

Cálculo Diferencial e Integral II

Cursos: LEAN, MEAero (2014/15, Semestre 2)

Apontamentos das aulas teóricas

0. INTRODUÇÃO

Este texto consiste numa transcrição razoavelmente fiel das apresentações feitas nas aulas teóricas. Não se destina a substituir um livro de texto. Para um tratamento coerente dos assuntos desta cadeira recomendamos [P1] e [P2]. Um excelente texto para os alunos mais interessados é [S].

Cada secção corresponde a uma aula teórica.

1. APRESENTAÇÃO E FUNCIONAMENTO DA CADEIRA

2. NOÇÕES TOPOLÓGICAS EM \mathbb{R}^n

O espaço Euclidiano de dimensão n é o conjunto de n -tuplos de números reais

$$\mathbb{R}^n = \{(x_1, \dots, x_n) : x_i \in \mathbb{R}\}.$$

Para $n = 1, 2, 3$ os elementos de \mathbb{R}^n correspondem a pontos numa linha, plano ou espaço tridimensional munidos de eixos coordenados.

Recorde-se que a *norma* de um vector $x = (x_1, \dots, x_n) \in \mathbb{R}^n$ é o real não negativo dado pela expressão

$$\|x\| = \sqrt{x_1^2 + \dots + x_n^2}$$

e que o seu significado geométrico é o comprimento do vector x . Tendo em conta o significado geométrico da adição de vectores, segue-se que a *distância* entre dois pontos x, y de \mathbb{R}^n é $\|x - y\|$.

Definição 2.1. *Seja $x_0 \in \mathbb{R}^n$ e $r > 0$. A bola aberta de centro em x_0 e raio r é o conjunto formado por todos os pontos de \mathbb{R}^n cuja distância a x_0 é inferior a r :*

$$B_r(x_0) = \{x \in \mathbb{R}^n : \|x - x_0\| < r\}$$

Para $n = 1$, a bola de raio r é um intervalo aberto de comprimento $2r$, para $n = 2$ é a região do plano limitada por uma circunferência de raio r excluindo a circunferência, e para $n = 3$, a região do espaço limitada por uma superfície esférica de raio r centrada em x_0 (excluindo a própria superfície).

Definição 2.2. *Seja A um subconjunto de \mathbb{R}^n . Um ponto $x \in \mathbb{R}^n$ diz-se*

- (i) interior a A se existe $r > 0$ tal que $B_r(x) \subset A$,
- (ii) exterior a A se existe $r > 0$ tal que $B_r(x) \cap A = \emptyset$ (alternativamente se x é interior ao complementar de A , $A^c = \mathbb{R}^n \setminus A$),
- (iii) fronteiro a A se não é interior nem exterior (alternativamente, se toda a bola aberta $B_r(x)$ tem intersecção não vazia com A e com A^c),

(iv) aderente a A se é interior ou fronteiro a A (alternativamente, se toda a bola aberta $B_r(x)$ intersecta A).

O conjunto dos pontos interiores a A designa-se por interior de A e denota-se por $\text{int } A$. O conjunto dos pontos exteriores a A designa-se por exterior de A e denota-se por $\text{ext } A$. O conjunto dos pontos fronteiros designa-se por fronteira de A e denota-se por $\text{front } A$. O conjunto dos pontos aderentes a A designa-se por aderência de A ou fecho de A e denota-se por \bar{A} .

Note-se que qualquer conjunto $A \subset \mathbb{R}^n$ divide \mathbb{R}^n em três conjuntos disjuntos: o interior, o exterior e a fronteira de A . Intuitivamente, os pontos do fecho de A são os pontos que estão "infinitamente próximos" de A , ou, equivalentemente, os pontos cuja distância a A é 0.

Exemplo 2.3. (a) Seja $A = \{(x, y) \in \mathbb{R}^2 : x \geq 0\}$. Então $\text{int } A = \{(x, y) \in \mathbb{R}^2 : x > 0\}$, $\text{ext } A = \{(x, y) \in \mathbb{R}^2 : x < 0\}$, $\text{front } A = \{(x, y) \in \mathbb{R}^2 : x = 0\}$ e $\bar{A} = A$.
 (b) Seja $A = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 = 1\}$. Então $\text{int } A = \emptyset$, $\text{ext } A = A^c = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 \neq 1\}$, $\text{front } A = A$, $\bar{A} = A$.
 (c) Seja $A = \{(x, y, z) \in \mathbb{R}^3 : z > x^2 + y^2\}$. Então $\text{int } A = A$, $\text{ext } A = \{(x, y, z) \in \mathbb{R}^3 : z < x^2 + y^2\}$, $\text{front } A = \{(x, y, z) \in \mathbb{R}^3 : z = x^2 + y^2\}$, e $\bar{A} = \{(x, y, z) \in \mathbb{R}^3 : z \geq x^2 + y^2\}$.
 (d) Seja $A = \{(x, y) \in \mathbb{R}^2 : x \in \mathbb{Q}\}$. Então $\text{int } A = \text{ext } A = \emptyset$ e $\text{front } A = \bar{A} = \mathbb{R}^2$.

Definição 2.4. Um conjunto $A \subset \mathbb{R}^n$ diz-se

- (i) aberto se $A = \text{int } A$,
- (ii) fechado se $A = \bar{A}$ (equivalentemente, se A^c é aberto),
- (iii) limitado se A está contido nalguma bola aberta (equivalentemente, se a função distância à origem é limitada em A),
- (iv) compacto se A é limitado e fechado.

Como iremos ver, os conjuntos compactos desempenham um papel fundamental no Cálculo de várias variáveis (essencialmente pela mesma razão que o fazem também no Cálculo de uma variável).

Exemplo 2.5. Para os conjuntos do Exemplo 2.3 temos

- (a) A é fechado. Não é aberto. Não é limitado e portanto também não é compacto.
- (b) A é fechado. Não é aberto. É limitado e compacto.
- (c) A não é fechado, é aberto, não é limitado nem compacto.
- (d) A não é aberto, nem fechado, nem limitado, nem compacto.

Note-se que um conjunto pode ser aberto e fechado. É possível provar que os únicos subconjuntos de \mathbb{R}^n que são abertos e fechados são \mathbb{R}^n e \emptyset .

2.6. Sucessões em \mathbb{R}^n . Uma sucessão em \mathbb{R}^n é uma função $\mathbb{N} \rightarrow \mathbb{R}^n$ que a cada natural $k \in \mathbb{N}$ associa um vector $x_k \in \mathbb{R}^n$. É costume denotar uma tal função por (x_k) ou $(x_k)_{k \in \mathbb{N}}$. Geometricamente uma sucessão pode ser representada por um conjunto de pontos no espaço euclidiano \mathbb{R}^n . Escrevendo

$$x_k = (x_{1,k}, x_{2,k}, \dots, x_{n,k})$$

obtemos n sucessões reais $x_{i,k}$ que se designam por *sucessões coordenadas*.

Exemplo 2.7. A expressão $x_k = \left(\frac{1}{k}, e^k\right)$ define uma sucessão em \mathbb{R}^2 . A primeira sucessão coordenada é $x_{1,k} = \frac{1}{k}$ e a segunda sucessão coordenada é $x_{2,k} = e^k$.

Definição 2.8. Uma sucessão (x_k) em \mathbb{R}^n converge para $y \in \mathbb{R}^n$ se para todo o $\epsilon > 0$ existe um natural N tal que, para $k > N$ se tem $x_k \in B_\epsilon(y)$. Escreve-se então $\lim x_k = y$ ou $\lim_{k \rightarrow \infty} x_k = y$.

A definição anterior traduz de forma precisa a ideia que os vectores x_k se estão a aproximar de y à medida que k aumenta. Recomenda-se a verificação da seguinte equivalência:

$$\lim x_k = y \quad \Leftrightarrow \quad \|x_k - y\| \rightarrow 0$$

A convergência de uma sucessão em \mathbb{R}^n reduz-se assim à convergência para 0 da sucessão numérica $\|x_k - y\|$.

Exemplo 2.9. A sucessão do Exemplo 2.7 não converge.

3. SUCESSÕES EM \mathbb{R}^n . LIMITES DE FUNÇÕES

Proposição 3.1. Seja (x_k) uma sucessão em \mathbb{R}^n . Então a sucessão x_k converge sse as suas sucessões coordenadas convergem e $\lim x_k = (y_1, \dots, y_n) \in \mathbb{R}^n$ se e só se $\lim x_{i,k} = y_i$

Dem. Se $\lim x_k = y$ então $\|x_k - y\| \rightarrow 0$. Para cada $i = 1, \dots, n$ temos $|x_{i,k} - y_i| \leq \|x_k - y\|$ logo as sucessões coordenadas $x_{i,k}$ convergem para y_i . Reciprocamente, se $\lim x_{i,k} = y_i$ para cada i , então $|x_{i,k} - y_i| \rightarrow 0$ para cada i e portanto

$$\|x_k - y\| = \sqrt{(x_{1,k} - y_1)^2 + \dots + (x_{n,k} - y_n)^2} \rightarrow 0$$

logo x_k converge para y . □

Exemplo 3.2. (i) Vemos novamente que a sucessão do Exemplo 2.7 não converge uma vez que a sua segunda sucessão coordenada não converge.

(ii) Tem-se

$$\lim_{k \rightarrow \infty} \left(e^{\frac{1}{k}}, \left(1 - \frac{1}{k}\right)^k \right) = \left(1, \frac{1}{e}\right)$$

uma vez que $\lim e^{\frac{1}{k}} = e^0 = 1$ e $\lim \left(1 - \frac{1}{k}\right)^k = e^{-1} = \frac{1}{e}$.

Definição 3.3. Uma sucessão (x_k) em \mathbb{R}^n diz-se limitada se existe $L > 0$ tal que $\|x_k\| < L$ para todo o k . Equivalentemente (x_k) é limitada se o conjunto $\{x_k : k \in \mathbb{N}\} \subset \mathbb{R}^n$ é limitado.

Proposição 3.4. Uma sucessão (x_k) em \mathbb{R}^n é limitada sse as suas sucessões coordenadas são limitadas.

Dem. Exercício. □

Note-se que em consequência das Proposições 3.1 e 3.4, uma sucessão convergente é limitada: uma sucessão convergente tem sucessões coordenadas convergentes, portanto limitadas, e é portanto limitada. É também fácil deduzir este resultado directamente das definições de sucessão convergente e limitada (exercício).

Exemplo 3.5. (i) A sucessão $\left(\cos k, \sin k, \frac{k^2+k}{3k^2+5}\right)$ é limitada mas não convergente.

(ii) A sucessão $\left(\frac{\log k}{k+1}, \frac{k^2+3}{3k^2+2k}, \frac{k}{2^k}\right)$ converge para $(0, \frac{1}{3}, 0)$ e é portanto também limitada.

3.6. Limite de uma função num ponto. Chegamos agora a um dos conceitos fundamentais do Cálculo.

Definição 3.7. Seja $A \subset \mathbb{R}^n$, $f: A \rightarrow \mathbb{R}^m$ uma função, e $a \in \bar{A}$ um ponto do fecho de A . Diz-se que $f(x)$ tende para $y \in \mathbb{R}^m$ quando x tende para a , e escreve-se

$$\lim_{x \rightarrow a} f(x) = y$$

se para toda a sucessão (x_k) de termos em A que converge para a se tem que $\lim f(x_k) = y$.

Intuitivamente, f tem limite y em a se, à medida que x se aproxima de a , o valor $f(x)$ de f se aproxima de y . Para que x se possa aproximar de a de modo a que faça sentido calcular o valor de $f(x)$ é necessário que a esteja no fecho do domínio de f . É aliás fácil de demonstrar que um ponto pertence ao fecho de A sse existe uma sucessão de termos em A que converge para a (exercício).

É um bom exercício demonstrar a equivalência da definição anterior com a seguinte:

$$\lim_{x \rightarrow a} f(x) = y \Leftrightarrow \forall \epsilon > 0 \exists \delta > 0 \ x \in A \text{ e } \|x - a\| < \delta \Rightarrow \|f(x) - y\| < \epsilon$$

Exemplo 3.8. Seja $A = \mathbb{R}^2 \setminus \{(0, 0)\}$ e $f: A \rightarrow \mathbb{R}$ a função definida pela expressão

$$f(x, y) = \frac{x + y}{\sqrt{x^2 + y^2}}$$

Então

$$\lim_{(x,y) \rightarrow (1,1)} f(x, y) = \frac{2}{\sqrt{2}} = \sqrt{2}$$

De facto, se x_k e y_k são sucessões reais que convergem para 1, com $(x_k, y_k) \neq (0, 0)$ então

$$\lim f(x_k, y_k) = \lim \frac{x_k + y_k}{\sqrt{x_k^2 + y_k^2}} = \frac{1 + 1}{\sqrt{1 + 1}} = \sqrt{2}$$

pelas propriedades dos limites das sucessões reais.

Note-se que $(0, 0) \in \bar{A}$. Existirá o limite de $f(x, y)$ em $(0, 0)$? O método que usámos para o cálculo do limite anterior não se aplica, uma vez que ao tentar calcular o limite da sucessão $f(x_k, y_k)$ obtemos uma indeterminação que não sabemos resolver sem mais informação acerca da sucessão (x_k, y_k) .

Podemos tentar obter informação calculando alguns exemplos: Tomando por exemplo $(x_k, y_k) = (\frac{1}{k}, \frac{1}{k})$ temos

$$f(x_k, y_k) = \frac{\frac{2}{k}}{\sqrt{\frac{2}{k^2}}} = \sqrt{2}$$

que, sendo uma sucessão constante, converge para $\sqrt{2}$. Isto diz-nos que, se o limite $\lim_{(x,y) \rightarrow (0,0)} f(x,y)$ existir, terá de ser igual a $\sqrt{2}$. No entanto, calculando o limite de $f(x_k, y_k)$ com $(x_k, y_k) = (\frac{1}{k}, 0)$ obtemos

$$f(x_k, y_k) = \frac{\frac{1}{k}}{\frac{1}{k}} = 1 \rightarrow 1$$

Conclui-se então que f não tem limite em $(0, 0)$.

4. CONTINUIDADE

Definição 4.1. Seja $A \subset \mathbb{R}^n$, $f: A \rightarrow \mathbb{R}^m$ uma função e $a \in A$. A função f diz-se contínua em a se $\lim_{x \rightarrow a} f(x) = f(a)$. Se B é um subconjunto de A , f diz-se contínua em B se é contínua em todos os pontos de B .

Intuitivamente, uma função é contínua em a se “ $f(x)$ está próximo de $f(a)$ sempre que x esteja suficientemente próximo de a ”. Uma maneira alternativa de escrever a definição de continuidade de f num ponto é (exercício)

$$\forall \epsilon > 0 \exists \delta > 0 x \in A \text{ e } \|x - a\| < \delta \Rightarrow \|f(x) - f(a)\| < \epsilon.$$

Exemplo 4.2. (i) A função $f: \mathbb{R}^2 \rightarrow \mathbb{R}$ definida por $f(x, y) = 2x + 3y$ é contínua em todos os pontos do seu domínio. Por exemplo, para verificar continuidade no ponto $(2, 3)$ tomamos uma sucessão arbitrária (x_k, y_k) a convergir para $(2, 3)$ e calculamos

$$\lim f(x_k, y_k) = \lim 2x_k + 3y_k = 4 + 9 = 13 = f(2, 3).$$

(ii) A função $f: \mathbb{R}^2 \rightarrow \mathbb{R}$ definida por

$$f(x, y) = \begin{cases} 1 & \text{se } x > 0 \\ 0 & \text{se } x \leq 0 \end{cases}$$

é contínua nos pontos com coordenada $x \neq 0$ e é descontínua nos pontos do eixo dos yy . Por exemplo, considerando a sucessão $(\frac{(-1)^k}{k}, 0)$ (que converge para $(0, 0)$), temos

$$f\left(\frac{(-1)^k}{k}, 0\right) = \begin{cases} 0 & \text{se } k \text{ é ímpar,} \\ 1 & \text{se } k \text{ é par.} \end{cases}$$

que não converge. Isto mostra que f não é contínua em $(0, 0)$.

Definição 4.3. Seja $A \subset \mathbb{R}^n$ e $f: A \rightarrow \mathbb{R}^m$ uma função. Escrevendo

$$f(x_1, \dots, x_n) = (f_1(x_1, \dots, x_n), \dots, f_m(x_1, \dots, x_n)),$$

a função $f_i: A \rightarrow \mathbb{R}$ chama-se a i -ésima função coordenada de f .

Exemplo 4.4. A função $f: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ definida por $f(x, y) = (2x + e^y, \text{sen}(xy))$, tem primeira função coordenada $f_1: \mathbb{R}^2 \rightarrow \mathbb{R}$ definida por $f_1(x, y) = 2x + e^y$ e segunda função coordenada $f_2: \mathbb{R}^2 \rightarrow \mathbb{R}$ definida por $f_2(x, y) = \text{sen}(xy)$.

Proposição 4.5 (Propriedades das funções contínuas).

- (i) As funções constantes são contínuas.
- (ii) As funções lineares (isto é, as transformações lineares) são contínuas.
- (iii) Seja $A \subset \mathbb{R}^n$ e $f: A \rightarrow \mathbb{R}^m$. Então f é contínua em $a \in A$ se e só se as funções coordenadas $f_1, \dots, f_m: A \rightarrow \mathbb{R}$ são contínuas em a .
- (iv) Sejam $A \subset \mathbb{R}^n$ e $f, g: A \rightarrow \mathbb{R}$ funções contínuas em $a \in A$. Então $f + g, f \cdot g: A \rightarrow \mathbb{R}$ são contínuas em $a \in A$ e, se $g(a) \neq 0$, $\frac{f}{g}$ é contínua em $a \in A$.
- (v) A composta de funções contínuas é contínua. Isto é, sejam $A \subset \mathbb{R}^n$, $B \subset \mathbb{R}^m$, $f: A \rightarrow \mathbb{R}^m$ e $g: B \rightarrow \mathbb{R}^p$. Se f é contínua em $a \in A$, $f(a) \in B$ e g é contínua em $f(a)$, então $g \circ f$ é contínua em a .

Dem. As demonstrações são todas imediatas a partir da definição de limite e continuidade e ficam como exercício. Vejamos por exemplo a demonstração da última afirmação. Temos a mostrar que se x_k é uma sucessão com termos no domínio da composta $g \circ f$ e $x_k \rightarrow a$, então $g(f(x_k)) \rightarrow g(f(a))$. Mas claramente, x_k pertence ao domínio A de f , logo, sendo f contínua em a , temos $f(x_k) \rightarrow f(a)$. Uma vez que x_k pertence ao domínio da composta, $f(x_k)$ pertence ao domínio de g . Como g é contínua em $f(a)$, $g(f(x_k)) \rightarrow g(f(a))$. \square

As propriedades acima garantem a continuidade de funções cujas coordenadas são dadas por expressões algébricas envolvendo funções elementares no domínio dessas expressões.

Exemplo 4.6. A função f do Exemplo 4.4 é contínua no seu domínio \mathbb{R}^2 . De facto, a função $(x, y) \mapsto y$ é contínua, uma vez que é linear. A composta com a função seno, $(x, y) \mapsto \text{sen}(y)$ é também contínua pois é a composição de funções contínuas. Uma vez que a função $(x, y) \mapsto 2x$ é contínua por ser linear conclui-se que a soma das duas últimas funções $(x, y) \mapsto 2x + \text{sen } y$, que é a primeira função coordenada de f , é contínua. Analogamente se vê que a segunda função coordenada de f é contínua, portanto f é contínua.

4.7. Cálculo de limites. Ao contrário do que sucedia para funções de uma variável, não há qualquer regra para o cálculo de limites de funções de várias variáveis num ponto (a não ser, claro, que saibamos que a função é contínua nesse ponto - a definição de continuidade no ponto é então uma regra para o cálculo do limite). É no entanto possível obter informação sobre limites calculando limites ao longo de curvas em \mathbb{R}^n que passam pelo ponto em questão. Estes limites reduzem-se a limites de funções de uma variável que podem ser calculados com as técnicas aprendidas em Cálculo 1. Vejamos alguns exemplos.

Exemplo 4.8. (i) Seja $f: \mathbb{R}^2 \setminus \{(0, 0)\} \rightarrow \mathbb{R}^2$ a função definida por $f(x, y) = \frac{xy}{x^2 + y^2}$. É imediato calcular o limite de f em qualquer ponto do seu domínio. Por exemplo

$$\lim_{(x,y) \rightarrow (1,1)} f(x, y) = f(1, 1) = \frac{1}{2}$$

porque f é contínua em $(1, 1)$ (como segue da Proposição 4.5).

Para obter informação relativa ao limite de f no ponto $(0,0)$ podemos calcular os limites direcionais de f no ponto $(0,0)$, isto é o limite de $f(x,y)$ em $(0,0)$ quando (x,y) se aproxima da origem ao longo de uma recta. Considerando uma recta com declive m que passa pela origem temos

$$\lim_{\substack{(x,y) \rightarrow (0,0) \\ y=mx}} \frac{xy}{x^2 + y^2} = \lim_{x \rightarrow 0} \frac{mx^2}{x^2 + m^2x^2} = \lim_{x \rightarrow 0} \frac{m}{1 + m^2} = \frac{m}{1 + m^2}$$

Estes limites dependem da direção em que nos aproximamos da origem (por exemplo o limite é 0 ao longo do eixo dos xx mas é $\frac{1}{2}$ ao longo da recta $y = x$). Conclui-se que o limite de f na origem não existe.

(ii) Seja $f: \mathbb{R}^2 \rightarrow \mathbb{R}$ a função definida por $f(x,y) = \frac{x^3y^2}{x^6+y^4}$. A função f tem limite em $(0,0)$? Os limites direcionais são todos nulos como se verifica facilmente:

$$\lim_{\substack{(x,y) \rightarrow (0,0) \\ y=mx}} \frac{x^3y^2}{x^6 + y^4} = \lim_{x \rightarrow 0} \frac{m^2x^5}{x^6 + m^4x^4} = \lim_{x \rightarrow 0} \frac{m^2x}{x^2 + m^4} = 0$$

e uma vez que a $f(x,0) = 0$, o limite ao longo do eixo dos yy é também nulo. No entanto, a função não tem limite na origem! Se fizermos (x,y) aproximar-se de $(0,0)$ ao longo da curva $y = x^{\frac{3}{2}}$ temos

$$\lim_{\substack{(x,y) \rightarrow (0,0) \\ y=x^{\frac{3}{2}}}} \frac{x^3y^2}{x^6 + y^4} = \lim_{x \rightarrow 0} \frac{x^6}{x^6 + x^6} = \frac{1}{2}.$$

5. CÁLCULO DE LIMITES. O TEOREMA DE WEIERSTRASS

Vimos na secção anterior como obter informação sobre o limite de uma função num ponto calculando limites ao longo de curvas que passam num ponto. Este método permite mostrar que o limite não existe, ou calcular o valor do limite desde que este exista, mas não permite provar a existência do limite. Para verificar que $\lim_{x \rightarrow a} f(x) = b$ quando não se pode invocar a continuidade de f , tem necessariamente que se usar a definição, isto é mostrar que a $\|f(x) - b\|$ tende para 0 quando $\|x - a\|$ tende para 0. Usualmente isto faz-se comparando (através de manipulações algébricas, estimativas conhecidas de estudo prévio, etc.) a quantidade $\|f(x) - b\|$ com uma função contínua de $\|x - a\|$ que se anule em $x = a$. Vejamos alguns exemplos:

Exemplo 5.1. (i) Determinar $\lim_{(x,y) \rightarrow (0,0)} \frac{x^3+x^2y+y^3}{x^2+y^2}$ se este existir.

No numerador temos um polinómio de terceiro grau, que à partida deverá ser muito menor que o denominador (que é um polinómio do segundo grau). O limite deverá portanto ser 0 (e é imediato verificar que os limites direcionais em $(0,0)$ são nulos). Isto é fácil de verificar recorrendo às desigualdades

$$|x| \leq \sqrt{x^2 + y^2}, \quad |y| \leq \sqrt{x^2 + y^2}$$

De facto,

$$\left| \frac{x^3 + x^2y + y^3}{x^2 + y^2} \right| \leq \frac{|x^3 + x^2y + y^3|}{x^2 + y^2} \leq \frac{|x^3| + |x^2y| + |y^3|}{x^2 + y^2}$$

e uma vez que

$$|x^3| = |x|^3 \leq \left(\sqrt{x^2 + y^2}\right)^3, \quad |x^2y| = |x|^2|y| \leq \left(\sqrt{x^2 + y^2}\right)^3, \quad |y^3| \leq \left(\sqrt{x^2 + y^2}\right)^3$$

conclui-se que

$$\left| \frac{x^3 + x^2y + y^3}{x^2 + y^2} \right| \leq \frac{3 \left(\sqrt{x^2 + y^2}\right)^3}{x^2 + y^2} = 3\sqrt{x^2 + y^2}$$

e portanto, $\frac{x^3 + x^2y + y^3}{x^2 + y^2} \rightarrow 0$ quando $(x, y) \rightarrow (0, 0)$.

(ii) Determinar $\lim_{(x,y) \rightarrow (0,0)} \frac{x \operatorname{sen}^2 y}{x^2 + y^2}$ se este existir.

É imediato verificar que os limites direccionais são nulos. Para verificar que o limite é nulo temos que estimar o tamanho do numerador. Para isso é útil observar que $|\operatorname{sen} y| \leq |y|$ (isto é imediato a partir da interpretação geométrica da função seno, ou, alternativamente, pode ser demonstrado usando o Teorema de Lagrange: $\operatorname{sen} y - \operatorname{sen} 0 = \cos \xi(y - 0)$ para algum ponto ξ intermédio entre 0 e y , logo $|\operatorname{sen} y| = |\cos \xi||y| \leq |y|$). Assim

$$\left| \frac{x \operatorname{sen}^2 y}{x^2 + y^2} \right| = \frac{|x| |\operatorname{sen} y|^2}{x^2 + y^2} \leq \frac{|x| |y|^2}{x^2 + y^2} \leq \frac{\left(\sqrt{x^2 + y^2}\right)^3}{x^2 + y^2} = \sqrt{x^2 + y^2}$$

Concluimos que

$$\lim_{(x,y) \rightarrow (0,0)} \frac{x \operatorname{sen}^2 y}{x^2 + y^2} = 0.$$

5.2. O Teorema de Weierstrass. Vamos agora ver uma das propriedades fundamentais das funções contínuas (análogo ao resultado com o mesmo nome estudado em Cálculo 1).

Teorema 5.3 (Teorema de Weierstrass). *Seja $A \subset \mathbb{R}^n$ compacto e $f: A \rightarrow \mathbb{R}$ uma função contínua. Então f tem máximo e mínimo em A . Isto é, existem $a \in A$ e $b \in A$ tais que $f(a) \leq f(x)$ para todo o $x \in A$ e $f(b) \geq f(x)$ para todo o x em A .*

Recorde-se que o ponto a do enunciado se diz um *ponto de mínimo* enquanto que $f(a)$ é o mínimo de f em A . Analogamente b é um *ponto de máximo* e $f(b)$ é o máximo de f em A . A demonstração do Teorema de Weierstrass é uma consequência simples do seguinte resultado.

Proposição 5.4 (Bolzano-Weierstrass). *Seja $A \subset \mathbb{R}^n$ um conjunto compacto. Então toda a sucessão de termos em A tem uma subsucessão convergente com limite pertencente a A .*

Dem. Recorde-se de Cálculo 1, que toda a sucessão limitada em \mathbb{R} tem uma subsucessão convergente. Vamos explicar a demonstração da Proposição no caso $n = 2$. A generalização

ao caso quando $n > 2$ não oferece dificuldades e fica como exercício. Supomos portanto que A é um subconjunto compacto de \mathbb{R}^2 e que (x_k, y_k) é uma sucessão de termos em A .

Uma vez que A é limitado, as sucessões coordenadas x_k e y_k são limitadas. Podemos portanto escolher uma subsucessão convergente $(x_{k_m})_{m \in \mathbb{N}}$ de x_k . A sucessão $(y_{k_m})_{m \in \mathbb{N}}$ é também limitada e portanto tem uma subsucessão convergente $(y_{k_{m_p}})_{p \in \mathbb{N}}$. Mas então

$$(x_{k_{m_p}}, y_{k_{m_p}})_{p \in \mathbb{N}}$$

é uma subsucessão de (x_k, y_k) que é convergente pois, por construção, ambas as sucessões coordenadas são convergentes. Uma vez que A é fechado, o limite desta subsucessão pertence necessariamente a A (o limite não pode ser um ponto v exterior a A uma vez que v é o centro de uma bola de raio positivo que não intersecta A e portanto v está distante de todos os termos da sucessão). \square

Nota 5.5. *O resultado anterior pode ser formulado de várias maneiras essencialmente equivalentes:*

(a) *Se $A \subset \mathbb{R}^n$ é limitado, toda a sucessão de termos em A tem uma subsucessão convergente.*

(b) *Toda a sucessão limitada em \mathbb{R}^n tem uma subsucessão convergente.*

De facto a Proposição 5.4 implica a formulação (a) acima uma vez que o fecho de um conjunto limitado é compacto. Claramente (a) implica (b) pois uma sucessão é limitada sse o conjunto formado pelos seus termos é um conjunto limitado. Finalmente a formulação (b) juntamente com a observação que o limite de uma sucessão de termos em A pertence necessariamente a \bar{A} implicam a Proposição 5.4.

6. DERIVADAS SEGUNDO UM VECTOR

Dem. do Teorema de Weierstrass 5.3. Basta demonstrar a existência de mínimo. A demonstração no caso do máximo é inteiramente análoga (ou alternativamente segue do facto que $\max_{x \in A} \{f(x)\} = -\min_{x \in A} \{-f(x)\}$). Seja $\ell = \inf\{f(x) : x \in A\} \in [-\infty, \infty[$. Por definição de ínfimo, podemos escolher para cada natural k , um elemento $x_k \in A$ tal que

(i) $f(x_k) \leq \ell + \frac{1}{k}$, no caso em que $\ell > -\infty$ (isto é, no caso em que o conjunto $\{f(x) : x \in A\}$ é limitado inferiormente)

(ii) $f(x_k) \leq -k$, no caso em que $\ell = -\infty$.

A Proposição 5.4 garante que (x_k) tem uma subsucessão $(x_{k_m})_{m \in \mathbb{N}}$ que converge para um ponto $a \in A$. Uma vez que f é contínua em a , temos

$$\lim_{m \rightarrow \infty} f(x_{k_m}) = f(a)$$

Em particular o caso (ii) em que $\ell = -\infty$ não pode ocorrer pois então a sucessão $f(x_{k_m})$ não seria limitada. Conclui-se portanto que a imagem de A por f é um conjunto limitado inferiormente, ou seja, que $\ell \in \mathbb{R}$. Como

$$\ell \leq f(x_{k_m}) \leq \ell + \frac{1}{k_m}$$

a sucessão $f(x_{k_m})$ converge para ℓ . Como também converge para $f(a)$ temos que $f(a) = \ell$ é o mínimo de f em A , o que conclui a demonstração. \square

6.1. A derivada de uma função segundo um vector. Vamos agora iniciar o estudo do cálculo diferencial em \mathbb{R}^n com o conceito de derivada de uma função segundo um vector.

Definição 6.2. *Seja $A \subset \mathbb{R}^n$, $f: A \rightarrow \mathbb{R}$ e a um ponto interior a A . Dado $v \in \mathbb{R}^n$, defina-se a derivada de f em a segundo o vector v como sendo o limite*

$$D_v f(a) = \lim_{t \rightarrow 0} \frac{f(a + tv) - f(a)}{t} = \frac{d}{dt} (f(a + tv))|_{t=0}.$$

Na definição anterior, a exigência que o ponto a seja interior a A garante que, para t próximo de 0, o ponto $a + tv \in A$, de forma que a expressão da qual se está a tomar o limite tem sentido.

O significado da derivada segundo um vector v é o seguinte (para $v \neq 0$): o vector v determina uma recta com direcção v que passa por a , que é descrita pela equação paramétrica $a + tv, t \in \mathbb{R}$. Restringindo a função f a essa recta obtemos uma função da variável real t

$$g(t) = f(a + tv)$$

e, a derivada segundo v é então

$$D_v f(a) = g'(0),$$

a derivada da função g em $t = 0$ ($t = 0$ é o parâmetro que corresponde ao ponto a na recta). A derivada segundo v é portanto uma medida da taxa de crescimento da função f na direcção de v (ver no entanto a Nota 6.4 abaixo).

Exemplo 6.3. *Seja $f: \mathbb{R}^2 \rightarrow \mathbb{R}$ a função definida por $f(x, y) = x^2 + y^2$. Tomando por exemplo $a = (1, 1)$ e $v = (2, 3)$ temos*

$$\begin{aligned} D_{(2,3)} f(1, 1) &= \frac{d}{dt} (f(1 + 2t, 1 + 3t))|_{t=0} \\ &= \frac{d}{dt} ((1 + 2t)^2 + (1 + 3t)^2)|_{t=0} \\ &= (4(1 + 2t) + 6(1 + 3t))|_{t=0} = 10 \end{aligned}$$

Este resultado é consistente com o facto de, no ponto $a = (1, 1)$, a função f crescer na direcção do vector $v = (2, 3)$.

Nota 6.4. *A derivada segundo um vector v depende, não apenas da direcção de v mas também do seu comprimento. Se $\alpha \in \mathbb{R}$, é fácil verificar (exercício) que*

$$D_{\alpha v} f(a) = \alpha D_v f(a).$$

O caso em que o vector v é um dos vectores da base canónica de \mathbb{R}^n é particularmente importante e recebe um nome específico.

Definição 6.5. *Seja $A \subset \mathbb{R}^n$, $f: A \rightarrow \mathbb{R}$ e $a \in \text{int } A$. Para $i = 1, \dots, n$, a i -ésima derivada parcial de f no ponto a é $D_{e_i}f(a)$, onde $e_i = (0, \dots, 0, 1, 0, \dots, 0)$ é o i -ésimo vector da base canónica de \mathbb{R}^n . As notações habituais para a i -ésima derivada parcial são*

$$\frac{\partial f}{\partial x_i}(a) \quad \text{ou} \quad D_i f(a).$$

Note-se que substituindo na definição temos

$$\frac{\partial f}{\partial x_i}(x_1, \dots, x_n) = \lim_{t \rightarrow 0} \frac{f(x_1, \dots, x_{i-1}, x_i + t, x_{i+1}, \dots, x_n) - f(x_1, \dots, x_n)}{t}$$

Isto significa que a i -ésima derivada parcial num ponto (x_1, \dots, x_n) é exactamente a derivada no ponto x_i da função que se obtém a partir de f fixando todas as variáveis excepto x_i . Esta interpretação diz-nos como calcular derivadas parciais na prática: para calcular a i -ésima derivada parcial aplica-se as regras de derivação habituais encarando todas as variáveis excepto a i -ésima como constantes.

Exemplo 6.6. (i) *Se $f: \mathbb{R}^2 \rightarrow \mathbb{R}$ é definida por $f(x, y) = 2x + e^{xy}$ temos*

$$\frac{\partial f}{\partial x}(x, y) = 2 + ye^{xy} \quad \frac{\partial f}{\partial y}(x, y) = xe^{xy}$$

(ii) *Temos*

$$\frac{\partial}{\partial z} \left(x \log \left(\frac{y}{z} \right) \right) = x \frac{-\frac{y}{z^2}}{\frac{y}{z}} = -\frac{x}{z}$$

para todos os x, y, z no domínio da função que estamos a derivar.

7. DIFERENCIABILIDADE

A definição de derivada segundo um vector e de derivada parcial para funções com valores vectoriais é dada exactamente pela mesma fórmula que usámos anteriormente no caso das funções escalares. Se $A \subset \mathbb{R}^n$, a é um ponto interior a A e $f: A \rightarrow \mathbb{R}^m$ é uma função, define-se a derivada de f segundo o vector v no ponto a por

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

e a i -ésima derivada parcial por

$$D_i f(a) = \frac{\partial f}{\partial x_i}(a) \stackrel{\text{def}}{=} D_{e_i} f(a)$$

onde e_i denota o i -ésimo vector da base canónica de \mathbb{R}^n . Uma vez que os limites de funções vectoriais se calculam coordenada a coordenada, escrevendo

$$f = (f_1, \dots, f_m)$$

tem-se

$$D_v f(a) = (D_v f_1(a), \dots, D_v f_m(a)) \quad \frac{\partial f}{\partial x_i}(a) = \left(\frac{\partial f_1}{\partial x_i}(a), \dots, \frac{\partial f_m}{\partial x_i}(a) \right)$$

Exemplo 7.1. Sendo $f: \mathbb{R}^2 \rightarrow \mathbb{R}^3$ a função definida por $f(x, y) = (x^2 + y, xe^{xy}, \text{sen}(xy))$ tem-se

$$\frac{\partial f}{\partial x}(x, y) = (2x, e^{xy} + xye^{xy}, y \cos(xy))$$

7.2. A definição de derivada. Seja f uma função real de uma variável real. O significado geométrico da diferenciabilidade de f num ponto a do domínio é que o gráfico de f tem uma recta tangente no ponto $(a, f(a))$. A derivada é então o declive dessa recta tangente. Mais precisamente, a derivada de f em a é o (único) escalar m tal que a função afim

$$x \mapsto f(a) + m(x - a)$$

(cujo gráfico é a recta de declive m que passa pelo ponto $(a, f(a))$) aproxima muito bem a função $f(x)$ perto de a . O sentido preciso da última afirmação é que

$$\frac{f(x) - f(a) - m(x - a)}{x - a} = \frac{f(x) - f(a)}{x - a} - m$$

tende para 0 quando x tende para a . O numerador da fração à esquerda deve ser visto como o erro cometido ao aproximar a função $f(x)$ pela função afim $f(a) + m(x - a)$. A definição de derivada diz que esse erro é tão pequeno, que mesmo dividido pelo infinitésimo $x - a$, ainda tende para 0.

As funções lineares (ou mais geralmente, as funções afins) são as funções mais simples. Uma função é diferenciável se "não é muito diferente de uma função afim" perto do ponto em questão.

Definição 7.3. Seja $A \subset \mathbb{R}^n$, $f: A \rightarrow \mathbb{R}^m$ e $a \in \text{int } A$. A função f diz-se diferenciável em a se existe uma transformação linear $L: \mathbb{R}^n \rightarrow \mathbb{R}^m$ tal que

$$\lim_{x \rightarrow a} \frac{f(x) - f(a) - L(x - a)}{\|x - a\|} = 0$$

A transformação linear L chama-se a derivada de f em a e denota-se por $Df(a)$.

A definição anterior diz que f é diferenciável num ponto a se, perto desse ponto, pode ser "muito bem aproximada por uma função afim" no seguinte sentido preciso: se escrevermos

$$\epsilon(x) = f(x) - f(a) - L(x - a)$$

para o erro cometido ao aproximar f pela função afim $x \mapsto f(a) + L(x - a)$, a definição diz que o erro $\epsilon(x)$ é um infinitésimo de ordem superior a $\|x - a\|$, isto é, que tende para 0, mesmo dividido por esta quantidade (que também tende para 0). Vejamos algumas propriedades das funções diferenciáveis.

Proposição 7.4. Se f é diferenciável em a , então f é contínua em a .

Dem. Como o limite da expressão $\frac{f(x) - f(a) - L(x - a)}{\|x - a\|}$ em $x = a$ é nulo, o numerador tem também que ter limite nulo. A função $x \mapsto L(x - a)$ é afim, logo é contínua e portanto $\lim_{x \rightarrow a} L(x - a) = L(a - a) = L0 = 0$. Conclui-se então que $\lim_{x \rightarrow a} f(x) - f(a) = 0$, isto é, que f é contínua em a . \square

Proposição 7.5. *Seja $A \subset \mathbb{R}^n$, $f: A \rightarrow \mathbb{R}^m$ e $a \in \text{int } A$. Se f é diferenciável em a , então a derivada de f segundo o vector v no ponto a existe para todo o $v \in \mathbb{R}^n$ e tem-se*

$$D_v f(a) = Df(a)v.$$

Dem. Escrevendo

$$\epsilon(x) = f(x) - f(a) - Df(a)(x - a)$$

para o erro cometido ao aproximar $f(x)$ pela função afim determinada pela sua derivada, temos

$$f(x) = f(a) + Df(a)(x - a) + \epsilon(x), \quad \lim_{x \rightarrow a} \frac{\epsilon(x)}{x - a} = 0.$$

Dado $v \in \mathbb{R}^n$ temos

$$\frac{f(a + tv) - f(a)}{t} = \frac{f(a) + Df(a)(tv) + \epsilon(a + tv) - f(a)}{t} = Df(a)v + \frac{\epsilon(a + tv)}{t}$$

(onde usámos que $Df(a)(tv) = tDf(a)v$ uma vez que $Df(a)$ é uma transformação linear). Uma vez que (para $v \neq 0$) temos

$$\frac{\epsilon(a + tv)}{t} = \frac{\epsilon(a + tv)}{\|tv\|} \|v\| \rightarrow 0 \quad \text{quando } t \rightarrow 0$$

conclui-se que

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

□

Recorde-se de Álgebra Linear que a representação matricial de uma transformação linear L se obtém da seguinte forma: A i -ésima coluna da matriz é formada pelas coordenadas da imagem pela transformação linear do i -ésimo vector da base. Aplicando a Proposição 7.5 isso diz-nos que a i -ésima coluna da matriz que representa $Df(a)$ (relativamente às bases canónicas de \mathbb{R}^n e \mathbb{R}^m) é o vector $Df(a)e_i$, onde e_i é o i -ésimo vector da base canónica de \mathbb{R}^n , ou seja, a i -ésima derivada parcial de f em a .

Definição 7.6. *Seja $A \subset \mathbb{R}^n$, $a \in \text{int } A$ e $f: A \rightarrow \mathbb{R}^m$ uma função com derivadas parciais em a . A matriz Jacobiana de f em a é a matriz $m \times n$ que tem na i -ésima coluna a i -ésima derivada parcial de f em a . Escrevendo $f = (f_1, \dots, f_m)$ a matriz tem entrada ij dada por $\frac{\partial f_i}{\partial x_j}(a)$, isto é, a matriz tem a forma*

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

Corolário 7.7. *Seja $A \subset \mathbb{R}^n$, $a \in \text{int } A$ e $f: A \rightarrow \mathbb{R}^m$ uma função diferenciável em a . A matriz que representa $Df(a)$ relativamente às bases canónicas de \mathbb{R}^m e \mathbb{R}^n é a matriz Jacobiana de f em a .*

Exemplo 7.8. Seja $f: \mathbb{R}^2 \rightarrow \mathbb{R}^3$ a função definida por $f(x, y) = (x^2 + y, \sin(x + y), xe^{xy})$. A matriz Jacobiana de f num ponto (x, y) é dada por

$$\begin{bmatrix} 2x & 1 \\ \cos(x + y) & \cos(x + y) \\ e^{xy} + xy e^{xy} & x^2 e^{xy} \end{bmatrix}$$

8. FUNÇÕES DE CLASSE C^1

A definição de diferenciabilidade pode ser utilizada para demonstrar a diferenciabilidade de algumas funções, embora em geral este método não seja muito prático.

