysis”, página 253 da terceira edição, afirma⁶⁰:

“It is a curious feature of Stokes’ theorem that the only thing that is difficult about it is the elaborate structure of definitions that are needed for its statement. These definitions concern differential forms, their derivatives, boundaries, and orientation. Once these concepts are understood, the statement of the theorem is very brief and succinct, and its proof presents no difficulty.”

Passamos à demonstração do teorema, que é dividida em dois casos: o primeiro em que a vizinhança coordenada não intersecciona o bordo de M e o segundo em que intersecciona.

Caso 1: $\phi(U) \cap \partial\mathbb{H}^n = \emptyset$.

Neste caso, temos $i^*\omega = 0$, pois o suporte de ω está em $\text{int } M$. Logo, o lado direito é zero. No lado esquerdo, para cada $i = 1, 2, \dots, n$, utilizamos (mudança de ordem de integração e) o Teorema Fundamental do Cálculo para concluir que

$$\int_{c_i}^{d_i} \frac{\partial a_i}{\partial x^i}(x) dx^i = a_i(x^1, \dots, d_i, \dots, x^n) - a_i(x^1, \dots, c_i, \dots, x^n) = 0. \quad (76)$$

A última igualdade segue porque, neste caso, $a_i \equiv 0$ em ∂R .

Caso 2: $\phi(U) \cap \partial\mathbb{H}^n \neq \emptyset$.

Isto significa que, no bloco R , devemos ter $c_n = 0$. Notamos que, para $i = 1, 2, \dots, n-1$, chegamos à mesma conclusão de (76). Por outro lado, pode ser

$$\int_{c_n}^{d_n} \frac{\partial a_n}{\partial x^n}(x) dx^n = -a_n(x^1, x^2, \dots, x^{n-1}, 0) \neq 0.$$

Segue que o lado esquerdo de (75) é igual a

$$- \int_{c_1}^{d_1} \int_{c_2}^{d_2} \dots \int_{c_{n-1}}^{d_{n-1}} a_n(x^1, x^2, \dots, x^{n-1}, 0) dx^1 dx^2 \dots dx^{n-1}. \quad (77)$$

Resta calcular o lado direito. As funções coordenada $a_i \circ \phi$, nas faces que formam o bordo retangular, são todas nulas, exceto possivelmente aquela em que $x^n = 0$. Dessa maneira, as funções ζ_i são todas nulas exceto o caso $i = n$, onde temos

$$\zeta_n(x^1, \dots, x^{n-1}) = -a_n(x^1, \dots, x^{n-1}, 0).$$

É então a escolha de orientação induzida no bordo que garante que também o lado direito deve ser igual à expressão em (77), encerrando a demonstração. \square

Para variedades diferenciáveis sem bordo, tem-se $\partial M = \emptyset$ e logo temos a seguinte consequência (alternativamente, isto foi o que provamos no “Caso 1” da demonstração do Teorema de Stokes).

Corolário 114. *Seja M uma variedade sem bordo de dimensão n e $\omega \in \Omega^{n-1}(M)$ uma forma de suporte compacto. Então*

$$\int_M d\omega = 0.$$

Uma forma η que é a derivada exterior de alguma forma ω , isto é, no caso $\eta = d\omega$, é dita uma forma exata. O corolário acima, nesta nova terminologia, afirma que toda forma exata tem integral nula em uma variedade sem bordo.

⁶⁰Em tradução livre: “É uma característica curiosa do Teorema de Stokes que a única coisa difícil sobre ele é a estrutura elaborada de definições que são necessárias para o seu enunciado. Essas definições dizem respeito a formas diferenciais, suas derivadas, bordos e orientação. Uma vez que esses conceitos são compreendidos, o enunciado do teorema é muito breve e sucinto, e sua demonstração não apresenta dificuldades.”

26.1 O Teorema da Divergência de Gauss-Ostrogradsky

Seja $U \subset \mathbb{R}^3$ um aberto limitado cujo bordo é uma hiperfície orientável S . Dadas funções escalares $F_1, F_2, F_3 : \mathbb{R}^3 \rightarrow \mathbb{R}$, consideramos a 2-forma

$$\omega = F_1 dy \wedge dz + F_2 dz \wedge dx + F_3 dx \wedge dy.$$

Nós poderíamos pensar nas três funções escalares como um vetor, mas decidimos pensar como uma 2-forma, o que em \mathbb{R}^3 , é possível. Temos

$$d\omega = (\operatorname{div} F) dx \wedge dy \wedge dz.$$

Por outro lado, dados $v = (v^1, v^2, v^3)$ e $w = (w^1, w^2, w^3)$, temos⁶¹

$$\begin{aligned} \omega(x)[v, w] &= F_1(x) \det \begin{bmatrix} v^2 & v^3 \\ w^2 & w^3 \end{bmatrix} - F_2(x) \det \begin{bmatrix} v^1 & v^3 \\ w^1 & w^3 \end{bmatrix} + F_3(x) \det \begin{bmatrix} v^1 & v^2 \\ w^1 & w^2 \end{bmatrix} \\ &= \det \begin{bmatrix} F_1(x) & F_2(x) & F_3(x) \\ v^1 & v^2 & v^3 \\ w^1 & w^2 & w^3 \end{bmatrix} = \langle F(x), v \times w \rangle = \left\langle F(x), \frac{v \times w}{|v \times w|} \right\rangle |v \times w| \\ &= \langle F(x), \nu(x) \rangle \cdot (\text{área orientada do paralelogramo gerado por } v \text{ e } w) \\ &= \langle F(x), \nu(x) \rangle d\sigma(x)[v, w]. \end{aligned}$$

Logo, o seguinte é consequência direta do Teorema de Stokes.

Teorema 115. *Seja $U \subset \mathbb{R}^3$ um aberto limitado cujo bordo é uma hiperfície orientável S . Para $F : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ um campo vetorial de classe C^1 , tem-se*

$$\int_U (\operatorname{div} F) dx \wedge dy \wedge dz = \int_S \langle F, \nu \rangle d\sigma,$$

onde ν é o vetor normal unitário exterior a S .

26.2 O Teorema do Rotacional de Kelvin-Stokes

Deixamos a verificação deste como exercício. O enunciado está abaixo.

Exercício 53. *Seja $F : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ um campo vetorial de classe C^1 e seja $S \subset \mathbb{R}^3$ uma superfície com bordo, limitada e orientável de classe C^∞ cuja fronteira é uma curva fechada simples C com a orientação induzida. O Teorema de Stokes (clássico) afirma que*

$$\int_S (\operatorname{rot} F \cdot \nu) d\sigma = \oint_C F.$$

Obtenha o resultado acima a partir do Teorema de Stokes para formas diferenciais.

Lembre que o rotacional de F (se pensarmos como vetor) é o campo vetorial definido por

$$\operatorname{rot} F(x) := \left(\frac{\partial F_3}{\partial y} - \frac{\partial F_2}{\partial z}, \frac{\partial F_1}{\partial z} - \frac{\partial F_3}{\partial x}, \frac{\partial F_2}{\partial x} - \frac{\partial F_1}{\partial y} \right).$$

⁶¹Era possível não falar no produto vetorial. O determinante 3×3 acima é o volume (orientado) do paralelepípedo gerado pelos vetores F , v e w . Logo, é igual a área da base – paralelogramo gerado por v e w – vezes a altura, que é $\langle F, \nu \rangle$.

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.