Proposição 8.1. (1) Se $f: \mathbb{R}^n \rightarrow \mathbb{R}^m$ é uma função constante, então f é diferenciável em \mathbb{R}^n e $Df(a) = 0$ para todo o $a \in \mathbb{R}^n$.
 (2) Se $f: \mathbb{R}^n \rightarrow \mathbb{R}^m$ é uma transformação linear, então f é diferenciável em \mathbb{R}^n e $Df(a) = f$ para todo o $a \in \mathbb{R}^n$.

Dem. (1) Sendo f constante, temos $\frac{f(x) - f(a) - 0(x - a)}{\|x - a\|} = \frac{0}{\|x - a\|} = 0$, que claramente tende para 0 quando $x \rightarrow a$. Isto mostra que $Df(a) = 0$.

(2) Se f é linear $f(x - a) = f(x) - f(a)$, logo

$$\frac{f(x) - f(a) - f(x - a)}{\|x - a\|} = \frac{0}{\|x - a\|} = 0$$

tende para 0 quando $x \rightarrow a$, o que mostra que f é a derivada de f em $x = a$, conforme pretendido. □

Uma consequência imediata da definição é que uma função com valores em \mathbb{R}^m é diferenciável num ponto sse cada uma das suas funções coordenadas é diferenciável nesse ponto.

Proposição 8.2. Seja $A \subset \mathbb{R}^n$, $a \in \text{int } A$ e $f: A \rightarrow \mathbb{R}^m$. Então f é diferenciável em a sse cada uma das suas funções coordenadas $f_i: A \rightarrow \mathbb{R}$ com $i = 1, \dots, m$ é diferenciável em a .

Dem. Tendo em conta o Corolário 7.7, f é diferenciável em a sse o quociente

$$\frac{\begin{bmatrix} f_1(x) \\ \vdots \\ f_m(x) \end{bmatrix} - \begin{bmatrix} f_1(a) \\ \vdots \\ f_m(a) \end{bmatrix} - \begin{bmatrix} \frac{\partial f_1}{\partial x_1}(a) & \cdots & \frac{\partial f_1}{\partial x_n}(a) \\ \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial x_1}(a) & \cdots & \frac{\partial f_m}{\partial x_n}(a) \end{bmatrix} \begin{bmatrix} x_1 - a_1 \\ \vdots \\ x_n - a_n \end{bmatrix}}{\|x - a\|}$$

tende para 0 quando $x \rightarrow a$. Uma vez que o limite se calcula coordenada a coordenada, este limite é nulo sse para cada $i = 1, \dots, m$ tivermos

$$\lim_{x \rightarrow a} \frac{f_i(x) - f_i(a) - \begin{bmatrix} \frac{\partial f_i}{\partial x_1}(a) & \cdots & \frac{\partial f_i}{\partial x_n}(a) \end{bmatrix} \begin{bmatrix} x_1 - a_1 \\ \vdots \\ x_n - a_n \end{bmatrix}}{\|x - a\|} = 0$$

ou seja, se cada uma das funções coordenadas f_i é diferenciável em a . □

A demonstração anterior chama a atenção para o significado das linhas da matriz Jacobiana: representam as derivadas das funções coordenadas de f .

Exemplo 8.3. Verificar se a função $f: \mathbb{R}^2 \rightarrow \mathbb{R}$ definida por

$$f(x, y) = \begin{cases} \frac{\cos(xy)-1}{x^2+y^2} & \text{se } (x, y) \neq (0, 0) \\ 0 & \text{se } (x, y) = (0, 0) \end{cases}$$

é diferenciável em $(0, 0)$.

Se f for diferenciável, a sua derivada é dada pela matriz $\left[\frac{\partial f}{\partial x}(0, 0) \quad \frac{\partial f}{\partial y}(0, 0) \right]$. Como $f(x, 0) = 0 = f(0, y)$ (isto é a função f anula-se ao longo dos eixos), é claro que as derivadas parciais na origem são nulas. Assim f é diferenciável em $(0, 0)$ sse

$$\lim_{(x,y) \rightarrow (0,0)} \frac{\frac{\cos(xy)-1}{x^2+y^2} - 0 - 0x - 0y}{\sqrt{x^2 + y^2}} = 0$$

Pela fórmula de Taylor com resto de 4ª ordem para a função coseno temos

$$\cos u = 1 - \frac{u^2}{2} + \cos(\xi) \frac{u^4}{4!} \quad \text{com } \xi \text{ entre } 0 \text{ e } u$$

Em particular, para u suficientemente pequeno (basta $u < 1$ por exemplo), temos

$$|\cos u - 1| = \left| -\frac{u^2}{2} + \cos(\xi) \frac{u^4}{4!} \right| \leq \frac{u^2}{2} + \frac{u^2}{2} = u^2$$

Segue-se que para (x, y) próximo de $(0, 0)$ temos

$$|\cos(xy) - 1| \leq x^2 y^2 \leq (x^2 + y^2)(x^2 + y^2)$$

e portanto

$$\left| \frac{\frac{\cos(xy)-1}{x^2+y^2} - 0 - 0x - 0y}{\sqrt{x^2 + y^2}} \right| = \left| \frac{\cos(xy) - 1}{(x^2 + y^2)^{\frac{3}{2}}} \right| \leq \sqrt{x^2 + y^2}$$

Conclui-se que f é diferenciável em $(0, 0)$ com derivada $\begin{bmatrix} 0 & 0 \end{bmatrix}$.

Não é prático ter que usar a definição para verificar a diferenciabilidade de uma função. Felizmente há um critério simples de diferenciabilidade que se pode aplicar na maioria dos exemplos.

Definição 8.4. Seja $A \subset \mathbb{R}^n$ um conjunto e $f: A \rightarrow \mathbb{R}^m$ uma função. f diz-se de classe C^1 no ponto $a \in \text{int } A$, se as derivadas parciais existem numa bola aberta centrada em a e são contínuas em a .

Teorema 8.5. Seja $A \subset \mathbb{R}^n$, $f: A \rightarrow \mathbb{R}^m$ e $a \in \text{int } A$. Se f é de classe C^1 em a , então f é diferenciável em a .

Exemplo 8.6. A função do Exemplo 8.3 é diferenciável em $\mathbb{R}^2 \setminus \{(0, 0)\}$ uma vez que é de classe C^1 nesse conjunto (pelas regras de derivação e as propriedades das funções contínuas que vimos).

9. A REGRA DE DERIVAÇÃO DA FUNÇÃO COMPOSTA

Dem. do Teorema 8.5. Pela Proposição 8.2 (e a definição de função de classe C^1) basta considerar o caso em que $m = 1$. Vamos fazer a demonstração no caso em que $n = 2$. O caso geral é inteiramente análogo e fica como exercício.

Seendo $f(x, y)$ de classe C^1 em (a, b) , temos a mostrar que

$$(1) \quad \lim_{(x,y) \rightarrow (a,b)} \frac{f(x, y) - f(a, b) - \frac{\partial f}{\partial x}(a, b)(x - a) - \frac{\partial f}{\partial y}(a, b)(y - b)}{\|(x - a, y - b)\|} = 0.$$

Escrevendo

$$f(x, y) - f(a, b) = f(x, y) - f(x, b) + f(x, b) - f(a, b)$$

podemos escrever a fração na expressão (1) como uma soma de duas parcelas

$$\frac{f(x, y) - f(x, b) - \frac{\partial f}{\partial y}(a, b)(y - b)}{\|(x - a, y - b)\|} + \frac{f(x, b) - f(a, b) - \frac{\partial f}{\partial x}(a, b)(x - a)}{\|(x - a, y - b)\|}$$

Quanto à segunda parcela temos (minorando o denominador)

$$\left| \frac{f(x, b) - f(a, b) - \frac{\partial f}{\partial x}(a, b)(x - a)}{\|(x - a, y - b)\|} \right| \leq \left| \frac{f(x, b) - f(a, b) - \frac{\partial f}{\partial x}(a, b)(x - a)}{x - a} \right| \rightarrow 0 \text{ quando } x \rightarrow a$$

por definição de derivada no ponto a da função $g(x) = f(x, b)$ (derivada esta que é exatamente a derivada parcial de f em ordem a x no ponto (a, b)).

Quanto à primeira parcela, aplicando o Teorema de Lagrange à função $y \mapsto f(x, y)$ (o que é possível porque estamos a assumir que esta função de y é diferenciável para (x, y) suficientemente perto de (a, b)) obtemos

$$f(x, y) - f(x, b) = \frac{\partial f}{\partial y}(x, \xi_{x,y})(y - b)$$

onde $\xi_{x,y}$ é um ponto entre b e y (a notação recorda que este ponto depende de x e de y). Isto diz-nos que

$$\left| \frac{f(x, y) - f(x, b) - \frac{\partial f}{\partial y}(a, b)(y - b)}{\|(x - a, y - b)\|} \right| = \left| \frac{\left(\frac{\partial f}{\partial y}(x, \xi_{x,y}) - \frac{\partial f}{\partial y}(a, b) \right) (y - b)}{\|(x - a, y - b)\|} \right| \leq \left| \frac{\partial f}{\partial y}(x, \xi_{x,y}) - \frac{\partial f}{\partial y}(a, b) \right|$$

A continuidade de $\frac{\partial f}{\partial y}(x, y)$ em (a, b) garante então que esta quantidade tende para 0 quando $(x, y) \rightarrow (a, b)$, o que conclui a demonstração. \square

Uma outra propriedade fundamental da derivada é a regra de derivação para a função composta.

Teorema 9.1. *Sejam $A \subset \mathbb{R}^n$, $f: A \rightarrow \mathbb{R}^m$, e $a \in \text{int } A$, $B \subset \mathbb{R}^m$ tal que $f(a) \in \text{int } B$ e $g: B \rightarrow \mathbb{R}^p$. Se f é diferenciável em a e g é diferenciável em $f(a)$, então a função composta $g \circ f$ é diferenciável em a e*

$$D(g \circ f)(a) = Dg(f(a)) \circ Df(a).$$

Pensando na derivada $Df(a)$ como uma aproximação linear de f perto de a , a regra de derivação torna-se clara: a função linear que aproxima a composta é a composição das aproximações lineares. Recorde-se de Álgebra Linear que a composição de transformações lineares se traduz no produto de matrizes. Assim, a matriz Jacobiana da função composta $g \circ f$ é o produto das matrizes Jacobianas de g e f nos pontos correspondentes.

10. A REGRA DA CADEIA

Exemplo 10.1. *Sejam $f: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ e $g: \mathbb{R}^2 \rightarrow \mathbb{R}^3$ as funções definidas por $f(x, y) = (xy, x + 2y)$ e $g(u, v) = (u^2, uv, v^2)$. Estas funções são diferenciáveis uma vez que são de classe C^1 . Vamos calcular a derivada da função composta $g \circ f$ no ponto $(1, 1)$. Temos*

$$Dg(u, v) = \begin{bmatrix} 2u & 0 \\ v & u \\ 0 & 2v \end{bmatrix} \quad Df(x, y) = \begin{bmatrix} y & x \\ 1 & 2 \end{bmatrix}$$

Uma vez que $f(1, 1) = (1, 3)$, a regra de derivação da função composta diz que

$$D(g \circ f)(1, 1) = Dg(1, 3) \circ Df(1, 1) = \begin{bmatrix} 2 & 0 \\ 3 & 1 \\ 0 & 6 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix} = \begin{bmatrix} 2 & 2 \\ 4 & 5 \\ 6 & 12 \end{bmatrix}$$

Podemos verificar este cálculo determinando a expressão da função composta $(g \circ f)(x, y)$ e derivando. Em geral esse método é menos eficiente que a aplicação da regra da função composta que efectivamente decompõe o cálculo em cálculos mais simples.

Note-se que se estivermos interessados apenas em uma das entradas da matriz Jacobiana de $g \circ f$, não é necessário calcular todas as entradas das matrizes Jacobianas de g e de f . De facto, a entrada ij da matriz $D(g \circ f)$ obtém-se multiplicando a linha i da matriz Dg pela coluna j da matriz Df (nos pontos correspondentes). A título de exemplo, para as funções acima, a entrada 22 da matriz Jacobiana é dada por

$$(2) \quad \frac{\partial(g \circ f)_2}{\partial y}(1, 1) = \frac{\partial g_2}{\partial u}(1, 3) \frac{\partial f_1}{\partial y}(1, 1) + \frac{\partial g_2}{\partial v}(1, 3) \frac{\partial f_2}{\partial y}(1, 1)$$

Esta regra para a determinação das derivadas parciais da função composta chama-se a *regra da cadeia*. É frequente escrevê-la da seguinte forma. Sendo

$$(u, v) = (f_1(x, y), f_2(x, y)) \quad \text{e} \quad (t, w, z) = (g_1(u, v), g_2(u, v), g_3(u, v))$$

a função composta $g \circ f$ expressa a dependência das variáveis (t, w, z) das variáveis (x, y) . A equação (2) expressa as derivadas parciais de (t, w, z) em ordem a (x, y) em termos de derivadas parciais envolvendo as variáveis intermédias (u, v) :

$$\frac{\partial w}{\partial y}(1, 1) = \frac{\partial w}{\partial u}(1, 3) \frac{\partial u}{\partial y}(1, 1) + \frac{\partial w}{\partial v}(1, 3) \frac{\partial v}{\partial y}(1, 1)$$

Por palavras, é necessário derivar w em ordem a cada uma das variáveis intermédias e multiplicar pela derivada dessa variável em ordem à variável em ordem à qual se está a

derivar. Enunciamos agora a regra geral que é uma consequência imediata do Teorema 9.1 e da regra para a multiplicação de matrizes.

Corolário 10.2 (Regra da cadeia). *Nas condições do Teorema 9.1, escrevendo $y = f(x)$ e $z = g(y)$, temos para cada $1 \leq i \leq p$ e $1 \leq j \leq n$*

$$\frac{\partial(g \circ f)_i}{\partial x_j}(a) = \frac{\partial g_i}{\partial y_1}(f(a)) \frac{\partial f_1}{\partial x_j}(a) + \frac{\partial g_i}{\partial y_2}(f(a)) \frac{\partial f_2}{\partial x_j}(a) + \dots + \frac{\partial g_i}{\partial y_m}(f(a)) \frac{\partial f_m}{\partial x_j}(a)$$

ou, equivalentemente,

$$\frac{\partial z_i}{\partial x_j}(a) = \frac{\partial z_i}{\partial y_1}(f(a)) \frac{\partial y_1}{\partial x_j}(a) + \frac{\partial z_i}{\partial y_2}(f(a)) \frac{\partial y_2}{\partial x_j}(a) + \dots + \frac{\partial z_i}{\partial y_m}(f(a)) \frac{\partial y_m}{\partial x_j}(a).$$

Exemplo 10.3. *Seja $g: \mathbb{R}^3 \rightarrow \mathbb{R}$ uma função diferenciável, vamos determinar uma expressão para a derivada parcial em ordem x da função $g(xy + x^2, y^3x, e^{xy})$ em termos das derivadas parciais de g :*

Escrevendo $g = g(u, v, w)$, a regra da cadeia diz que

$$\begin{aligned} & \frac{\partial}{\partial x} (g(xy + x^2, y^3x, e^{xy})) \\ &= \frac{\partial g}{\partial u}(xy + x^2, y^3x, e^{xy}) \frac{\partial u}{\partial x} + \frac{\partial g}{\partial v}(xy + x^2, y^3x, e^{xy}) \frac{\partial v}{\partial x} + \frac{\partial g}{\partial w}(xy + x^2, y^3x, e^{xy}) \frac{\partial w}{\partial x} \\ &= \frac{\partial g}{\partial u}(xy + x^2, y^3x, e^{xy}) (y + 2x) + \frac{\partial g}{\partial v}(xy + x^2, y^3x, e^{xy}) y^3 + \frac{\partial g}{\partial w}(xy + x^2, y^3x, e^{xy}) ye^{xy} \end{aligned}$$

O Teorema 9.1 tem como consequência que as operações algébricas habituais preservam a diferenciabilidade.

Proposição 10.4. *Seja $A \subset \mathbb{R}^n$. Se $f, g: A \rightarrow \mathbb{R}$ são diferenciáveis em $a \in \text{int } A$, então $f + g$ e fg são diferenciáveis em a . Se $g(a) \neq 0$, então f/g é diferenciável em a .*

Dem. Vejamos por exemplo o caso do quociente. A função quociente $f/g: A \rightarrow \mathbb{R}$ é a composta das funções $k: \{(x, y) \in \mathbb{R}^2: y \neq 0\} \rightarrow \mathbb{R}$ definida por $k(x, y) = x/y$ e $h: A \rightarrow \mathbb{R}^2$ definida por $h(u) = (f(u), g(u))$. A função h é diferenciável em a porque cada uma das suas funções coordenadas é diferenciável e a função k é diferenciável no seu domínio porque é de classe C^1 . O Teorema 9.1 garante então que f/g é diferenciável em a . \square

Dem. do Teorema 9.1. Precisamos de mostrar que o quociente

$$(3) \quad \frac{g(f(x)) - g(f(a)) - Dg(f(a))Df(a)(x - a)}{\|x - a\|}$$

tende para 0 quando x tende para a . Escrevendo

$$\epsilon_a(x) = f(x) - f(a) - Df(a)(x - a)$$

para o erro cometido ao aproximar $f(x)$ pela função afim determinada pela derivada em a , temos

$$(4) \quad Df(a)(x - a) = f(x) - f(a) - \epsilon_a(x)$$

e, por definição de derivada, temos

$$\lim_{x \rightarrow a} \frac{\epsilon_a(x)}{\|x - a\|} = 0$$

Substituindo (4) em (3) obtemos

$$\frac{g(f(x)) - g(f(a)) - Dg(f(a))(f(x) - f(a))}{\|x - a\|} + Dg(f(a)) \frac{\epsilon_a(x)}{\|x - a\|}$$

Uma vez que $Dg(f(a))$ é uma função contínua (é uma transformação linear) a segunda parcela tende para 0. Quanto à primeira parcela, podemos escrevê-la na forma

$$\frac{g(f(x)) - g(f(a)) - Dg(f(a))(f(x) - f(a))}{\|f(x) - f(a)\|} \frac{\|f(x) - f(a)\|}{\|x - a\|}$$

Como f é contínua em a (porque é diferenciável), temos que $f(x) \rightarrow f(a)$ quando $x \rightarrow a$, logo

$$\lim_{x \rightarrow a} \frac{g(f(x)) - g(f(a)) - Dg(f(a))(f(x) - f(a))}{\|f(x) - f(a)\|} = 0$$

por definição de derivada de g no ponto $f(a)$. Para concluir a demonstração resta observar que a diferenciabilidade de f em a garante que o quociente

$$\frac{\|f(x) - f(a)\|}{\|x - a\|}$$

é limitado numa vizinhança de a . De facto, substituindo (4) nesta expressão obtemos

$$Df(a) \frac{x - a}{\|x - a\|} + \frac{\epsilon_a(x)}{\|x - a\|}$$

A primeira parcela é claramente limitada (certamente as componentes do vector são menores ou iguais que a soma dos módulos das entradas da matriz que representa $Df(a)$) e a segunda parcela tende para 0 pelo que é também limitada. \square

11. APLICAÇÕES GEOMÉTRICAS DA DERIVADA

Se $f: \mathbb{R}^n \rightarrow \mathbb{R}$ é uma função diferenciável em a , a sua derivada é representada por uma matriz linha, que se pode portanto identificar com um vector de \mathbb{R}^n . O vector correspondente designa-se por *gradiente de f* no ponto a .

Definição 11.1. *Seja $A \subset \mathbb{R}^n$, $f: A \rightarrow \mathbb{R}$ e $a \in \text{int } A$. O gradiente de f em a é o vector*

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

Recorde-se que se f é diferenciável em a , temos a seguinte fórmula para a derivada de f em a segundo o vector v :

$$D_v f(a) = Df(a)v$$

Quando f é uma função escalar podemos re-escrever esta fórmula usando o gradiente:

$$D_v f(a) = Df(a)v = \frac{\partial f}{\partial x_1}(a)v_1 + \dots + \frac{\partial f}{\partial x_n}(a)v_n = \langle \nabla f(a), v \rangle$$

(onde $\langle \cdot, \cdot \rangle$ denota o produto interno usual de vectores de \mathbb{R}^n). Recorde-se que se v e w são vectores de \mathbb{R}^n e $\theta(v, w)$ é o ângulo entre v e w temos

$$\langle v, w \rangle = \|v\| \|w\| \cos \theta(v, w)$$

Assim

$$D_v f(a) = \|\nabla f(a)\| \|v\| \cos \theta$$

e portanto a derivada direcional de f segundo v é máxima quando v tem a mesma direcção que o gradiente. Além disso se v é um vector unitário com a mesma direcção que o gradiente, a taxa de crescimento segundo v é precisamente a norma do gradiente. Obtemos assim a seguinte *interpretação geométrica do gradiente*: $\nabla f(a)$ é um vector que aponta na direcção em que o valor de f cresce mais rapidamente a partir de a . O seu comprimento é a taxa de crescimento de f nessa direcção.

Exemplo 11.2. Consideremos a função $f: \mathbb{R}^2 \rightarrow \mathbb{R}$ definida por $f(x, y) = x^2 + y^2$. Temos

$$\nabla f(x, y) = (2x, 2y).$$

O gráfico de f é um parabolóide. O cálculo do gradiente diz-nos que em cada ponto $(x_0, y_0) \neq (0, 0)$ do plano a função f cresce mais rapidamente na direcção oposta à origem, e que a taxa de crescimento aumenta à medida que nos afastamos da origem.

Definição 11.3. Um conjunto de nível de uma função escalar $f: A \rightarrow \mathbb{R}$ (com $A \subset \mathbb{R}^n$) é um conjunto da forma

$$\{(x_1, \dots, x_n) \in A: f(x_1, \dots, x_n) = c\}$$

para algum $c \in \mathbb{R}$. É usual designar o conjunto acima por $f^{-1}(c)$.

Exemplo 11.4. Para a função $f(x, y) = x^2 + y^2$ do exemplo anterior, os conjuntos do nível são

$$f^{-1}(c) = \begin{cases} \emptyset & \text{se } c < 0 \\ \{0\} & \text{se } c = 0 \\ \{(x, y) \in \mathbb{R}^2: x^2 + y^2 = c\} \text{ (a circunferência de raio } \sqrt{c} \text{)} & \text{se } c > 0 \end{cases}$$

No exemplo anterior $\nabla f(x, y)$ é um vector perpendicular à linha de nível que passa por (x, y) . Intuitivamente é claro que isto deve acontecer. Caso contrário, a derivada segundo um vector tangente à linha de nível seria não nulo e isso impediria a função de ser constante ao longo da curva de nível. Iremos muito em breve formalizar esta intuição.

11.5. Caminhos em \mathbb{R}^n .

Definição 11.6. Um caminho em \mathbb{R}^n é uma função contínua $\gamma: I \rightarrow \mathbb{R}^n$ com I um intervalo de \mathbb{R} .

Intuitivamente podemos pensar que $\gamma(t)$ descreve a posição de uma partícula no espaço no instante t . A imagem do caminho γ é então a trajectória descrita pela partícula. Note-se no entanto que o caminho contém mais informação do que a trajectória uma vez que diz também como é que esta é descrita em função do parâmetro t .

Se um caminho γ é diferenciável em $t_0 \in I$, então

$$D\gamma(t_0) = \gamma'(t_0) = \lim_{t \rightarrow t_0} \frac{\gamma(t) - \gamma(t_0)}{t - t_0}$$

Intuitivamente, à medida que t se aproxima de t_0 , a direcção da razão incremental aproxima-se da direcção tangente à trajectória. Esta intuição é confirmada pela definição de derivada, que diz que o erro cometido ao aproximar a função $\gamma(t)$ pela função afim

$$\gamma(t_0) + \gamma'(t_0)(t - t_0)$$

em torno de $t = t_0$ é muito pequeno. O significado geométrico da derivada de um caminho num instante t_0 é portanto a de *um vector tangente à curva descrita por γ no ponto $\gamma(t_0)$* .

Pensando em $\gamma(t)$ como o vector de posição de uma partícula, $\gamma'(t_0)$ é, por definição, a velocidade instantânea da partícula em t_0 .

A interpretação geométrica da derivada de um caminho permite-nos achar equações de rectas tangentes a curvas assim como dos planos ortogonais a estas.

Exemplo 11.7. Seja $\gamma: \mathbb{R} \rightarrow \mathbb{R}^3$ o caminho definido por

$$\gamma(t) = (\cos t, \sin t, t).$$

γ descreve uma hélice em torno do eixo dos zz , que se projeta na circunferência de raio 1 no plano xy . Temos $\gamma'(t) = (-\sin t, \cos t, 1)$. A recta tangente à curva descrita por γ em $\gamma(0) = (1, 0, 0)$ é descrita pela equação paramétrica

$$\gamma(0) + t\gamma'(0) = (1, 0, 0) + t(0, 1, 1) = (1, t, t), \text{ com } t \in \mathbb{R}.$$

O plano perpendicular à curva descrita por γ em $(1, 0, 0)$ é o plano que passa por esse ponto e é perpendicular ao vector tangente $(0, 1, 1)$. É portanto dado pela equação

$$\langle (0, 1, 1), (x - 1, y - 0, z - 0) \rangle = 0 \Leftrightarrow y + z = 0.$$

12. APLICAÇÕES GEOMÉTRICAS DA DERIVADA (CONTINUAÇÃO). DERIVADAS PARCIAIS DE ORDEM SUPERIOR.

12.1. O gradiente de uma função é perpendicular aos conjuntos de nível dessa função. Com a noção de caminho introduzida acima podemos precisar a ideia de que o gradiente de uma função escalar de várias variáveis é perpendicular aos conjuntos de nível da função. Seja $f: \mathbb{R}^n \rightarrow \mathbb{R}$ uma função diferenciável e $\gamma: I \rightarrow \mathbb{R}^n$ um caminho

diferenciável com imagem contida no conjunto de nível c de f , $f^{-1}(c) = \{(x_1, \dots, x_n) \in \mathbb{R}^n : f(x_1, \dots, x_n) = c\}$. Então a função composta

$$t \mapsto f(\gamma(t))$$

é, por definição, constante igual a c . Derivando a expressão

$$f(\gamma(t)) = c$$

obtemos

$$\frac{d}{dt}(f(\gamma(t))) = 0$$

e aplicando a regra de derivação da função composta, a última equação pode escrever-se

$$Df(\gamma(t))\gamma'(t) = 0 \Leftrightarrow \langle \nabla f(\gamma(t)), \gamma'(t) \rangle = 0.$$

A equação anterior diz que ∇f é perpendicular à tangente a qualquer curva contida no conjunto de nível $f^{-1}(c)$. É esta a formalização da ideia que o gradiente é perpendicular aos conjuntos de nível.

12.2. Plano tangente e recta normal a uma superfície. Uma aplicação da ortogonalidade do gradiente aos conjuntos de nível é a determinação de planos tangentes e rectas normais a superfícies.

Exemplo 12.3. *Seja $C = \{(x, y, z) \in \mathbb{R}^3 : x^2 + y^2 = z^2\}$. Determinar o plano tangente e a recta normal no ponto $(1, 1, \sqrt{2})$.*

O conjunto C é um cone de duas folhas em torno do eixo Oz com uma abertura de 45 graus. De facto, considerando as coordenadas z e $r = \sqrt{x^2 + y^2}$ num qualquer semiplano vertical limitado pelo eixo dos zz , vemos que a intersecção de C com o semi-plano é definida pela equação $r^2 = z^2 \Leftrightarrow |z| = |r| \Leftrightarrow z = \pm r$. C é portanto a superfície que se obtém rodando as semi-retas $z = \pm r$ em torno do eixo dos zz .

C é por definição o conjunto de nível 0 da função $f: \mathbb{R}^3 \rightarrow \mathbb{R}$ definida por $f(x, y, z) = x^2 + y^2 - z^2$. Um vector perpendicular a C no ponto $(1, 1, \sqrt{2})$ é portanto dado por $\nabla f(1, 1, \sqrt{2})$. Como

$$\nabla f(x, y, z) = (2x, 2y, -2z)$$

temos que um vector normal a C em $(1, 1, \sqrt{2})$ é $(2, 2, -2\sqrt{2})$. Segue-se que a recta normal a C em $(1, 1, \sqrt{2})$ tem equação paramétrica

$$(1, 1, \sqrt{2}) + t(2, 2, -2\sqrt{2}) = (1 + 2t, 1 + 2t, \sqrt{2} - 2\sqrt{2}t), \quad t \in \mathbb{R}$$

e que o plano tangente é dado pela equação cartesiana

$$\langle (2, 2, -2\sqrt{2}), (x - 1, y - 1, z - \sqrt{2}) \rangle = 0 \Leftrightarrow x + y - \sqrt{2}z = 0.$$

Note-se que o resultado anterior é consistente com o facto de o plano tangente ao cone conter a recta $z = r$ no plano vertical definida pelo ponto $(1, 1, \sqrt{2})$ e o eixo dos zz (este plano passa claramente pela origem).

Alternativamente, o plano tangente a C pode ser obtido da forma seguinte: perto de $(1, 1, \sqrt{2})$, o conjunto C é o gráfico da função $g(x, y) = \sqrt{x^2 + y^2}$. O plano tangente a este

gráfico no ponto $(1, 1, g(1, 1))$ é o gráfico da melhor aproximação afim de $g(x, y)$ em torno de $(1, 1)$, que é determinada pela derivada $Dg(1, 1)$. Uma vez que

$$\frac{\partial g}{\partial x} = \frac{\partial}{\partial x} \left((x^2 + y^2)^{\frac{1}{2}} \right) = \frac{1}{2}(x^2 + y^2)^{-\frac{1}{2}} 2x = \frac{x}{\sqrt{x^2 + y^2}}$$

e portanto, por simetria,

$$\frac{\partial g}{\partial y} = \frac{y}{\sqrt{x^2 + y^2}}$$

temos que a aproximação linear a $g(x, y)$ junto a $(1, 1)$ é dada por

$$g(1, 1) + \frac{\partial g}{\partial x}(1, 1)(x - 1) + \frac{\partial g}{\partial y}(1, 1)(y - 1) = \sqrt{2} + \frac{1}{\sqrt{2}}(x - 1) + \frac{1}{\sqrt{2}}(y - 1) = \frac{1}{\sqrt{2}}x + \frac{1}{\sqrt{2}}y$$

O plano tangente é portanto dado pela equação

$$z = \frac{1}{\sqrt{2}}x + \frac{1}{\sqrt{2}}y$$

A partir desta equação podemos obter um vector perpendicular a C em $(1, 1, \sqrt{2})$ e assim recuperar uma equação para a recta perpendicular nesse ponto.

12.4. Derivadas parciais de ordem superior. Se $f: \mathbb{R}^n \rightarrow \mathbb{R}$ é uma função escalar e a derivada parcial $\frac{\partial f}{\partial x_i}$ existe em todos os pontos, ela define uma função escalar $\frac{\partial f}{\partial x_i}: \mathbb{R}^n \rightarrow \mathbb{R}$. Esta função pode por sua vez ter derivadas parciais, que se chamam derivadas parciais de segunda ordem de f . É usual escrever

$$\frac{\partial^2 f}{\partial x_j \partial x_i} \text{ para } \frac{\partial}{\partial x_j} \left(\frac{\partial f}{\partial x_i} \right) \text{ quando } i \neq j$$

e

$$\frac{\partial^2 f}{\partial x_i^2} \text{ para } \frac{\partial}{\partial x_i} \left(\frac{\partial f}{\partial x_i} \right)$$

e analogamente para derivadas parciais destas funções.

Exemplo 12.5. Seja $f: \mathbb{R}^2 \rightarrow \mathbb{R}$ a função definida por $f(x, y) = xe^{xy^2}$. Então

$$\frac{\partial^2 f}{\partial x \partial y} = \frac{\partial}{\partial x} \left(\frac{\partial f}{\partial y} \right) = \frac{\partial}{\partial x} \left(2x^2 ye^{xy^2} \right) = 4xye^{xy^2} + 2x^2 y^3 e^{xy^2}$$

13. O LEMA DE SCHWARZ

Definição 13.1. Seja f uma função escalar de n variáveis reais. Uma derivada parcial de ordem k de f é uma expressão da forma

$$\frac{\partial}{\partial x_{i_1}} \left(\frac{\partial}{\partial x_{i_2}} \left(\dots \frac{\partial f}{\partial x_{i_k}} \right) \right)$$

com i_1, \dots, i_k entre 1 e n . Abrevia-se a expressão anterior por

$$\frac{\partial^k f}{\partial x_{i_1} \dots \partial x_{i_k}}$$

Se houver derivações consecutivas em ordem à mesma variável substitui-se a expressão $\partial x_i \cdots \partial x_i$ no denominador por $(\partial x_i)^m$ (com m o número de ∂x_i 's).

Por exemplo escrevemos

$$\frac{\partial}{\partial z} \frac{\partial}{\partial y} \frac{\partial}{\partial y} \frac{\partial f}{\partial x} = \frac{\partial^4 f}{\partial z (\partial y)^2 \partial x}.$$

Exemplo 13.2. Seja $f: \mathbb{R}^2 \rightarrow \mathbb{R}$ a função definida por $f(x, y) = x^4 y^2 + xy^2$. Temos

$$\frac{\partial f}{\partial x} = 4x^3 y^2 + y^2 \quad \frac{\partial f}{\partial y} = 2x^4 y + 2xy$$

e

$$\frac{\partial^2 f}{\partial x^2} = 12x^2 y^2 \quad \frac{\partial^2 f}{\partial y \partial x} = 8x^3 y + 2y \quad \frac{\partial^2 f}{\partial x \partial y} = 8x^3 y + 2y \quad \frac{\partial^2 f}{\partial y^2} = 2x^4 + 2x$$

Temos ainda, por exemplo,

$$\frac{\partial^3 f}{\partial y \partial x^2} = 24x^2 y$$

A igualdade entre $\frac{\partial^2 f}{\partial x \partial y}$ e $\frac{\partial^2 f}{\partial y \partial x}$ no exemplo anterior não é coincidência como iremos agora ver.

Definição 13.3. Seja $A \subset \mathbb{R}^n$, $a \in \text{int } A$ e $f: A \rightarrow \mathbb{R}$ uma função. A função f diz-se de classe C^k em a (com $1 \leq k \leq \infty$) se todas as derivadas parciais de f até à ordem k estão definidas numa bola aberta contendo a e são contínuas em a .

É uma consequência das regras de derivação e das propriedades das funções contínuas que expressões algébricas elementares das variáveis compostas com funções elementares (como seno, exponencial, etc...) definem funções de classe C^k no seu domínio. Por exemplo a função do Exemplo 13.2 é de classe C^∞ .

Teorema 13.4 (Lema de Schwarz). Seja $A \subset \mathbb{R}^n$, $a \in \text{int } A$, e $f: A \rightarrow \mathbb{R}$ uma função de classe C^2 em a . Então para todos os i, j entre 1 e n , temos

$$\frac{\partial^2 f}{\partial x_i \partial x_j}(a) = \frac{\partial^2 f}{\partial x_j \partial x_i}(a)$$

Corolário 13.5. Se f é de classe C^k em a (com $k \geq 1$) então as derivadas parciais de f de ordem menor ou igual a k dependem apenas do número de vezes que f é derivada em ordem a cada uma das variáveis (e não da ordem pela qual a derivação é executada).

Este resultado demonstra-se por indução a partir do Lema de Schwarz. A ideia da demonstração é tornada clara pelo seguinte exemplo. Seja $f: \mathbb{R}^3 \rightarrow \mathbb{R}$ uma função de classe C^4 . Para verificar que

$$\frac{\partial^4 f}{\partial x \partial z \partial x \partial y} = \frac{\partial^4 f}{(\partial x)^2 \partial z \partial y}$$

basta notar que a função $\frac{\partial f}{\partial y}$ é de classe C^2 (em geral uma derivada parcial de ordem i de uma função de classe C^k é claramente de classe C^{k-i}) e que portanto

$$\frac{\partial}{\partial z} \frac{\partial}{\partial x} \left(\frac{\partial f}{\partial y} \right) = \frac{\partial}{\partial x} \frac{\partial}{\partial z} \left(\frac{\partial f}{\partial y} \right)$$

pelo Lema de Schwarz.

Dem. do Lema de Schwarz. Vamos fazer a demonstração apenas no caso de uma função de duas variáveis $f = f(x, y)$. A demonstração do caso geral é igual, mas com a notação um pouco mais complicada.

Seja $(a, b) \in \mathbb{R}^2$ um ponto do interior do domínio de f em que f é de classe C^2 . Por definição de derivadas parciais temos

$$\begin{aligned} \frac{\partial^2 f}{\partial x \partial y}(a, b) &= \lim_{h \rightarrow 0} \frac{\frac{\partial f}{\partial y}(a+h, b) - \frac{\partial f}{\partial y}(a, b)}{h} \\ &= \lim_{h \rightarrow 0} \left(\frac{\lim_{k \rightarrow 0} \frac{f(a+h, b+k) - f(a+h, b)}{k} - \frac{f(a, b+k) - f(a, b)}{k}}{h} \right) \\ &= \lim_{h \rightarrow 0} \left(\frac{\lim_{k \rightarrow 0} \frac{f(a+h, b+k) - f(a+h, b) - f(a, b+k) + f(a, b)}{hk}}{h} \right) \end{aligned}$$

e, de forma inteiramente análoga, vemos que

$$\frac{\partial^2 f}{\partial y \partial x}(a, b) = \lim_{k \rightarrow 0} \left(\frac{\lim_{h \rightarrow 0} \frac{f(a+h, b+k) - f(a+h, b) - f(a, b+k) + f(a, b)}{hk}}{k} \right)$$

Escrevendo

$$\Delta_{a,b}(h, k) = f(a+h, b+k) - f(a+h, b) - f(a, b+k) + f(a, b)$$

para o denominador das expressões acima, concluímos que as derivadas parciais cuja igualdade queremos demonstrar são os dois *limites iterados*

$$\lim_{h \rightarrow 0} \left(\lim_{k \rightarrow 0} \frac{\Delta_{a,b}(h, k)}{hk} \right) \quad \text{e} \quad \lim_{k \rightarrow 0} \left(\lim_{h \rightarrow 0} \frac{\Delta_{a,b}(h, k)}{hk} \right)$$

Ambos estes limites iterados são iguais aos limites de sucessões

$$\lim_{n \rightarrow \infty} \frac{\Delta_{a,b}(h_n, k_n)}{h_n k_n}$$

onde h_n, k_n são certas sucessões de termos em \mathbb{R}^+ que convergem para 0. Por exemplo, para determinar um par de sucessões (h_n, k_n) que calcula o limite iterado da esquerda podemos escolher para h_n uma sucessão arbitrária (de termos positivos a convergir para 0). Basta depois escolher, para cada n , um número $k_n > 0$ tão pequeno que

$$\left| \frac{\Delta_{a,b}(h_n, k_n)}{h_n k_n} - \frac{\frac{\partial f}{\partial y}(a+h_n, b) - \frac{\partial f}{\partial y}(a, b)}{h_n} \right| < \frac{1}{n}$$

o que é possível porque para cada n ,

$$\lim_{k \rightarrow 0} \frac{\Delta_{a,b}(h_n, k)}{h_n k} = \frac{\frac{\partial f}{\partial y}(a + h_n, b) - \frac{\partial f}{\partial y}(a, b)}{h_n}.$$

De facto, com essa escolha de k_n teremos

$$\lim_{n \rightarrow \infty} \frac{\Delta_{a,b}(h_n, k_n)}{h_n k_n} = \lim_{n \rightarrow \infty} \frac{\frac{\partial f}{\partial y}(a + h_n, b) - \frac{\partial f}{\partial y}(a, b)}{h_n} = \frac{\partial^2 f}{\partial x \partial y}(a, b).$$

Conclui-se que, para demonstrar o Lema de Schwarz, basta provar que o limite a duas variáveis

$$\lim_{(h,k) \rightarrow (0,0)} \frac{\Delta_{a,b}(h, k)}{hk}$$

existe. Isso garantirá a existência e igualdade de ambos os limites iterados. Para verificar a existência deste limite considere-se a função $\varphi: [0, 1] \rightarrow \mathbb{R}$ definida por

$$\phi(t) = f(a + th, b + k) - f(a + th, b)$$

e note-se que

$$\Delta_{a,b}(h, k) = \phi(1) - \phi(0).$$

Uma vez que f é de classe C^1 , a função ϕ é diferenciável e portanto, pelo Teorema de Lagrange, existe $t_0 \in [0, 1]$ tal que

$$\phi(1) - \phi(0) = \phi'(t_0) \cdot (1 - 0) = \phi'(t_0).$$

Mas

$$\phi'(t) = \left(\frac{\partial f}{\partial x}(a + th, b + k) - \frac{\partial f}{\partial x}(a + th, b) \right) h.$$

logo

$$\Delta_{a,b}(h, k) = \phi'(t_0) = \left(\frac{\partial f}{\partial x}(a + t_0 h, b + k) - \frac{\partial f}{\partial x}(a + t_0 h, b) \right) h$$

Uma vez que a função

$$y \mapsto \frac{\partial f}{\partial x}(a + t_0 h, y)$$

é diferenciável (porque f é C^2) o Teorema de Lagrange (e a definição de derivada parcial em ordem a y) garantem que existe $t_1 \in [0, 1]$ tal que

$$\frac{\partial f}{\partial x}(a + t_0 h, b + k) - \frac{\partial f}{\partial x}(a + t_0 h, b) = \frac{\partial^2 f}{\partial y \partial x}(a + t_0 h, b + t_1 k) k$$

e portanto

$$\frac{\Delta_{a,b}(h, k)}{hk} = \frac{\left(\frac{\partial^2 f}{\partial y \partial x}(a + t_0 h, b + t_1 k) k \right) h}{hk} = \frac{\partial^2 f}{\partial y \partial x}(a + t_0 h, b + t_1 k)$$

Como $\frac{\partial^2 f}{\partial y \partial x}$ é contínua, conclui-se que

$$\lim_{(h,k) \rightarrow (0,0)} \frac{\Delta_{a,b}(h, k)}{hk} = \lim_{(h,k) \rightarrow (0,0)} \frac{\partial^2 f}{\partial y \partial x}(a + t_0 h, b + t_1 k) = \frac{\partial^2 f}{\partial y \partial x}(a, b)$$

o que conclui a demonstração. \square

Nota 13.6. *Note-se que na demonstração do Lema de Schwarz não foi realmente utilizado que todas as derivadas parciais de segunda ordem existiam e eram contínuas. Analisando a demonstração vemos que para que $\frac{\partial^2 f}{\partial x_i \partial x_j}(a) = \frac{\partial^2 f}{\partial x_j \partial x_i}(a)$ basta que as derivadas parciais (de primeira ordem) em ordem a x_i e x_j existam junto a a e que uma delas tenha derivada parcial em ordem à outra variável, contínua em a . Segue-se então da demonstração que as duas derivadas parciais de segunda ordem existem e são iguais.*

14. A FÓRMULA DE TAYLOR

Recorde-se de Cálculo 1, que se $I \subset \mathbb{R}$ é um intervalo, a é um ponto interior a I e $f: I \rightarrow \mathbb{R}$ é uma função $(k+1)$ vezes diferenciável em I , temos a fórmula de Taylor de ordem k com resto de Lagrange:

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

onde ξ é um ponto entre a e $a+h$. Cada termo não nulo na expressão acima domina os restantes quando h se aproxima de 0. O polinómio de Taylor de ordem k é o polinómio de grau menor ou igual a k que melhor aproxima a função f perto de a . O erro cometido ao aproximar f pelo polinómio Taylor de ordem k é de tal forma pequeno que tende para 0, mesmo quando dividido por h^k , quando h tende para 0.

A fórmula de Taylor para funções de uma variável permite derivar uma fórmula análoga para funções de várias variáveis escalares. Os polinómios de Taylor são agora polinómios de várias variáveis.

Seja $U \subset \mathbb{R}^n$ um aberto e $a \in U$. Dada uma função $f: U \rightarrow \mathbb{R}$ e um vector $h \in \mathbb{R}^n$ consideramos a função

$$\phi(t) = f(a+th)$$

que se obtém compondo f com a função $t \mapsto a+th$. A função $\phi(t)$ é, essencialmente, a restrição da função f à recta que passa por a que tem a direcção do vector h . Uma vez que U é aberto, a função $\phi(t)$ está definida num intervalo aberto que contém $t=0$.

Assumindo que f é diferenciável e aplicando a regra da cadeia, obtemos a seguinte expressão para $\phi'(t)$:

$$\begin{aligned} \phi'(t) &= \frac{d}{dt}(f(a_1+th_1, \dots, a_n+th_n)) = \frac{\partial f}{\partial x_1}(a+th)h_1 + \dots + \frac{\partial f}{\partial x_n}(a+th)h_n \\ &= \sum_{i=1}^n \frac{\partial f}{\partial x_i}(a+th)h_i = \langle \nabla f(a+th), h \rangle \end{aligned}$$

Assumindo que f é de classe C^2 , podemos continuar a derivar esta expressão.

$$\phi''(t) = \frac{d}{dt} \left(\sum_{i=1}^n \frac{\partial f}{\partial x_i}(a+th)h_i \right)$$

Pela regra da cadeia temos

$$\frac{d}{dt} \left(\frac{\partial f}{\partial x_i}(a + th) \right) = \sum_{j=1}^n \frac{\partial^2 f}{\partial x_j \partial x_i}(a + th) h_j$$

logo

$$(5) \quad \phi''(t) = \sum_{i=1}^n \sum_{j=1}^n \frac{\partial^2 f}{\partial x_i \partial x_j}(a + th) h_i h_j.$$

Esta expressão pode ser escrita na forma matricial

$$\phi''(t) = \begin{bmatrix} h_1 & \cdots & h_n \end{bmatrix} \begin{bmatrix} \frac{\partial^2 f}{(\partial x_1)^2}(a + th) & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_n}(a + th) \\ \vdots & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x_n \partial x_1}(a + th) & \cdots & \frac{\partial^2 f}{(\partial x_n)^2}(a + th) \end{bmatrix} \begin{bmatrix} h_1 \\ \vdots \\ h_n \end{bmatrix} = h^T H f(a) h$$

A matriz na expressão acima chama-se a *matriz Hessiana de f* no ponto $a + th$.

Definição 14.1. A matriz Hessiana de uma função real de n variáveis no ponto a é a matriz $n \times n$

$$H f(a) = \left[\frac{\partial^2 f}{\partial x_i \partial x_j}(a) \right]$$

que tem por entrada ij a derivada parcial de segunda ordem $\frac{\partial^2 f}{\partial x_i \partial x_j}(a)$.

Nota 14.2. Note-se que o Lema de Schwarz garante que em qualquer ponto onde f seja uma função de classe C^2 , a matriz Hessiana de f é uma matriz simétrica (isto é, as entradas ij e ji coincidem).

A expressão (5) obtida acima para $\phi''(t)$ é um polinómio de segundo grau homogéneo nas componentes do vector h . Para concretizar a correspondência entre matrizes simétricas $n \times n$ e polinómios homogéneos de grau 2 em várias variáveis considere-se o seguinte exemplo.

Exemplo 14.3. Temos

$$f(x, y) = \begin{bmatrix} x & y \end{bmatrix} \begin{bmatrix} 2 & 3 \\ 3 & 4 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x & y \end{bmatrix} \begin{bmatrix} 2x + 3y \\ 3x + 4y \end{bmatrix} = 2x^2 + 3xy + 3xy + 4y^2 = 2x^2 + 6xy + 4y^2$$

Voltando ao cálculo das derivadas da função

$$\phi(t) = f(a + th)$$

fica como exercício verificar que, para f de classe C^k se tem

$$\phi^{(k)}(t) = \sum_{i_1=1}^n \cdots \sum_{i_k=1}^n \frac{\partial^k f}{\partial x_{i_1} \cdots \partial x_{i_k}}(a + th) h_{i_1} \cdots h_{i_k}$$

Teorema 14.4 (Fórmula de Taylor de ordem 2). *Seja $U \subset \mathbb{R}^n$ um aberto, $f: U \rightarrow \mathbb{R}$ uma função de classe C^3 e $a \in U$. Então para $h \in \mathbb{R}^n$ suficientemente próximo de 0, temos*

$$f(a+h) = f(a) + \langle \nabla f(a), h \rangle + \frac{1}{2} h^T Hf(a)h + R_2(h)$$

onde

$$\lim_{h \rightarrow 0} \frac{R_2(h)}{\|h\|^2} = 0.$$

Dem. Seja $h \in \mathbb{R}^n$ tal que o segmento de recta que une a a $a+h$ está inteiramente contido em U (isto acontece necessariamente se h estiver suficientemente próximo de 0 uma vez que U é aberto). A função $\phi(t) = f(a+th)$ está então definida e é três vezes diferenciável para $t \in [0, 1]$. A fórmula de Taylor de segunda ordem com resto de Lagrange para $\phi(t)$ diz-nos que

$$\phi(1) = \phi(0) + \phi'(0) + \frac{1}{2} \phi''(0) + \frac{1}{3!} \phi'''(\xi)$$

para algum $\xi \in [0, 1]$. Substituindo na expressão anterior as fórmulas para as derivadas de ϕ obtidas acima temos

$$f(a+h) = f(a) + \langle \nabla f(a), h \rangle + \frac{1}{2} h^T Hf(a)h + \frac{1}{3!} \sum_{i=1}^n \sum_{j=1}^n \sum_{k=1}^n \frac{\partial^3 f}{\partial x_i \partial x_j \partial x_k} (a + \xi h) h_i h_j h_k$$

Se $r > 0$ é tal que o compacto $K = \{x \in \mathbb{R}^n : \|x - a\| \leq r\}$ está contido em U , o Teorema de Weierstrass garante que o módulo de todas as derivadas parciais de terceira ordem de f é limitado em K . Conclui-se que existe uma constante $C > 0$ tal que, para todo $\|h\| \leq r$

$$R_2(h) = \frac{1}{3!} \sum_{i=1}^n \sum_{j=1}^n \sum_{k=1}^n \frac{\partial^3 f}{\partial x_i \partial x_j \partial x_k} (a + \xi h) h_i h_j h_k \leq C \|h\|^3$$

Isto conclui a demonstração. □

Exemplo 14.5. *Considere-se a função $f: \mathbb{R}^2 \rightarrow \mathbb{R}$ definida por $f(x, y) = e^{x+3y}$. Temos*

$$\nabla f(x, y) = (e^{x+3y}, 3e^{x+3y})$$

e

$$Hf(x, y) = \begin{bmatrix} \frac{\partial^2}{\partial x^2} & \frac{\partial^2 f}{\partial x \partial y} \\ \frac{\partial^2 f}{\partial y \partial x} & \frac{\partial^2}{\partial y^2} \end{bmatrix} = \begin{bmatrix} e^{x+3y} & 3e^{x+3y} \\ 3e^{x+3y} & 9e^{x+3y} \end{bmatrix}$$

Como

$$\nabla f(0, 0) = (1, 3) \quad e \quad Hf(0, 0) = \begin{bmatrix} 1 & 3 \\ 3 & 9 \end{bmatrix}$$

A fórmula de Taylor em $(0, 0)$ diz-nos que

$$f(x, y) = 1 + x + 3y + x^2 + 6xy + 9y^2 + R(x, y)$$

onde

$$\lim_{(x,y) \rightarrow (0,0)} \frac{R_2(x, y)}{x^2 + y^2} = 0$$

Nota 14.6. *Mostra-se facilmente que o polinómio de Taylor de ordem k de uma função escalar f de classe C^k é único. O coeficiente de um monómio nas componentes do acréscimo h expressa-se em termos de uma das derivadas parciais de ordem k de f (e, na realidade, fornece uma interpretação possível para essa derivada parcial). Esta unicidade pode ser explorada para obter desenvolvimentos de Taylor. Por exemplo, uma vez que*

$$e^u = 1 + u + u^2 + \dots$$

obtemos

$$e^{x+3y} = 1 + (x + 3y) + (x + 3y)^2 + \dots = 1 + x + 3y + x^2 + 6xy + 9y^2 + \dots$$

conforme a expressão obtida no exemplo anterior.

15. EXTREMOS

Vamos agora aplicar a fórmula de Taylor ao estudo de extremos de uma função de várias variáveis. Começamos por definir alguma terminologia. Dada uma função $f: A \rightarrow \mathbb{R}$ com $A \subset \mathbb{R}^n$, dizemos que $a \in A$ é um *ponto de máximo* (relativo) de f , se existe uma bola aberta B centrada em a tal que para todo o $x \in B$, se tem $f(x) \leq f(a)$. Se $f(x) \leq f(a)$ para todo o $x \in A$ o ponto a diz-se um *ponto de máximo absoluto*. O valor de f num ponto de máximo diz-se um *máximo* de f . Se $f(x) < f(a)$ para $x \neq a$, o máximo diz-se *estrito* (quer no caso relativo, quer no absoluto). Analogamente definimos mínimo e ponto de mínimo, relativo ou absoluto, estrito ou não. Finalmente um *extremo de f* significa um máximo ou um mínimo de f .

Recorde-se de Cálculo 1, que se I é um intervalo de \mathbb{R} e $f: I \rightarrow \mathbb{R}$ é diferenciável, para que $a \in \text{int } I$ seja um ponto de extremo é necessário que $f'(a) = 0$ (de facto se a derivada é não nula, têm de existir pontos arbitrariamente próximos de a nos quais $f(x) > f(a)$ e outros em que $f(x) < f(a)$). Isto continua a ser verdade para funções de várias variáveis, exactamente pela mesma razão.

Proposição 15.1. *Seja $A \subset \mathbb{R}^n$ e $f: A \rightarrow \mathbb{R}$ uma função diferenciável. Se $a \in \text{int } A$ é um ponto de extremo de f então $\nabla f(a) = 0$.*

Dem. Se $\nabla f(a) \neq 0$, a derivada de f segundo o vector $\nabla f(a)$ é positiva:

$$D_{\nabla f(a)} f(a) = \langle \nabla f(a), \nabla f(a) \rangle = \|\nabla f(a)\|^2 > 0$$

Consequentemente, para $t > 0$ suficientemente pequeno temos $f(a + t\nabla f(a)) > f(a)$ e $f(a - t\nabla f(a)) < f(a)$ o que mostra que a não é um ponto de extremo. \square

Definição 15.2. *Quando $\nabla f(a) = 0$, o ponto a diz-se um ponto de estacionariedade ou um ponto crítico de f .*

A Proposição 15.1 diz que um ponto de extremo de uma função diferenciável no interior do domínio é necessariamente um ponto de estacionariedade. Esta condição necessária pode por vezes ser usada juntamente com o Teorema de Weierstrass para achar os extremos de uma função de várias variáveis como é ilustrado no seguinte exemplo.

Exemplo 15.3. Seja $T = \{(x, y) \in \mathbb{R}^2: x \geq 0, y \geq 0, x + y \leq 1\}$ e $f: T \rightarrow \mathbb{R}$ a função definida por $f(x, y) = xy(1 - x - y)$. Determinar os extremos de f em T .

O conjunto T é compacto e a função f é contínua pelo que o Teorema de Weierstrass garante que f tem um máximo e um mínimo em T . Uma vez que f é diferenciável, os pontos de extremo podem ocorrer ou na fronteira de T ou em pontos de estacionariedade no interior. Na fronteira de T , a função f anula-se e no interior de T temos $f(x, y) > 0$ (uma vez que os três factores x , y e $1 - x - y$ são positivos no interior de T) logo o mínimo de f em T é 0, sendo todos os pontos da fronteira os pontos de mínimo. O máximo de f terá portanto que ocorrer no interior de T num ponto de estacionariedade.

Resta-nos determinar os pontos de estacionariedade no interior de T . Temos

$$\nabla f(x, y) = (y(1 - x - y) - xy, x(1 - x - y) - xy) = (y(1 - 2x - y), x(1 - x - 2y))$$

Os pontos de estacionariedade são as soluções de

$$\begin{cases} y(1 - 2x - y) = 0 \\ x(1 - x - 2y) = 0 \end{cases} \Leftrightarrow \begin{cases} y = 0 \text{ ou } 2x + y = 1 \\ x = 0 \text{ ou } x + 2y = 1 \end{cases}$$

A única solução do sistema anterior que não pertence à fronteira de T é o ponto $(x, y) = (\frac{1}{3}, \frac{1}{3})$. Conclui-se que este é o ponto onde f atinge o seu máximo absoluto, que é

$$f\left(\frac{1}{3}, \frac{1}{3}\right) = \frac{1}{27}.$$

Recorde-se de Cálculo 1, que a segunda derivada de uma função f de uma variável pode ser utilizada para decidir se um ponto a onde f' se anula é um ponto de máximo ou de mínimo: se $f''(a) > 0$, o ponto a é um ponto de mínimo (relativo estrito), enquanto que, se $f''(a) < 0$, a é um ponto de máximo (relativo estrito). Esta afirmação é uma consequência imediata da fórmula de Taylor de segunda ordem no ponto a que garante, nessas condições, que o gráfico de f junto a a se parece com uma parábola virada para baixo ou para cima respectivamente (com a correspondendo ao vértice). O mesmo raciocínio pode agora invocado para funções de várias variáveis. A única diferença é que o termo de segunda ordem na fórmula de Taylor é mais complicado (é uma forma quadrática) pelo que há mais situações a considerar.

Definição 15.4. Um ponto de estacionariedade diz-se um ponto de sela se não é um ponto de extremo.

O nome “ponto de sela” refere-se ao formato do gráfico de uma função que é um protótipo da situação em questão: a função $f: \mathbb{R}^2 \rightarrow \mathbb{R}$ definida por $f(x, y) = x^2 - y^2$. É imediato verificar que $(0, 0)$ é um ponto de estacionariedade de f que não é um ponto de máximo nem de mínimo (uma vez que as restrições de f ao eixo dos xx e dos yy têm respectivamente um mínimo e um máximo na origem). O aspecto do gráfico de f é o de uma sela (o cavalo estaria orientado segundo o eixo xx).

Proposição 15.5. Seja $U \subset \mathbb{R}^n$ um aberto e $f: U \rightarrow \mathbb{R}$ uma função de classe C^3 . Seja a um ponto de estacionariedade de f e $q(h) = h^T H f(a) h$ a forma quadrática definida pela matriz Hessiana de f no ponto a . Então

- (i) Se $q(h) > 0$ para todo o $h \neq 0$, então a é um ponto de mínimo relativo estrito.
- (ii) Se $q(h) < 0$ para todo o $h \neq 0$, então a é um ponto de máximo relativo estrito.
- (iii) Se existem $h_1, h_2 \in \mathbb{R}^n$ tais que $q(h_1) < 0$ e $q(h_2) > 0$, então a é um ponto de sela.
- (iv) Se a é um ponto de mínimo, então $q(h) \geq 0$ para todo o $h \in \mathbb{R}^n$.
- (v) Se a é um ponto de máximo, então $q(h) \leq 0$ para todo o $h \in \mathbb{R}^n$.

Nota 15.6. *Recorde-se de Álgebra Linear que nos casos (i) a (v) da Proposição anterior a forma quadrática $q(h)$ se diz respectivamente*

- (i) definida positiva e que isto é equivalente a dizer que todos os valores próprios da matriz $Hf(a)$ são positivos.
- (ii) definida negativa e que isto é equivalente a dizer que todos os valores próprios da matriz $Hf(a)$ são negativos.
- (iii) indefinida e que isto é equivalente a dizer que $Hf(a)$ tem valores próprios com ambos os sinais.
- (iv) semi-definida positiva e que isto é equivalente a dizer que todos os valores próprios de $Hf(a)$ são maiores ou iguais a 0.
- (v) semi-definida negativa e que isto é equivalente a dizer que todos os valores próprios de $Hf(a)$ são menores ou iguais a 0.

16. EXTREMOS (CONT.)

Dem. da Proposição 15.5. Para verificar as afirmações (iii)-(v), considere-se as funções $\phi(t) = f(a + th)$ onde $h \in \mathbb{R}^n$ é um vector fixo. Calculámos acima as derivadas de ϕ : temos $\phi'(0) = \langle \nabla f(a), h \rangle = 0$ (uma vez que $\nabla f(a) = 0$ por hipótese) e (cf. (5))

$$\phi''(0) = h^T Hf(a)h = q(h).$$

Se existem h_1 e h_2 tais que $q(h_1) > 0$ e $q(h_2) < 0$, então a função $t \mapsto f(a + th_1)$ tem um mínimo em $t = 0$ enquanto que a função $t \mapsto f(a + th_2)$ tem um máximo em $t = 0$. Conclui-se que f não tem um extremo em a , ou seja, que a é um ponto de sela.

Por outro lado, se f tem um máximo em a , então a função $\phi(t) = f(a + th)$ tem necessariamente um máximo em 0 para todo o $h \in \mathbb{R}^n$. Portanto $\phi''(0) = q(h) \leq 0$ para todo o $h \in \mathbb{R}^n$. Analogamente se f tem um mínimo em a , temos $q(h) \geq 0$ para todo o $h \in \mathbb{R}^n$.

Resta-nos demonstrar a afirmação (i) (uma vez que (ii) é inteiramente análoga a esta). Suponhamos então que $q(h) > 0$ para todo o $h \neq 0$. Recorde-se de Álgebra Linear que, sendo simétrica, a matriz $Hf(a)$ tem apenas valores próprios reais e é diagonalizável por meio de uma matriz de mudança de base S ortogonal (isto significa, recorde-se, que as colunas de S formam uma base ortonormal de \mathbb{R}^n , ou seja que $S^t S = \text{Id}$). As colunas de S são vectores próprios de $Hf(a)$. Sendo $\lambda_1, \dots, \lambda_n$ os valores próprios de $Hf(a)$ e escrevendo $\Lambda = \text{diag}(\lambda_1, \dots, \lambda_n)$ para a matriz diagonal com λ_i na diagonal, então

$$Hf(a) = S\Lambda S^{-1} = S\Lambda S^T$$

e portanto $q(h) = (S^T h)^T \Lambda (S^T h)$. Claramente $q(h) > 0$ para $h \neq 0$ sse todos os valores próprios λ_i de $Hf(a)$ são positivos. Seja μ o menor destes valores próprios. Como S é

ortogonal, $\|S^T h\| = \|h\|$ e verifica-se então facilmente (exercício) que

$$q(h) \geq \mu \|h\|^2$$

Como a é um ponto de estacionariedade de f , a fórmula de Taylor de segunda ordem em a é dada por

$$f(a+h) = f(a) + \langle \nabla f(a), h \rangle + \frac{1}{2}q(h) + R_2(h) = f(a) + 0 + \frac{1}{2}q(h) + R_2(h)$$

Uma vez que

$$\lim_{h \rightarrow 0} \frac{R_2(h)}{\|h\|^2} = 0$$

segue-se que $f(a+h) \geq f(a) + 0.99999\mu\|h\|^2 > f(a)$ para $h \neq 0$ suficientemente pequeno, o que mostra que a é um ponto de mínimo relativo estrito. \square

Exemplo 16.1. *Determinar e classificar os pontos de estacionariedade das seguintes funções.*

(i) $f: \mathbb{R}^2 \rightarrow \mathbb{R}$ definida por $f(x, y) = x^2 + xy + y^2 + x - y$. Os pontos de estacionariedade de f são as soluções do sistema

$$\begin{cases} \frac{\partial f}{\partial x} = 2x + y + 1 = 0 \\ \frac{\partial f}{\partial y} = x + 2y - 1 = 0 \end{cases} \Leftrightarrow \begin{cases} x = -1 \\ y = 1 \end{cases}$$

O único ponto de estacionariedade de f é portanto $(-1, 1)$. A matriz Hessiana de f é dada por

$$Hf(x, y) = \begin{bmatrix} \frac{\partial^2 f}{\partial x^2} & \frac{\partial^2 f}{\partial x \partial y} \\ \frac{\partial^2 f}{\partial y \partial x} & \frac{\partial^2 f}{\partial y^2} \end{bmatrix} = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$$

(neste caso esta matriz é independente do ponto (x, y)). É fácil determinar os valores próprios desta matriz A resolvendo a equação $\det(A - \lambda \text{Id}) = 0$. Obtemos $\lambda_1 = 1$ e $\lambda_2 = 3$, o que mostra que $(-1, 1)$ é um ponto de mínimo (relativo, estrito) de f .

Note-se que para matrizes 2×2 , não é necessário achar os valores próprios quando queremos apenas determinar o seu sinal. De facto, se os valores próprios são λ_1, λ_2 então o traço da matriz é $\text{tr } A = \lambda_1 + \lambda_2$ e o determinante é $\det A = \lambda_1 \lambda_2$. Isto permite determinar o sinal dos valores próprios de A sem os calcular: se $\det A < 0$, os valores próprios têm sinais opostos. Se $\det A > 0$ então os valores próprios têm o mesmo sinal que é o sinal de $\text{tr } A$. No exemplo que estamos a tratar temos $\det Hf(-1, 1) = 3 > 0$ e $\text{tr } Hf(-1, 1) = 4$ donde concluímos novamente que ambos os valores próprios de $Hf(-1, 1)$ são positivos.

(ii) $f: \mathbb{R}^3 \rightarrow \mathbb{R}$ a função definida por $f(x, y, z) = x^2 - yz + z^2$. Temos

$$\nabla f(x, y, z) = (2x, -z, -y + 2z)$$

pelo que $(0, 0, 0)$ é o único ponto de estacionariedade. A matriz Hessiana de f é dada por

$$Hf(x, y, z) = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & -1 & 2 \end{bmatrix}$$

(novamente a matriz Hessiana não depende de (x, y, z)). Os valores próprios são as soluções da equação

$$\det(Hf(0, 0, 0) - \lambda \text{Id}) = 0 \Leftrightarrow (2 - \lambda)(-\lambda(2 - \lambda) - 1) = 0$$

que são

$$\lambda = 2 \text{ ou } \lambda = 1 \pm \sqrt{2}$$

Uma vez que há valores próprios com ambos os sinais, concluímos que $(0, 0, 0)$ é um ponto de sela.

(iii) $f: \mathbb{R}^2 \rightarrow \mathbb{R}$ definida por $f(x, y) = x^2 - y^6$. O único ponto de estacionariedade de f é $(0, 0)$ e temos

$$Hf(x, y) = \begin{bmatrix} 2 & 0 \\ 0 & -30y^4 \end{bmatrix}$$

e portanto

$$Hf(0, 0) = \begin{bmatrix} 2 & 0 \\ 0 & 0 \end{bmatrix}$$

Esta matriz tem valores próprios 2 e 0 e é portanto semi-definida positiva. Usando a matriz Hessiana podemos apenas concluir que $(0, 0)$ se trata de um ponto de sela ou de um ponto de mínimo (não pode ser um ponto de máximo porque a função f tem um mínimo ao longo da direção própria do valor próprio 2). Ao longo da direção própria do valor próprio 0 (que é o eixo dos yy), a segunda derivada é nula pelo que não temos informação suficiente. Podemos no entanto estudar a função ao longo desta direção: a expressão

$$f(0, y) = -y^6$$

mostra que a restrição de f ao eixo dos yy tem um máximo na origem, pelo que $(0, 0)$ é um ponto de sela.

Alternativamente podemos estudar o sinal de f perto de $(0, 0)$. De facto f terá um mínimo em $(0, 0)$ sse $f(x, y) - f(0, 0) = f(x, y)$ for positivo para (x, y) suficientemente próximo de $(0, 0)$. Uma vez que

$$f(x, y) = x^2 - y^6 = (x - y^3)(x + y^3)$$

analisando os sinais dos dois factores vemos que f é positiva à direita da curva $x = |y|^3$ e à esquerda da curva $x = -|y|^3$ e que f é negativa entre estas duas curvas. Segue-se que há pontos arbitrariamente próximos de $(0, 0)$ em que f toma valores maiores e menores que $f(0, 0) = 0$ logo $(0, 0)$ não é um ponto de extremo.

17. DEFINIÇÃO DE INTEGRAL

Recorde-se de Cálculo 1, que se $I = [a, b] \subset \mathbb{R}$ é um intervalo e $f: I \rightarrow \mathbb{R}$ é uma função, o integral de f em I ,

$$\int_a^b f(x)dx$$

calcula a “área entre o gráfico de f e o eixo dos xx ”. As aspas referem-se ao facto que, na realidade, a área é contada positivamente quando f é positiva e negativamente quando $f < 0$. O significado do integral de uma função escalar de várias variáveis é análogo.

Definição 17.1. Um intervalo em \mathbb{R}^n é um produto cartesiano de intervalos de \mathbb{R} .

Assim, um intervalo em \mathbb{R} é um intervalo usual, um intervalo em \mathbb{R}^2 é um rectângulo e um intervalo em \mathbb{R}^3 é um paralelepípedo.

Se $I \subset \mathbb{R}^2$ é um intervalo limitado e $f: I \rightarrow \mathbb{R}$ é uma função limitada, queremos definir $\int_I f$ de forma a que o integral calcule o “volume debaixo do gráfico de f ” (contado com sinais como no caso de uma variável). Se a função depende de mais de duas variáveis, a ideia é a mesma mas o volume será agora um “volume $(n + 1)$ -dimensional” da região compreendida entre o gráfico da função (que é um subconjunto de \mathbb{R}^{n+1}) e o intervalo $I \subset \mathbb{R}^n \times \{0\}$.

Para definir rigorosamente o integral usamos a mesma técnica que em Cálculo 1. É fácil definir o integral para funções em escada (funções que são constantes nos subintervalos de uma partição do domínio) e podemos usar os integrais destas funções para aproximar o valor do integral.

Definição 17.2. Uma partição de um intervalo $I = [a_1, b_1] \times \cdots \times [a_n, b_n] \subset \mathbb{R}^n$ é uma subdivisão de I em subintervalos determinada por partições dos factores $[a_i, b_i]$. Uma função $s: I \rightarrow \mathbb{R}$ diz-se uma função em escada se é limitada e constante no interior de cada um dos intervalos de alguma partição de I .

Definição 17.3. O volume n -dimensional (ou simplesmente volume) do intervalo $I = [a_1, b_1] \times \cdots \times [a_n, b_n]$ é definido por

$$\text{vol}(I) = (b_1 - a_1) \cdots (b_n - a_n)$$

Note-se que a definição anterior é consistente com a nossa intuição: o volume 1-dimensional dá o comprimento do intervalo, o volume 2-dimensional dá a área do rectângulo e o volume 3-dimensional dá o volume usual do paralelepípedo.

Definição 17.4. Seja $I \subset \mathbb{R}^n$ um intervalo limitado e $s: I \rightarrow \mathbb{R}$ uma função em escada. Seja $\{I_j\}$ o conjunto dos subintervalos de uma partição de I tal que s é constante igual a s_j em $\text{int } I_j$. Define-se o integral de s em I por

$$\int_I s = \sum_j s_j \text{vol}(I_j).$$

A definição anterior é compatível com a ideia do integral como “volume debaixo do gráfico”. De facto o “volume debaixo do gráfico” de uma função em escada s é o volume de uma união de intervalos de \mathbb{R}^{n+1} que têm por base os interiores dos intervalos I_j e por “altura” o valor s_j que s assume nesses pontos. Este ‘volume com sinal’ é exactamente $s_j \text{vol}(I_j)$. Note-se que uma função em escada pode tomar valores arbitrários na fronteira dos intervalos de uma partição do domínio. Esses valores não têm qualquer influência no integral, o que é consistente com a ideia que a união das fronteiras dos intervalos de uma partição devem ter volume n -dimensional nulo.

Exemplo 17.5. *Seja $s: [0, 1] \times [0, 2] \rightarrow \mathbb{R}$ a função definida por*

$$s(x, y) = \begin{cases} 1 & \text{se } x \leq \frac{1}{2} \\ 0 & \text{se } x > \frac{1}{2} \text{ e } y < 1 \\ -3 & \text{se } x > \frac{1}{2} \text{ e } y \geq 1 \end{cases}$$

Então s é uma função em escada, uma vez que é constante no interior dos intervalos

$$[0, \frac{1}{2}] \times [0, 1], \quad [0, \frac{1}{2}] \times [1, 2], \quad [\frac{1}{2}, 1] \times [0, 1], \quad [\frac{1}{2}, 1] \times [1, 2]$$

que formam uma partição de $[0, 1] \times [0, 2]$. O integral de s é dado por

$$\begin{aligned} \int_I s &= 1 \cdot \text{vol}([0, \frac{1}{2}] \times [0, 1]) + 1 \cdot \text{vol}([0, \frac{1}{2}] \times [1, 2]) + \\ &\quad + 0 \cdot \text{vol}([\frac{1}{2}, 1] \times [0, 1]) + (-3) \cdot \text{vol}([\frac{1}{2}, 1] \times [1, 2]) \\ &= 1 \cdot \frac{1}{2} \cdot 1 + 1 \cdot \frac{1}{2} \cdot 1 + 0 \cdot \frac{1}{2} \cdot 1 - 2 \cdot \frac{1}{2} \cdot 1 \\ &= -\frac{1}{2} \end{aligned}$$

Em rigor, é necessário verificar que a Definição 17.4 faz sentido. De facto, se $s: I \rightarrow \mathbb{R}$ é uma função em escada, há mais do que uma partição do domínio I que satisfaz a condição na definição e é necessário verificar que o valor obtido para o integral de s não depende da escolha de partição sujeita a essas condições. Não é difícil levar a cabo esta verificação notando que, dadas quaisquer duas partições P_1 e P_2 de um intervalo I , existe uma partição P que refina P_1 e P_2 , no sentido em que, qualquer subintervalo de uma das partições P_i é uma união de subintervalos de P . A conclusão do argumento fica como exercício para os alunos interessados.

Podemos agora definir integral de forma inteiramente análoga ao que foi feito em Cálculo 1 para funções de uma variável.

Definição 17.6. *Seja $I \subset \mathbb{R}^n$ um intervalo limitado e $f: I \rightarrow \mathbb{R}$ uma função limitada. O integral inferior de f é definido por*

$$\underline{\int}_I f = \sup \left\{ \int_I s : s \text{ é uma função em escada e } s \leq f \right\}$$

O integral superior de f é definido por

$$\overline{\int}_I f = \inf \left\{ \int_I t : t \text{ é uma função em escada e } t \leq f \right\}$$

A função f diz-se integrável se os integrais superior e inferior coincidem e, nesse caso, define-se o integral de f por

$$\int_I f = \underline{\int}_I f = \overline{\int}_I f$$

O integral de uma função f é portanto o valor comum das aproximações ao volume debaixo do gráfico de f que se obtêm considerando funções em escada por cima e por baixo de f .

Põe-se agora a questão de como calcular o integral de uma função. Para funções de uma variável, tínhamos a regra de Barrow (ou Teorema Fundamental do Cálculo):

$$\int_a^b f(x) dx = F(b) - F(a)$$

com $F(x)$ uma primitiva de $f(x)$. O cálculo do integral de uma função de várias variáveis reduz-se ao cálculo de vários integrais de funções de uma variável usando o seguinte resultado.

Teorema 17.7 (Teorema de Fubini - versão preliminar para funções de duas variáveis). *Seja $I = [a, b] \times [c, d] \subset \mathbb{R}^2$ um intervalo e $f: I \rightarrow \mathbb{R}$ uma função integrável tal que para cada $x \in [a, b]$, a função $f_x: [c, d] \rightarrow \mathbb{R}$ definida por $f_x(y) = f(x, y)$ é integrável. Então escrevendo*

$$A(x) = \int_c^d f_x(y) dy$$

temos

$$\int_I f = \int_a^b A(x) dx = \int_a^b \left(\int_c^d f(x, y) dy \right) dx$$

O resultado anterior reduz o cálculo do integral de uma função de duas variáveis ao cálculo de dois integrais de Cálculo 1. Traduz a ideia intuitiva que o volume debaixo do gráfico de f se pode obter ‘somando’ (isto é, integrando) as áreas ($A(x)$ na expressão acima) das regiões que se obtêm intersectando a região debaixo do gráfico de f com os planos verticais paralelos ao plano yz com abcissa x .

Exemplo 17.8. *Seja $f: [0, 1] \times [0, 1] \rightarrow \mathbb{R}$ a função definida por $f(x, y) = x + y$. Admitindo que f é integrável (o que será demonstrado em breve), o Teorema de Fubini diz que*

$$\int_{[0,1] \times [0,1]} f = \int_0^1 A(x) dx = \int_0^1 \left(\int_0^1 x + y dy \right) dx$$

O integral dentro de parentesis é calculado com $x \in [0, 1]$ fixo, pelo que x desempenha o papel de uma constante. Assim

$$A(x) = x + \frac{y^2}{2} \Big|_{y=0}^{y=1} = x + \frac{1}{2}$$

e

$$\int_{[0,1] \times [0,1]} f = \int_0^1 x + \frac{1}{2} dx = 1.$$

18. O TEOREMA DE FUBINI

Na secção anterior definimos o integral de uma função $f: I \rightarrow \mathbb{R}$ definida num intervalo limitado $I \subset \mathbb{R}^n$. É frequente usar a notação

$$\int \cdots \int_I f(x_1, \dots, x_n) dx_1 \cdots dx_n$$

para um tal integral, que se chama um integral múltiplo. Assim, um integral de uma função de duas variáveis costuma escrever-se

$$\int \int_I f(x, y) dx dy$$

e o de uma função de três variáveis por

$$\int \int \int_I f(x, y, z) dx dy dz$$

Dissemos também que o cálculo de um tal integral se reduz, pelo Teorema de Fubini a um *integral iterado*, isto é a cálculos sucessivos de integrais de funções de uma variável. Por exemplo, para uma função integrável $f: [a, b] \times [c, d] \rightarrow \mathbb{R}$ (tal que para cada $x \in [a, b]$ a função $f_x: [c, d] \rightarrow \mathbb{R}$ definida por $f_x(y) = f(x, y)$ é integrável), temos

$$\int \int_{[a,b] \times [c,d]} f(x, y) dx dy = \int_a^b \left(\int_c^d f(x, y) dy \right) dx$$

No integral dentro de parentesis, x está fixo e é tratado como uma constante.

Um integral deve ser visto como “uma soma dos valores da função $f(x, y)$ quando (x, y) percorre os pontos de um rectângulo”. O Teorema de Fubini traduz então a *propriedade associativa* da soma: esta pode ser calculada fazendo primeiro as somas de f sobre cada “fatia vertical” em que x está fixo; estas somas dependerão (em princípio) de x ; somando as somas respeitantes a todas as fatias verticais (para todos os valores possíveis de x) produz a soma sobre todo o rectângulo. Vejamos outro exemplo.

Exemplo 18.1. Seja $f: \mathbb{R}^2 \rightarrow \mathbb{R}$ a função definida por $f(x, y) = x^2 y$. Assumindo a integrabilidade de f , calcular

$$\int \int_{[0,1] \times [0,1]} f(x, y) dx dy$$

Pelo Teorema de Fubini, este integral é dado por

$$\int_0^1 \left(\int_0^1 x^2 y dy \right) dx = \int_0^1 x^2 \frac{y^2}{2} \Big|_{y=0}^{y=1} dx = \int_0^1 \frac{x^2}{2} dx = \frac{1}{6}$$

Vejam agora o enunciado geral do Teorema de Fubini.

Teorema 18.2 (Teorema de Fubini). *Sejam $I \subset \mathbb{R}^n$ e $J \subset \mathbb{R}^m$ intervalos limitados e $f: I \times J \rightarrow \mathbb{R}$ uma função integrável. Para cada $x \in I$ seja $f_x: J \rightarrow \mathbb{R}$ a função definida por $f_x(y) = f(x, y)$ e para cada $y \in J$ seja $f_y: I \rightarrow \mathbb{R}$ a função definida por $f_y(x) = f(x, y)$. Então as funções de I para \mathbb{R} definidas por*

$$x \mapsto \int_J f_x \quad e \quad x \mapsto \overline{\int_J f_x}$$

e as funções de J para \mathbb{R} definidas por

$$y \mapsto \int_I f_y \quad e \quad y \mapsto \overline{\int_I f_y}$$

são integráveis e

$$\int_{I \times J} f = \int_I \left(\int_J f_x \right) = \int_I \left(\overline{\int_J f_x} \right) = \int_J \left(\int_I f_y \right) = \int_J \left(\overline{\int_I f_y} \right).$$

Nota 18.3. *Na maioria dos casos em que se aplica este Teorema, as funções f_x e f_y são integráveis para todos os valores de x e de y . Nesse caso os integrais inferiores e superiores coincidem e o Teorema diz que*

$$\int_{I \times J} f = \int_I \left(\int_J f_x \right) = \int_J \left(\int_I f_y \right)$$

ou, com notação mais sugestiva

$$\int \int_{I \times J} f(x, y) dx dy = \int_I \left(\int_J f(x, y) dy \right) dx = \int_J \left(\int_I f(x, y) dx \right) dy$$

Aplicando várias vezes o Teorema 18.2, qualquer integral múltiplo sobre um retângulo se reduz ao cálculo iterado de integrais de funções de uma variável. Por exemplo

$$\int_{[0,1] \times [0,2] \times [0,3]} f = \int_0^1 \left(\int_{[0,2] \times [0,3]} f(x, y, z) dy dz \right) dx = \int_0^1 \left(\int_0^2 \left(\int_0^3 f(x, y, z) dz \right) dy \right) dx$$

A ordem de integração é arbitrária. O Teorema de Fubini dá seis maneiras equivalentes de calcular o integral acima como um integral iterado.

Definição 18.4. *Seja $A \subset \mathbb{R}^n$ um conjunto limitado, e $f: A \rightarrow \mathbb{R}$ uma função limitada. Defina-se*

$$\int_A f = \int_I \tilde{f}$$

onde I é um intervalo limitado de \mathbb{R}^n que contém A e $\tilde{f}: I \rightarrow \mathbb{R}$ é a função definida por

$$\tilde{f}(x) = \begin{cases} f(x) & \text{se } x \in A \\ 0 & \text{caso contrário.} \end{cases}$$

A definição anterior é consistente com a ideia que o volume debaixo do gráfico de f sobre o conjunto A deve ser igual ao volume debaixo do gráfico de \tilde{f} (que é o prolongamento por 0 de f a I) sobre o intervalo I . Fica como exercício a verificação que a definição anterior faz sentido, isto é, que a noção de integrabilidade em A resultante e o valor do integral em A são independentes da escolha do intervalo I contendo A .

Vejam agora alguns exemplos de cálculo com o Teorema de Fubini. Assumiremos durante os exemplos que as funções em causa são integráveis, o que será justificado mais tarde.

Exemplo 18.5. (1) Seja $T = \{(x, y) \in \mathbb{R}^2: x \geq 0, y \geq 0, x + y \leq 1\}$. Calcular $\int \int_T xy \, dx dy$.

De acordo com a Definição 18.4, temos que integrar a função $\tilde{f}: [0, 1] \times [0, 1] \rightarrow \mathbb{R}$ definida por

$$\tilde{f}(x, y) = \begin{cases} xy & \text{se } (x, y) \in T \\ 0 & \text{caso contrário.} \end{cases}$$

Podemos calcular este integral de duas maneiras diferentes (organizando a soma primeiro em fatias horizontais ou verticais) mas dada a simetria da função integranda e da região, as duas são inteiramente análogas. Calculemos o integral somando primeiro sobre as fatias verticais (isto é com x constante). Para cada $x \in [0, 1]$, a função $\tilde{f}_x: [0, 1] \rightarrow \mathbb{R}$ é definida por

$$\tilde{f}_x(y) = \begin{cases} xy & \text{se } (x, y) \in T \Leftrightarrow 0 \leq y \leq 1 - x \\ 0 & \text{se } 1 - x \leq y \leq 1 \end{cases}$$

O Teorema de Fubini diz então que

$$\begin{aligned} \int \int_T xy \, dx dy &= \int_0^1 \left(\int_0^1 \tilde{f}_x(y) dy \right) dx \\ &= \int_0^1 \left(\int_0^{1-x} xy dy + \int_{1-x}^1 0 dy \right) dx \\ &= \int_0^1 \left(\int_0^{1-x} xy dy \right) dx \\ &= \int_0^1 x \frac{y^2}{2} \Big|_{y=0}^{y=1-x} dx \\ &= \int_0^1 \frac{1}{2} x (1-x)^2 dx = \frac{1}{24} \end{aligned}$$

Na prática não é costume escrever o prolongamento por 0 da função a integrar, nem o termo correspondente à integração da função 0 que apareceu na segunda linha do cálculo anterior. Os extremos de integração do integral iterado ficam então a depender das variáveis que já estão fixas naquele estágio do cálculo, como na terceira linha do cálculo acima (onde se indica que a fatia vertical com abscissa x sobre a qual temos que somar a função integranda é a fatia $0 \leq y \leq 1 - x$).

(2) Calcular $\int \int_S x dx dy$ onde $S = \{(x, y) \in \mathbb{R}^2 : x \geq 0, 2x^2 \leq y \leq 1 + x^2\}$.

As parábolas $y = 1 + x^2$ e $y = 2x^2$ intersectam-se no ponto $(1, 2)$, pelo que a coordenada x varia entre 0 e 1 quando (x, y) percorre a região S . Assim, há tantas fatias verticais quantos $x \in [0, 1]$. Para cada um destes x , o intervalo de integração é definido pelas desigualdades $2x^2 \leq y \leq 1 + x^2$ logo, o Teorema de Fubini diz que

$$\begin{aligned} \int \int_S x dx dy &= \int_0^1 \left(\int_{2x^2}^{1+x^2} x dy \right) dx \\ &= \int_0^1 x (1 + x^2 - 2x^2) dx = \int_0^1 x - x^3 dx = \frac{1}{4} \end{aligned}$$

Também podemos calcular o integral na outra ordem de integração, em que organizamos a soma por fatias horizontais. Para isso temos primeiro que determinar o número de fatias horizontais, isto é o domínio de variação de y quando (x, y) percorre a região de integração S . Esboçando a região S vemos imediatamente que $0 \leq y \leq 2$ (o que também pode ser visto directamente das desigualdades $0 \leq x \leq 1$ e $2x^2 \leq y \leq 1 + x^2$). Note-se no entanto que a expressão para a fatia horizontal em função de y não é sempre igual quando y varia entre 0 e 2: quando y está entre 0 e 1, a fatia é limitada à esquerda pelo eixo dos yy e à direita pela parábola $y = 2x^2$, ou seja a fatia é definida por $0 \leq x \leq \sqrt{\frac{y}{2}}$, enquanto que para $y \geq 1$, a fatia é limitada à esquerda pela parábola $y = 1 + x^2$ e à direita pela parábola $y = 2x^2$ de forma que $\sqrt{y-1} \leq x \leq \sqrt{\frac{y}{2}}$. Conclui-se assim que o integral é dado por

$$\begin{aligned} \int \int_S x dx dy &= \int_0^1 \left(\int_0^{\sqrt{\frac{y}{2}}} x dx \right) dy + \int_1^2 \left(\int_{\sqrt{y-1}}^{\sqrt{\frac{y}{2}}} x dx \right) dy \\ &= \int_0^1 \frac{y}{4} dy + \int_1^2 \frac{y}{4} - \frac{y-1}{2} dy = \frac{1}{4} \end{aligned}$$

(3) Calcular

$$\int \int \int_T 1 dx dy dz$$

com $T = \{(x, y, z) \in \mathbb{R}^3 : x \geq 0, y \geq 0, z \geq 0, x + y + z \leq 1\}$.

Dada a simetria da região e da função a integrar, todas as ordens de integração são análogas. Vamos por exemplo dividir o sólido T em fatias horizontais. Claramente há uma fatia para cada $z \in [0, 1]$. Fixando $z \in [0, 1]$ nas equações que

definam T obtemos uma expressão para a fatia:

$$T_z = \{(x, y) \in \mathbb{R}^2 : x \geq 0, y \geq 0, x + y \leq 1 - z\}$$

O Teorema de Fubini diz que

$$\begin{aligned} \int \int \int_T 1 dx dy dz &= \int_0^1 \left(\int_{T_z} 1 dx dy \right) dz \\ &= \int_0^1 \left(\int_0^{1-z} \left(\int_0^{1-z-y} 1 dx \right) dy \right) dz \end{aligned}$$

onde na segunda igualdade usamos que a região de integração T_z é inteiramente análoga ao exemplo (1) acima (trata-se de um triângulo cujos vértices são a origem, e os pontos $(1 - z, 0)$ e $(0, 1 - z)$).

19. MAIS EXEMPLOS. DEMONSTRAÇÃO DO TEOREMA DE FUBINI.

Os integrais múltiplos têm diversas aplicações.

Definição 19.1. O volume n -dimensional de $A \subset \mathbb{R}^n$ é definido como $\text{vol}(A) = \int_A 1$.

Esta definição é consistente com a ideia que o integral de 1 é o volume $(n+1)$ -dimensional debaixo do gráfico da função constante igual a 1 sobre o conjunto A . De facto este volume deveria ser igual ao volume n -dimensional da “base” A vezes a altura do gráfico (que é 1).

Se $f: A \rightarrow \mathbb{R}$ é uma função, define-se a *média* de f em A por

$$\bar{f} = \frac{\int_A f}{\text{vol}(A)}$$

Um exemplo particularmente útil é quando f é uma das funções coordenadas x_i . O número \bar{x}_i diz-se então a i -ésima coordenada do *centróide* de A (o centro de massa de um corpo homogéneo com a forma de A). Por exemplo, para $A \subset \mathbb{R}^3$, o centróide é o ponto $(\bar{x}, \bar{y}, \bar{z})$ onde

$$\bar{x} = \frac{\int \int \int x dx dy dz}{\int \int \int 1 dx dy dz}, \quad \bar{y} = \frac{\int \int \int y dx dy dz}{\int \int \int 1 dx dy dz}, \quad \bar{z} = \frac{\int \int \int z dx dy dz}{\int \int \int 1 dx dy dz}$$

Sempre que possível deve explorar-se a simetria no cálculo destas quantidades (e dos integrais em geral). Por exemplo, um sólido simétrico em relação ao plano xy terá necessariamente $\bar{z} = 0$

Se $f: A \rightarrow \mathbb{R}$ designar a densidade de massa (ou de carga) por unidade de volume/área em A , então a massa (carga) total pode ser calculada através do integral de f em A .

As fórmulas da mecânica para as coordenadas do centro de massa e momento de inércia em relação a um eixo L de um sistema de partículas indexadas por i são

$$\vec{r}_{CM} = \sum_i m_i \vec{r}_i, \quad I_L = \sum_i m_i (d_i)^2$$

(com \vec{r}_i as posições das partículas, m_i as respectivas massas e d_i a distância da partícula i ao eixo L). Estas fórmulas generalizam-se da forma evidente ao cálculo do centro de massa ou do momento de inércia de uma placa $A \subset \mathbb{R}^2$ ou de um sólido $A \subset \mathbb{R}^3$ com *densidade*

de massa $f: A \rightarrow \mathbb{R}$ (por unidade de área ou volume respectivamente). Por exemplo a coordenada x do centro de massa é dada no caso de um sólido por

$$x_{CM} = \frac{\int \int \int_A x f(x, y, z) \, dx dy dz}{\text{vol}(A)}$$

e o momento de inércia em torno de um eixo L é dado por

$$I_L = \int \int \int_A f(x, y, z) d_L(x, y, z)^2 \, dx dy dz$$

onde $d_L(x, y, z)$ denota a distância do ponto (x, y, z) ao eixo L .

Exemplo 19.2. Escrever um integral triplo que calcula o volume do sólido

$$V = \{(x, y, z): x \geq 0, y \geq 0, x + y \leq 1, 0 \leq z \leq 1 + x^2 + y^2\}$$

Este sólido é a região por debaixo do gráfico da função $1 + x^2 + y^2$ sobre o triângulo no plano xy definido pelas equações $x \geq 0, y \geq 0, x + y \leq 1$. O Teorema de Fubini permite-nos organizar a soma sobre V fazendo-a primeiro ao longo dos segmentos verticais $0 \leq z \leq 1 + x^2 + y^2$ determinados por um ponto (x, y) no triângulo, seguida de integração no triângulo:

$$\begin{aligned} \text{vol}(V) &= \int \int \int_T 1 \, dx dy dz \\ &= \int \int_{\{(x,y): x \geq 0, y \geq 0, x+y \leq 1\}} \left(\int_0^{1+x^2+y^2} 1 \, dz \right) dx dy \\ &= \int_0^1 \left(\int_0^{1-x} \left(\int_0^{1+x^2+y^2} 1 \, dz \right) dy \right) dx \end{aligned}$$

A título de exemplo, escrevamos o mesmo integral na ordem de integração correspondente a dividir o sólido T em fatias horizontais. Como para $x, y \geq 0$, temos que $x + y \leq 1 \Rightarrow x^2 + y^2 \leq 1$ vemos que $0 \leq z \leq 2$. Para cada $z \in [0, 2]$ a fatia correspondente é definida pelas condições

$$x \geq 0, \quad y \geq 0, \quad x + y \leq 1, \quad x^2 + y^2 \geq z - 1$$

A última condição é automática quando $z \leq 1$, pelo que a fatia correspondente é o triângulo definido pelas condições $x \geq 0, y \geq 0, x + y \leq 1$. A contribuição destas fatias para o volume calcula-se através do integral

$$\int_0^1 \left(\int_0^1 \left(\int_0^{1-x} 1 \, dy \right) dx \right) dz$$

Para $z \geq 1$, a região de integração é a interseção do triângulo acima com o exterior de uma circunferência de raio $\sqrt{z-1}$. Para escrever os limites de integração desta região temos que considerar separadamente os casos $\sqrt{z-1} \leq \frac{1}{\sqrt{2}} \Leftrightarrow z \leq \frac{3}{2}$ em que a circunferência não intersecta a hipotenusa do triângulo (a quantidade $\frac{1}{\sqrt{2}}$ é a distância da hipotenusa à

origem) e os casos em que $\frac{3}{2} \leq z \leq 2$. Uma vez que, neste último caso, a interseção da circunferência e da hipotenusa é dada por

$$\begin{cases} x^2 + y^2 = 1 - z \\ x + y = 1 \end{cases} \Leftrightarrow \begin{cases} x = \frac{1 \pm \sqrt{2z-3}}{2} \\ y = \frac{1 \mp \sqrt{2z-3}}{2} \end{cases}$$

obtemos as seguintes expressões para o volume da porção de V compreendida entre os planos $z = 1$ e $z = \frac{3}{2}$:

$$\int_1^{\frac{3}{2}} \left(\int_0^{\sqrt{z-1}} \left(\int_0^{1-x} 1 dy \right) dx + \int_{\sqrt{z-1}}^1 \left(\int_{\sqrt{z-1-x^2}}^{1-x} 1 dy \right) dx \right) dz$$

e a porção contida entre os planos $z = \frac{3}{2}$ e $z = 2$:

$$\int_{\frac{3}{2}}^2 \left(\int_0^{\frac{1-\sqrt{2z-3}}{2}} \left(\int_{\sqrt{z-1-x^2}}^{1-x} 1 dy \right) dx + \int_{\frac{1+\sqrt{2z-3}}{2}}^{\sqrt{z-1}} \left(\int_{\sqrt{z-1-x^2}}^{1-x} 1 dy \right) dx + \int_{\sqrt{z-1}}^1 \left(\int_0^{1-x} 1 dy \right) dx \right) dz$$

A expressão para o volume de T obtém-se somando todos estes integrais iterados.

Dem. do Teorema 18.2. Começamos por notar que o Teorema é válido quando f é uma função em escada. Note-se que, mesmo para uma função em escada, a integrabilidade das funções f_x e f_y não é garantida para todos os valores de x e de y . De facto, na fronteira dos intervalos de uma partição adequada a f a única restrição imposta a uma função em escada é que seja limitada, e uma função limitada pode não ser integrável. No entanto, quando x pertence ao interior de um intervalo da partição de I , a função $f_x: J \rightarrow \mathbb{R}$ é ela própria uma função em escada, e é imediato verificar que as funções $I \rightarrow \mathbb{R}$ definidas por

$$x \mapsto \int_{\underline{J}} f_x \quad \text{e} \quad x \mapsto \int_{\overline{J}} f_x$$

são elas próprias funções em escada (que possivelmente diferem na fronteira dos intervalos que constituem a partição de I). Deixa-se como exercício demonstrar que o integral destas funções em I é igual ao integral de f em $I \times J$ (e analogamente para a outra ordem de integração).

Assumindo a validade do Teorema para funções em escada vamos agora ver que, quando f é integrável, a função $x \mapsto \int_{\underline{J}} f_x$ é integrável em I e que o seu integral coincide com $\int_{I \times J} f$ (os outros casos a demonstrar são inteiramente análogos).

Observe-se primeiro que, sendo $K \subset \mathbb{R}^d$ um intervalo, uma função $g: K \rightarrow \mathbb{R}$ é integrável sse para cada $\epsilon > 0$ existem funções em escada $s, t: K \rightarrow \mathbb{R}$ tais que $s(x) \leq f(x) \leq t(x)$ para todo o $x \in K$ e

$$\int_K t - \int_K s < \epsilon$$

De facto isto acontece sse os integrais inferior e superior de g em K diferem em menos de ϵ para todo o $\epsilon > 0$, ou seja se são iguais.

Uma vez que f é integrável em $I \times J$, dado $\epsilon > 0$, existem funções em escada $s, t: I \times J \rightarrow \mathbb{R}$ com $s \leq f \leq t$ e $\int_{I \times J} t - \int_{I \times J} s < \epsilon$. A desigualdade $s \leq f \leq t$ implica que, para cada $x \in I$, temos

$$\int_{\underline{J}} s(x, y) dy \leq \int_{\underline{J}} f(x, y) dy \leq \int_{\underline{J}} t(x, y) dy$$

Mas, como observámos acima, as expressões à esquerda e direita nestas desigualdades são funções em escada em I . Pelo Teorema de Fubini para funções em escada temos

$$\int_I \left(\int_{\underline{J}} s(x, y) dy \right) dx = \int_{I \times J} s, \quad \int_I \left(\int_{\underline{J}} t(x, y) dy \right) dx = \int_{I \times J} t$$

e portanto os integrais destas funções em escada diferem em menos de ϵ . Isto mostra que a função

$$x \mapsto \int_{\underline{J}} f(x, y) dy$$

é integrável em I .

O seu integral está necessariamente entre os integrais $\int_{I \times J} s$ e $\int_{I \times J} t$ e só há um número que satisfaz esta condição para todos os s, t , nomeadamente, $\int_{I \times J} f$. Isto conclui a demonstração. \square

20. PROPRIEDADES DO INTEGRAL. INTEGRABILIDADE

O integral múltiplo tem as seguintes propriedades análogas às já conhecidas para o integral de funções de uma variável (que de uma maneira ou outra vêm de “o integral ser uma soma”).

Proposição 20.1 (Propriedades do integral). *Seja $I \subset \mathbb{R}^n$ um intervalo limitado e $f, g: I \rightarrow \mathbb{R}$ funções integráveis. Então*

1. (Linearidade) *Dados $\alpha, \beta \in \mathbb{R}$, a função $\alpha f + \beta g$ é integrável em I e*

$$\int_I \alpha f + \beta g = \alpha \int_I f + \beta \int_I g$$

2. (Monotonia) *Se $f \leq g$, então $\int_I f \leq \int_I g$.*

3. (Desigualdade triangular) *A função $|f|$ é integrável e $|\int_I f| \leq \int_I |f|$*

Sejam $A, B \subset \mathbb{R}^n$ conjuntos disjuntos e $f: A \cup B \rightarrow \mathbb{R}$ uma função. Então

4. (Aditividade) *Se f é integrável em A e em B , então f é integrável em $A \cup B$ e*

$$\int_{A \cup B} f = \int_A f + \int_B f$$

Dem. As demonstrações de todas estas afirmações são imediatas a partir da definição de integral e as propriedades análogas da soma. A título de exemplo suponhamos que f, g são

integráveis. Então dado $\epsilon > 0$, existem $s_1, s_2, t_1, t_2: I \rightarrow \mathbb{R}$ em escada tais que $s_1 \leq f \leq t_1$ e $s_2 \leq g \leq t_2$ e

$$\int_I t_1 - \int_I s_1 < \frac{\epsilon}{2} \quad \text{e} \quad \int_I t_2 - \int_I s_2 < \frac{\epsilon}{2}$$

Mas então $s_1 + s_2$ e $t_1 + t_2$ são funções em escada, $s_1 + s_2 \leq f + g \leq t_1 + t_2$ e

$$\int_I (t_1 + t_2) - \int_I (s_1 + s_2) < \epsilon$$

pelo que $f + g$ é integrável e vemos imediatamente que $\int_I (f + g) = \int_I f + \int_I g$.

A propriedade 4. é uma consequência da propriedade que acabamos de demonstrar e do facto de a extensão por 0 da função f a um intervalo I que contenha $A \cup B$ ser a soma das extensões por 0 a I das restrições de f a A e B respectivamente.

A demonstração das restantes propriedades deixa-se como exercício. \square

Até agora temos estado a aplicar o Teorema de Fubini sem nunca verificar a sua hipótese que é a integrabilidade das funções envolvidas. Vamos agora ver critérios que garantem a integrabilidade (em conjuntos limitados) de essencialmente todas as funções (limitadas) que ocorrem na prática.

Teorema 20.2. *Seja $I \subset \mathbb{R}^n$ um intervalo compacto e $f: I \rightarrow \mathbb{R}$ uma função contínua. Então f é integrável em I .*

Dem. Vamos explicar como achar sucessões $s_n, t_n: I \rightarrow \mathbb{R}$ de funções em escada satisfazendo

$$s_0 \leq s_1 \leq \dots \leq s_n \leq \dots \leq f \leq \dots \leq t_n \leq \dots \leq t_1 \leq t_0$$

e tais que

$$\lim_{n \rightarrow \infty} \left(\int_I t_n - \int_I s_n \right) = 0$$

Isto garante a integrabilidade de f .

Para cada $n \geq 0$, seja P_n a partição de $I = [a_1, b_1] \times \dots \times [a_n, b_n]$ que se obtém subdividindo cada um dos intervalos $[a_i, b_i]$ em 2^n partes iguais (isto é passamos de P_n a P_{n+1} dividindo ao meio cada aresta de cada subintervalo de P_n). Sendo $\{I_j\}$ os subintervalos da partição P_n definimos

$$s_n(x) = \min\{s(y) : y \in I_j \text{ onde } I_j \text{ é um subintervalo contendo } x\}$$

e analogamente,

$$t_n(x) = \max\{s(y) : y \in I_j \text{ onde } I_j \text{ é um subintervalo contendo } x\}$$

Assim, no interior do subintervalo I_j da partição P_n , as funções s_n (respectivamente t_n) tomam como valores o mínimo (respectivamente o máximo) da função f no intervalo compacto I_j (que existem pelo Teorema de Weierstrass). Quando x pertence a mais de um intervalo da partição, tomamos o máximo ou o mínimo de f na união dos intervalos em questão.

Claramente s_n e t_n são funções em escada e, uma vez que os subintervalos da partição P_{n+1} estão contidos em subintervalos da partição P_n , temos

$$s_n \leq s_{n+1} \quad \text{e} \quad t_{n+1} \leq t_n$$

Vamos agora ver que

$$(6) \quad \lim_{n \rightarrow \infty} \max\{t_n(x) - s_n(x) : x \in I\} = 0$$

Isto é suficiente para provar a integrabilidade de f uma vez que

$$\int_I t_n - \int_I s_n \leq \text{vol}(I) \cdot \max\{t_n(x) - s_n(x) : x \in I\}.$$

Suponhamos que (6) é falso. Então existe $l > 0$, uma sucessão crescente de naturais $n_1 \leq n_2 \leq \dots \leq n_k \leq \dots$ e uma sucessão x_k de pontos em I tais que

$$t_{n_k}(x_k) - s_{n_k}(x_k) > l$$

Uma vez que I é compacto, a sucessão x_k tem uma subsucessão convergente. Refazendo a nossa escolha da sucessão crescente n_k podemos assumir sem perda de generalidade que a própria sucessão x_k converge. Seja $y = \lim x_k$. Como f é contínua em y , existe $r > 0$ tal que

$$\max\{f(x) : x \in B_r(y)\} - \min\{f(x) : x \in B_r(y)\} < l$$

Mas para k suficientemente grande não só os pontos x_k , como todos os intervalos da partição P_{n_k} que os contêm estão contidos em $B_r(y)$ e portanto

$$t_{n_k}(x_k) - s_{n_k}(x_k) < l$$

o que é uma contradição. Isto conclui a demonstração. \square

21. INTEGRABILIDADE (CONT.). MUDANÇA DE VARIÁVEIS

O critério de integrabilidade dado pelo Teorema 20.2 não é suficiente para verificar as condições do Teorema de Fubini na maioria dos exemplos vistos acima. De facto, quando a região de integração não é um intervalo, o integral é definido extendendo a função integranda por 0 a um intervalo que contém a região de integração. Mesmo que a função integranda seja contínua na região de integração original, o prolongamento não será em geral contínuo no intervalo e portanto não poderemos aplicar o Teorema 20.2. Vamos agora ver que a integrabilidade num intervalo (limitado) fica garantida desde que a função (limitada) que estamos a integrar não seja "demasiado descontínua", isto é, desde que as descontinuidades da função ocorram num conjunto de "volume n -dimensional negligível", conceito que iremos agora definir.

Definição 21.1. *Um conjunto $A \subset \mathbb{R}^n$ tem conteúdo nulo se para todo o $\epsilon > 0$ existe uma família finita I_1, \dots, I_N de intervalos limitados de \mathbb{R}^n tal que*

- (i) $A \subset I_1 \cup \dots \cup I_N$
- (ii) $\sum_{k=1}^N \text{vol}(I_k) < \epsilon$

A ideia da definição anterior é que o conjunto A tem volume n -dimensional $\leq \epsilon$ para todo $\epsilon > 0$ e portanto deverá ter volume n -dimensional 0.

Nota 21.2. *Note-se que o conceito de conteúdo nulo depende do espaço ambiente. Por exemplo um rectângulo aberto em \mathbb{R}^2 não tem conteúdo nulo (tente demonstrá-lo!) em \mathbb{R}^2 , mas o mesmo rectângulo visto como subconjunto de \mathbb{R}^3 tem conteúdo nulo, como se verifica facilmente.*

Proposição 21.3 (Propriedades dos conjuntos de conteúdo nulo). (i) *Se $A \subset B$ e B tem conteúdo nulo, então A tem conteúdo nulo.*

(ii) *Uma união finita de conjuntos com conteúdo nulo tem conteúdo nulo.*

(iii) *Se $I \subset \mathbb{R}^n$ é um intervalo limitado e $f: I \rightarrow \mathbb{R}$ é uma função integrável, então o gráfico de f , isto é o conjunto $\{(x, f(x)): x \in I\} \subset \mathbb{R}^{n+1}$, tem conteúdo nulo.*

Dem. A demonstração das duas primeiras afirmações fica como exercício. Dado $\epsilon > 0$, sejam $s, t: I \rightarrow \mathbb{R}$ funções em escada tais que $s(x) \leq f(x) \leq t(x)$ para todo o $x \in I$ e $\int_I t - \int_I s < \frac{\epsilon}{2}$ (que existem porque f é integrável). Sendo $\{I_i\}$ os subintervalos de uma partição de I tal que s, t são constantes no interior dos intervalos I_i . Temos então que o gráfico de f sobre o conjunto $\cup_i \text{int } I_i$ está contido em

$$\{(x, y) \in \mathbb{R}^{n+1} : x \in \cup_i \text{int } I_i, \text{ e } s(x) \leq y \leq t(x)\}$$

que é uma união de intervalos disjuntos com volume total menor que $\frac{\epsilon}{2}$.

Resta-nos cobrir o resto do gráfico de f (a parte sobre a fronteira dos subintervalos I_i) por uma família $\{J_k\}$ de intervalos de \mathbb{R}^{n+1} com $\sum_k \text{vol}(J_k) < \frac{\epsilon}{2}$. Como f é limitada, existe $R > 0$ tal que $|f(x)| \leq R$ para todo o $x \in I$. Sendo $F = I \setminus \cup_i \text{int } I_i$ a união das fronteiras dos subintervalos I_i , fica como exercício descrever uma família $\{M_k\}$ de intervalos de \mathbb{R}^n tais que $F \subset \cup_k M_k$ e $\sum_k \text{vol}(M_k) < \frac{\epsilon}{4R}$. Podemos então tomar $J_k = M_k \times [-R, R]$ \square

Exemplo 21.4. (i) *O conjunto $A = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 = 1\}$ tem conteúdo nulo em \mathbb{R}^2 . De facto A é a união dos gráficos das funções contínuas (e portanto integráveis) $f, g: [-1, 1] \rightarrow \mathbb{R}$, definidas por*

$$f(x) = \sqrt{1 - x^2}, \quad g(x) = -\sqrt{1 - x^2}$$

(ii) *O conjunto $B = \{(x, y, z) \in \mathbb{R}^3 : x^2 + y^2 + z^2 = 1\}$ tem conteúdo nulo, pois está contido na união dos gráficos das funções $f, g: [-1, 1] \times [-1, 1] \rightarrow \mathbb{R}$ definidas por*

$$f(x, y) = \begin{cases} \sqrt{1 - x^2 - y^2} & \text{se } x^2 + y^2 \leq 1 \\ 0 & \text{se } x^2 + y^2 > 1 \end{cases} \quad g(x, y) = \begin{cases} \sqrt{1 - x^2 - y^2} & \text{se } x^2 + y^2 \leq 1 \\ 0 & \text{se } x^2 + y^2 > 1 \end{cases}$$

que são contínuas e portanto integráveis (pelo Teorema 20.2).

Intuitivamente, qualquer subconjunto de uma família finita de linhas compactas em \mathbb{R}^2 ou de superfícies compactas em \mathbb{R}^3 tem conteúdo nulo. De facto um tal conjunto pode ser coberto por uma família finita de gráficos de funções contínuas definidas em intervalos, como no exemplo anterior.

É possível (mas não fácil - tente-o!) demonstrar que um intervalo de \mathbb{R}^n com interior não vazio não tem conteúdo nulo em \mathbb{R}^n , tal como seria de esperar. Segue-se da Proposição 21.3(i) que se $\text{int } A \neq \emptyset$, então A não tem conteúdo nulo.

Podemos agora enunciar um critério de integrabilidade que é suficiente para os nossos propósitos.

Teorema 21.5. *Seja $I \subset \mathbb{R}^n$ um intervalo e $f: I \rightarrow \mathbb{R}$ uma função limitada. Se o conjunto dos pontos de descontinuidade de f tem conteúdo nulo, então f é integrável em I .*

Ideia da Demonstração. A ideia da demonstração é muito semelhante à do Teorema 20.2 e Proposição 21.3(iii): Dado $\epsilon > 0$, precisamos de explicar como achar funções em escada $s, t: I \rightarrow \mathbb{R}$ com $s(x) \leq f(x) \leq t(x)$ e $\int_I t - \int_I s < \epsilon$. Como f é limitada em I , existe $R > 0$ tal que $|f(x)| \leq R$ para todo o $x \in I$. Como o conjunto D dos pontos de descontinuidade de f tem conteúdo nulo, podemos achar uma partição de I tal que a soma dos volumes dos subintervalos cujo interior intersecta D é menor que $\frac{\epsilon}{4R}$. Nesses subintervalos definimos $s(x) = -R$ e $t(x) = R$. O integral de $t - s$ sobre a união destes subintervalos é portanto $< \frac{\epsilon}{2}$. No complementar do interior desses subintervalos, a função f é contínua e o argumento na demonstração do Teorema 20.2 mostra como escolher as funções s, t (sendo possivelmente necessário refinar a partição de I) de tal forma que a diferença entre os integrais de t e s nos rectângulos em que f é contínua seja menor que $\frac{\epsilon}{2}$. \square

Exemplo 21.6. *Seja $S = \{(x, y) \in \mathbb{R}^2: x^2 + y^2 \leq 1\}$. Qualquer função contínua $f: S \rightarrow \mathbb{R}$ é integrável em S . De facto, f é automaticamente limitada pelo Teorema de Weierstrass e o conjunto dos pontos de descontinuidade do prolongamento por 0 de f a um intervalo que contenha S está contido no conjunto $\{(x, y) \in \mathbb{R}^2: x^2 + y^2 = 1\}$ (a fronteira de S) que tem conteúdo nulo pelo Exemplo 21.4.*

Mais geralmente, o argumento do exemplo anterior mostra que se $S \subset \mathbb{R}^n$ é um conjunto cuja fronteira tem conteúdo nulo, toda a função contínua e limitada em S é integrável em S .

Nota 21.7. *O Critério de integrabilidade de Lebesgue diz que uma função limitada $f: I \rightarrow \mathbb{R}$ (com $I \subset \mathbb{R}^n$ um intervalo limitado) é integrável sse o conjunto dos pontos de descontinuidade de f tem medida nula. Um conjunto $A \subset \mathbb{R}^n$ tem medida nula se para todo o $\epsilon > 0$ existe uma família contável de intervalos limitados I_k tal que $A \subset \cup_k I_k$ e $\sum_{k=1}^{\infty} \text{vol}(I_k) < \epsilon$. Recomenda-se aos alunos interessados que consultem [S] para uma demonstração deste resultado.*

21.8. Mudança de variáveis no integral. Em Cálculo 1, uma dos métodos básicos para o cálculo de integrais é a substituição, ou mudança de variáveis. Recorde-se que a substituição $x = g(t)$ transforma o integral

$$\int_a^b f(x) dx$$

em

$$\int_{g^{-1}(a)}^{g^{-1}(b)} f(g(t))g'(t) dt$$

onde g é uma função diferenciável, com $g'(t) \neq 0$, definida num intervalo cuja imagem contém $[a, b]$. A mudança de variável $x = g(t) \Leftrightarrow t = g^{-1}(x)$ deve ser vista como um “dicionário” que nos permite traduzir qualquer coisa escrita em termos de x em algo escrito em termos de t (e vice-versa). Na fórmula de mudança de variável o intervalo compreendido entre $g^{-1}(a)$ e $g^{-1}(b)$ é o intervalo $[a, b]$ escrito em termos de t , a função $f(g(t))$ é a função $f(x)$ escrita em termos de t e, finalmente, $g'(t)dt = \frac{dx}{dt}dt$ é “ dx escrito em termos da variável t ”.

Vamos agora estudar mudança de variáveis em integrais em múltiplos. Estas são ainda mais importantes que em Cálculo 1. As mudanças de variáveis são agora utilizadas para simplificar não só a função integranda mas também (e até especialmente) a *região de integração* que já não é necessariamente um intervalo. Começamos por ver um exemplo antes de formular precisamente o conceito de mudança de variável e enunciar o Teorema de mudança de variáveis no integral.

Exemplo 21.9. Sendo $S = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 \leq 1\}$ vamos calcular o integral

$$\int \int_S \sqrt{x^2 + y^2} dx dy$$

fazendo a mudança de variável definida pela expressão

$$\begin{cases} x = r \cos \theta \\ y = r \sin \theta \end{cases}$$

(portanto r é a distância à origem de um ponto do plano, e θ é o ângulo que esse ponto faz com o eixo dos xx). A região S escrita em termos das coordenadas r, θ é simplesmente $0 \leq r \leq 1$ e $0 \leq \theta < 2\pi$. Substituindo nas expressões para x e y em função de r, θ , vemos que a função integranda escrita nas novas variáveis é a função

$$\sqrt{r^2 \cos^2 \theta + r^2 \sin^2 \theta} = r$$

Finalmente, tal como no caso da mudança de variável em integrais de funções de uma variável, é necessário multiplicar a função integranda por uma função - o “factor de conversão de áreas” - que é dado por

$$\left| \frac{\partial(x, y)}{\partial(r, \theta)} \right| = \left| \begin{vmatrix} \frac{\partial x}{\partial r} & \frac{\partial x}{\partial \theta} \\ \frac{\partial y}{\partial r} & \frac{\partial y}{\partial \theta} \end{vmatrix} \right| = \left| \begin{vmatrix} \cos \theta & -r \sin \theta \\ \sin \theta & r \cos \theta \end{vmatrix} \right| = r$$

O integral nas variáveis (x, y) transforma-se no seguinte integral muito simples nas variáveis (r, θ) :

$$\int_0^{2\pi} \int_0^1 r r dr d\theta = 2\pi \left. \frac{r^3}{3} \right|_0^1 = \frac{2\pi}{3}$$

É instrutivo tentar calcular o integral em coordenadas cartesianas e ver quão mais trabalhoso é!

22. EXEMPLOS DE TRANSFORMAÇÃO DE COORDENADAS. O TEOREMA DE MUDANÇA DE VARIÁVEIS.

Definição 22.1. Uma transformação de coordenadas (ou mudança de coordenadas, ou mudança de variável) em \mathbb{R}^n é uma aplicação $\varphi: U \rightarrow \mathbb{R}^n$ com U um aberto de \mathbb{R}^n , tal que

- (1) φ é de classe C^1 ,
- (2) φ é injectiva,
- (3) $\det D\varphi(x) \neq 0$ para todo o $x \in U$.

A função

$$\det D\varphi(x)$$

chama-se o *Jacobiano* da transformação φ .

Nota 22.2. Veremos mais tarde que a definição anterior é equivalente à seguinte, que é mais natural mas menos prática: $\varphi: U \rightarrow \mathbb{R}^n$ é uma transformação de coordenadas se é de classe C^1 , injectiva, $V = \varphi(U)$ é aberto e a função inversa $\varphi^{-1}: V \rightarrow U$ é de classe C^1 .

Uma mudança de variável $\varphi: U \rightarrow V = \varphi(U)$ deve ser vista como um “dicionário” entre U e V . Em termos de variáveis $x \in U \subset \mathbb{R}^n$ e $y \in V \subset \mathbb{R}^n$ temos

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

e as regras φ e φ^{-1} dizem-nos como traduzir qualquer objecto definido em Cálculo (por exemplo uma função, um limite, um integral,...) das variáveis x para y e vice-versa.

Vejamos alguns exemplos importantes.

Exemplo 22.3. (1) **Coordenadas polares.** Trata-se da transformação de coordenadas definida por $\varphi:]0, +\infty[\times]0, 2\pi[\rightarrow \mathbb{R}^2$ definida pela expressão

$$\varphi(r, \theta) = (r \cos \theta, r \sin \theta)$$

Escrevendo $(x, y) = \varphi(r, \theta)$ temos

$$\begin{cases} x = r \cos \theta \\ y = r \sin \theta \end{cases}$$

e a trigonometria elementar torna claro que se (x, y) são as coordenadas euclidianas usuais, então r corresponde à distância de (x, y) à origem, enquanto que θ é o ângulo que o vector (x, y) faz com o semi-eixo positivo dos xx .

Claramente φ é de classe C^1 e a interpretação geométrica de r, θ tornam claro que a função φ é injectiva (e que a sua imagem é o aberto $\mathbb{R}^2 \setminus \{(x, 0) : x \geq 0\}$). Resta-nos verificar a condição sobre o Jacobiano de φ :

$$\det D\varphi(r, \theta) = \begin{vmatrix} \frac{\partial x}{\partial r} & \frac{\partial x}{\partial \theta} \\ \frac{\partial y}{\partial r} & \frac{\partial y}{\partial \theta} \end{vmatrix} = \begin{vmatrix} \cos \theta & -r \sin \theta \\ \sin \theta & r \cos \theta \end{vmatrix} = r(\cos^2 \theta + \sin^2 \theta) = r$$

que de facto não se anula no domínio de φ . Este valor do Jacobiano da transformação é necessário para mudar um integral duplo para coordenadas polares e portanto deve ser memorizado. O mesmo se aplica ao valor do Jacobiano nos exemplos que se seguem.

As coordenadas (r, θ) chamam-se coordenadas polares e são particularmente úteis quando é necessário descrever situações em que há simetria de rotação em torno da origem. Por exemplo, o conjunto

$$S = \{(x, y) \in \mathbb{R}^2 : 1 \leq x^2 + y^2 < 4, 0 < y \leq x\}$$

corresponde em coordenadas polares ao intervalo

$$\varphi^{-1}(S) = \{(r, \theta) \in]0, +\infty[\times]0, 2\pi[: 1 \leq r < 2, 0 < \theta \leq \frac{\pi}{4}\} = [1, 2[\times]0, \frac{\pi}{4}]$$

Isto é mais facilmente verificável através da interpretação geométrica das coordenadas r, θ mas pode também ser visto através de substituição: nas coordenadas polares a condição $1 \leq x^2 + y^2 < 4$ transforma-se em

$$1 \leq (r \cos \theta)^2 + (r \sin \theta)^2 < 4 \Leftrightarrow 1 < r^2 < 4 \Leftrightarrow 1 < r < 2$$

e a condição $0 < y \leq x$ transforma-se em

$$0 < r \sin \theta \leq r \cos \theta \Leftrightarrow 0 < \sin \theta \leq \cos \theta \Leftrightarrow 0 < \theta \leq \frac{\pi}{4}$$

- (2) **Coordenadas cilíndricas.** Trata-se das coordenadas obtidas em \mathbb{R}^3 mantendo uma das variáveis e passando as duas restantes para coordenadas polares. Por exemplo, as coordenadas cilíndricas em torno do eixo dos zz são definidas pela transformação de coordenadas

$$\varphi:]0, +\infty[\times]0, 2\pi[\times \mathbb{R} \rightarrow \mathbb{R}^3$$

definida por

$$\varphi(r, \theta, z) = (r \cos \theta, r \sin \theta, z)$$

ou, em termos de coordenadas, por

$$\begin{cases} x = r \cos \theta \\ y = r \sin \theta \\ z = z \end{cases}$$

Novamente, a trigonometria elementar identifica r com a distância ao eixo dos zz e θ com o ângulo formado pela projeção do ponto (x, y, z) no plano xy com o semi-eixo positivo dos xx .

É imediato que φ é de classe C^1 e a interpretação geométrica torna claro que φ é uma função injectiva com imagem $\mathbb{R}^3 \setminus \{(x, 0, z) : x \geq 0, z \in \mathbb{R}\}$. O Jacobiano da transformação φ é dado por

$$\det D\varphi(r, \theta, z) = \begin{vmatrix} \cos \theta & -r \sin \theta & 0 \\ \sin \theta & r \cos \theta & 0 \\ 0 & 0 & 1 \end{vmatrix} = r$$

e portanto não se anula no domínio de φ .

As coordenadas cilíndricas adequam-se particularmente à descrição de situações em que há simetria de rotação em torno de um eixo, que se manifesta na eliminação

da variável θ quando descrevemos o objecto em questão em coordenadas cilíndricas. Por exemplo, a região

$$S = \{(x, y, z) \in \mathbb{R}^3 : x^2 + y^2 \leq z \leq 2 - \sqrt{x^2 + y^2}\}$$

em coordenadas cilíndricas fica

$$\varphi^{-1}(S) = \{(r, \theta, z) \in]0, +\infty[\times]0, 2\pi[\times \mathbb{R} : r^2 \leq z \leq 2 - r\}$$

Isto significa que a intersecção de S com um semiplano vertical onde θ é constante, em termos das coordenadas (r, z) nesse semi-plano é a região

$$R = \{(r, z) \in]0, +\infty[\times \mathbb{R} : r^2 \leq z \leq 2 - r\}$$

limitada pelo eixo dos zz , a parábola $z = r^2$ e a recta $z = 2 - r$. Uma vez que esta intersecção não depende de θ , conclui-se que o sólido S se obtém rodando a região R num semiplano vertical em torno do eixo dos zz , ou seja que é a região do espaço limitada inferiormente pelo parabolóide $z = x^2 + y^2$ e superiormente pelo cone $z = 2 - \sqrt{x^2 + y^2}$.

- (3) **Coordenadas esféricas.** Estas coordenadas consistem na distância à origem juntamente com as coordenadas usuais em geografia, nomeadamente a latitude e a longitude (embora as convenções sejam geralmente diferentes das usadas nos mapas). A transformação de coordenadas é

$$\varphi:]0, +\infty[\times]0, 2\pi[\times]0, \pi[\rightarrow \mathbb{R}^3$$

definida por

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

ou, em coordenadas,

$$\begin{cases} x = \rho \cos \phi \cos \theta \\ y = \rho \cos \phi \sin \theta \\ z = \rho \sin \phi \end{cases}$$

Geometricamente, ρ corresponde à distância do ponto (x, y, z) à origem, ϕ é o ângulo (no intervalo $]0, \pi[$) entre o vector (x, y, z) e o semi-eixo positivo dos zz e θ é o ângulo da projecção do vector (x, y, z) no plano xy com o semi-eixo positivo dos xx (no intervalo $]0, 2\pi[$). A coordenada θ tem portanto o mesmo significado que nas coordenadas cilíndricas em torno do eixo dos zz e corresponde à longitude (embora nos mapas o intervalo de variação do ângulo seja normalmente $]-\pi, \pi[$). A coordenada ϕ corresponde à latitude (embora aqui o ângulo seja medido a partir do polo norte e não do equador como é habitual nos mapas).

Claramente φ é de classe C^1 e novamente a trigonometria elementar torna clara a injectividade de φ juntamente com o facto de a sua imagem consistir em $\mathbb{R}^3 \setminus$

$\{(x, 0, z): x \geq 0, z \in \mathbb{R}\}$. Quanto ao Jacobiano, temos

$$\begin{aligned} \det D\varphi(\rho, \theta, \phi) &= \begin{vmatrix} \sin \phi \cos \theta & -\rho \sin \phi \sin \theta & \rho \cos \phi \cos \theta \\ \sin \phi \sin \theta & \rho \sin \phi \cos \theta & \rho \cos \phi \sin \theta \\ \cos \phi & 0 & -\rho \sin \phi \end{vmatrix} \\ &= -\rho^2 (\sin^3 \phi \cos^2 \theta + \sin \phi \cos^2 \phi \sin^2 \theta + \cos^2 \phi \sin \phi \cos^2 \theta + \sin^3 \phi \sin^2 \theta) \\ &= -\rho^2 (\sin \phi \cos^2 \theta + \sin \phi \sin^2 \theta) = -\rho^2 \sin \phi \end{aligned}$$

que não se anula no domínio de φ .

Estas coordenadas são particularmente apropriadas para descrever situações em que há simetria esférica, isto é, simetria de rotação em torno da origem. Por exemplo, a região

$$S = \{(x, y, z): 1 \leq x^2 + y^2 + z^2 \leq 4, x, y, z \geq 0\}$$

transforma-se no intervalo

$$\varphi^{-1}(S) = \{(\rho, \theta, \varphi) \in [0, +\infty[\times]0, 2\pi[\times]0, \pi[: 1 \leq \rho \leq 2, 0 < \theta \leq \frac{\pi}{2}, 0 < \varphi \leq \frac{\pi}{2}\}$$

Isto é claro a partir da interpretação geométrica das coordenadas mas pode também ser visto por substituição das expressões que definem as coordenadas (exercício).

Nota 22.4. Por vezes é útil considerar pequenas variações das coordenadas anteriores. Podemos por exemplo considerar intervalos diferentes para a variação do ângulo θ (de comprimento 2π claro), trocar o papel dos eixos (isto já foi mencionado para as coordenadas cilíndricas mas pode também ser útil para as esféricas), fazer uma translação da origem (se a simetria é de rotação em torno de um ponto que não é a origem), re-escalar as coordenadas, etc...

Podemos agora enunciar o Teorema sobre mudança de variáveis nos integrais múltiplos.

Teorema 22.5 (Teorema de mudança de variáveis). *Sejam U, V abertos limitados de \mathbb{R}^n e $\varphi: U \rightarrow V$ uma transformação de coordenadas. Seja $A \subset V$ um conjunto e $f: A \rightarrow \mathbb{R}$ uma função integrável em A . Então a função $f \circ \varphi | \det D\varphi |$ é integrável em $\varphi^{-1}(A)$ e*

$$\int_{\varphi^{-1}(A)} f \circ \varphi | \det D\varphi | = \int_A f.$$

A fórmula de mudança de variáveis pode ser escrita usando variáveis da seguinte forma. Escrevendo

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

temos

$$\int_{\varphi^{-1}(A)} f(\varphi(x)) \left| \frac{\partial(y_1, \dots, y_n)}{\partial(x_1, \dots, x_n)} \right| dx = \int_A f(y) dy$$

onde se usa a notação sugestiva

$$\left| \frac{\partial(y_1, \dots, y_n)}{\partial(x_1, \dots, x_n)} \right| \stackrel{def}{=} | \det(D\varphi(x)) |$$

para o factor de conversão de volumes n -dimensional.

Uma mnemónica útil para esta fórmula é a seguinte:

- $\varphi^{-1}(A)$ é a região de integração A escrita em termos das variáveis x
- $f(\varphi(x))$ é a função integranda $f(y)$ escrita em termos das variáveis x
- $\left| \frac{\partial(y_1, \dots, y_n)}{\partial(x_1, \dots, x_n)} \right| dx$ é dy escrito em termos das variáveis x .

A demonstração do Teorema 22.5 é bastante elaborada. Mais tarde iremos dar uma breve explicação da fórmula (em particular da ocorrência do factor de conversão de volumes) mas não iremos sequer esboçar a demonstração. Referimos os alunos interessados a [S] para uma demonstração deste resultado.

23. EXEMPLOS DE CÁLCULO DE INTEGRAIS COM MUDANÇA DE VARIÁVEIS

Vejam os alguns exemplos de aplicação do Teorema 22.5

Exemplo 23.1. *Calcular o momento de inércia em relação ao eixo dos xx (ver Secção 19) da placa*

$$S = \{(x, y) \in \mathbb{R}^2 : 0 \leq x \leq y, 1 \leq x^2 + y^2 \leq 4\}$$

que tem densidade de massa dada pela função $f(x, y) = 1$.

A distância ao eixo dos xx é a função $(x, y) \mapsto y^2$ pelo que temos a calcular

$$I_x = \int \int_S y^2 f(x, y) dx dy = \int \int_S y^2 dx dy$$

Em coordenadas polares a região de integração escreve-se $\frac{\pi}{4} \leq \theta \leq \frac{\pi}{2}$ e $1 \leq r \leq 2$ logo o integral transforma-se em

$$I_x = \int_{\frac{\pi}{4}}^{\frac{\pi}{2}} \int_1^2 r^2 \sin^2 \theta r d\theta dr = \int_{\frac{\pi}{4}}^{\frac{\pi}{2}} \frac{1 - \cos(2\theta)}{2} d\theta \int_1^2 r^3 dr = \frac{\pi + 1}{4} \cdot \frac{15}{4}$$

Exemplo 23.2. *Escrever uma expressão para*

$$\int \int_S f(x, y) dx dy$$

em coordenadas polares, onde

$$S = \{(x, y) \in \mathbb{R}^2 : 0 \leq x \leq 1, -x \leq y \leq x\}$$

Temos que escrever a região S em coordenadas polares. Geometricamente é claro que o ângulo θ varia entre $-\frac{\pi}{4}$ e $\frac{\pi}{4}$ (estamos a tomar $]-\pi, \pi[$ para intervalo de variação do ângulo θ de forma a não ter que dividir o integral em dois). Para cada θ fixo, a distância à origem em S varia entre 0 e a distância a origem do ponto da recta $x = 1$ que faz o ângulo θ com o semi-eixo positivo dos xx . Substituindo na equação temos

$$x = 1 \Leftrightarrow r \cos \theta = 1 \Leftrightarrow r = \frac{1}{\cos \theta}$$

logo o integral em questão escreve-se

$$\int_{-\frac{\pi}{4}}^{\frac{\pi}{4}} \int_0^{\frac{1}{\cos \theta}} f(r \cos \theta, r \sin \theta) r dr d\theta$$

Vamos escrever o mesmo integral na ordem de integração oposta. Em S a coordenada r varia entre 0 e a distância à origem dos pontos $(1, \pm 1)$ que é $\sqrt{2}$. É no entanto necessário separar a soma sobre os pontos de S em dois integrais, porque o domínio de variação de θ para r fixo depende do valor de r . Quando r está em $]0, 1]$, o ângulo θ varia entre $-\frac{\pi}{4}$ e $\frac{\pi}{4}$. Mas para r entre 1 e $\sqrt{2}$, o ângulo varia de $-\frac{\pi}{4}$ à linha $x = 1$ e novamente desta linha até $\frac{\pi}{4}$. Temos que achar os pontos da linha $x = 1$ que estão a distância r da origem (para r entre 1 e $\sqrt{2}$):

$$x = 1 \Leftrightarrow r \cos \theta = 1 \Leftrightarrow \theta = \pm \arccos \frac{1}{r}$$

Conclui-se assim que o integral é

$$\int_0^1 \int_{-\frac{\pi}{4}}^{\frac{\pi}{4}} f(r \cos \theta, r \sin \theta) r d\theta dr + \int_1^{\sqrt{2}} \left(\int_{-\frac{\pi}{4}}^{-\arccos \frac{1}{r}} f(r \cos \theta, r \sin \theta) r d\theta + \int_{\arccos \frac{1}{r}}^{\frac{\pi}{4}} f(r \cos \theta, r \sin \theta) r d\theta \right) dr$$

Exemplo 23.3. Calcular o centróide (ver Secção 19) do sólido

$$V = \{(x, y, z) \in \mathbb{R}^3 : \sqrt{3(x^2 + z^2)} \leq y \leq \sqrt{1 - x^2 - z^2}\}$$

Uma vez que a expressão que define V depende apenas de y e de $\sqrt{x^2 + z^2}$, que é a distância do ponto (x, y, z) ao eixo dos yy , é conveniente descrever V usando coordenadas cilíndricas em torno do eixo dos yy :

$$\begin{cases} x = r \cos \theta \\ y = y \\ z = r \sin \theta \end{cases}$$

Nestas coordenadas V é definido pelas condições

$$\sqrt{3}r \leq y \leq \sqrt{1 - r^2}$$

e portanto V obtém-se rodando em torno do eixo dos yy a região compreendida entre a semirecta $y = \sqrt{3}r$ e o arco de circunferência $y = \sqrt{1 - r^2}$. A trigonometria elementar mostra que estas linhas se intersectam no ponto $(r, y) = (\frac{1}{2}, \frac{\sqrt{3}}{2})$ (alternativamente podemos igualar as equações das duas linhas).

Por simetria é claro que o centróide de V estará no eixo dos yy , isto é que $\bar{x} = \bar{z} = 0$. Resta-nos portanto calcular

$$\bar{y} = \frac{\int \int \int_V y dx dy dz}{\int \int \int_V 1 dx dy dz}$$

A expressão para o volume de V (o denominador na expressão anterior) em coordenadas cilíndricas é

$$\begin{aligned} \text{vol}(V) &= \int_0^{2\pi} \int_0^{\frac{1}{2}} \int_{\sqrt{3}r}^{\sqrt{1-r^2}} r dy dr d\theta \\ &= 2\pi \int_0^{\frac{1}{2}} r \sqrt{1-r^2} - \sqrt{3}r^2 dr \\ &= \pi \left(\frac{2\sqrt{2}-1}{3\sqrt{2}} - \frac{1}{4\sqrt{3}} \right) \end{aligned}$$

Vamos calcular o numerador na expressão para \bar{y} usando coordenadas esféricas (tomando para ângulo ϕ , o ângulo com o eixo dos yy)

$$\begin{cases} x = \rho \text{sen } \phi \cos \theta \\ y = \rho \cos \phi \\ z = \rho \text{sen } \phi \text{sen } \theta \end{cases}$$

A expressão para V em coordenadas esféricas é muito fácil de obter a partir da expressão em coordenadas cilíndricas. De facto a coordenada θ é igual em ambos os sistemas de coordenadas enquanto que as restantes variáveis estão relacionadas pelas expressões

$$\begin{cases} r = \rho \text{sen } \phi \\ y = \rho \cos \phi \end{cases}$$

Portanto nestas coordenadas esféricas, V é descrito por

$$0 \leq \rho \leq 1, \quad 0 \leq \phi \leq \frac{\pi}{6}$$

(uma vez que $\frac{\pi}{6} = \arctan \frac{1}{\sqrt{3}}$) e portanto

$$\begin{aligned} \int \int \int_V y \, dx dy dz &= \int_0^{2\pi} \int_0^1 \int_0^{\frac{\pi}{6}} \rho \cos \phi \rho^2 \text{sen } \phi d\phi d\rho d\theta \\ &= 2\pi \int_0^1 \rho^3 d\rho \int_0^{\frac{\pi}{6}} \frac{1}{2} \text{sen } 2\phi d\phi \\ &= 2\pi \cdot \frac{1}{4} \cdot \frac{1 - \frac{1}{2}}{4} = \frac{\pi}{16} \end{aligned}$$

Conclui-se assim que o centróide de V é o ponto

$$(\bar{x}, \bar{y}, \bar{z}) = \left(0, \frac{1}{16 \left(\frac{2\sqrt{2}-1}{3\sqrt{2}} - \frac{1}{4\sqrt{3}} \right)}, 0 \right)$$

Exemplo 23.4. Sendo $S = \{(x, y) \in \mathbb{R}^2 : 1 \leq xy \leq 2, x \leq y \leq 4x, x > 0, y > 0\}$, calcular

$$\int \int_S \text{sen}(xy) dx dy$$

usando uma mudança de coordenadas apropriada.

A região de integração pode escrever-se na forma

$$1 \leq xy \leq 2, \quad 1 \leq \frac{y}{x} \leq 4$$

o que sugere a mudança de variável

$$\begin{cases} u = xy \\ v = \frac{y}{x} \end{cases}$$

A função $\varphi(x, y) = (xy, \frac{y}{x})$ é claramente de classe C^1 e é fácil ver que é injectiva como função do primeiro quadrante para o primeiro quadrante. De facto temos

$$uv = y^2 \Leftrightarrow y = \sqrt{uv}$$

e portanto

$$x = \frac{u}{y} = \sqrt{\frac{u}{v}}$$

logo $\varphi^{-1}(u, v) = (\sqrt{\frac{u}{v}}, \sqrt{uv})$. O Jacobiano de φ é dado por

$$\det D\varphi(x, y) = \begin{vmatrix} y & x \\ -\frac{y}{x^2} & \frac{1}{x} \end{vmatrix} = \frac{2y}{x}$$

logo não se anula no primeiro quadrante (isto é também uma consequência de φ ter a inversa de classe C^1 que determinámos acima, juntamente com a regra de derivação da função composta). Conclui-se que φ é uma mudança de variável do primeiro quadrante de \mathbb{R}^2 para o primeiro quadrante de \mathbb{R}^2 . A região S escrita nas variáveis u, v é o intervalo

$$\varphi^{-1}(S) = \{(u, v) \in \mathbb{R}^2 : 1 \leq u \leq 2, 1 \leq v \leq 4\}$$

logo o Teorema 22.5 implica que

$$\int \int_S \sin(xy) dx dy = \int_1^4 \left(\int_1^2 \sin(u) \left| \frac{\partial(x, y)}{\partial(u, v)} \right| du \right) dv$$

Podemos calcular o Jacobiano da transformação usando a expressão obtida acima para (x, y) em função de (u, v) , ou, alternativamente, notar que, pela regra de derivação da função composta,

$$\left| \frac{\partial(x, y)}{\partial(u, v)} \right| = \frac{1}{\left| \frac{\partial(u, v)}{\partial(x, y)} \right|}$$

Vimos já que

$$\left| \frac{\partial(u, v)}{\partial(x, y)} \right| = \frac{2y}{x} = 2v$$

logo

$$\int \int_S \sin(xy) dx dy = \int_1^4 \left(\int_1^2 \sin(u) \frac{1}{2v} du \right) dv = (\cos 2 - \cos 1) \log 2$$

24. MUDANÇA DE VARIÁVEIS (CONCLUSÃO). A REGRA DE LEIBNIZ

Começamos por dar uma breve explicação da fórmula de mudança de variáveis (Teorema 22.5). Recorde-se primeiro de Álgebra Linear, que o significado geométrico do determinante de uma matriz $n \times n$, A , é o volume do paralelepípedo n -dimensional gerado pelas colunas de A (com sinal positivo se o referencial correspondente tiver a mesma orientação que a base canónica e negativo caso contrário). Assim,

$$|\det A| = \text{volume do paralelepípedo em } \mathbb{R}^n \text{ gerado pelas colunas de } A$$

Nota 24.1. O significado preciso do parágrafo anterior é o seguinte. Se escrevermos

$$[v_1, \dots, v_n] = \{t_1 v_1 + \dots + t_n v_n \in \mathbb{R}^n : t_1, \dots, t_n \in [0, 1]\}$$

para o paralelepípedo n -dimensional cujas arestas são os vectores $v_1, \dots, v_n \in \mathbb{R}^n$, esperaríamos que qualquer noção razoável de volume n -dimensional tivesse as seguintes propriedades:

- (i) Sendo $\{e_1, \dots, e_n\}$ a base canónica de \mathbb{R}^n , $\text{vol}([e_1, \dots, e_n]) = 1$
- (ii) $\text{vol}([v_1, \dots, v_i, \dots, v_j, \dots, v_n]) = \text{vol}([v_1, \dots, v_j, \dots, v_i, \dots, v_n])$
- (iii) Dado $\alpha \in \mathbb{R}$, tem-se $\text{vol}([\alpha v_1, \dots, v_n]) = |\alpha| \text{vol}([v_1, \dots, v_n])$
- (iv) $\text{vol}([v_1, \dots, v_i + \alpha v_j, \dots, v_j, \dots, v_n]) = \text{vol}([v_1, \dots, v_i, \dots, v_j, \dots, v_n])$

(Para se convencer da última propriedade faça um desenho em \mathbb{R}^2). Ora, por um lado, é imediato das propriedades do determinante estudadas em Álgebra Linear que a função $\mathbb{R}^n \times \dots \times \mathbb{R}^n \rightarrow \mathbb{R}_0^+$ que a um n -tuplo (v_1, \dots, v_n) de vectores de \mathbb{R}^n associa o módulo do determinante da matriz que tem v_i como i -ésima coluna tem as propriedades (i)-(iv) acima. Por outro lado, é possível demonstrar que existe uma única função com estas propriedades.

Seja então $\varphi: U \rightarrow V$ uma transformação de coordenadas em \mathbb{R}^n , $A \subset V$ e $f: A \rightarrow \mathbb{R}$ uma função contínua. Suponhamos que podemos aproximar o conjunto $\varphi^{-1}(A)$ por uma união de intervalos I_1, \dots, I_n disjuntos. As imagens destes intervalos por φ aproximam o conjunto A e portanto, pela aditividade do integral temos

$$\int_A f(y) dy \sim \sum_{i=1}^n \int_{\varphi(I_i)} f(y) dy$$

Se os intervalos I_i forem muito pequenos, o mesmo sucederá com os conjuntos $\varphi(I_i)$. Sendo a função f contínua, ela é aproximadamente constante em em cada conjunto $\varphi(I_i)$ e então a soma acima é aproximadamente

$$(7) \quad \sum_{i=1}^n \text{vol}(\varphi(I_i)) f(y_i)$$

com $y_i \in \varphi(I_i)$. Sendo $x_i = \varphi^{-1}(y_i)$ o ponto de I_i correspondente a y_i , o conjunto $\varphi(I_i)$ deve (desde que I_i seja muito pequeno) ser aproximado pela imagem de I_i através da aproximação linear de φ perto de x_i dada pela derivada de φ em x_i , que é,

$$x \mapsto L_{x_i}(x) = y_i + D\varphi(x_i)(x - x_i)$$

O conjunto $L_{x_i}(I_i)$ é um paralelepípedo com arestas paralelas às colunas da matriz $D\varphi(x_i)$. A relação entre o módulo do determinante e o volume de paralelepípedos que referimos antes diz-nos que devemos ter

$$\text{vol}(L_{x_i}(I_i)) = |\det D\varphi(x_i)| \text{vol}(I_i)$$

Obtemos assim a seguinte aproximação para (7):

$$\sum_{i=1}^n \text{vol}(I_i) |\det D\varphi(x_i)| f(\varphi(x_i))$$

É natural esperar que, à medida que o tamanho dos intervalos I_i tende para 0, esta soma se aproxime de

$$\int_{\varphi^{-1}(A)} f(\varphi(x)) |\det D\varphi(x)| dx$$

Isto conclui o nosso argumento de plausibilidade para a fórmula do Teorema 22.5. Novamente sugere-se aos alunos interessados que consultem [S] para uma demonstração do Teorema.

24.2. O volume da bola unitária em \mathbb{R}^n . Para terminar a nossa discussão do Teorema de Mudança de Variáveis vamos fazer um cálculo útil e algo surpreendente.

Exemplo 24.3 (Cálculo do volume da bola de raio 1 em \mathbb{R}^n). *Seja*

$$B_n = \{(x_1, \dots, x_n) \in \mathbb{R}^n : x_1^2 + \dots + x_n^2 \leq 1\}$$

a bola de raio 1 em \mathbb{R}^n . Por definição o volume n -dimensional é dado por

$$\text{vol}(B_n) = \int_{B_n} 1$$

A projeção de B_n no plano x_1x_2 é o disco unitário $\{(x_1, x_2) \in \mathbb{R}^2 : x_1^2 + x_2^2 \leq 1\}$ e o conjunto dos pontos de B_n que têm primeiras duas coordenadas (x_1, x_2) neste disco é

$$\{(x_1, x_2, x_3, \dots, x_n) \in \mathbb{R}^n : x_3^2 + \dots + x_n^2 \leq 1 - x_1^2 - x_2^2\}$$

logo o Teorema de Fubini garante que

$$\int_{B_n} 1 = \int_{\{(x_1, x_2) \in \mathbb{R}^2 : x_1^2 + x_2^2 \leq 1\}} \left(\int_{(x_3, \dots, x_n) \in \mathbb{R}^{n-2} : x_3^2 + \dots + x_n^2 \leq 1 - x_1^2 - x_2^2} 1 dx_3 \cdots dx_n \right) dx_1 dx_2$$

O integral dentro de parêntesis é o volume da bola de raio $\sqrt{1 - x_1^2 - x_2^2}$ em \mathbb{R}^{n-2} . É fácil ver (efectuando a mudança de variável $x \mapsto \lambda x$) que se $\lambda \in \mathbb{R}^+$, $A \subset \mathbb{R}^k$ e $\lambda A = \{\lambda x : x \in A\}$ então $\text{vol}(\lambda A) = \lambda^k \text{vol}(A)$. Donde se conclui que

$$\text{vol}(B_n) = \int_{\{(x_1, x_2) \in \mathbb{R}^2 : x_1^2 + x_2^2 \leq 1\}} \text{vol}(B_{n-2}) (1 - x_1^2 - x_2^2)^{\frac{n}{2}-1} dx_1 dx_2$$

e mudando para coordenadas polares

$$\begin{cases} x_1 = r \cos \theta \\ x_2 = r \sin \theta \end{cases}$$

este integral transforma-se em

$$\text{vol}(B_n) = \int_0^{2\pi} \int_0^1 \text{vol}(B_{n-2}) r (1-r^2)^{\frac{n}{2}-1} dr d\theta = 2\pi \text{vol}(B_{n-2}) \frac{(1-r^2)^{\frac{n}{2}}}{-2 \left(\frac{n}{2}\right)} \Bigg|_{r=0}^{r=1} = \frac{2\pi}{n} \text{vol}(B_{n-2})$$

A fórmula anterior permite calcular os volumes de B_n por recorrência a partir de $\text{vol}(B_1) = 2$ e $\text{vol}(B_2) = \pi$. Para os primeiros valores de n obtemos

n	1	2	3	4	5	6
$\text{vol}(B_n)$	2	π	$\frac{4\pi}{3}$	$\frac{\pi^2}{2}$	$\frac{8\pi^2}{15}$	$\frac{\pi^3}{6}$

Não é difícil demonstrar por indução a validade da fórmula geral:

$$\text{vol}(B_n) = \begin{cases} \frac{\pi^{n/2}}{(n/2)!} & \text{se } n \text{ é par,} \\ \frac{2^{(n+1)/2} \pi^{(n-1)/2}}{n!!} & \text{se } n \text{ é ímpar.} \end{cases}$$

onde $n!! = n(n-2) \cdots 3 \cdot 1$. Note-se que

$$\lim_{n \rightarrow \infty} \text{vol}(B_n) = 0$$

Isto diz, em particular que, à medida que n tende para infinito, o volume de uma bola n -dimensional de raio 1 é negligível face ao do (hiper)cubo $[0, 1] \times \cdots \times [0, 1]$. Isto é à primeira vista surpreendente, mas torna-se menos surpreendente se pensarmos que a diagonal do hipercubo unitário em \mathbb{R}^n mede \sqrt{n} .

24.4. A regra de Leibniz. Ocorrem frequentemente nas aplicações funções definidas por integrais que dependem de parâmetros, isto é por expressões da forma

$$\varphi(t) = \int_A f(x, t) dx$$

Uma tal expressão diz-se um *integral paramétrico*. Vamos agora ver condições suficientes para a continuidade e diferenciabilidade de tais funções.

Proposição 24.5. *Seja $A \subset \mathbb{R}^n$ um compacto tal que front A tem conteúdo nulo, e $U \subset \mathbb{R}^m$ um aberto. Seja $f: A \times U \rightarrow \mathbb{R}$ uma função contínua e $\varphi: U \rightarrow \mathbb{R}$ a função definida pela expressão*

$$\varphi(t) = \int_A f(x, t) dx$$

Então

(i) A função φ é contínua em U .

(ii) **Regra de Leibniz:** Se $\frac{\partial f}{\partial t_i}$ existe e é contínua em $A \times U$, então $\frac{\partial \varphi}{\partial t_i}$ existe em U e

$$\frac{\partial \varphi}{\partial t_i}(t) = \int_A \frac{\partial f}{\partial t_i}(x, t) dx$$

Note-se que todas os integrais que aparecem no enunciado anterior estão bem definidos pois o Teorema 21.5 garante que as funções contínuas são integráveis em A .

Exemplo 24.6. Seja $\varphi: \mathbb{R} \rightarrow \mathbb{R}$ a função definida por

$$\varphi(t) = \int_0^1 e^{tx^2} dx$$

Então a regra de Leibniz garante que φ é diferenciável e

$$\varphi'(t) = \int_0^1 x^2 e^{tx^2} dx.$$

Exemplo 24.7. Seja $\varphi: \mathbb{R} \rightarrow \mathbb{R}$ a função definida por

$$\varphi(t) = \int \int_{\{(x,y) \in \mathbb{R}^2: x^2+y^2 \leq t^2\}} e^{t(x^2+y^2)} dx dy$$

Vamos achar uma expressão para a derivada de φ .

O integral pode ser expresso em coordenadas polares da seguinte forma

$$\varphi(t) = \int_0^t \left(\int_0^{2\pi} e^{tr^2} r d\theta \right) dr = 2\pi \int_0^t r e^{tr^2} dr$$

Neste integral paramétrico há dependência do parâmetro não apenas na função integranda mas também na região de integração, pelo que não se pode aplicar directamente a regra de Leibniz. Este (e outros problemas semelhantes) podem ser resolvidos recorrendo à regra da cadeia. Definimos a função

$$\psi(u, v) = 2\pi \int_0^u r e^{vr^2} dr$$

que “separa” a dependência do parâmetro em duas componentes distintas - a região de integração por um lado, e a função integranda por outro. Claramente

$$\varphi(t) = \psi(t, t)$$

logo, desde que ψ seja diferenciável temos, pela regra da cadeia,

$$(8) \quad \varphi'(t) = \frac{\partial \psi}{\partial u}(t, t) + \frac{\partial \psi}{\partial v}(t, t)$$

A derivada parcial de ψ em ordem a u existe pelo Teorema Fundamental do Cálculo:

$$\frac{\partial \psi}{\partial u}(u, v) = 2\pi u e^{vu^2}$$

enquanto que a derivada parcial de ψ em ordem a v existe e pode ser calculada pela regra de Leibniz:

$$\frac{\partial \psi}{\partial v}(u, v) = 2\pi \int_0^u r^3 e^{vr^2} dr$$

Como ambas estas funções são contínuas (para a continuidade da segunda usamos a continuidade do integral paramétrico), a função ψ é diferenciável e conclui-se então que φ é diferenciável e, por (8),

$$\varphi'(t) = 2\pi te^{t^3} + 2\pi \int_0^t r^3 e^{tr^2} dr$$

25. O TEOREMA DA FUNÇÃO INVERSA

A demonstração da regra de Leibniz é uma consequência simples do seguinte resultado fundamental que está também na base da demonstração dos Teoremas 20.2 e 21.5.

Teorema 25.1 (Teorema de Heine-Cantor). *Seja $A \subset \mathbb{R}^n$ compacto e $f: A \rightarrow \mathbb{R}$ contínua. Então f é uniformemente contínua em A , isto é, para todo $\epsilon > 0$ existe $\delta > 0$ tal que para todos os $x, y \in A$ com $\|x - y\| < \delta$ se tem $|f(x) - f(y)| < \epsilon$.*

Dem. Suponhamos que a conclusão não é válida. Então existe $\epsilon > 0$ tal que para todo $\delta > 0$ existem pontos $x, y \in A$ com $\|x - y\| < \delta$ e $|f(x) - f(y)| \geq \epsilon$. Tomando $\delta = \frac{1}{k}$ podemos assim escolher sucessões x_k, y_k em A com $\|x_k - y_k\| < \frac{1}{k}$ e $|f(x_k) - f(y_k)| \geq \epsilon$. Como A é compacto, a sucessão x_k tem uma subsucessão x_{k_m} convergente para um ponto $a \in A$. Uma vez que a distância entre x_k e y_k tende para 0, conclui-se que a subsucessão y_{k_m} converge também para a . Mas f é contínua em a logo temos

$$\lim_{m \rightarrow \infty} \|f(x_{k_m}) - f(y_{k_m})\| = \|f(a) - f(a)\| = 0$$

o que contradiz a desigualdade $|f(x_{k_m}) - f(y_{k_m})| \geq \epsilon$ válida para todo o m . \square

Nota 25.2. *É instrutivo comparar a definição de continuidade uniforme de f num conjunto A com a continuidade de f em todos os pontos de $x \in A$. A primeira condição é mais forte. Dado $\epsilon > 0$, a continuidade para todo o x permite encontrar δ , dependendo de x tal que $\|x - y\| < \delta \Rightarrow |f(x) - f(y)| < \epsilon$. A continuidade uniforme permite encontrar δ que é independente de x . A continuidade uniforme diz assim que a função “é contínua da mesma maneira em todos os pontos de A ”.*

Um exemplo de uma função que é contínua mas não uniformemente contínua é a função $f: \mathbb{R} \rightarrow \mathbb{R}$ definida por $f(x) = x^2$. Mais geralmente, uma função diferenciável num intervalo de \mathbb{R} é uniformemente contínua nesse intervalo sse a sua derivada é limitada nesse intervalo (exercício).

Dem. da Proposição 24.5. (i) Queremos ver que o integral paramétrico φ é contínuo num ponto qualquer $t \in U$ dado. Podemos escolher $r > 0$ tal que o compacto $\overline{B_r}(t) \subset U$. A função $f(x, t)$ é contínua no compacto $A \times \overline{B_r}(t)$ logo, pelo Teorema de Heine-Cantor, para cada $\epsilon > 0$ podemos escolher $\delta > 0$ tal que para todos os $x \in A$ e $s \in \overline{B_r}(t)$ temos

$$|f(x, t) - f(x, s)| < \epsilon$$

Mas então

$$|\varphi(t) - \varphi(s)| = \left| \int_A f(x, t) dx - \int_A f(x, s) dx \right| \leq \int_A |f(x, t) - f(x, s)| dx \leq \int_A \epsilon = \epsilon \text{ vol}(A)$$

o que mostra que φ é contínua.

- (ii) A derivada parcial $\frac{\partial \varphi}{\partial t_i}$ do integral paramétrico em $t \in U$ é, por definição, escrevendo e_i para o i -ésimo vector da base canónica de \mathbb{R}^m ,

$$\begin{aligned} \frac{\partial \varphi}{\partial t_i}(t) &= \lim_{h \rightarrow 0} \frac{\varphi(t + he_i) - \varphi(t)}{h} \\ &= \lim_{h \rightarrow 0} \int_A \frac{f(x, t + he_i) - f(x, t)}{h} dx \end{aligned}$$

Para provar a regra de Leibniz precisamos de comparar os integrais na última linha com $\int_A \frac{\partial f}{\partial t_i}(x, t) dx$. Pelo Teorema de Lagrange, para cada $x \in A$, existe ξ_x entre 0 e h tal que

$$\frac{f(x, t + he_i) - f(x, t)}{h} = \frac{\frac{\partial f}{\partial t_i}(x, t + \xi_x e_i)h}{h} = \frac{\partial f}{\partial t_i}(x, t + \xi_x e_i)$$

logo

$$\left| \frac{f(x, t + he_i) - f(x, t)}{h} - \frac{\partial f}{\partial t_i}(x, t) \right| = \left| \frac{\partial f}{\partial t_i}(x, t + \xi_x e_i) - \frac{\partial f}{\partial t_i}(x, t) \right|$$

Como $\frac{\partial f}{\partial t_i}$ é contínua, o Teorema de Heine-Cantor garante que, para h suficientemente pequeno, podemos tornar esta última quantidade menor que um qualquer ϵ dado. Para esses valores de h teremos

$$\left| \int_A \frac{f(x, t + he_i) - f(x, t)}{h} dx - \int_A \frac{\partial f}{\partial t_i}(x, t) dx \right| \leq \epsilon \text{vol}(A)$$

e portanto,

$$\lim_{h \rightarrow 0} \int_A \frac{f(x, t + he_i) - f(x, t)}{h} dx = \int_A \frac{\partial f}{\partial t_i}(x, t) dx$$

conforme queríamos demonstrar. □

25.3. O Teorema da Função Inversa. Vamos agora regressar ao estudo do Cálculo Diferencial em \mathbb{R}^n para explicar um dos teoremas fundamentais, que dá um critério baseado no cálculo diferencial para que uma função seja (localmente) invertível.

Primeiro recorde-se que dados dois conjuntos quaisquer X, Y , uma função $f: X \rightarrow Y$, diz-se *injectiva* se leva pontos diferentes de X para pontos diferentes de Y ou, equivalentemente, se

$$f(x_1) = f(x_2) \Rightarrow x_1 = x_2$$

Recorde-se ainda que f é injectiva sse f é *invertível*, isto é, se existe $g: f(X) \rightarrow X$ (onde $f(X) = \{f(x): x \in X\} \subset Y$ é a imagem de f) tal que

- (i) $g \circ f = \text{Id}_X$, isto é, $g(f(x)) = x$ para todo o $x \in X$,
- (ii) $f \circ g = \text{Id}_{f(X)}$, isto é, $f(g(y)) = y$ para todo o $y \in f(X)$.

Além disso a *função inversa* $g: f(X) \rightarrow X$, que se denota habitualmente por f^{-1} , é dada por

$$f^{-1}(y) = (\text{único } x \text{ tal que } f(x) = y).$$

Exemplo 25.4. Tomando $X = Y = \mathbb{R}$, a função $f(x) = x^2$ não é injectiva (por exemplo $f(1) = f(-1) = 1$). A função $f(x) = x^3$ é injectiva, sendo a função inversa $g(y) = \sqrt[3]{y}$.

Teorema 25.5 (Teorema da Função Inversa). *Seja $U \subset \mathbb{R}^n$ um aberto, x um ponto de U , e $f: U \rightarrow \mathbb{R}^n$ uma função de classe C^1 . Se $\det Df(x) \neq 0$, então existe um aberto V contendo x tal que*

- (1) *A restrição de f a V é injectiva.*
- (2) *$f(V)$ é aberto*
- (3) *A função inversa $f^{-1}: f(V) \rightarrow V$ é de classe C^1*

Além disso, temos a seguinte regra de derivação para a função inversa:

$$Df^{-1}(f(x)) = Df(x)^{-1}$$

As condições (i)-(iii) no Teorema abreviam-se dizendo que f é *localmente invertível* perto de x ou que f tem uma inversa local junto a x .

Nota 25.6. *Recorde-se de Álgebra Linear, que a condição $\det Df(x) \neq 0$ é equivalente a dizer que a derivada $Df(x): \mathbb{R}^n \rightarrow \mathbb{R}^n$ é invertível. Assim, em termos gerais, o Teorema 25.5 diz que uma função de classe C^1 é invertível (ou injectiva) junto a um ponto em que a sua aproximação linear, dada pela derivada, o seja. Isto é bastante plausível mas há alguma subtilidade neste resultado. Ele seria falso se não exigíssemos a continuidade da derivada como é possível mostrar através de um exemplo mesmo quando $n = 1$.*

Nota 25.7. *Note-se que a fórmula para a derivada é uma consequência imediata da regra de derivação da função composta. De facto se f tem uma inversa diferenciável g na vizinhança de um ponto x , então*

$$g \circ f = \text{Id}_V \Leftrightarrow g(f(x)) = x \text{ para todo o } x \in V$$

Tendo em conta que a derivada da identidade é a identidade (em todos os pontos) temos, pela regra de derivação da função composta:

$$Dg(f(x))Df(x) = \text{Id}$$

o que mostra que

$$Dg(f(x)) = Df(x)^{-1}$$

Em particular, a condição $\det Df(x) \neq 0$ é necessária para a existência de uma inversa diferenciável perto de x .

Não vamos demonstrar o Teorema 25.5 porque a demonstração é muito longa (apesar de ser interessante e este ser um dos Teoremas mais importantes do Cálculo Diferencial). Sugere-se aos alunos interessados que resolvam a ficha suplementar sobre o Teorema da Função Inversa, disponível na página da cadeira, que contém uma demonstração guiada deste importante Teorema.

26. O TEOREMA DA FUNÇÃO INVERSA (CONCLUSÃO). O TEOREMA DA FUNÇÃO IMPLÍCITA.

Se $f: I \rightarrow \mathbb{R}$ é uma função diferenciável num intervalo $I \subset \mathbb{R}$ e a derivada f' não se anula, então f é uma função estritamente monótona e portanto injectiva. Quando $n > 1$, a injectividade da derivada em cada ponto de um aberto já não garante a injectividade global como mostra o seguinte exemplo importante:

Exemplo 26.1 (A exponencial complexa). *Seja $f: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ a função definida por*

$$f(x, y) = (e^x \cos y, e^x \operatorname{sen} y)$$

Claramente a função f é de classe C^1 . Uma vez que

$$Df(x, y) = \begin{bmatrix} e^x \cos y & -e^x \operatorname{sen} y \\ e^x \operatorname{sen} y & e^x \cos y \end{bmatrix}$$

temos

$$\det Df(x, y) = e^x(\cos^2 y + \operatorname{sen}^2 y) = e^x$$

Como o Jacobiano nunca se anula, o Teorema da Função Inversa garante que f é localmente invertível em torno de cada $(x, y) \in \mathbb{R}^2$. No entanto é claro que f não é injectiva, pois é periódica com período 2π em y : $f(x, y + 2\pi) = f(x, y)$. Por exemplo $f(0, 0) = f(0, 2\pi) = (1, 0)$.

Usando coordenadas polares no conjunto de chegada é fácil entender geometricamente o efeito da transformação f . Cada recta horizontal com ordenada y , é enviada na semi-recta que faz um ângulo y com o semi-eixo dos xx positivo. Assim, f “enrola” cada faixa horizontal de largura 2π no plano (x, y) em torno da origem no plano imagem. A imagem de f é $\mathbb{R}^2 \setminus \{(0, 0)\}$ e a função é injectiva em qualquer faixa horizontal aberta de largura $\leq 2\pi$.

Nota 26.2. *A função do exemplo anterior pode ser encarada como uma função de \mathbb{C} para \mathbb{C} usando a identificação usual do conjunto dos números complexos com o plano. Temos então $f(x + iy) = e^x \cos y + ie^x \operatorname{sen} y$. A expressão anterior define a exponencial do número complexo $x + iy$ e, como tal, f é uma das funções mais importantes em Matemática. Será estudada detalhadamente no próximo semestre.*

Exemplo 26.3. *Mostrar que a função $f: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ definida por $f(x, y) = (xy, x + y^2)$ é localmente invertível numa vizinhança do ponto $(1, 2)$ e calcular $Df^{-1}(f(1, 2))$ (onde f^{-1} é a inversa local em torno de $(1, 2)$).*

Claramente a função f é de classe C^1 em \mathbb{R}^2 . Temos

$$Df(x, y) = \begin{bmatrix} y & x \\ 1 & 2y \end{bmatrix}$$

logo

$$\det Df(x, y) = 2y^2 - x$$

Em $(x, y) = (1, 2)$ temos $\det Df(1, 2) = 7 \neq 0$ logo o Teorema da Função Inversa garante a invertibilidade em torno do ponto $(1, 2)$. Como $f(1, 2) = (2, 5)$, a regra de derivação da função inversa diz que

$$Df^{-1}(2, 5) = Df(1, 2)^{-1} = \begin{bmatrix} 2 & 1 \\ 1 & 4 \end{bmatrix}^{-1} = \frac{1}{7} \begin{bmatrix} 4 & -1 \\ -1 & 2 \end{bmatrix}$$

26.4. O Teorema da Função Implícita. Uma função $(u, v) = f(x, y)$ é invertível sse o sistema

$$\begin{cases} u = f_1(x, y) \\ v = f_2(x, y) \end{cases}$$

pode ser resolvido de forma única em ordem às variáveis x, y . Podemos portanto interpretar o Teorema da Função Inversa como dando condições mediante as quais é possível resolver um sistema como o indicado acima em ordem às variáveis x, y . O Teorema da função implícita que agora iremos estudar dá condições análogas (baseadas no cálculo diferencial) para que se possa resolver uma equação (ou um sistema de equações) em ordem a uma variável (ou variáveis). Começemos por discutir o caso de uma equação de duas variáveis.

A equação

$$x^2 + xy - 2 = 0$$

pode facilmente ser resolvida em ordem a y . Obtemos

$$y = \frac{2 - x^2}{x}$$

A maior parte das equações não pode ser resolvida analiticamente. Por exemplo, não é possível resolver a equação

$$(9) \quad e^y + xy - 2 = 0$$

explicitamente em ordem a y . No entanto seria de esperar que para cada valor de x exista um valor de y , pelo menos se estivermos próximo de uma solução como por exemplo $(x, y) = (0, \log 2)$. No caso de uma equação linear como

$$ax + by = 3$$

(com a, b parâmetros reais) sabemos que a equação pode ser resolvida em ordem a y desde que $b \neq 0$:

$$y = \frac{1}{b} (3 - ax)$$

O Teorema da Função Implícita diz que uma equação não linear da forma

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

pode ser resolvida em ordem a y perto de uma solução (x_0, y_0) desde que a equação linear que melhor aproxima esta equação, nomeadamente

$$\frac{\partial f}{\partial x}(x_0, y_0)(x - x_0) + \frac{\partial f}{\partial y}(x_0, y_0)(y - y_0) = 0$$

possa ser resolvida em ordem a y , isto é, desde que $\frac{\partial f}{\partial y}(x_0, y_0) \neq 0$. No caso da equação (9) acima temos

$$\frac{\partial}{\partial x}(e^y + xy - 2) = y, \quad \frac{\partial}{\partial y}(e^y + xy - 2) = e^y + x$$

logo a equação linear que melhor a aproxima perto de $(0, \log 2)$ é

$$(10) \quad \log(2)x + 2(y - \log 2) = 0$$

que pode ser resolvida em ordem a y . O Teorema da Função Implícita garante que a equação (9) pode também ser resolvida em ordem y . Isto, é existe uma função $g(x)$ definida perto de $x = 0$ tal que, para (x, y) perto de $(0, \log 2)$,

$$e^y + xy - 2 = 0 \Leftrightarrow y = g(x)$$

Esta função $g(x)$ é *definida implicitamente* pela equação (9) e chama-se a *função implícita*.

Teorema 26.5 (Teorema da função implícita para funções de \mathbb{R}^2 para \mathbb{R}). *Seja $U \subset \mathbb{R}^2$ um conjunto aberto, $f: U \rightarrow \mathbb{R}$ uma função de classe C^1 e $(x_0, y_0) \in U$ tal que $f(x_0, y_0) = 0$. Se $\frac{\partial f}{\partial y}(x_0, y_0) \neq 0$, então existem um aberto $V \subset \mathbb{R}$ contendo x_0 , um aberto W de \mathbb{R} contendo y_0 , com $V \times W \subset U$, e uma função $g: V \rightarrow W$ de classe C^1 tal que*

$$(x, y) \in V \times W \text{ e } f(x, y) = 0 \Leftrightarrow y = g(x)$$

Nota 26.6. *Claro que a variável y não desempenha nenhum papel especial nesta história. Podemos trocar x e y e obter um critério para que seja possível resolver a equação em ordem a x .*

O valor 0 para o conjunto de nível também não desempenha nenhum papel especial. O conjunto de nível c de uma função $f(x, y)$ é o conjunto de nível 0 da função $f(x, y) - c$ pelo que o enunciado do Teorema permanece válido se substituirmos o valor 0 por qualquer outra constante real.

Geometricamente, o conjunto definido por uma equação

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

é uma curva no plano. Se ela passar no ponto (x_0, y_0) , isto é, se $f(x_0, y_0) = 0$, então a recta de equação

$$\frac{\partial f}{\partial x}(x_0, y_0)(x - x_0) + \frac{\partial f}{\partial y}(x_0, y_0)(y - y_0) = 0$$

é tangente à curva no ponto (x_0, y_0) . Resolver a equação $f(x, y) = 0$ em ordem a y de forma a obter $y = g(x)$ significa descrever o conjunto de nível $f(x, y) = 0$ como o gráfico de uma função de x :

$$\{(x, y) \in V \times W : f(x, y) = 0\} = \{(x, g(x)) : x \in V\}$$

A condição $\frac{\partial f}{\partial y}(x_0, y_0) \neq 0$ é que a segunda componente de $\nabla f(x_0, y_0)$ não se anule, isto é que o vector perpendicular ao conjunto de nível em (x_0, y_0) não seja horizontal, ou equivalentemente, que a tangente ao conjunto de nível em (x_0, y_0) não seja uma reta vertical (o que é claramente desejável para garantir que o conjunto de nível é o gráfico de uma função de x).

As derivadas de uma função definida implicitamente por uma equação podem ser calculadas recorrendo à regra da derivação da função composta. De facto, nas condições do enunciado do Teorema 26.5 temos, para todo o $x \in V$,

$$f(x, g(x)) = 0$$

e derivando esta equação em ordem a x , obtemos uma equação que permite calcular $g'(x)$:

$$\frac{d}{dx}(f(x, g(x))) = 0 \Leftrightarrow \frac{\partial f}{\partial x}(x, g(x)) + \frac{\partial f}{\partial y}(x, g(x))g'(x) = 0 \Leftrightarrow g'(x) = -\frac{\frac{\partial f}{\partial x}(x, g(x))}{\frac{\partial f}{\partial y}(x, g(x))}$$

Avaliando esta expressão em $x = x_0$, (donde $g(x_0) = y_0$) temos

$$g'(x_0) = -\frac{\frac{\partial f}{\partial x}(x_0, y_0)}{\frac{\partial f}{\partial y}(x_0, y_0)}$$

Exemplo 26.7. Voltando ao exemplo (9). A função $f(x, y) = e^y + xy - 2$ é de classe C^1 , temos $f(0, \log 2) = 0$ e $\frac{\partial f}{\partial y}(0, \log 2) = 2 \neq 0$ pelo que o Teorema da Função Implícita garante a existência de uma função $g(x)$ definida numa vizinhança de 0, tal que para (x, y) perto de $(0, \log 2)$ se tem

$$e^y + xy - 2 = 0 \Leftrightarrow y = g(x)$$

A expressão para a derivada obtida acima diz que

$$g'(0) = -\frac{\log 2}{2}$$

27. REVISÕES PARA O TESTE

28. O TEOREMA DA FUNÇÃO IMPLÍCITA

Na aula passada discutimos o Teorema da Função Implícita no caso de funções de \mathbb{R}^2 para \mathbb{R} . O exemplo seguinte (em que conseguimos fazer todas as contas explicitamente) é um bom exemplo a ter em mente quando precisamos de nos lembrar do enunciado do Teorema.

Exemplo 28.1. Seja $F: \mathbb{R}^2 \rightarrow \mathbb{R}$ a função definida por $F(x, y) = x^2 + y^2 - 1$, cujo conjunto de nível 0 é a circunferência de raio 1.

$$F(x, y) = 0 \Leftrightarrow x^2 + y^2 = 1$$

Podemos resolver a equação em ordem explicitamente:

$$x^2 + y^2 = 1 \Leftrightarrow y = \pm\sqrt{1 - x^2}$$

e vemos que em torno de cada ponto da circunferência com ordenada não nula, a circunferência é o gráfico de uma função de classe C^1 (a função $x \mapsto \sqrt{1 - x^2}$ se a ordenada é positiva e $x \mapsto -\sqrt{1 - x^2}$ se a ordenada é negativa). Por outro lado em torno dos pontos $(\pm 1, 0)$, a circunferência não é o gráfico de uma função de x (para cada $x < 1$ suficientemente próximo de 1 há dois pontos na circunferência com abcissa x). Os pontos $(\pm 1, 0)$

são precisamente os pontos da circunferência nos quais $\frac{\partial F}{\partial y} = 2y$ se anula e, portanto, não se verificam as hipóteses do Teorema 26.5.

Este exemplo é esclarecedor se não nos lembrarmos qual é a derivada parcial que não se deve anular no enunciado do Teorema 26.5.

Note-se que, analogamente, os pontos em torno dos quais a circunferência não é o gráfico de uma função de y (equivalentemente, em que a equação $x^2 + y^2 = 1$ não pode ser resolvida de forma unívoca em ordem a x) são os pontos em que $\frac{\partial F}{\partial x} = 2x$ se anula.

Vejam agora a versão geral do Teorema da Função Implícita relativo à resolução de sistemas de várias equações a várias incógnitas. Para motivar o enunciado consideremos o caso de um sistema linear de 2 equações a 3 variáveis. Por exemplo,

$$\begin{cases} 2x + 3y + z = 2 \\ -x + 2y + 4z = 1 \end{cases}$$

O que podemos esperar de um tal sistema é que seja possível exprimir duas das variáveis (tantas quantas as equações) em termos das restantes. Por exemplo, para exprimir x, y em função de z podemos primeiro considerar o sistema equivalente

$$\begin{cases} 2x + 3y = 2 - z \\ -x + 2y = 1 - 4z \end{cases}$$

A Álgebra Linear dá-nos um critério para que se possa achar x, y únicos, dado z . Isto acontece sse o determinante da matriz

$$\begin{bmatrix} 2 & 3 \\ -1 & 2 \end{bmatrix}$$

dos coeficientes das variáveis x, y é diferente de 0 (ou equivalentemente se esta matriz é invertível). Um sistema não linear de m equações a n incógnitas, pode ser escrito na forma

$$F(x) = 0$$

onde x denota um vector de \mathbb{R}^n e F é uma função definida em \mathbb{R}^n (ou num seu subconjunto) que toma valores em \mathbb{R}^m . Perto de uma solução x_0 do sistema, o sistema $F(x) = 0$ pode ser aproximado por um sistema linear, nomeadamente,

$$DF(x_0)(x - x_0) = 0$$

e a Álgebra Linear diz-nos quando é que este sistema linear pode ser resolvido em ordem a m das componentes de x (quando o determinante da matriz $m \times m$ dos coeficientes destas componentes é diferente de 0). O Teorema da Função Implícita diz-nos que, quando isto acontece, isto é, quando a aproximação linear do sistema $F(x) = 0$ pode ser resolvida em ordem a m das variáveis, o mesmo é verdade para o sistema não linear desde que estejamos suficientemente perto da solução x_0 .

Teorema 28.2 (Teorema da Função Implícita). *Seja $U \subset \mathbb{R}^n$ um aberto e $F: U \rightarrow \mathbb{R}^m$ uma função de classe C^1 . Escrevemos os vectores de \mathbb{R}^n na forma (x, y) com x um vector*

de \mathbb{R}^{n-m} formado por quaisquer $(n-m)$ coordenadas (não necessariamente as primeiras) e y o vector de \mathbb{R}^m formado pelas restantes coordenadas.

Dado $(x_0, y_0) \in U$ com $F(x_0, y_0) = 0$ e $\det \frac{\partial F}{\partial y}(x_0, y_0) \neq 0$ (onde $\frac{\partial F}{\partial y}$ denota a matriz $m \times m$ formada pelas derivadas parciais de F em ordem às coordenadas de y), existem abertos $V \subset \mathbb{R}^{n-m}$ contendo x_0 e $W \subset \mathbb{R}^m$ contendo y_0 e uma função $g: V \rightarrow W$ de classe C^1 tais que $V \times W \subset U$ e

$$((x, y) \in V \times W \text{ e } F(x, y) = 0) \Leftrightarrow y = g(x)$$

Nota 28.3. Geometricamente, o enunciado do Teorema diz que perto do ponto (x_0, y_0) , o conjunto de nível definido pelo sistema $F(x, y) = 0$ é da forma

$$\{(x, g(x)): x \in V\}$$

com V um aberto de \mathbb{R}^{n-m} e g de classe C^1 , ou seja, que o conjunto de nível é o gráfico de uma função de classe C^1 de $n-m$ das variáveis.

Nota 28.4. A condição $\det \frac{\partial F}{\partial y} \neq 0$ no Teorema 28.2 pode ser descrita da seguinte forma: “o determinante das derivadas parciais da função F em ordem às variáveis que queremos expressar em termos das restantes tem de ser diferente de 0”.

Nota 28.5. As derivadas da função implícita $g(x)$ cuja existência o Teorema 28.2 garante podem ser calculadas recorrendo à regra da cadeia. De facto temos, por definição da função $g(x)$, que ela satisfaz a condição

$$F(x, g(x)) = 0$$

Derivando esta relação em ordem às componentes de x obtêm-se equações envolvendo as derivadas parciais de g que as determinam completamente.

Exemplo 28.6. Mostrar que o sistema

$$\begin{cases} x^2 - y^2 + z^2 = 1 \\ 3x + y^2 + z = 3 \end{cases}$$

define (x, y) como uma função de classe C^1 de z numa vizinhança do ponto $(1, 1, -1)$. Sendo $(x, y) = g(z)$, calcular $g'(-1)$.

Seja $F: \mathbb{R}^3 \rightarrow \mathbb{R}^2$ a função definida por

$$F(x, y, z, w) = (x^2 - y^2 + z^2 - 1, 3x + y^2 + z - 3)$$

O sistema acima pode ser escrito na forma $F(x, y, z) = (0, 0)$. Verificamos que $F(1, 1, -1) = (0, 0)$. Claramente F é de classe C^1 e

$$\frac{\partial F}{\partial(x, y)} = \begin{bmatrix} \frac{\partial F_1}{\partial x} & \frac{\partial F_1}{\partial y} \\ \frac{\partial F_2}{\partial x} & \frac{\partial F_2}{\partial y} \end{bmatrix} = \begin{bmatrix} 2x & -2y \\ 3 & 2y \end{bmatrix}$$

No ponto $(1, 1, -1)$ temos

$$\det \frac{\partial F}{\partial(x, y)}(1, 1, -1) = \begin{vmatrix} 2 & -2 \\ 3 & 2 \end{vmatrix} = 10 \neq 0$$

logo o Teorema 28.2 garante que o sistema define uma função $(x, y) = g(z) = (g_1(z), g_2(z))$ de classe C^1 numa vizinhança de $z = -1$.

Podemos obter uma expressão para $g(z)$ derivando as equações que definem a função g . Temos

$$\begin{cases} g_1(z)^2 - g_2(z)^2 + z^2 = 1 \\ 3g_1(z) + g_2(z)^2 + z = 3 \end{cases}$$

e derivando estas equações em ordem a z obtemos

$$\begin{cases} 2g_1(z)g_1'(z) - 2g_2(z)g_2'(z) + 2z = 0 \\ 3g_1'(z) + 2g_2(z)g_2'(z) + 1 = 0 \end{cases}$$

Uma vez que $g(-1) = (1, 1)$, isto é, que $g_1(-1) = 1 = g_2(-1)$ obtemos

$$\begin{cases} 2g_1'(-1) - 2g_2'(-1) = 2 \\ 3g_1'(-1) + 2g_2'(-1) = -1 \end{cases} \Leftrightarrow \begin{cases} g_1'(-1) = \frac{1}{5} \\ g_2'(-1) = -\frac{4}{5} \end{cases}$$

29. O TEOREMA DA FUNÇÃO IMPLÍCITA (CONCLUSÃO)

Dem. do Teorema 28.2. Vamos demonstrar o Teorema 28.2 a partir do Teorema da Função Inversa. Começamos por ilustrar a ideia da demonstração explicando como se pode resolver um sistema linear de 2 equações a 3 incógnitas invertendo uma matriz 3×3 . Consideremos o sistema

$$\begin{cases} 2x + 3y - z = 1 \\ x + y - 2z = 2 \end{cases}$$

e vejamos como o podemos resolver em ordem a x, y invertendo uma função. O sistema anterior é equivalente a

$$\begin{cases} 2x + 3y - z = 1 \\ x + y - 2z = 2 \\ z = c \end{cases}$$

para algum $c \in \mathbb{R}$. Suponhamos que conseguimos inverter o sistema

$$\begin{cases} 2x + 3y - z = a \\ x + y - 2z = b \\ z = c \end{cases} \Leftrightarrow \begin{cases} x = f(a, b, c) \\ y = g(a, b, c) \\ z = c \end{cases}$$

Então fazendo $a = 1, b = 2$, obtemos a solução do problema inicial, nomeadamente:

$$\begin{cases} 2x + 3y - z = 1 \\ x + y - 2z = 2 \\ z = c \end{cases} \Leftrightarrow \begin{cases} x = f(1, 2, c) \\ y = g(1, 2, c) \\ z = c \end{cases} \Leftrightarrow \begin{cases} x = f(1, 2, z) \\ y = g(1, 2, z) \end{cases}$$

Vamos agora fazer o mesmo em abstrato: Sendo F como no enunciado do Teorema 28.2, seja $f: U \rightarrow \mathbb{R}^n$ a função definida por

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

A matriz Jacobiana de f pode ser escrita como uma matriz por blocos

$$Df(x, y) = \begin{bmatrix} \text{Id} & \frac{\partial F}{\partial x} \\ 0 & \frac{\partial F}{\partial y} \end{bmatrix}$$

e como tal tem determinante

$$\det Df(x, y) = 1 \cdot \det \frac{\partial F}{\partial y}(x, y)$$

Conclui-se que se F satisfaz as hipóteses do Teorema da Função Implícita no ponto (x_0, y_0) , então f satisfaz as hipóteses do Teorema da Função Inversa no mesmo ponto. Portanto existe um aberto A contendo (x_0, y_0) tal que a restrição de f a A tem uma inversa de classe C^1 . Para $(x, y) \in A$ temos

$$F(x, y) = 0 \Leftrightarrow (x, F(x, y)) = (x, 0) \Leftrightarrow f(x, y) = (x, 0) \Leftrightarrow (x, y) = f^{-1}(x, 0)$$

A função inversa de $f(x, y)$ é da forma $f^{-1}(x, y) = (x, h(x, y))$ para alguma função h de classe C^1 , logo a última equação é equivalente a

$$\begin{cases} x = x \\ y = h(x, 0) \end{cases} \Leftrightarrow y = h(x, 0)$$

Obtemos assim a expressão para a função implícita g em termos da inversa de f :

$$g(x) = h(x, 0).$$

Fica como exercício mostrar que para o aberto W do enunciado do Teorema 28.2 se pode tomar qualquer aberto contendo y_0 tal que $\{x_0\} \times \overline{W} \subset A$ e que então $V = g^{-1}(W)$. \square

Nota 29.1. *Deduzimos acima o Teorema da Função Implícita a partir do Teorema da Função Inversa. É um bom exercício fazer o recíproco e deduzir o Teorema da Função Inversa a partir do Teorema da Função Implícita. Estes dois teoremas fundamentais do Cálculo Diferencial são portanto equivalentes.*

Nota 29.2. *Na demonstração anterior, f é uma mudança de variáveis em que as últimas variáveis são as componentes da função F . Uma interpretação para o Teorema 28.2 é portanto, que em torno de um ponto no qual a matriz DF tem característica máxima, as componentes de F podem ser usadas como coordenadas. Essas coordenadas transformam o conjunto de nível $F(x, y) = 0$ numa porção do plano $y = 0$, que é obviamente um gráfico (o gráfico da função nula). A Função implícita é a função que se obtém desta função nula compondo com a mudança de variáveis f^{-1} .*

Já observámos que se pode calcular a derivada da função implícita derivando a fórmula que define a função implícita usando a regra de derivação da função composta. Vamos agora deduzir uma fórmula para a derivada. Assumindo que o sistema $F(x, y) = 0$ se pode resolver na forma $y = g(x)$ com g diferenciável temos então $F(x, g(x)) = 0$ e derivando esta função composta obtemos

(11)

$$\frac{\partial F}{\partial x}(x, g(x)) \cdot \text{Id} + \frac{\partial F}{\partial y}(x, g(x)) Dg(x) = 0 \Leftrightarrow Dg(x) = - \left(\frac{\partial F}{\partial y}(x, g(x)) \right)^{-1} \frac{\partial F}{\partial x}(x, g(x))$$

Se F é uma função escalar, y é um escalar e portanto a matriz $\frac{\partial F}{\partial y}$ e a função g são também escalares. A fórmula anterior pode então escrever-se na forma:

$$(12) \quad \frac{\partial y}{\partial x_i} = \frac{\partial g}{\partial x_i} = -\frac{\frac{\partial F}{\partial x_i}(x, g(x))}{\frac{\partial F}{\partial y}(x, g(x))}$$

Exemplo 29.3. Determinar se o subconjunto de \mathbb{R}^3 definido pelas condições

$$\begin{cases} x + yz = 1 \\ x^3z + x - y = 0 \end{cases}$$

é o gráfico de uma função de uma variável numa vizinhança do ponto $(1, 1, 0)$.

A função $F(x, y, z) = (x + yz - 1, x^3z + x - y = 0)$ é de classe C^1 e $F(1, 1, 0) = (0, 0)$. Temos

$$DF(x, y, z) = \begin{bmatrix} 1 & z & y \\ 3x^2z + 1 & -1 & x^3 \end{bmatrix}$$

e no ponto $(1, 1, 0)$

$$DF(1, 1, 0) = \begin{bmatrix} 1 & 0 & 1 \\ 1 & -1 & 1 \end{bmatrix}$$

Assim

$$\det \frac{\partial F}{\partial(x, y)}(1, 1, 0) = -1, \quad \det \frac{\partial F}{\partial(x, z)}(1, 1, 0) = 0, \quad \det \frac{\partial F}{\partial(y, z)}(1, 1, 0) = 1$$

Isto significa que podemos resolver o sistema em ordem às variáveis (x, y) ou (y, z) e portanto que o subconjunto definido pelas equações é o gráfico de uma função de classe C^1 de z e também o gráfico de uma função de x numa vizinhança do ponto $(1, 1, 0)$. Escrevendo por exemplo $(x, y) = g(z)$ e aplicando a fórmula (11) temos

$$g'(z) = \begin{bmatrix} g'_1(z) \\ g'_2(z) \end{bmatrix} = -\frac{\partial F}{\partial(x, y)}^{-1} \frac{\partial F}{\partial z}$$

e portanto

$$g'(0) = -\begin{bmatrix} 1 & 0 \\ 1 & -1 \end{bmatrix}^{-1} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} -1 \\ 0 \end{bmatrix}$$

Exemplo 29.4. Mostrar que a equação

$$xy + e^{yz} = 1$$

define y como função de (x, z) numa vizinhança do ponto $(1, 0, 2)$ e calcular $\frac{\partial^2 y}{\partial x \partial z}(1, 2)$. A função $F(x, y, z) = xy + e^{yz} - 1$ é de classe C^1 e $F(1, 0, 2) = 0$. Temos

$$\frac{\partial F}{\partial y} = x + ze^{yz}$$

e portanto $\frac{\partial F}{\partial y}(1, 0, 2) = 3 \neq 0$. O Teorema da Função Implícita garante que a equação define uma função $y = g(x, z)$ numa vizinhança de $(1, 0, 2)$. Aplicando a fórmula (12) temos

$$\frac{\partial y}{\partial z} = \frac{\partial g}{\partial z} = -\frac{\frac{\partial F}{\partial z}}{\frac{\partial F}{\partial y}} = -\frac{ye^{yz}}{x + ze^{yz}}$$

E portanto (recordando que y é uma função de x e z)

$$\frac{\partial^2 y}{\partial x \partial z} = \frac{\partial}{\partial x} \left(-\frac{ye^{yz}}{x + ze^{yz}} \right) = -\frac{\left(\frac{\partial y}{\partial x} e^{yz} + yz \frac{\partial y}{\partial x} e^{yz} \right) (x + ze^{yz}) - ye^{yz} (1 + z^2 \frac{\partial y}{\partial x} e^{yz})}{(x + ze^{yz})^2}$$

Para usar esta fórmula precisamos do valor de $\frac{\partial y}{\partial x}(1, 2)$. Temos

$$\frac{\partial y}{\partial x} = -\frac{\frac{\partial F}{\partial x}}{\frac{\partial F}{\partial y}} = -\frac{y}{x + ze^{yz}}$$

Substituindo em $(x, z) = (1, 2)$ (e portanto $y = g(1, 2) = 0$) temos

$$\frac{\partial y}{\partial x}(1, 2) = 0, \quad \frac{\partial y}{\partial z}(1, 2) = 0, \quad \frac{\partial^2 y}{\partial x \partial z}(1, 2) = 0$$

Nota 29.5. O Teorema 28.2 garante apenas que a função implícita é de classe C^1 . No entanto, se F é de classe C^k com $k \geq 1$, e as condições do Teorema 28.2 se verificam, é uma consequência da fórmula (11) que a função g é também de classe C^k . Só é necessário verificar que as expressões que se obtêm derivando sucessivamente a fórmula (como no Exemplo anterior) são contínuas. A demonstração detalhada deixa-se como exercício.

30. A DEFINIÇÃO DE VARIEDADE

Uma variedade de dimensão k , ou variedade- k , é um subconjunto de \mathbb{R}^n que “é localmente como um aberto de \mathbb{R}^k ”. Ou seja é o análogo k -dimensional de uma “curva” (que corresponde ao caso em que $k = 1$) e de “superfície” (que corresponde a $k = 2$).

Por exemplo, uma circunferência é uma variedade-1 em \mathbb{R}^2 , mas a união dos dois eixos não é. Apesar de ser um conjunto de dimensão 1, perto do ponto onde os eixos se intersectam, o conjunto formado pela união dos dois eixos não se parece com um aberto de \mathbb{R} . Da mesma forma, uma superfície esférica em \mathbb{R}^3 é uma variedade-2, mas a superfície cônica definida pela equação $x^2 + y^2 = z^2$ não é uma variedade-2 (perto de $(0, 0, 0)$ não se parece com um aberto de \mathbb{R}^2). Se retirarmos este ponto problemático ao conjunto definido por $x^2 + y^2 = z^2$ obtemos uma variedade-2. O equador da superfície esférica unitária, definida por

$$x^2 + y^2 + z^2 = 1, \quad z = 0$$

é uma variedade-1 em \mathbb{R}^3 . Note-se que começamos com três graus de liberdade correspondentes às variáveis x, y, z e ao impôr 2 restrições ficamos com apenas $1 = 3 - 2$ grau de liberdade.

Em geral, esperamos que uma variedade- k seja definida por $(n - k)$ equações *independentes*, que diminuirão em $(n - k)$ os graus de liberdade, de forma a produzir um conjunto de dimensão k .

Recorde-se de Álgebra Linear que a *característica de uma matriz* A é o número máximo de linhas linearmente independentes de A , ou equivalentemente, o número máximo de colunas linearmente independentes de A .

Definição 30.1. *Seja $1 \leq k < n$. Um subconjunto $S \subset \mathbb{R}^n$ diz-se uma variedade de dimensão k se para cada $x \in S$, existe um aberto $U \subset \mathbb{R}^n$ contendo x , e uma função $F: U \rightarrow \mathbb{R}^{n-k}$ de classe C^1 tal que*

- $S \cap U = \{x \in U: F(x) = 0\}$
- *Para todo o $x \in S \cap U$ a característica de $DF(x)$ é máxima (isto é, é igual a $n - k$).*

Por convenção, uma variedade-0 em \mathbb{R}^n é um conjunto de pontos isolados em \mathbb{R}^n , e uma variedade- n em \mathbb{R}^n é um aberto de \mathbb{R}^n .

Nota 30.2. *A condição relativa à característica da matriz F diz que os gradientes das funções coordenadas de F são linearmente independentes, o que garante a “independência” das equações $F_i = 0$, para $i = 1, \dots, n - k$.*

Nota 30.3. *O Teorema da Função Implícita garante que, localmente, uma variedade- k S é o gráfico de uma função de classe C^1 de k das variáveis. Este é um sentido preciso no qual S é localmente “como um aberto de \mathbb{R}^k ”.*

Exemplo 30.4. *A circunferência $S = \{(x, y) \in \mathbb{R}^2: x^2 + y^2 = 1\}$ é uma variedade-1 em \mathbb{R}^2 .*

O conjunto S é o conjunto de nível 0 da função de classe C^1 , $F: \mathbb{R}^2 \rightarrow \mathbb{R}$, definida por $F(x, y) = x^2 + y^2 - 1$. A matriz Jacobiana desta função é

$$DF(x, y) = \begin{bmatrix} 2x & 2y \end{bmatrix}$$

Esta matriz tem característica máxima (igual a 1) sse não se anula. O único ponto onde a característica é nula, é a origem $(0, 0)$ que não é um ponto de S . Conclui-se que a característica é máxima em todos os pontos de S , logo S é uma variedade de dimensão $2 - 1 = 1$.

Exemplo 30.5. *O conjunto $X = \{(x, y) \in \mathbb{R}^2: xy = 0\}$ não é uma variedade-1. De facto, não existe nenhuma vizinhança U de $(0, 0)$ tal que $U \cap X$ seja o gráfico de uma função de uma variável. De facto, em qualquer vizinhança U existem infinitos pontos da forma $(0, y) \in S$ com $y \neq 0$ pelo que $U \cap X$ não é o gráfico de uma função de x , e analogamente, existem infinitos pontos da forma $(x, 0) \in U \cap X$ pelo que $U \cap X$ não é o gráfico de uma função de y .*

Nota 30.6. *Para verificar que o conjunto de nível S de uma função F não é uma variedade, não é suficiente verificar que a característica de DF não é máxima em alguns pontos de F . A propriedade de ser ou não uma variedade é uma propriedade do subconjunto de \mathbb{R}^n e não de um dos (muitos) sistemas de equações que o definem. A título de exemplo, o conjunto de nível 0 da função $F(x, y) = (x^2 + y^2 - 1)^2$ é a variedade S do Exemplo 30.4 mas é fácil verificar que $DF(x, y) = 0$ em todos os pontos de S .*

Exemplo 30.7. A superfície esférica $S = \{(x, y, z) \in \mathbb{R}^3: x^2 + y^2 + z^2 = 1\}$ é uma variedade-2 em \mathbb{R}^3 . De facto, $F(x, y, z) = x^2 + y^2 + z^2 - 1$ é uma função de classe C^1 em \mathbb{R}^3 cujo conjunto de nível 0 é S . Uma vez que

$$DF(x, y, z) = [2x \quad 2y \quad 2z]$$

só tem característica 0 em $(0, 0, 0)$ que não pertence a S , conclui-se que S é uma variedade de dimensão $3 - 1 = 2$.

Exemplo 30.8. O conjunto $S = \{(x, y, z) \in \mathbb{R}^3: x^2 + y^2 + z^2 = 3, x - y + 2z = 1\}$ é uma variedade-1 em \mathbb{R}^3 (é a intersecção de um plano com uma superfície esférica, logo é uma circunferência).

Sendo $F: \mathbb{R}^3 \rightarrow \mathbb{R}^2$ a função definida por $F(x, y, z) = (x^2 + y^2 + z^2 - 3, x - y + 2z - 1)$, temos que

$$S = \{(x, y, z) \in \mathbb{R}^3: F(x, y, z) = (0, 0)\}$$

Claramente F é de classe C^1 e

$$DF(x, y, z) = \begin{bmatrix} 2x & 2y & 2z \\ 1 & -1 & 2 \end{bmatrix}$$

O conjunto dos pontos onde a característica desta matriz é menor que 2 é o conjunto dos pontos onde todos os pares de colunas da matriz são linearmente dependentes, ou seja, o conjunto dos pontos onde

$$\begin{vmatrix} 2x & 2y \\ 1 & -1 \end{vmatrix} = 0, \quad \begin{vmatrix} 2y & 2z \\ -1 & 2 \end{vmatrix} = 0, \quad \begin{vmatrix} 2x & 2z \\ 1 & 2 \end{vmatrix} = 0$$

ou seja o conjunto dos pontos que satisfaz o sistema

$$\begin{cases} -2x - 2y = 0 \\ 4y + 2z = 0 \\ 4x - 2z = 0 \end{cases} \Leftrightarrow \begin{cases} y = -x \\ z = -2y \\ 0 = 0 \end{cases}$$

ou seja os pontos da forma $(x, -x, 2x)$ com $x \in \mathbb{R}$. Substituindo estes pontos nas equações que definem S obtemos

$$\begin{cases} x^2 + x^2 + 4x^2 = 3 \\ x + x + 4x = 1 \end{cases}$$

que é um sistema impossível. Isto mostra que a característica de $DF(x, y, z)$ é máxima em todos os pontos de S , pelo que S é uma variedade de dimensão $3 - 2 = 1$.

31. PARAMETRIZAÇÕES. ESPAÇOS TANGENTE E NORMAL

Exemplo 31.1. Seja $S = \{(x, y, z) \in \mathbb{R}^3: z = x^2 + y^2\}$. Então S é uma variedade de dimensão 2. De facto, a função $F(x, y, z) = z - x^2 + y^2$ é de classe C^1 , S é o conjunto de nível 0 de F e

$$DF(x, y, z) = [-2x \quad -2y \quad 1]$$

nunca se anula.

O conjunto S do exemplo anterior é o gráfico da função $f: \mathbb{R}^2 \rightarrow \mathbb{R}$ de classe C^1 definida por $f(x, y) = x^2 + y^2$. Mais geralmente, qualquer gráfico de uma função de classe C^1 definida num aberto é uma variedade. De facto, seja $U \subset \mathbb{R}^k$ um aberto e $f: U \rightarrow \mathbb{R}^{n-k}$ uma função de classe C^1 . O gráfico de f é o conjunto

$$\{(x, f(x)) \in \mathbb{R}^n : x \in U\}$$

que é precisamente o conjunto de nível 0 da função $F: U \times \mathbb{R}^{n-k} \rightarrow \mathbb{R}^{n-k}$ definida por

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

Esta função é de classe C^1 e tem matriz jacobiana

$$DF(x, y) = \begin{bmatrix} -Df(x) & \text{Id} \end{bmatrix}$$

que claramente tem característica máxima.

O argumento anterior juntamente com a observação, já feita anteriormente, que o Teorema da Função Implícita e a definição de variedade implicam imediatamente que uma variedade é localmente o gráfico de uma função de classe C^1 de k das variáveis, podem ser resumidos no seguinte resultado.

Proposição 31.2. *Um conjunto $S \subset \mathbb{R}^n$ é uma variedade de dimensão k sse é localmente o gráfico de uma função de classe C^1 definida num aberto de \mathbb{R}^k , isto é, se para cada $x \in S$, existe um aberto $U \subset \mathbb{R}^n$ contendo x , um aberto $V \subset \mathbb{R}^k$ e uma função $f: V \rightarrow \mathbb{R}^{n-k}$ de classe C^1 tal que*

$$S \cap U = \{(t, f(t)) : t \in V\}$$

(onde escrevemos um vector de \mathbb{R}^n na forma (t, s) com t um vector de \mathbb{R}^k formado por uma escolha de k das n coordenadas e s o vector de \mathbb{R}^{n-k} formado pelas restantes coordenadas).

Nota 31.3. *Note-se que se um conjunto $S \subset \mathbb{R}^n$ é o gráfico de uma função de algumas das coordenadas, essa função é única. De facto, S é um gráfico sse existe um conjunto de coordenadas de \mathbb{R}^n (que escrevemos num vector $t \in \mathbb{R}^k$) tal que, para cada t , existe no máximo um elemento da forma $(t, s) \in S$. A função da qual S é o gráfico é então definida pela expressão*

$$g(t) = \text{único } s \text{ tal que } (t, s) \in S$$

(e o seu domínio é o conjunto dos t s para os quais existe um tal s).

Exemplo 31.4. *O conjunto*

$$S = \{(x, y, z) \in \mathbb{R}^3 : z = \sqrt{x^2 + y^2}\}$$

não é uma variedade-2 em \mathbb{R}^3 .

De facto, numa vizinhança de $(0, 0, 0)$ não é o gráfico de uma função de (x, z) nem de (y, z) . É o gráfico da função $f(x, y) = \sqrt{x^2 + y^2}$ mas esta função não é de classe C^1 .

O conjunto $S \setminus \{(0, 0, 0)\}$ é uma variedade-2, uma vez que é o gráfico da função $g: \mathbb{R}^2 \setminus \{(0, 0)\} \rightarrow \mathbb{R}$ definida por $g(x, y) = \sqrt{x^2 + y^2}$ (que é de classe C^1).

Para fazer cálculos em variedades (o que iremos passar o resto do semestre a fazer) precisamos de ter uma maneira eficiente de nos referirmos aos pontos da variedade. Uma variedade- k é um conjunto de dimensão k , e portanto em princípio deveríamos precisar de apenas k parâmetros reais para especificar a posição de um ponto na variedade. Um exemplo familiar é a das coordenadas cartográficas - um ponto à superfície da Terra é geralmente especificado dando dois números - a longitude e a latitude. Em geral, os pontos de uma variedade de dimensão k podem (pelo menos localmente) ser especificados por k parâmetros reais.

Definição 31.5. *Seja $S \subset \mathbb{R}^n$ uma variedade de dimensão k . Uma parametrização de (uma porção de) S é uma função $g: V \rightarrow S$ em que*

- V é um aberto de \mathbb{R}^k
- g é de classe C^1 e injectiva
- $Dg(t)$ tem característica máxima (igual a k) para todo $t \in V$

A condição sobre a matriz Jacobiana da parametrização é uma condição de injectividade “infinitesimal” - diz que a derivada de g é injectiva em cada ponto do domínio. É um bom exercício verificar usando o Teorema da Função Inversa que quaisquer duas parametrizações de uma variedade diferem por uma mudança de coordenadas em \mathbb{R}^k (entre os conjuntos correspondentes à interseção das imagens das parametrizações).

Seja $U \subset \mathbb{R}^k$ um aberto e $f: U \rightarrow \mathbb{R}^{n-k}$ uma função de classe C^1 . Então o gráfico de f

$$S = \{(x, f(x)): x \in U\} \subset \mathbb{R}^n$$

é uma variedade de classe C^1 . A função $g: U \rightarrow \mathbb{R}^n$ definida por

$$g(x) = (x, f(x))$$

é uma parametrização de S : claramente g é injectiva e de classe C^1 e tem imagem S . Por outro lado, a matriz jacobiana de g é

$$Dg(x) = \begin{bmatrix} \text{Id} \\ Df(x) \end{bmatrix}$$

que claramente tem característica máxima. A Proposição 31.2 garante então que podemos parametrizar uma variedade perto de qualquer um dos seus pontos.

Exemplo 31.6. *Seja $S = \{(x, y) \in \mathbb{R}^2: x^2 + y^2 = 1\}$ a circunferência (Exemplo 30.4). A função $g:]0, 2\pi[\rightarrow S$ definida por*

$$g(\theta) = (\cos \theta, \sin \theta)$$

é uma parametrização (com imagem $S \setminus \{(1, 0)\}$). De facto g é injectiva, de classe C^1 e

$$Dg(\theta) = \begin{bmatrix} -\sin \theta \\ \cos \theta \end{bmatrix}$$

nunca se anula (e portanto tem sempre característica 1).

A circunferência S pode também ser parametrizada localmente como um gráfico de uma das funções $x \mapsto \pm\sqrt{1-x^2}$ ou $y \mapsto \pm\sqrt{1-y^2}$ mas a imagem de cada uma destas parametrizações cobre apenas metade da circunferência.

Exemplo 31.7. Uma possível parametrização do parabolóide $S = \{(x, y, z) \in \mathbb{R}^3 : z = x^2 + y^2\}$ do Exemplo 34.1 é a função $g: \mathbb{R}^2 \rightarrow \mathbb{R}^3$ definida por

$$g(x, y) = (x, y, x^2 + y^2)$$

uma vez que S é o gráfico da função $x^2 + y^2$. Uma outra parametrização possível é a função $h:]0, +\infty[\times]0, 2\pi[\rightarrow \mathbb{R}^3$ definida por

$$h(r, \theta) = (r \cos \theta, r \sin \theta, r^2)$$

que usa duas das coordenadas cilíndricas como parâmetros. De facto, a função h é de classe C^1 , injectiva e a sua imagem é $S \setminus \{(x, 0, x^2) : x \geq 0\}$. A matriz jacobiana

$$Dh(r, \theta) = \begin{bmatrix} \cos \theta & -r \sin \theta \\ \sin \theta & r \cos \theta \\ 2r & 0 \end{bmatrix}$$

tem sempre característica 2 uma vez que o determinante da matriz formada pelas duas primeiras linhas é r e portanto nunca se anula.

Definição 31.8. Seja S uma variedade- k em \mathbb{R}^n . Um vector $v \in \mathbb{R}^n$ diz-se tangente a S no ponto $x \in S$ se existe uma função $\gamma:]-\epsilon, \epsilon[\rightarrow S$ de classe C^1 , com $\gamma(0) = x$ e $\gamma'(0) = v$.

O espaço tangente a S em x é o conjunto $T_x S$ de todos os vectores tangentes a x em S . O espaço normal a S em x é o espaço vectorial $(T_x S)^\perp = \{w \in \mathbb{R}^n : w \perp v \text{ para todo o } v \in T_x S\}$

Portanto um vector é tangente a S num ponto x se é tangente em x a uma curva contida em S que passa por x . Os espaços tangente e normal podem ser calculados a partir de equações que definam a variedade, ou a partir de parametrizações da seguinte forma.

Proposição 31.9. Seja S uma variedade- k em \mathbb{R}^n e $x \in S$. Então

- (1) $T_x S$ é um espaço vectorial de dimensão k .
- (2) Se $g: V \rightarrow S$ é uma parametrização com $g(t) = x$, então

$$T_x S = \text{espaço das colunas de } Dg(t) = \{Dg(t)w : w \in \mathbb{R}^k\} \subset \mathbb{R}^n$$

ou seja é o conjunto das combinações lineares de colunas de $Dg(t)$.

- (3) Se $F: U \rightarrow \mathbb{R}^{n-k}$ é uma função definida num aberto U contendo x tal que $S \cap U = \{w \in \mathbb{R}^n : F(w) = 0\}$ com F de classe C^1 e $DF(w)$ de característica máxima para $w \in S$, então

$$T_x S = N(DF(x)) = \{v \in \mathbb{R}^n : DF(x)v = 0\}$$

é o núcleo da matriz $DF(x)$. Além disso, as linhas da matriz $DF(x)$ formam uma base para $(T_x S)^\perp$.

O espaço normal $(T_x S)^\perp$ é o complemento ortogonal do espaço tangente $T_x S$ e pode ser calculado a partir deste pelos métodos estudados em Álgebra Linear.

32. EXEMPLOS DE CÁLCULO DE ESPAÇOS TANGENTES E NORMAIS.

Note-se que a última afirmação da Proposição 31.9 é uma consequência imediata das definições. De facto o núcleo de $DF(x)$ consiste precisamente no conjunto dos vectores de \mathbb{R}^n que são perpendiculares às linhas da matriz $DF(x)$, logo as linhas da matriz (que são linearmente independentes porque a característica de $DF(x)$ é máxima) formam uma base para o complemento ortogonal do núcleo, que é o espaço normal $(T_x S)^\perp$.

Uma outra perspectiva (equivalente) sobre a última afirmação da Proposição 31.9 é a seguinte: as linhas da matriz $DF(x)$ são os gradientes das funções coordenadas de F . Para cada $i = 1, \dots, n - k$, o gradiente $\nabla F_i(x)$ é perpendicular no ponto x ao conjunto de nível $F_i(x) = 0$. A variedade S é a *interseção* de todos estes conjuntos de nível, logo os vectores $\nabla F_1(x), \dots, \nabla F_{n-k}(x)$ são todos ortogonais a S . Como são $(n - k)$ vectores linearmente independentes no espaço vectorial $(T_x S)^\perp$ (que tem dimensão $n - k$), formam uma base para este espaço.

Exemplo 32.1. *Seja*

$$S = \{(x, y, z) \in \mathbb{R}^3 : (x - 1)^2 + y^2 = 1, z = x^2 + y^2\}$$

Calcular o espaço normal e tangente a S em $(1, 1, 2)$, assim como a recta tangente e o plano normal a S nesse ponto.

Sendo $F(x, y, z) = ((x - 1)^2 + y^2 - 1, z - x^2 - y^2)$, o conjunto S é igual a $\{(x, y, z) \in \mathbb{R}^3 : F(x, y, z) = (0, 0)\}$. F é de classe C^1 e

$$DF(x, y, z) = \begin{bmatrix} 2(x - 1) & 2y & 0 \\ -2x & -2y & 1 \end{bmatrix}$$

Os pontos onde a característica de F é menor que 2 verificam

$$\begin{vmatrix} 2(x - 1) & 0 \\ -2x & 1 \end{vmatrix} = 2(x - 1) = 0, \quad \begin{vmatrix} 2y & 0 \\ -2y & 1 \end{vmatrix} = 2y = 0$$

Uma vez que os pontos da forma $(1, 0, z)$ não pertencem a S , conclui-se que DF tem característica máxima nos pontos de S e portanto S é uma variedade de dimensão 1.

De acordo com a Proposição 31.9, o espaço tangente é o núcleo da matriz

$$DF(1, 1, 2) = \begin{bmatrix} 0 & 2 & 0 \\ -2 & -2 & 1 \end{bmatrix}$$

isto é, é o conjunto dos vectores $(u, v, w) \in \mathbb{R}^3$ tais que

$$\begin{cases} 2v = 0 \\ -2u - 2v + w = 0 \end{cases} \Leftrightarrow (u, v, w) = (u, 0, 2u) = u(1, 0, 2)$$

logo

$$T_{(1,1,2)}S = \{u(1, 0, 2) : u \in \mathbb{R}\}$$

A recta tangente no ponto $(1, 1, 2)$ é uma recta paralela ao espaço tangente que passa pelo ponto $(1, 1, 2)$. A equação paramétrica da recta tangente é

$$(x, y, z) = (1, 1, 2) + t(1, 0, 2), \quad t \in \mathbb{R}$$

O plano normal no ponto $(1, 1, 2)$ é um plano paralelo ao espaço normal que passa pelo ponto $(1, 1, 2)$. A direção perpendicular ao plano é $(1, 0, 2)$ logo a equação cartesiana é

$$(x - 1, y - 1, z - 2) \cdot (1, 0, 2) = 0 \Leftrightarrow (x - 1) + 2(z - 2) = 0 \Leftrightarrow x + 2z = 5$$

Alternativamente, podemos calcular o espaço tangente e normal (e a reta tangente e o plano normal) usando uma parametrização. A projeção de S no plano xy é a circunferência com equação $(x - 1)^2 + y^2 = 1$, que é descrita por $x = 1 + \cos \theta$, $y = \sin \theta$, com $0 < \theta < 2\pi$. Uma vez que $z = x^2 + y^2$ sobre S , temos que $g:]0, 2\pi[\rightarrow \mathbb{R}^3$ definida por

$$g(\theta) = (1 + \cos \theta, \sin \theta, (1 + \cos \theta)^2 + \sin^2 \theta) = (1 + \cos \theta, \sin \theta, 2 + 2 \cos \theta)$$

é uma parametrização de S . O parâmetro θ correspondente ao ponto $(1, 1, 2)$ é $\theta = \frac{\pi}{2}$ (isto é, $g(\frac{\pi}{2}) = (1, 1, 2)$). Temos

$$g'(\theta) = (-\sin \theta, \cos \theta, -2 \sin \theta)$$

logo $g'(\frac{\pi}{2}) = (-1, 0, -2)$ é um vector tangente a S em $(1, 1, 2)$. Portanto obtemos novamente

$$T_{(1,1,2)}S = \{(u, 0, 2u) : u \in \mathbb{R}\}$$

A partir daqui obtemos a equação para o espaço normal, que é formado pelos vectores perpendiculares a $(1, 0, 2)$:

$$(T_{(1,1,2)}S)^\perp = \{(u, v, w) \in \mathbb{R}^3 : u + 2w = 0\}.$$

Exemplo 32.2. Seja

$$S = \{(x, y, z) \in \mathbb{R}^3 : \tan z = \frac{y}{x}\}$$

Calcular o plano tangente e a recta perpendicular no ponto $(1, 1, \frac{\pi}{4})$.

A função $F(x, y, z) = \tan z - \frac{y}{x}$ é de classe C^1 no seu domínio, que é,

$$\{(x, y, z) \in \mathbb{R}^3 : x \neq 0 \text{ e } z \neq \frac{\pi}{2} + k\pi \text{ com } k \in \mathbb{Z}\}$$

Temos

$$DF(x, y, z) = \left[-\frac{y}{x^2} \quad \frac{1}{x} \quad \frac{1}{\cos^2 z} \right]$$

que não se anula no domínio. Conclui-se que S é uma variedade de dimensão 2 em \mathbb{R}^3 . O vector $\nabla F(1, 1, \frac{\pi}{4}) = (-1, 1, 2)$ é perpendicular a S no ponto $(1, 1, \frac{\pi}{4})$ logo o plano tangente é definido por

$$\langle (x - 1, y - 1, z - \frac{\pi}{4}), (-1, 1, 2) \rangle = 0 \Leftrightarrow -x + y + 2z = \frac{\pi}{2}$$

e a recta perpendicular tem equação paramétrica

$$(x, y, z) = (1, 1, \frac{\pi}{4}) + t(-1, 1, 2).$$

Note-se que podemos entender geometricamente a superfície S usando coordenadas cilíndricas. Como

$$\frac{y}{x} = \frac{r \sin \theta}{r \cos \theta} = \tan \theta$$

a equação que define S diz que z difere de θ num múltiplo de π . A interseção de S com um plano horizontal em que fixamos z ($\neq \frac{\pi}{2} + k\pi$) é uma recta que faz o ângulo z com o eixo

dos xx (exceptuando a interseção com o eixo zz). A superfície S é assim obtida rodando uma reta em torno do eixo dos zz sendo o ângulo da reta dado por z (uma "rampa em caracol dupla" que é o formato das moléculas de ADN). Uma parametrização para S é por exemplo

$$g(z, t) = (t \cos z, t \sin z, z), \quad \text{com } t \neq 0, z \neq \frac{\pi}{2} + k\pi$$

Se retirarmos a condição sobre z obtemos uma parametrização de uma variedade T um pouco maior que S , que inclui também as retas (perfuradas pelo eixo dos zz) paralelas ao eixo dos yy às alturas $z = \frac{\pi}{2} + k\pi$. Perto destas retas, a equação

$$\tan z = \frac{y}{x}$$

não está definida, mas é fácil verificar que a equação

$$\cotg z = \frac{x}{y}$$

define a variedade T numa vizinhança dessas retas.

33. VARIEDADES (CONCLUSÃO).

Dem. da Proposição 31.9. Seja C o espaço das colunas de $Dg(t)$ e N o núcleo de $DF(x)$. Vamos ver que

$$(13) \quad C \subset T_x S \subset N$$

Uma vez que tanto C como N são subespaços vectoriais de \mathbb{R}^n de dimensão k (no caso de C porque as colunas de $Dg(t)$ são linearmente independentes, e no caso de N porque $DF(x)$ tem característica $n - k$) a inclusão $C \subset N$ implica igualdade entre estes espaços. Mas então o conjunto $T_x S$ tem de coincidir com C e N e, em particular, tem de ser um espaço vectorial.

Para verificar a primeira inclusão em (13), temos que ver que dado um vector $w \in \mathbb{R}^k$, o vector $Dg(t)w$ pertence a $T_x S$. Seja $\gamma:] - \epsilon, \epsilon[\rightarrow S$ o caminho definido por

$$\gamma(s) = g(t + sw)$$

(onde ϵ é suficientemente pequeno para que $t + sw$ pertença ao domínio da parametrização g quando $s \in] - \epsilon, \epsilon[$). Então γ é de classe C^1 , toma valores em S , e $\gamma(0) = x$, logo $\gamma'(0)$ é um vector tangente a S em x . Mas pela regra de derivação da função composta temos

$$\gamma'(s) = Dg(t + sw)w \quad \Rightarrow \quad \gamma'(0) = Dg(t)w$$

Logo $Dg(t)w \in T_x S$. Uma vez que $w \in \mathbb{R}^k$ é arbitrário, conclui-se que $C \subset T_x S$.

Para ver a segunda inclusão em (13), seja agora $\gamma:] - \epsilon, \epsilon[\rightarrow S$ uma função de classe C^1 com $\gamma(0) = x$ (de maneira que $\gamma'(0)$ é, por definição, um elemento de $T_x S$). Então $F(\gamma(s)) = 0$ para todo o $s \in] - \epsilon, \epsilon[$ e portanto $\frac{d}{ds}(F(\gamma(s))) = 0$. Aplicando a regra de derivação da função composta obtemos

$$DF(\gamma(s))\gamma'(s) = 0 \quad \Rightarrow \quad DF(\gamma(0))\gamma'(0) = 0 \quad \Leftrightarrow \quad DF(x)\gamma'(0) = 0$$

Logo $\gamma'(0) \in N$, e uma vez que $\gamma'(0)$ é um elemento arbitrário de $T_x S$, conclui-se que $T_x S \subset N$. Isto conclui a demonstração. \square

Exemplo 33.1. A orientação de um corpo rígido é determinada pela posição de um referencial ortonormado orientado $\{v_1, v_2, v_3\}$ em \mathbb{R}^3 . Recordamos que orientado significa que a orientação do terceiro vector do referencial é determinada a partir da posição relativa entre v_1 e v_2 pela “regra da mão direita” ou, equivalentemente, que $v_3 = v_1 \times v_2$.

Escrevendo

$$v_1 = (x, y, z), \quad v_2 = (u, v, w), \quad v_3 = (a, b, c)$$

as equações que traduzem o facto de $\{v_1, v_2, v_3\}$ ser um referencial ortonormado são

$$\begin{aligned} \|v_1\|^2 = 1 &\Leftrightarrow x^2 + y^2 + z^2 = 1 \\ \|v_2\|^2 = 1 &\Leftrightarrow u^2 + v^2 + w^2 = 1 \\ \|v_3\|^2 = 1 &\Leftrightarrow a^2 + b^2 + c^2 = 1 \\ \langle v_1, v_2 \rangle = 0 &\Leftrightarrow xu + yv + zw = 0 \\ \langle v_1, v_3 \rangle = 0 &\Leftrightarrow xa + yb + zc = 0 \\ \langle v_2, v_3 \rangle = 0 &\Leftrightarrow ua + vb + wc = 0 \end{aligned}$$

É um bom exercício verificar que as 6 equações anteriores definem uma variedade-3 em \mathbb{R}^9 . Os pontos desta variedade correspondem aos referenciais ortonormados (com ambas as orientações possíveis). Esta variedade tem duas componentes cada uma das quais corresponde a uma orientação possível (a situação é inteiramente análoga ao caso da variedade-1 em \mathbb{R}^2 definida pela equação $xy = 1$ que tem duas componentes contidas, respectivamente, no primeiro e terceiro quadrantes).

Uma maneira mais eficiente de descrever a variedade dos referenciais ortonormados é utilizar notação matricial. As matrizes reais 3×3 identificam-se com \mathbb{R}^9 e escrevendo

$$A = \begin{bmatrix} x & u & a \\ y & v & b \\ z & w & c \end{bmatrix}$$

as 6 equações acima são equivalentes à equação matricial

$$(14) \quad A^T A = \text{Id}$$

(apesar da igualdade matricial consistir em 9 equações escalares, a matriz $A^T A$ é simétrica pelo que apenas 6 das equações são distintas). Recorde-se de Álgebra Linear que uma matriz que satisfaz (14) se diz ortogonal.

A orientação do referencial formado pelas colunas da matriz A é positivo se $\det A > 0$ e negativo caso contrário. Vamos agora tentar “visualizar” a variedade de dimensão 3

$$R = \{A \in M_{3 \times 3}(\mathbb{R}) : A^T A = \text{Id}, \det A > 0\}$$

cujos pontos correspondem aos referenciais ortonormados de orientação positiva. A chave para essa visualização é perceber que uma matriz $A \in R$ representa uma rotação em \mathbb{R}^3 . De facto, é sabido de Álgebra Linear que uma matriz ortogonal (portanto normal) é diagonalizável, sendo os valores próprios complexos de módulo 1. Os valores próprios $\lambda_1, \lambda_2, \lambda_3$

de A são as raízes do polinómio característico $p(\lambda) = \det(A - \lambda I)$. Uma vez que este polinómio é real, as suas raízes complexas ocorrem necessariamente em pares conjugados. Escrevendo $\lambda_2 = \cos \theta + i \operatorname{sen} \theta$ e $\lambda_3 = \cos \theta - i \operatorname{sen} \theta$, a equação

$$\det A = \lambda_1 \lambda_2 \lambda_3 = 1$$

implica que $\lambda_1 = 1$. Fica como exercício verificar que se w_1 é um vector próprio de λ_1 a matriz A representa uma rotação de θ em torno do eixo w_1 .

Há alguma ambiguidade na determinação dos vectores w_1 e ângulo θ acima, mesmo que insistamos que o ângulo esteja entre 0 e 2π e que w_1 tenha comprimento 1 . De facto, uma rotação de θ no sentido antihorário do ponto de vista de w_1 é exactamente o mesmo que uma rotação de $2\pi - \theta$ no sentido antihorário do ponto de vista de $-w_1$. Esta ambiguidade pode ser quase eliminada insistindo que o ângulo da rotação esteja no intervalo $[0, \pi]$. Desta forma a ambiguidade fica “concentrada” nas rotações de ângulo π - para estas, o vector w_1 só fica determinado a menos de um sinal.

A argumentação anterior dá-nos um modelo para o espaço R : podemos codificar a rotação determinada pelo vector w_1 e ângulo $\theta \in [0, \pi]$ no vector

$$\theta w_1 \in \{(x, y, z) \in \mathbb{R}^3 : x^2 + y^2 + z^2 \leq \pi^2\}$$

Note-se que esta codificação é consistente com o facto de rotações de um ângulo 0 serem a identidade e de, nesse caso, o eixo ser irrelevante. Para obter uma correspondência biunívoca entre o conjunto acima e as rotações temos no entanto que “identificar” um vector na superfície esférica

$$x^2 + y^2 + z^2 = \pi^2$$

com o seu simétrico. Isto permite-nos finalmente visualizar o espaço tridimensional R como o “espaço que se obtém da bola de raio π centrada na origem identificando pontos antípodas na superfície da esfera de raio π ”. Se estivermos dentro deste espaço e tentarmos “atravessar” a superfície esférica, emergimos no ponto antípoda.

Em termos técnicos, a função definida no aberto $U = \{(x, y, z) \in \mathbb{R}^3 : x^2 + y^2 + z^2 < \pi^2\}$ que associa a cada ponto deste aberto a rotação de um ângulo $\sqrt{x^2 + y^2 + z^2}$ em torno do eixo gerado por (x, y, z) é uma parametrização de R que cobre todos os pontos de R exceto as rotações de ângulo π . Uma escolha natural de parâmetros são os dois ângulos das coordenadas esféricas para descrever o eixo de rotação, juntamente com o ângulo da rotação. É um bom exercício escrever uma fórmula para esta parametrização de R .

Recomenda-se aos alunos interessados que utilizem o programa Curved Spaces de Jeffrey Weeks (a ligação para fazer download está disponível na secção “Materiais de Estudo” na página da cadeira) para visualizar o que seria viajar numa nave espacial através de várias variedades de dimensão 3 . O espaço R que temos estado a discutir é um dos espaços disponíveis para navegação. Aparece em “Spherical Spaces” - “Projective 3-space”.

34. O MÉTODO DOS MULTIPLICADORES DE LAGRANGE

Frequentemente é necessário otimizar quantidades sujeitas a certas restrições. Consideremos os seguintes exemplos.

Exemplo 34.1. Determinar os pontos P da parábola com equação $y = x^2$ e Q da reta com equação $y = -2 + x$ que se encontram mais próximos.

Este problema pode ser resolvido achando o ponto da parábola que minimiza a função que dá a distância de um ponto (x, y) à recta com equação $y = -2 + x$.

Exemplo 34.2. Achar as dimensões que uma caixa com base rectangular sem tampa e com área total 6 deve ter para que o volume do interior da caixa seja máximo.

Se x, y, z são as dimensões da caixa (com z a altura), o volume do interior é dado pela expressão $V(x, y, z) = xyz$. A restrição que a área da caixa é 6 pode ser escrita na forma

$$xy + 2xz + 2yz = 6$$

pelo que o problema consiste em minimizar a função $V(x, y, z)$ sujeita à restrição dada pela equação anterior.

Problemas como os anteriores podem por vezes ser resolvidos usando o seguinte resultado:

Proposição 34.3 (Método dos multiplicadores de Lagrange). *Seja $U \subset \mathbb{R}^n$ um aberto e $f: U \rightarrow \mathbb{R}$ uma função de classe C^1 . Seja $F: U \rightarrow \mathbb{R}^m$ uma função de classe C^1 , $S = \{x \in U: F(x) = 0\}$ e suponhamos que a característica de $DF(x)$ é máxima para todo o $x \in S$.*

Então, se a restrição de f a S tem um extremo num ponto $x \in S$, existem $\lambda_1, \dots, \lambda_m \in \mathbb{R}$ (chamados multiplicadores de Lagrange) tais que a função

$$g(x) = f(x) + \lambda_1 F_1(x) + \dots + \lambda_m F_m(x)$$

(onde $F_i(x)$ denota a i -ésima função coordenada de F) tem um ponto de estacionariedade em x .

Este resultado permite encontrar candidatos a pontos de extremo de uma função f sujeita às restrições $F_1 = \dots = F_m = 0$ resolvendo o sistema

$$\begin{cases} \nabla g(x) = 0 \\ F(x) = 0 \end{cases}$$

que é um sistema de $n + m$ equações para as $n + m$ incógnitas x_1, \dots, x_n e $\lambda_1, \dots, \lambda_m$. Genericamente esperamos obter um conjunto finito de soluções, que são os “pontos de estacionariedade” da restrição de f a S . Frequentemente podemos aplicar o Teorema de Weierstrass para garantir a existência de um extremo e é então fácil encontrar o extremo entre os vários candidatos.

Nota 34.4. *Há um critério de classificação destes “pontos de estacionariedade” envolvendo derivadas de segunda ordem de f (e de F) mas não o iremos estudar por este ser pouco prático.*

Dem. da Proposição 34.3. Seja $\gamma:] - \epsilon, \epsilon[\rightarrow S$ um caminho de classe C^1 com $\gamma(0) = 0$. Se x é um extremo da restrição $f|_S$ de f a S , então a função $f(\gamma(t))$ tem um extremo em $t = 0$. Consequentemente

$$\frac{d}{dt} (f(\gamma(t))) = \langle \nabla f(\gamma(t)), \gamma'(t) \rangle$$

anula-se em $t = 0$, ou seja,

$$\langle \nabla f(\gamma(0)), \gamma'(0) \rangle = \langle \nabla f(x), \gamma'(0) \rangle = 0$$

Ora $\gamma'(0)$ é um elemento arbitrário de $T_x S$, pelo que a equação anterior diz que, num ponto de extremo,

$$\nabla f(x) \in (T_x S)^\perp$$

Mas os vectores $\nabla F_1(x), \dots, \nabla F_m(x)$ formam uma base para $(T_x S)^\perp$ (pela Proposição 31.9(3)), logo têm que existir $\lambda_1, \dots, \lambda_m \in \mathbb{R}$ tais que

$$\nabla f(x) + \lambda_1 \nabla F_1(x) + \dots + \lambda_m \nabla F_m(x) = 0$$

Isto conclui a demonstração. \square

Nota 34.5. Como observámos na demonstração anterior, a condição $\nabla g(x) = 0$ na Proposição 34.3 é equivalente à condição $\nabla f(x) \in (T_x S)^\perp$. Esta condição necessária para a existência de um extremo de $f|_S$ é intuitiva se tivermos em conta o significado geométrico do gradiente: se $\nabla f(x)$ não é ortogonal a S num dado ponto $x \in S$, a projecção de $\nabla f(x)$ no espaço tangente a x dá uma direcção em S ao longo da qual f cresce. Na direcção oposta f decresce logo f não tem um extremo em x .

Exemplo 34.6. Vamos aplicar a Proposição 34.3 para resolver o problema enunciado no Exemplo 34.1. A distância de um ponto $(x_0, y_0) \in \mathbb{R}^2$ acima da reta $y = -x + 2$ a esta reta é dada pela expressão

$$\langle (x_0, y_0), \frac{1}{\sqrt{2}}(-1, 1) \rangle + \sqrt{2}$$

(a primeira parcela é a projecção do vector (x_0, y_0) na direcção perpendicular à reta e a segunda parcela é a distância da origem à reta - faça um desenho!)

Conclui-se que a função que dá a distância de um ponto (x, y) do plano acima da reta $y = x - 2$ a esta reta é

$$f(x, y) = -\frac{1}{\sqrt{2}}x + \frac{1}{\sqrt{2}}y + \sqrt{2}$$

Vamos portanto considerar a função auxiliar

$$g(x, y) = -\frac{1}{\sqrt{2}}x + \frac{1}{\sqrt{2}}y + \sqrt{2} + \lambda(y - x^2)$$

e resolver o sistema

$$\begin{cases} \frac{\partial g}{\partial x} = 0 \\ \frac{\partial g}{\partial y} = 0 \\ y = x^2 \end{cases} \Leftrightarrow \begin{cases} -\frac{1}{\sqrt{2}} - 2\lambda x = 0 \\ \frac{1}{\sqrt{2}} + \lambda = 0 \\ y = x^2 \end{cases} \Leftrightarrow \begin{cases} x = \frac{1}{2} \\ \lambda = -\frac{1}{\sqrt{2}} \\ y = \frac{1}{4} \end{cases}$$

O único candidato a extremo de f sobre a parábola é o ponto $(\frac{1}{2}, \frac{1}{4})$. Intuitivamente é claro que este tem de ser o ponto na parábola que minimiza f . Para o justificar, podemos usar o Teorema de Weierstrass: o conjunto

$$C = \{(x, y) \in \mathbb{R}^2 : y = x^2, f(x, y) \leq 50\}$$

dos pontos da parábola que estão a distância ≤ 50 da reta é compacto. Como f é contínua, tem máximo e mínimo em C . Claramente o máximo é 50 atingido nos dois pontos que da

interseção da reta $f(x, y) = 50$ com a parábola. O único ponto onde pode ocorrer o mínimo é portanto $(\frac{1}{2}, \frac{1}{4})$. Finalmente, é claro que o mínimo de f sobre C é igual ao mínimo de f sobre toda a parábola.

Podemos determinar o ponto Q da reta que está mais próximo de $P = (\frac{1}{2}, \frac{1}{4})$ achando a interseção da recta que passa por P e é ortogonal à reta de equação $y = -2 + x$ com esta última. A equação paramétrica da reta ortogonal é

$$(\frac{1}{2}, \frac{1}{4}) + t(1, -1) \Leftrightarrow \begin{cases} x = \frac{1}{2} + t \\ y = \frac{1}{4} - t \end{cases}$$

e substituindo em $y = -2 + x$ obtemos

$$\frac{1}{4} - t = -2 + \frac{1}{2} + t \Leftrightarrow 2t = \frac{7}{4} \Leftrightarrow t = \frac{7}{8}$$

pelo que

$$Q = (\frac{1}{2}, \frac{1}{4}) + \frac{7}{8}(1, -1) = (\frac{11}{8}, -\frac{5}{8}).$$

Nota 34.7. O exemplo anterior podia resolver-se alternativamente minimizando a função real de variável real $f(x, x^2)$ onde

$$f(x, y) = -\frac{1}{\sqrt{2}}x + \frac{1}{\sqrt{2}}y + \sqrt{2}$$

é a função distância à reta (e na realidade este método seria mais eficiente).

Exemplo 34.8. Para resolver o problema enunciado no Exemplo 34.2 devemos considerar a função

$$g(x, y, z) = xyz + \lambda(xy + 2xz + 2yz - 5)$$

e resolver o sistema

$$(15) \quad \begin{cases} \frac{\partial g}{\partial x} = 0 \\ \frac{\partial g}{\partial y} = 0 \\ \frac{\partial g}{\partial z} = 0 \\ xy + 2xz + 2yz = 5 \end{cases} \Leftrightarrow \begin{cases} yz + \lambda(y + 2z) = 0 \\ xz + \lambda(x + 2z) = 0 \\ xy + 2\lambda(x + y) = 0 \\ xy + 2xz + 2yz = 6 \end{cases}$$

As primeiras duas equações podem-se escrever na forma

$$\begin{cases} y(z + \lambda) = -2\lambda z \\ x(z + \lambda) = -2\lambda z \end{cases}$$

Se $z + \lambda = 0$ então $-2\lambda z = 0$ e portanto $\lambda = 0$ ou $z = 0$. Mas então $z = \lambda = 0$ e a terceira equação em (15) implica que $xy = 0$ o que é impossível pela última equação em (15). Conclui-se que $z + \lambda \neq 0$ e portanto

$$x = y = -\frac{2\lambda z}{z + \lambda}$$

Substituindo na terceira equação em (15) obtemos

$$x^2 + 4\lambda x = 0 \Leftrightarrow x(x + 4\lambda) = 0 \Leftrightarrow x = 0 \text{ ou } x = -4\lambda$$

Novamente a última equação em (15) mostra que $x = 0$ é impossível. Portanto $x = y = -4\lambda$. Substituindo numa das primeiras equações em (15) obtemos

$$-4\lambda(z + \lambda) = -2\lambda z \Leftrightarrow -2\lambda z - 4\lambda^2 = 0 \Leftrightarrow -2\lambda(z + 2\lambda) = 0 \Leftrightarrow \lambda = 0 \text{ ou } z = -2\lambda$$

Se $\lambda = 0$ então as três primeiras equações em (15) implicam que $yz = xz = xy = 0$ o que contradiz a quarta equação. Conclui-se finalmente que

$$x = -4\lambda, \quad y = -4\lambda, \quad z = -2\lambda$$

e portanto

$$y = x, \quad z = \frac{x}{2}$$

Substituindo na equação

$$xy + 2xz + 2yz = 5$$

obtem-se

$$x^2 + x^2 + x^2 = 6 \Leftrightarrow x = \sqrt{2}$$

e portanto as dimensões da caixa de área 6 que maximizam o volume são:

$$x = \sqrt{2}, \quad y = \sqrt{2}, \quad z = \frac{1}{\sqrt{2}}$$

sendo o volume máximo $\sqrt{2}$.

Fica como exercício a aplicação do Teorema de Weierstrass para garantir que o ponto encontrado efectivamente maximiza o volume.

Nota 34.9. O método dos multiplicadores de Lagrange pode também ser utilizado para achar extremos de funções em conjuntos de \mathbb{R}^n limitados por variedades. Por exemplo, imagine-se que queremos achar os extremos de uma função f de classe C^1 no conjunto $B = \{(x, y, z) \in \mathbb{R}^3 : x^2 + y^2 + z^2 \leq 1\}$. Se um ponto de extremo ocorrer no interior da esfera, terá de ocorrer num ponto de estacionariedade de f . Se ocorrer na fronteira, terá de ser, em particular, um ponto de extremo da restrição de f à superfície da esfera de raio 1 e portanto terá de satisfazer as condições da Proposição 34.3. Podemos portanto encontrar os pontos de extremo (que têm de existir pelo Teorema de Weierstrass) entre os pontos de estacionariedade de f no interior de B e os pontos da fronteira de B em que ∇f é perpendicular à fronteira. Para achar o máximo e mínimo absoluto de f em B basta calcular o valor de f nos pontos achados e escolher o maior e menor valores respectivamente.

35. INTEGRAIS DE CAMPOS ESCALARES EM VARIEDADES

Seja S uma variedade- k em \mathbb{R}^n e $f: S \rightarrow \mathbb{R}$ uma função. A função f diz-se um *campo escalar* em S . Dado um conjunto $A \subset S$ pretendemos definir

$$\int_A f$$

com a intenção que este integral seja a “soma dos valores que f toma nos pontos de A ”. Para calcular esta soma iremos usar uma parametrização $g: V \rightarrow S$ (com $V \subset \mathbb{R}^k$) um aberto, tal que $A \subset g(V)$.

Usando a parametrização como dicionário, a soma pretendida é transformada numa soma sobre o conjunto dos parâmetros correspondentes a A , ou seja, sobre $g^{-1}(A)$, dos valores tomados pela função integranda nos pontos correspondentes, isto é, da função $f(g(t))$. Isto sugere a fórmula $\int_{g^{-1}(A)} f(g(t))dt$ para o integral acima. A fórmula anterior não é a correta porque o integral não é verdadeiramente uma soma sobre todos os pontos da região de integração do valor da função integranda. É antes um limite de somas cujos termos são áreas de pequenos paralelepípedos k -dimensionais multiplicadas pelo valor que a função toma num ponto do paralelepípedo. Um paralelepípedo $I \subset g^{-1}(A)$ corresponde a um “paralelepípedo curvo” $g(I) \subset A$. A diferença entre as áreas de I e $g(I)$ tem de ser compensada na fórmula para o integral. Desde que I seja muito pequeno, $g(I)$ é muito bem aproximado pela imagem de I pela aproximação linear L de g determinada por $Dg(t)$ num ponto t do paralelepípedo. Se a i -ésima aresta de I tem comprimento h_i , a i -ésima aresta do paralelogramo k -dimensional $L(I)$ será $\frac{\partial g}{\partial t_i}(t)h_i$. Conclui-se que a razão entre os volumes k -dimensionais de $L(I)$ e de I é o volume de um paralelogramo k -dimensional com arestas $\frac{\partial g}{\partial t_i}(t)$. Explicaremos abaixo que o volume k -dimensional de um paralelogramo k -dimensional com arestas $v_1, \dots, v_k \in \mathbb{R}^n$ é dado pela expressão

$$\sqrt{\det A^T A}$$

onde A é a matriz $n \times k$ que tem v_i na coluna i . Note-se que a entrada ij da matriz $A^T A$ é o produto interno $\langle v_i, v_j \rangle$. Isto motiva a seguinte definição.

Definição 35.1. *Seja $S \subset \mathbb{R}^n$ uma variedade- k em \mathbb{R}^n , $A \subset S$ e $f: A \rightarrow \mathbb{R}$. Seja $V \subset \mathbb{R}^k$ um aberto e $g: V \rightarrow S$ uma parametrização de S cuja imagem contém A . Defina-se*

$$(16) \quad \int_A f = \int_{g^{-1}(A)} f(g(t)) \sqrt{\det Dg(t)^T Dg(t)} dt$$

(onde o integral no termo direito é o integral de uma função escalar numa região de \mathbb{R}^k definido e estudado anteriormente).

No termo direito da igualdade (16) $g^{-1}(A)$ é a região de integração A escrita em termos dos parâmetros t , $f(g(t))$ é a função integranda escrita em termos dos parâmetros t e $\sqrt{\det Dg(t)^T Dg(t)}$ é o factor de conversão de volumes k -dimensionais.

Nota 35.2. *Quando $k = 1$, $Dg(t)$ é uma matriz coluna pelo que a expressão $Dg(t)^T Dg(t)$ na fórmula (16) é a matriz 1×1 com entrada $\|g'(t)\|^2$. Assim, no caso em que $k = 1$, a expressão (16) toma a forma mais simples*

$$\int_A f = \int_{g^{-1}(A)} f(g(t)) \|g'(t)\| dt$$

Nota 35.3. *É frequente usar-se a notação*

$$\int_C f ds$$

para o integral de um campo escalar numa variedade-1 (curva) contida em \mathbb{R}^n (ds representa o “comprimento de arco”) e

$$\int \int_S f \, dS$$

para o integral de um campo escalar numa variedade-2 (uma superfície) em \mathbb{R}^3 (dS diz-se um “elemento de superfície”).

O integral de campos escalares em variedades tem todas as aplicações habituais dos integrais. Pode ser usado para calcular volumes k -dimensionais (comprimentos e áreas quando $k = 1$ e 2 respetivamente), médias de funções, o centróide, a massa ou carga total integrando uma densidade de massa ou carga, o momento de inércia em torno de um eixo, o centro de massa, etc... As fórmulas são as habituais. Por exemplo, o volume k -dimensional de um subconjunto A de uma variedade- k é

$$\text{vol}(A) = \int_A 1$$

A coordenada i do centróide é dada por

$$\bar{x}_i = \frac{\int_A x_i}{\text{vol}(A)}$$

O momento de inércia em torno do eixo L é dado por

$$I_L = \int_A f d_L^2$$

onde f é a densidade de massa e d_L é a distância ao eixo L . Etc...

Exemplo 35.4. Calcular a massa do arco de parábola

$$C = \{(x, y) \in \mathbb{R}^2 : y = x^2, 0 \leq x \leq 1\}$$

sabendo que a densidade de massa é dada por $f(x, y) = x$.

Uma parametrização para C é dada por $g: [0, 1] \rightarrow \mathbb{R}^2$ definida por $g(x) = (x, x^2)$. Temos $g'(x) = (1, 2x)$ e $\|g'(x)\| = \sqrt{1 + 4x^2}$.

Logo a massa é

$$\int_C x = \int_0^1 x \sqrt{1 + 4x^2} dx = \frac{(1 + 4x^2)^{\frac{3}{2}}}{8 \cdot \frac{3}{2}} \Big|_0^1 = \frac{5\sqrt{5} - 1}{12}$$

Exemplo 35.5. Calcular o comprimento da curva parametrizada por $g(\theta) = (\cos \theta, \sin \theta, \theta)$ com $0 \leq \theta \leq 4\pi$.

Temos $g'(\theta) = (-\sin \theta, \cos \theta, 1)$ e $\|g'(\theta)\| = \sqrt{\cos^2 \theta + \sin^2 \theta + 1} = \sqrt{2}$ logo o comprimento é dado por

$$\int_C 1 = \int_0^{4\pi} \|g'(\theta)\| d\theta = \int_0^{4\pi} \sqrt{2} d\theta = 4\pi\sqrt{2}.$$

Este valor pode ser confirmado pensando da seguinte forma: o comprimento é o dobro do comprimento da curva descrita quando θ varia entre 0 e 2π . Esta última é uma hélice

inscrita num cilindro de raio 1 em torno do eixo dos zz que dá uma volta completa em torno do eixo dos zz . Se desenrolarmos o cilindro de forma a obter um quadrado de lado 2π , a curva desenrola-se na diagonal do quadrado que tem comprimento $2\pi\sqrt{2}$.

Exemplo 35.6. Calcular o momento de inércia em torno do eixo dos zz da placa

$$S = \{(x, y, z) \in \mathbb{R}^3 : z = x^2 + y^2, x^2 + y^2 \leq 1\}$$

admitindo que a densidade de massa é dada pela função constante igual a 1.

Uma parametrização para o parabolóide S é dada por

$$g(r, \theta) = (r \cos \theta, r \sin \theta, r^2), \quad \text{com } 0 < r < 1, 0 < \theta < 2\pi$$

Temos

$$Dg(r, \theta) = \begin{bmatrix} \cos \theta & -r \sin \theta \\ \sin \theta & r \cos \theta \\ 2r & 0 \end{bmatrix}$$

e portanto

$$Dg(r, \theta)^T Dg(r, \theta) = \begin{bmatrix} 1 + 4r^2 & 0 \\ 0 & r^2 \end{bmatrix}$$

Uma vez que a distância ao eixo dos zz é dada por $d(x, y, z) = x^2 + y^2$ conclui-se que

$$I_z = \int_S x^2 + y^2 = \int_0^{2\pi} \int_0^1 r^2 \sqrt{r^2(1 + 4r^2)} dr d\theta = 2\pi \int_0^1 r^3 \sqrt{1 + 4r^2} dr$$

e integrando por partes obtemos

$$2\pi \left(r^2 \frac{(1 + 4r^2)^{\frac{3}{2}}}{8 \cdot \frac{3}{2}} \Big|_0^1 - \int_0^1 2r \frac{(1 + 4r^2)^{\frac{3}{2}}}{12} dr \right) = 2\pi \left(\frac{5\sqrt{5}}{12} - \frac{(1 + 4r^2)^{\frac{5}{2}}}{6 \cdot 8 \cdot \frac{5}{2}} \Big|_0^1 \right) = \pi \left(\frac{5\sqrt{5}}{12} + \frac{1}{60} \right)$$

36. INTEGRAIS DE CAMPOS ESCALARES EM VARIEDADES (CONT.)

Nota 36.1. O comprimento de uma linha C é dado por $\int_C 1$. Sendo $g: [t_0, t_1] \rightarrow C$ uma parametrização, o comprimento de C é dado pela expressão

$$\int_{t_0}^{t_1} \|g'(t)\| dt$$

Se pensarmos em $g(t)$ como a posição de uma partícula no instante t , $\|g'(t)\|$ é a magnitude da velocidade instântanea e o integral acima, que é o integral da velocidade em ordem ao tempo, dá o espaço percorrido pela partícula.

Na aula anterior demos a seguinte fórmula (que iremos em breve justificar) para o volume k -dimensional de um paralelogramo com arestas $v_1, \dots, v_k \in \mathbb{R}^n$:

$$\sqrt{\det A^T A}$$

onde A é a matriz que tem v_i como i -ésima coluna. No caso em que $k = 2$ e $n = 3$, é conhecida outra fórmula para a área do paralelogramo: a área é o módulo do produto

externo dos vectores v_1 e v_2 , ou seja $\|v_1\|\|v_2\|\sin\alpha$ com α o ângulo entre v_1 e v_2 . Vamos verificar que este valor coincide com o obtido através da nossa fórmula geral:

$$A^T A = \begin{bmatrix} \|v_1\|^2 & \langle v_1, v_2 \rangle \\ \langle v_2, v_1 \rangle & \|v_2\|^2 \end{bmatrix}$$

Tendo em conta que $\langle v_1, v_2 \rangle = \langle v_2, v_1 \rangle = \|v_1\|\|v_2\|\cos\alpha$ temos

$$\det A^T A = \|v_1\|^2\|v_2\|^2 - \|v_1\|^2\|v_2\|^2\cos^2\alpha = \|v_1\|\|v_2\|^2(1 - \cos^2\alpha)$$

e portanto

$$\sqrt{\det A^T A} = \|v_1\|\|v_2\|\sin\alpha$$

A fórmula para a área do paralelogramo em termos do produto externo das arestas dá-nos a seguinte fórmula alternativa a (16) para o integral ao longo de uma superfície em \mathbb{R}^3 de uma função escalar. Sendo $g(t, s)$ uma parametrização de S ,

$$\int_S f dS = \int \int_{g^{-1}(S)} f(g(t, s)) \left\| \frac{\partial g}{\partial t} \times \frac{\partial g}{\partial s} \right\| dt ds$$

Nota 36.2. A fórmula anterior (válida apenas para superfícies em \mathbb{R}^3) não é necessariamente mais fácil de utilizar que a fórmula geral (16). A razão pelo que a incluímos é que ela aparece frequentemente em livros de texto.

Exemplo 36.3. Calcular a área de

$$S = \{(x, y, z) \in \mathbb{R}^3 : z = \sqrt{x^2 + y^2}, 1 < z < 2, x > 0, y > 0\}$$

Em coordenadas cilíndricas em torno do eixo dos zz , a superfície é definida pela equação $z = r$ com $1 < z < 2$ e $0 < \theta < \frac{\pi}{2}$. Uma parametrização para S é portanto dada por

$$g(r, \theta) = (r \cos \theta, r \sin \theta, r), \quad \text{com } 1 < r < 2, 0 < \theta < \frac{\pi}{2}$$

Temos

$$\frac{\partial g}{\partial r} \times \frac{\partial g}{\partial \theta} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \cos \theta & \sin \theta & 1 \\ -r \sin \theta & r \cos \theta & 0 \end{vmatrix}$$

onde \mathbf{i} , \mathbf{j} e \mathbf{k} são os vectores da base canónica de \mathbb{R}^3 , logo

$$\frac{\partial g}{\partial r} \times \frac{\partial g}{\partial \theta} = -r \cos(\theta)\mathbf{i} - r \sin(\theta)\mathbf{j} + r\mathbf{k} = (-r \cos \theta, -r \sin \theta, r)$$

e

$$\left\| \frac{\partial g}{\partial r} \times \frac{\partial g}{\partial \theta} \right\| = \sqrt{r^2 \cos^2 \theta + r^2 \sin^2 \theta + r^2} = \sqrt{2}r$$

pelo que a área de S é

$$\int_S 1 = \int_1^2 \int_0^{\frac{\pi}{2}} \sqrt{2}r d\theta dr = \frac{3\pi}{2\sqrt{2}}.$$

36.4. Justificação da fórmula para o factor de conversão de volumes k -dimensionais.

Sejam v_1, \dots, v_k vectores de \mathbb{R}^n . Suponhamos primeiro que estes vectores são ortogonais dois a dois (isto é $\langle v_i, v_j \rangle = 0$, para $i \neq j$). Então o volume k -dimensional do paralelogramo com arestas v_i

$$P(v_1, \dots, v_k) = \{t_1 v_1 + \dots + t_k v_k \in \mathbb{R}^n : 0 \leq t_1, \dots, t_k \leq 1\}$$

deveria ser

$$\text{vol}(P) = \|v_1\| \|v_2\| \cdots \|v_k\|$$

Se escrevermos A para a matriz $n \times k$ que tem v_i na i -ésima coluna, a matriz $A^T A$ será uma matriz diagonal (a entrada ij com $i \neq j$ é $\langle v_i, v_j \rangle = 0$) e a i -ésima entrada da diagonal é $\langle v_i, v_i \rangle = \|v_i\|^2$. Ou seja,

$$A^T A = \begin{bmatrix} \|v_1\|^2 & \cdots & 0 \\ 0 & \ddots & 0 \\ 0 & \cdots & \|v_k\|^2 \end{bmatrix}$$

Conclui-se que a fórmula $\sqrt{\det A^T A}$ fornece o valor correcto para o volume de P .

O caso geral em que v_1, \dots, v_k são vectores quaisquer de \mathbb{R}^n pode ser reduzido ao caso especial anterior da seguinte forma. Seja w_k a projecção de v_k no espaço ortogonal ao espaço $L(\{v_1, \dots, v_{k-1}\})$ gerado pelos restantes vectores. É fácil convenceremo-nos que o volume- k dimensional dos paralelogramos $P(v_1, \dots, v_{k-1}, v_k)$ e $P(v_1, \dots, v_{k-1}, w_k)$ devem ser iguais (pela aditividade do volume - faça um desenho quando $k = 2$). Fazendo este raciocínio indutivamente, conclui-se que se w_1, \dots, w_k se obtêm de v_1, \dots, v_k pelo processo de ortogonalização de Gram-Schmidt, então

$$\text{vol } P(v_1, \dots, v_k) = \text{vol } P(w_1, \dots, w_k)$$

Basta-nos portanto verificar que se B é a matriz $n \times k$ que tem por colunas os vectores w_i , temos

$$\sqrt{\det A^T A} = \sqrt{\det B^T B}$$

Mas por definição do processo de ortogonalização de Gram-Schmidt, a relação entre B e A é dada pela equação

$$B = AC$$

onde C é uma matriz triangular superior com 1s na diagonal. Por exemplo, se $k = 3$, a matriz C é

$$C = \begin{bmatrix} 1 & -\frac{\langle v_2, v_1 \rangle}{\|v_1\|^2} & -\frac{\langle v_3, v_1 \rangle}{\|v_1\|^2} \\ 0 & 1 & -\frac{\langle v_3, v_2 \rangle}{\|v_2\|^2} \\ 0 & 0 & 1 \end{bmatrix}$$

Conclui-se que

$$\begin{aligned}\sqrt{\det B^T B} &= \sqrt{\det(AC)^T(AC)} \\ &= \sqrt{\det C^T A^T AC} \\ &= \sqrt{\det C^T \det A^T A \det C} \\ &= \sqrt{1 \cdot \det(A^T A) \cdot 1} = \sqrt{\det A^T A}.\end{aligned}$$

36.5. Propriedades do integral de campos escalares. O integral de campos escalares tem as propriedades habituais de um integral que agora relembramos. O último item na lista abaixo não é, em rigor, uma propriedade do integral mas sim a verificação de que a definição de integral faz sentido! De facto, se o número obtido na fórmula (16) dependesse da parametrização g não estaríamos verdadeiramente a definir um valor associado apenas a A e f .

Proposição 36.6 (Propriedades do integral de campos escalares em variedades). *Seja $S \subset \mathbb{R}^n$ uma variedade- k , $A \subset S$ um conjunto contido na imagem de alguma parametrização de S , e $f, k: A \rightarrow \mathbb{R}$ funções.*

- (i) (Linearidade) Para cada $\alpha, \beta \in \mathbb{R}$, temos $\int_A \alpha f + \beta k = \alpha \int_A f + \beta \int_A k$
- (ii) (Monotonia) Se $f \leq k$, então $\int_A f \leq \int_A k$
- (iii) (Desigualdade triangular) $|\int_A f| \leq \int_A |f|$
- (iv) (Aditividade) Se $A = A_1 \cup A_2$ e $A_1 \cap A_2 = \emptyset$ então $\int_A f = \int_{A_1} f + \int_{A_2} f$
- (v) (Invariância por reparametrização) Se $g: V \rightarrow \mathbb{R}^n$ e $h: W \rightarrow \mathbb{R}^n$ são parametrizações de S que contêm A na sua imagem, então

$$\int_{g^{-1}(A)} f(g(t)) \sqrt{\det Dg(t)^T Dg(t)} dt = \int_{h^{-1}(A)} f(h(s)) \sqrt{\det Dh(s)^T Dh(s)} ds$$

37. INTEGRAIS DE LINHA DE CAMPOS VECTORIAIS

Nota 37.1. *Só definimos integral num subconjunto A de uma variedade S que esteja contido na imagem de alguma parametrização. Embora esta restrição não tenha consequências práticas para o cálculo de integrais (porque é sempre possível encontrar uma parametrização de S cuja imagem cobre todo o S exceto um subconjunto com volume k -dimensional nulo - pense-se por exemplo na parametrização da circunferência dada por $g(\theta) = (\cos \theta, \sin \theta)$ com $0 < \theta < 2\pi$) ela não é muito elegante. Sugere-se aos alunos interessados que consultem [S] para ver como definir integral num subconjunto arbitrário de uma variedade.*

Dem. da Proposição 36.6(v). As parametrizações g e h são funções injectivas. A função que “traduz” dos parâmetros t para os parâmetros s é a função $\phi: g^{-1}(h(W)) \rightarrow W$ definida por

$$s = \phi(t) = h^{-1}(g(t))$$

Claramente ϕ é uma função injetiva (porque é a composta de duas funções injetivas), mas não é claro que seja de classe C^1 (não se pode invocar a regra de derivação da função composta porque h^{-1} só está definida num subconjunto da variedade, que não é um aberto de \mathbb{R}^n). Assim que demonstrarmos que ϕ é de classe C^1 , a sua função inversa $\phi^{-1} = g^{-1} \circ h$

também o será (é da mesma natureza) e portanto ϕ será uma transformação de coordenadas em \mathbb{R}^k . Vamos adiar a justificação que ϕ é C^1 e fazer a mudança de variável $s = \phi(t)$ no integral

$$\int_{h^{-1}(A)} f(h(s)) \sqrt{\det Dh(s)^T Dh(s)} ds$$

Como $\phi^{-1}(h^{-1}(A)) = g^{-1}(h(h^{-1}(A))) = g^{-1}(A)$ e $f(h(\phi(t))) = f(h(h^{-1}(g(t)))) = f(g(t))$ obtemos

$$(17) \quad \int_{g^{-1}(A)} f(g(t)) \sqrt{\det Dh(\phi(t))^T Dh(\phi(t))} |\det D\phi(t)| dt$$

Por outro lado, uma vez que $h \circ \phi = g$, a regra de derivação da função composta diz-nos que

$$Dg(t) = Dh(\phi(t))D\phi(t)$$

logo

$$Dg(t)^T Dg(t) = D\phi(t)^T Dh(\phi(t))^T Dh(\phi(t)) D\phi(t)$$

e

$$\begin{aligned} \sqrt{\det Dg(t)^T Dg(t)} &= \sqrt{\det D\phi(t) \det(Dh(\phi(t))^T Dh(\phi(t))) \det D\phi(t)} \\ &= |\det D\phi(t)| \sqrt{\det(Dh(\phi(t))^T Dh(\phi(t)))} \end{aligned}$$

Substituindo na equação (17) obtemos a expressão pretendida.

Resta-nos justificar que a função $\phi = h^{-1} \circ g$ é de classe C^1 . Isto é fácil de ver se h for uma parametrização em termos de k das coordenadas de \mathbb{R}^n : Escrevendo os vectores de \mathbb{R}^n como (x, y) com $x \in \mathbb{R}^k$ o vector formado pelas k coordenadas livres e y o vector de \mathbb{R}^{n-k} formado pelas restantes coordenadas, teríamos $h(x) = (x, k(x))$ com k uma função de classe C^1 definida num aberto de \mathbb{R}^k tomando valores em \mathbb{R}^{n-k} . Para uma tal parametrização

$$h^{-1}(x, y) = x$$

logo a expressão para a função $\phi = h^{-1} \circ g$ é dada por k das funções coordenadas de g e portanto ϕ é de classe C^1 .

Além disso, a função $\phi = h^{-1} \circ g$ tem inversa de classe C^1 pelo teorema da Função Inversa: De facto, a função $h^{-1}: h(W) \rightarrow W$ pode ser prolongada naturalmente à função $\pi: \mathbb{R}^n \rightarrow \mathbb{R}^k$ definida por $\pi(x, y) = x$. Uma vez que $\pi(h(x)) = x$ e π é de classe C^1 , a regra de derivação da função composta garante que $D\pi(h(x))Dh(x) = \text{Id}$. Logo a restrição de $D\pi = \pi$ (π é linear) a $T_{h(x)}S \subset \mathbb{R}^n$ é um isomorfismo de espaços vectoriais

$$\pi: T_{h(x)}S \rightarrow \mathbb{R}^k$$

Como $\phi = h^{-1} \circ g = \pi \circ g$ temos, pela regra de derivação da função composta,

$$D\phi(t) = D\pi(g(t)) \circ Dg(t) = \pi \circ Dg(t)$$

Mas $Dg(t)$ é um isomorfismo entre \mathbb{R}^k e $T_{g(t)}S$, logo $D\phi(t)$ é um isomorfismo de \mathbb{R}^k para \mathbb{R}^k . Conclui-se que $\det D\phi(t) \neq 0$ para todo o t no domínio de ϕ e portanto, pelo Teorema da Função Inversa, ϕ^{-1} é de classe C^1 .

Para terminar a demonstração, sejam agora g, h parametrizações arbitrárias. Dado t pertencente ao domínio de ϕ , seja f uma parametrização de uma vizinhança de $g(t)$ em S como o gráfico de uma função de k das variáveis. Então, numa vizinhança de t temos que

$$\phi = h^{-1} \circ g = (h^{-1} \circ f) \circ (f^{-1} \circ g) = (f^{-1} \circ h)^{-1} \circ (f^{-1} \circ g)$$

é a composta de duas funções de classe C^1 e é portanto de classe C^1 . \square

37.2. Integrais de linha de campos vectoriais. Um *campo vectorial* em \mathbb{R}^n é uma função $\vec{F}: U \rightarrow \mathbb{R}^n$ com $U \subset \mathbb{R}^n$. Uma maneira de visualizar um campo vectorial (quando $n = 2$ ou 3) é desenhar para alguns pontos $x \in U$ um vector $\vec{F}(x)$ com origem em x .

Os campos vectoriais surgem frequentemente na Física como modelos para campos de forças, como por exemplo os campos gravíticos, eléctricos ou magnéticos, para os quais a visualização indicada acima faz todo o sentido.

Vamos agora definir o integral de um campo vectorial ao longo de uma curva (variedade-1). A ideia geométrica da definição é que o integral calcula a “soma das contribuições de \vec{F} segundo a curva” ou, mais precisamente, é o integral do campo escalar que associa a cada ponto x da curva a componente tangente do vector $\vec{F}(x)$. Para calcular essa componente tangente é necessário escolher um sentido no qual a curva é percorrida (a componente tangente difere num sinal consoante o sentido). Se escrevermos $\vec{t}(x)$ para um vector tangente à curva em x , unitário, que aponta na direcção em que estamos a percorrer a curva, a componente tangente de $\vec{F}(x)$ é

$$\vec{F}(x) \cdot \vec{t}(x)$$

Seja $g: [t_0, t_1] \rightarrow C$ uma parametrização de C . Então

$$\vec{t}(g(t)) = \pm \frac{g'(t)}{\|g'(t)\|}$$

com sinal positivo ou negativo consoante $g(t)$ percorre a curva C no sentido pretendido ou no sentido contrário. Portanto

$$\int_C \vec{F}(x) \cdot \vec{t}(x) = \pm \int_{t_0}^{t_1} \vec{F}(g(t)) \cdot \frac{g'(t)}{\|g'(t)\|} \|g'(t)\| dt = \pm \int_{t_0}^{t_1} \vec{F}(g(t)) \cdot g'(t) dt$$

É esta a fórmula que utilizamos para definir o integral de um campo vectorial.

Definição 37.3. *Seja C uma variedade-1 em \mathbb{R}^n e suponhamos que foi escolhido um sentido no qual C deve ser percorrida. Se $g:]t_0, t_1[\rightarrow C$ é uma parametrização de toda a curva C , definimos*

$$(18) \quad \int_C \vec{F} \cdot d\vec{r} = \pm \int_{t_0}^{t_1} \vec{F}(g(t)) \cdot g'(t) dt$$

em que o sinal é $+$ se a parametrização percorre C no sentido pretendido e $-$ caso contrário.

Nota 37.4. *Para que a expressão no termo direito (18) faça sentido, não é necessário que g seja uma parametrização de uma variedade. Basta que $g: [t_0, t_1] \rightarrow \mathbb{R}^n$ seja uma função contínua e seccionalmente C^1 (o que significa que existe uma partição de $[t_0, t_1]$ numa*

família finita de subintervalos em cada um dos quais g é continuamente diferenciável). Uma tal função g diz-se um caminho seccionalmente regular e a sua imagem não é necessariamente uma variedade (pode auto interseccionar-se e conter pontos nos quais não há reta tangente). Define-se o integral de caminho de um campo vectorial \vec{F} ao longo do caminho seccionalmente regular g pela expressão

$$\int_{t_0}^{t_1} \vec{F}(g(t)) \cdot g'(t) dt$$

e é costume denotar este valor por

$$\int_C \vec{F} \cdot d\vec{r}$$

com C a imagem do caminho g . Esta notação é abusiva porque C não determina o caminho g (mesmo a menos de reparametrização) mas é standard.

Nota 37.5. O significado físico de $\int_C \vec{F} \cdot d\vec{r}$ é o trabalho realizado pelo campo de forças \vec{F} sobre uma partícula que percorre a trajectória C no sentido dado.

Exemplo 37.6. Sendo $C = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 = 1\}$ e $\vec{F}(x, y) = (x - y, x^2)$, calcular

$$\int_C \vec{F} \cdot d\vec{r}$$

onde C é percorrida no sentido horário.

Uma parametrização para C é dada por $g(t) = (\cos t, \sin t)$ com $0 \leq t \leq 2\pi$. Temos $g'(t) = (-\sin t, \cos t)$. A parametrização percorre C no sentido contrário ao pretendido portanto

$$\begin{aligned} \int_C (x - y, x^2) \cdot d\vec{r} &= - \int_0^{2\pi} (\cos t - \sin t, \cos^2 t) \cdot (-\sin t, \cos t) dt \\ &= - \int_0^{2\pi} -\sin t \cos t + \sin^2 t + \cos^3 t dt = -\pi. \end{aligned}$$

38. CAMPOS GRADIENTES E CAMPOS FECHADOS

Exemplo 38.1. Calcular o trabalho realizado pelo campo de forças $\vec{F}(x, y, z) = (2, z, 1)$ sobre uma partícula que percorre o segmento de reta que une $(0, 1, 0)$ a $(1, 2, 3)$.

Uma parametrização para o segmento de reta é dado pela expressão

$$g(t) = (0, 1, 0) + t((1, 2, 3) - (0, 1, 0)) = (t, 1 + t, 3t), \quad \text{com } 0 \leq t \leq 1$$

Temos

$$g'(t) = (1, 1, 3)$$

e portanto

$$\int_C \vec{F} \cdot d\vec{r} = \int_0^1 (2, 3t, 1) \cdot (1, 1, 3) dt = \int_0^1 2 + 3t + 3 dt = 5 + \frac{3}{2} = \frac{13}{2}$$

Vamos agora ver a primeira metade do Teorema Fundamental do Cálculo para integrais de linha de campos vectoriais. Embora fundamental, este resultado é uma consequência imediata da definição e da Regra de Barrow de Cálculo 1.

Teorema 38.2 (Regra de Barrow). *Seja $\varphi: U \rightarrow \mathbb{R}$ uma função de classe C^1 e $\vec{F} = \nabla\varphi$. Se $g: [t_0, t_1] \rightarrow U$ é um caminho seccionalmente regular com ponto inicial $g(t_0) = A$ e ponto final $g(t_1) = B$ então*

$$\int_C \vec{F} \cdot d\vec{r} = \varphi(B) - \varphi(A)$$

(onde C denota a imagem de g).

Dem. Por definição

$$\int_C \vec{F} \cdot d\vec{r} = \int_{t_0}^{t_1} \vec{F}(g(t)) \cdot g'(t) dt = \int_{t_0}^{t_1} \nabla\varphi(g(t)) \cdot g'(t) dt$$

Mas pela regra de derivação da função composta temos

$$\nabla\varphi(g(t)) \cdot g'(t) = \frac{d}{dt} (\varphi(g(t)))$$

logo o integral acima é

$$\int_{t_0}^{t_1} \frac{d}{dt} (\varphi(g(t))) dt = \varphi(g(t_1)) - \varphi(g(t_0)) = \varphi(B) - \varphi(A)$$

onde na primeira igualdade usamos a regra de Barrow de Cálculo 1. \square

Um campo vectorial \vec{F} que é o gradiente de uma função escalar φ diz-se um *campo gradiente* e uma função φ tal que $\vec{F} = \nabla\varphi$ diz-se um *potencial escalar* para \vec{F} . Em termos Físicos, o Teorema 38.2 diz que se \vec{F} é o gradiente de uma função potencial, o trabalho realizado por \vec{F} sobre uma partícula é igual à diferença de potencial entre os pontos final e inicial da trajectória.

O Teorema 38.2 diz-nos, em particular, que o integral de caminho de um campo gradiente depende apenas das extremidades do caminho em questão. Um campo vectorial com esta propriedade diz-se *conservativo*.

Exemplo 38.3. *Um campo vectorial constante num aberto de \mathbb{R}^n é conservativo. Consideremos por exemplo o campo vectorial $\vec{F}(x, y, z) = (1, 3, -2)$. Para ver que \vec{F} é conservativo basta achar um potencial escalar para \vec{F} , ou seja uma solução do sistema*

$$\begin{cases} \frac{\partial\varphi}{\partial x} = 1 \\ \frac{\partial\varphi}{\partial y} = 3 \\ \frac{\partial\varphi}{\partial z} = -2 \end{cases}$$

É imediato reconhecer que

$$\varphi(x, y, z) = x + 3y - 2z$$

é uma solução do sistema (em breve iremos ver como calcular um potencial de forma mais metódica) logo \vec{F} é conservativo.

Na realidade os conceitos de campo conservativo e gradiente são equivalentes. Enunciamos aqui este resultado (que forma parte do teorema fundamental do cálculo para integrais de linha) mas adiamos a demonstração até à próxima aula.

Teorema 38.4. *Seja $U \subset \mathbb{R}^n$ um aberto e $\vec{F}: U \rightarrow \mathbb{R}^n$ um campo vectorial contínuo. Então as seguintes afirmações são equivalentes.*

- (i) \vec{F} é um campo conservativo.
- (ii) $\int_C \vec{F} \cdot d\vec{r} = 0$ para todo o caminho seccionalmente regular fechado (isto é, com o ponto inicial igual ao ponto final) em U .
- (iii) \vec{F} é um campo gradiente.

Quando um campo vectorial \vec{F} é conservativo, o Teorema 38.2 simplifica o cálculo de integrais quer porque o reduz à determinação de um potencial escalar para \vec{F} , quer porque permite o cálculo do integral ao longo de um caminho mais conveniente com as mesmas extremidades. É por isso importante ter um critério eficiente para determinar se um campo vectorial é ou não um gradiente.

Uma condição necessária para que um campo seja gradiente é-nos dada pelo Lema de Schwarz (Teorema 13.4). Por exemplo, suponhamos que $\vec{F}: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ é um campo gradiente de classe C^1 . Isto significa que $(F_1, F_2) = \left(\frac{\partial \varphi}{\partial x}, \frac{\partial \varphi}{\partial y} \right)$ para alguma função potencial $\varphi: \mathbb{R}^2 \rightarrow \mathbb{R}$. Então, pelo Lema de Schwarz, temos

$$\frac{\partial F_1}{\partial y} = \frac{\partial^2 \varphi}{\partial y \partial x} = \frac{\partial^2 \varphi}{\partial x \partial y} = \frac{\partial F_2}{\partial x}$$

Definição 38.5. *Seja $U \subset \mathbb{R}^n$ um aberto. Um campo vectorial $\vec{F} = (F_1, \dots, F_n): U \rightarrow \mathbb{R}^n$ diz-se fechado se para todo o $1 \leq i < j \leq n$ se tem*

$$\frac{\partial F_i}{\partial x_j} = \frac{\partial F_j}{\partial x_i}$$

Proposição 38.6. *Seja $U \subset \mathbb{R}^n$ um aberto e $\vec{F}: U \rightarrow \mathbb{R}^n$ um campo vectorial de classe C^1 . Se \vec{F} é um campo gradiente, então \vec{F} é fechado.*

Dem. Se $\vec{F} = \nabla \varphi$ então temos

$$\frac{\partial F_i}{\partial x_j} = \frac{\partial}{\partial x_j} \left(\frac{\partial \varphi}{\partial x_i} \right) = \frac{\partial^2 \varphi}{\partial x_j \partial x_i} = \frac{\partial^2 \varphi}{\partial x_i \partial x_j} = \frac{\partial}{\partial x_i} \left(\frac{\partial \varphi}{\partial x_j} \right) = \frac{\partial F_j}{\partial x_i}$$

□

Nota 38.7. *O critério anterior dá-nos uma condição necessária para que um campo seja gradiente. Veremos mais adiante que não é suficiente, isto é, que existem campos fechados que não são gradientes.*

Note-se ainda que “praticamente nenhum” campo é fechado, no sentido em que se escrevermos uma expressão ao acaso para um campo, a probabilidade de produzir um campo fechado é extremamente pequena (seria 0 se conseguíssemos realmente escolher o campo ao acaso). Por maioria de razão, o mesmo se aplica a campos gradientes. Dito de outro

modo, os campos gradientes (ou seja, primitiváveis) são muito escassos e portanto “especiais. Não seria de esperar que se pudesse aplicar o Teorema 38.2 muitas vezes para calcular integrais mas, na realidade, muitos dos campos de forças mais importantes em Física (o campo gravitacional, o campo eléctrico) são conservativos pelo que a regra de Barrow é bastante útil.

Exemplo 38.8. Determinar quais dos seguintes campos vectoriais são conservativos e, em caso afirmativo, calcular um potencial escalar para o campo.

(1) $\vec{F}: \mathbb{R}^3 \rightarrow \mathbb{R}^3$ definido por $\vec{F}(x, y, z) = (y + z, x, x + yz)$.

Vamos ver se \vec{F} é fechado:

$$\begin{aligned}\frac{\partial F_1}{\partial y} &= \frac{\partial}{\partial y}(y + z) = 1 = \frac{\partial}{\partial x}(x) = \frac{\partial F_2}{\partial x} \\ \frac{\partial F_1}{\partial z} &= \frac{\partial}{\partial z}(y + z) = 1 = \frac{\partial}{\partial x}(x + yz) = \frac{\partial F_3}{\partial x} \\ \frac{\partial F_2}{\partial z} &= \frac{\partial}{\partial z}(y + z) = 1 \neq z = \frac{\partial}{\partial y}(x + yz) = \frac{\partial F_3}{\partial y}\end{aligned}$$

Conclui-se que \vec{F} não é fechado e portanto não é conservativo.

(2) $\vec{F}: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ definido por $\vec{F}(x, y) = (x^2 + y, x - y)$.

Uma vez que

$$\frac{\partial F_1}{\partial y} = 1 = \frac{\partial F_2}{\partial x}$$

o campo \vec{F} é fechado. O campo \vec{F} é um gradiente, sse é possível resolver o sistema de equações

$$(19) \quad \begin{cases} \frac{\partial \varphi}{\partial x} = x^2 + y \\ \frac{\partial \varphi}{\partial y} = x - y \end{cases}$$

Recorde-se que a derivada parcial em ordem a x é a derivada da função de x que se obtém fixando y (ou seja, “tomando y como uma constante”). Resolver a equação

$$\frac{\partial \varphi}{\partial x} = x^2 + y$$

consiste portanto em escolher, para cada y , uma primitiva para a função de x definida por $x \mapsto x^2 + y$. Uma solução é

$$\varphi(x, y) = \frac{x^3}{3} + xy$$

Há no entanto outras soluções porque para cada y podemos somar a esta primitiva uma constante. Não há no entanto razão para que a constante que somamos seja a mesma para todo o y . Escrevendo $A(y)$ para a constante de integração para um dado y obtemos

$$\frac{\partial \varphi}{\partial x} = x^2 + y \Leftrightarrow \varphi(x, y) = \frac{x^3}{3} + xy + A(y)$$

e substituindo na segunda equação de (19) obtemos

$$\frac{\partial}{\partial y} \left(\frac{x^3}{3} + xy + A(y) \right) = x - y \Leftrightarrow x + A'(y) = x - y \Leftrightarrow A'(y) = -y$$

Assim

$$A(y) = -\frac{y^2}{2} + C \quad \text{com } C \in \mathbb{R}$$

e concluímos que, por exemplo,

$$\varphi(x, y) = \frac{x^3}{3} + xy - \frac{y^2}{2}$$

é um potencial para \vec{F} , e portanto \vec{F} é conservativo.

39. PROPRIEDADES DO INTEGRAL

O integral de caminho de um campo vectorial tem as propriedades habituais de um integral.

Proposição 39.1 (Propriedades do integral de caminho). *Seja $U \subset \mathbb{R}^n$ um aberto, $\vec{F}, \vec{G}: U \rightarrow \mathbb{R}^n$ campos vectoriais contínuos, e C a imagem de um caminho seccionalmente regular $g: [a, b] \rightarrow U$. Então*

(i) (Linearidade) Dados $\alpha, \beta \in \mathbb{R}$ tem-se $\int_C (\alpha \vec{F} + \beta \vec{G}) \cdot d\vec{r} = \alpha \int_C \vec{F} \cdot d\vec{r} + \beta \int_C \vec{G} \cdot d\vec{r}$

(ii) (Aditividade) Se c é tal que $a < c < b$, g_1 é a restrição de g a $[a, c]$, g_2 é a restrição de g a $[c, b]$ e C_1, C_2 são as imagens de g_1 e g_2 (de forma que $C = C_1 \cup C_2$) então

$$\int_C \vec{F} \cdot d\vec{r} = \int_{C_1} \vec{F} \cdot d\vec{r} + \int_{C_2} \vec{F} \cdot d\vec{r}$$

(iii) (Desigualdade triangular) $\left| \int_C \vec{F} \cdot d\vec{r} \right| \leq \int_C \|\vec{F}\|$ (onde o integral do lado direito é o integral do campo escalar $\|\vec{F}\|$)

(iv) (Invariância por reparametrização) Se $\phi: [c, d] \rightarrow [a, b]$ é uma função bijectiva, de classe C^1 , cuja derivada não se anula, e $h: [c, d] \rightarrow U$ é o caminho definido por $h = g \circ \phi$, então

$$\int_a^b \vec{F}(g(t)) \cdot g'(t) dt = \pm \int_c^d \vec{F}(h(s)) \cdot h'(s) ds$$

onde o sinal é + se $\phi'(t) > 0$ (e portanto os caminhos g e h têm a mesma orientação) e - se $\phi'(t) < 0$ (e portanto g e h têm orientações opostas).

Dem: Exercício. Para provar a última propriedade basta fazer a mudança de variável $s = \phi(t)$ no integral respeitante ao caminho h . Note-se que se $\phi'(t) < 0$, temos $\phi^{-1}(c) = b$ e $\phi^{-1}(d) = a$ o que justifica o sinal negativo. \square

A aditividade do integral é frequentemente usada quando o caminho de integração se separa naturalmente em vários troços. Seria então aborrecido escrever uma parametrização para o caminho inteiro - seria uma função definida por ramos. Vejamos um exemplo.

Exemplo 39.2. Calcular o trabalho realizado por $\vec{F}(x, y) = (y + xy, x^2)$ ao longo da fronteira do triângulo $\{(x, y) : x, y \geq 0, x + y \leq 1\}$ percorrido no sentido anti-horário.

Escrevendo $C_1 = \{(x, 0) \in \mathbb{R}^2 : 0 \leq x \leq 1\}$, $C_2 = \{(x, y) \in \mathbb{R}^2 : x + y = 1, x \geq 0, y \geq 0\}$ e $C_3 = \{(0, y) \in \mathbb{R}^2 : 0 \leq y \leq 1\}$ para os três lados do triângulo, temos

$$\int_C \vec{F} \cdot d\vec{r} = \int_{C_1} \vec{F} \cdot d\vec{r} + \int_{C_2} \vec{F} \cdot d\vec{r} + \int_{C_3} \vec{F} \cdot d\vec{r}$$

Uma vez que $\vec{F}(x, 0) = (0, x^2)$ é perpendicular ao caminho C_1 , e $\vec{F}(0, y) = (y, 0)$ é perpendicular a C_3 vemos que

$$\int_{C_1} \vec{F} \cdot d\vec{r} = \int_{C_3} \vec{F} \cdot d\vec{r} = 0$$

Uma parametrização de C_2 é dada por $g(t) = (1, 0) + t((0, 1) - (1, 0)) = (1 - t, t)$ com $0 \leq t \leq 1$ logo o trabalho é

$$\begin{aligned} \int_C \vec{F} \cdot d\vec{r} &= \int_{C_2} \vec{F} \cdot d\vec{r} \\ &= \int_0^1 \vec{F}(1 - t, t) \cdot (-1, 1) dt = \int_0^1 (t + (1 - t)t, (1 - t)^2) \cdot (-1, 1) dt \\ &= \int_0^1 -2t + t^2 + 1 - 2t + t^2 dt = -\frac{1}{3} \end{aligned}$$

Exemplo 39.3. Mostrar que $\vec{F}: \mathbb{R}^3 \rightarrow \mathbb{R}^3$ definido por $\vec{F}(x, y, z) = (2x + ze^{xz}, z, y + xe^{xz})$ é um campo gradiente e aproveitar para calcular o integral de \vec{F} ao longo do caminho $g: [0, 2\pi] \rightarrow \mathbb{R}^3$ definido por $g(t) = (t \cos t, t \sin t, t)$.

Para mostrar que \vec{F} é um gradiente temos que achar um potencial escalar, ou seja, resolver o sistema

$$\begin{cases} \frac{\partial \varphi}{\partial x} = 2x + ze^{xz} \\ \frac{\partial \varphi}{\partial y} = z \\ \frac{\partial \varphi}{\partial z} = y + xe^{xz} \end{cases}$$

A primeira equação é equivalente a

$$\varphi(x, y, z) = x^2 + e^{xz} + A(y, z)$$

onde $A(y, z)$ é uma função qualquer de classe C^1 . Substituindo na segunda equação obtém-se

$$\frac{\partial}{\partial y} (x^2 + e^{xz} + A(y, z)) = z \Leftrightarrow \frac{\partial A}{\partial y} = z \Leftrightarrow A(y, z) = yz + B(z)$$

Finalmente substituindo

$$\varphi(x, y, z) = x^2 + e^{xz} + yz + B(z)$$

na última equação temos

$$\frac{\partial}{\partial z} (x^2 + e^{xz} + yz + B(z)) = y + xe^{xz} \Leftrightarrow xe^{xz} + y + B'(z) = y + xe^{xz} \Leftrightarrow B'(z) = 0$$

Conclui-se que um potencial escalar para \vec{F} é dado por

$$\varphi(x, y, z) = x^2 + e^{xz} + yz$$

(os outros potenciais diferem deste numa constante). Pela Regra de Barrow, o integral de \vec{F} ao longo do caminho g é dado por

$$\varphi(g(2\pi)) - \varphi(g(0)) = \varphi(2\pi, 0, 2\pi) - \varphi(0, 0, 0) = 4\pi^2 + e^{4\pi^2} - 1.$$

40. O TEOREMA FUNDAMENTAL DO CÁLCULO

Exemplo 40.1 (Um campo radial é conservativo). Um campo radial é um campo

$$\vec{F}: \mathbb{R}^n \setminus \{0\} \rightarrow \mathbb{R}^n$$

que no ponto (x_1, \dots, x_n) tem a direção da recta que une o ponto à origem (isto é a direção do vector (x_1, \dots, x_n)) e cuja magnitude depende apenas da distância à origem.

Escrevendo \vec{e}_r para o versor radial - um vector unitário que aponta na direção oposta à origem - temos a expressão

$$\vec{e}_r = \frac{(x_1, \dots, x_n)}{\|(x_1, \dots, x_n)\|} = \left(\frac{x_1}{\sqrt{x_1^2 + \dots + x_n^2}}, \dots, \frac{x_n}{\sqrt{x_1^2 + \dots + x_n^2}} \right)$$

e escrevendo $r = \sqrt{x_1^2 + \dots + x_n^2}$ para a distância à origem, concluímos que um campo radial é um campo vectorial $\vec{F}: \mathbb{R}^n \setminus \{0\}$ que pode ser escrito na forma

$$\vec{F}(x_1, \dots, x_n) = f(r)\vec{e}_r = f(r) \left(\frac{x_1}{r}, \dots, \frac{x_n}{r} \right)$$

Para verificar que um tal campo é conservativo podemos achar um potencial. Temos que resolver a equação

$$\nabla\phi = \vec{F}$$

para a incógnita $\phi: \mathbb{R}^n \setminus \{0\} \rightarrow \mathbb{R}$, ou seja o sistema

$$\begin{cases} \frac{\partial\phi}{\partial x_1} = f(r) \frac{x_1}{r} \\ \vdots \\ \frac{\partial\phi}{\partial x_n} = f(r) \frac{x_n}{r} \end{cases}$$

Uma vez que

$$\frac{\partial r}{\partial x_i} = \frac{1}{2} (x_1^2 + \dots + x_n^2)^{-\frac{1}{2}} 2x_i = \frac{x_i}{r}$$

cada equação do sistema acima é da forma

$$\frac{\partial\phi}{\partial x_i} = f(r) \frac{\partial r}{\partial x_i}$$

e sendo $V(r) = \int f(r)dr$ uma primitiva da função $f:]0, +\infty[\rightarrow \mathbb{R}$ temos

$$\frac{\partial}{\partial x_i} \left(V \left(\sqrt{x_1^2 + \dots + x_n^2} \right) \right) = V'(r) \frac{\partial r}{\partial x_i} = f(r) \frac{x_i}{r}$$

Conclui-se que

$$\phi(x_1, \dots, x_n) = V \left(\sqrt{x_1^2 + \dots + x_n^2} \right)$$

é um potencial para \vec{F} .

Um exemplo concreto é o campo gravitacional de uma massa pontual M , que de acordo com a lei da atração universal de Newton é dado pela expressão

$$\vec{F}(x, y, z) = -\frac{GM}{r^2} \vec{e}_r$$

(com G a constante de atração universal). Uma vez que

$$\int -\frac{GM}{r^2} dr = \frac{GM}{r}$$

temos a seguinte expressão para o potencial gravítico:

$$V(r) = \frac{GM}{r}.$$

40.2. Conjuntos conexos.

Definição 40.3. Seja A um subconjunto de \mathbb{R}^n . Um caminho em A é uma função contínua $g: [a, b] \rightarrow A$ com $a, b \in \mathbb{R}$ e $a < b$. Um conjunto aberto $U \subset \mathbb{R}^n$ diz-se conexo (por arcos) se para todos os $A, B \in U$ existe um caminho $g: [a, b] \rightarrow A$ com $g(a) = A$ e $g(b) = B$ (a e b são arbitrários).

Um conjunto conexo é um conjunto que, intuitivamente, é “formado por uma só peça”.

Exemplo 40.4. (i) \mathbb{R}^n é conexo para todo o n . Um caminho que une $A, B \in \mathbb{R}^n$ é por exemplo o segmento de reta $g(t) = A + t(B - A)$ com $0 \leq t \leq 1$.

(ii) $\mathbb{R}^2 \setminus \{(0, 0)\}$ é conexo.

(iii) $U = \mathbb{R}^2 \setminus \{(x, 0) : x \in \mathbb{R}\}$ não é conexo (não é possível ligar um ponto acima do eixo dos xx a um ponto abaixo do eixo dos xx - se $g(t) = (g_1(t), g_2(t))$ for um caminho que una os dois pontos, $g_2(t)$ tem de se anular pelo Teorema do Valor Intermediário, e portanto a imagem do caminho não está contida em U).

Qualquer subconjunto aberto $U \subset \mathbb{R}^n$ pode ser escrito como uma união disjunta de abertos conexos, que se chamam as *componentes conexas* de U . No exemplo 40.4(iii), as componentes conexas são $\{(x, y) \in \mathbb{R}^2 : y > 0\}$ e $\{(x, y) \in \mathbb{R}^2 : y < 0\}$.

40.5. O Teorema Fundamental do Cálculo.

Teorema 40.6 (Teorema Fundamental do Cálculo). Seja $U \subset \mathbb{R}^n$ um aberto e $\vec{F}: U \rightarrow \mathbb{R}^n$ um campo vetorial contínuo. As seguintes afirmações são equivalentes

(i) \vec{F} é conservativo

(ii) $\oint_C \vec{F} \cdot d\vec{r} = 0$ para todo o caminho fechado seccionalmente regular contido em U .

(iii) Existe $\varphi: U \rightarrow \mathbb{R}$ de classe C^1 tal que $\vec{F} = \nabla\varphi$.

Se as condições acima se verificam, V é uma componente conexa de U , e x_0 é um ponto de V qualquer fixado, um potencial escalar para \vec{F} em V é dado pelo integral indefinido

$$\varphi(x) = \int_{x_0}^x \vec{F} \cdot d\vec{r}$$

Dem. • (i) \Rightarrow (ii): Seja $g: [a, b] \rightarrow U$ um caminho seccionalmente regular fechado e $A = g(a) = g(b)$. Seja $h: [a, b] \rightarrow \mathbb{R}^n$ o caminho constante definido por $h(t) = A$. Como \vec{F} é conservativo temos

$$\int_a^b \vec{F}(g(t)) \cdot g'(t) dt = \int_a^b \vec{F}(h(t)) \cdot h'(t) dt = \int_a^b \vec{F}(h(t)) \cdot 0 dt = 0$$

- (ii) \Rightarrow (i): Sejam C_1 e C_2 as curvas descritas por dois caminhos seccionalmente regulares g_1 e g_2 com ponto inicial A e ponto final B . Escrevendo $-C_2$ para a curva C_2 percorrida no sentido inverso, temos que $C = C_1 \cup -C_2$ é a imagem de um caminho seccionalmente regular fechado. Por hipótese temos

$$\oint_C \vec{F} \cdot d\vec{r} = 0$$

Por aditividade do integral temos

$$\oint_C \vec{F} \cdot d\vec{r} = \int_{C_1} \vec{F} \cdot d\vec{r} - \int_{C_2} \vec{F} \cdot d\vec{r}$$

e concluimos portanto que

$$\int_{C_1} \vec{F} \cdot d\vec{r} = \int_{C_2} \vec{F} \cdot d\vec{r}$$

Portanto \vec{F} é conservativo.

- (iii) \Rightarrow (i): É uma consequência imediata da regra de Barrow (Teorema 38.2)
- (i) \Rightarrow (iii): Basta ver que para cada componente conexa V de U e $x_0 \in V$, o integral indefinido

$$\varphi(x) = \int_{x_0}^x \vec{F} \cdot d\vec{r}$$

é um potencial para \vec{F} em V . Note-se que $\varphi(x)$ está bem definido uma vez que V é conexo (logo existe um caminho em V que liga x_0 a x_1) e \vec{F} é conservativo (logo o valor do integral é independente do caminho escolhido).

Por definição temos

$$\frac{\partial \varphi}{\partial x_i}(x) = \lim_{h \rightarrow 0} \frac{\varphi(x + he_i) - \varphi(x)}{h}$$

com e_i o i -ésimo vetor da base canónica de \mathbb{R}^n . Para h suficientemente pequeno, o segmento de recta que une x a $x + he_i$ está contido em V e portanto, para calcular

$\varphi(x + he_i)$ podemos escolher um caminho qualquer de x_0 a x seguido do segmento de recta $[x, x + he_i]$. Então

$$\varphi(x + he_i) - \varphi(x) = \int_{x_0}^x \vec{F} \cdot d\vec{r} + \int_x^{x+he_i} \vec{F} \cdot d\vec{r} - \int_{x_0}^x \vec{F} \cdot d\vec{r} = \int_x^{x+he_i} \vec{F} \cdot d\vec{r}$$

e usando $g(t) = x + the_i$ com $0 \leq t \leq 1$ como parametrização do segmento de reta de x a $x + he_i$ temos

$$\frac{\varphi(x + he_i) - \varphi(x)}{h} = \frac{1}{h} \int_0^1 \vec{F}(x + the_i) \cdot he_i dt = \int_0^1 F_i(x + the_i) dt$$

Uma vez que a i -ésima componente F_i de F é uma função contínua, temos que

$$\lim_{h \rightarrow 0} \int_0^1 F_i(x + the_i) dt = F_i(x)$$

logo

$$\frac{\partial \varphi}{\partial x_i}(x) = F_i(x)$$

o que conclui a demonstração. □

41. O TEOREMA DE GREEN

Exemplo 41.1 (O ralo da banheira). *Seja $\vec{F}: \mathbb{R}^2 \setminus \{(0, 0)\} \rightarrow \mathbb{R}^2$ o campo vetorial definido por*

$$\vec{F}(x, y) = \left(-\frac{y}{x^2 + y^2}, \frac{x}{x^2 + y^2} \right)$$

Uma vez que o vetor $(-y, x)$ é perpendicular a (x, y) , a direção de \vec{F} é tangente às circunferências centradas na origem. No primeiro quadrante, a primeira componente de \vec{F} é negativa, logo \vec{F} aponta no sentido anti-horário. Finalmente,

$$\|\vec{F}(x, y)\| = \sqrt{\frac{x^2}{(x^2 + y^2)^2} + \frac{y^2}{(x^2 + y^2)^2}} = \frac{1}{\sqrt{x^2 + y^2}}$$

de forma que o comprimento de $\vec{F}(x, y)$ é inversamente proporcional à distância à origem. Há portanto alguma semelhança entre \vec{F} e o campo de velocidades de um fluido que se escoar através do ralo de uma banheira, pelo que é frequente referir-mo-nos a $\vec{F}(x, y)$ como o ralo da banheira.

É fácil verificar que $\vec{F}(x, y)$ é fechado:

$$\frac{\partial}{\partial y} \left(-\frac{y}{x^2 + y^2} \right) = -\frac{(x^2 + y^2) - 2y^2}{(x^2 + y^2)^2} = \frac{y^2 - x^2}{(x^2 + y^2)^2}$$

é igual a

$$\frac{\partial}{\partial x} \left(\frac{x}{x^2 + y^2} \right) = \frac{(x^2 + y^2) - 2x^2}{(x^2 + y^2)^2} = \frac{y^2 - x^2}{(x^2 + y^2)^2}$$

No entanto $\vec{F}(x, y)$ não é um gradiente: se calcularmos o integral de \vec{F} ao longo da circunferência C parametrizada por $g(\theta) = (\cos \theta, \sin \theta)$ com $0 \leq \theta \leq 2\pi$, obtemos

$$\begin{aligned} \oint_C \vec{F} \cdot d\vec{r} &= \int_0^{2\pi} \vec{F}(\cos \theta, \sin \theta) \cdot (-\sin \theta, \cos \theta) d\theta \\ &= \int_0^{2\pi} \left(\frac{-\sin \theta}{\cos^2 \theta + \sin^2 \theta}, \frac{\cos \theta}{\cos^2 \theta + \sin^2 \theta} \right) \cdot (-\sin \theta, \cos \theta) d\theta \\ &= \int_0^{2\pi} 1 d\theta = 2\pi \end{aligned}$$

O facto de existir um caminho fechado ao longo do qual o integral de \vec{F} não é nulo mostra que \vec{F} não é um gradiente (cf. Teorema 40.6(ii)). O cálculo anterior podia alternativamente ser efetuado recorrendo ao significado geométrico do integral de linha de um campo vetorial: como $\vec{F}(x, y)$ é tangente às circunferências de raio r centradas na origem e aponta no sentido anti-horário, a componente de \vec{F} tangente a uma circunferência de raio r percorrida no sentido direto é $\vec{F} \cdot \vec{t} = \frac{1}{r}$. Conclui-se que o integral de \vec{F} ao longo da circunferência é dado por $2\pi r \cdot \frac{1}{r} = 2\pi$.

Vejamos o que falha quando tentamos calcular um potencial para \vec{F} . Tentamos resolver o sistema

$$\begin{cases} \frac{\partial \varphi}{\partial x} = -\frac{y}{x^2+y^2} \\ \frac{\partial \varphi}{\partial y} = \frac{x}{x^2+y^2} \end{cases}$$

Para $x \neq 0$, a segunda equação é equivalente a

$$\frac{\partial \varphi}{\partial y} = \frac{\frac{1}{x}}{1 + \left(\frac{y}{x}\right)^2} \Leftrightarrow \varphi(x, y) = \arctan\left(\frac{y}{x}\right) + C(x)$$

para alguma função $C(x)$. Substituindo esta solução na primeira equação do sistema obtemos $C'(x) = 0$. Conclui-se que para $x \neq 0$, existem funções potencial $\varphi(x, y)$ para $\vec{F}(x, y)$ dadas pela expressão

$$\varphi(x, y) = \begin{cases} \arctan\left(\frac{y}{x}\right) + C_1 & \text{se } x > 0 \\ \arctan\left(\frac{y}{x}\right) + C_2 & \text{se } x < 0 \end{cases}$$

com $C_1, C_2 \in \mathbb{R}$ (as constantes de integração em cada componente conexas do domínio podem ser diferentes). Os potenciais que encontramos diferem do ângulo das coordenadas polares numa constante. É possível escolher as constantes de forma a que a função $\varphi(x, y)$ fique contínua (e na realidade - exercício - de classe C^∞) em metade do eixo dos yy (podemos tomar por exemplo $C_1 = 0, C_2 = \pi$) mas é impossível escolher C_1 e C_2 de forma a obter uma função $\varphi(x, y)$ contínua no domínio $\mathbb{R}^2 \setminus \{(0, 0)\}$ de $\vec{F}(x, y)$.

O exemplo anterior mostra que o facto de um campo vetorial ser ou não gradiente depende não apenas da expressão que define o campo mas do domínio em questão (e mais precisamente da forma desse domínio como veremos mais à frente).

O seguinte Teorema Fundamental do Cálculo para integrais de linha no plano ajudar-nos-á a perceber melhor o exemplo anterior.

Teorema 41.2 (Teorema de Green). *Seja $U \subset \mathbb{R}^2$ um aberto, e $(P, Q): U \rightarrow \mathbb{R}^2$ um campo vetorial de classe C^1 . Seja $A \subset \mathbb{R}^2$ um aberto limitado com $A = \text{int } \bar{A}$, e $\bar{A} \subset U$ tal que $\partial A = \text{front } A$ é uma variedade-1 em \mathbb{R}^2 . Então*

$$\int \int_A \frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} dx dy = \oint_{\partial A} (P, Q) \cdot d\vec{r}$$

em que ∂A é percorrida de tal forma que A está “sempre à esquerda”.

Como todos os teoremas fundamentais do cálculo, o Teorema anterior tem a forma geral seguinte: a soma dos valores de uma certa derivada de uma função numa região (neste caso a derivada é $\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}$) é igual à soma da função nas “extremidades” dessa região.

O Teorema dá uma igualdade entre dois integrais de tipos diferentes. Como veremos nos exemplos abaixo, uma das razões porque é útil é que permite substituir o cálculo de um integral (que pode ser difícil ou impossível de calcular diretamente) pelo cálculo do outro. Pode também ser aplicado de forma mais sofisticada, como iremos ver.

Nota 41.3. *Um aberto do plano que satisfaça as condições exigidas para A no enunciado do Teorema de Green chama-se um domínio regular. A ideia da definição é que A é uma região do plano limitada por curvas regulares. O Teorema de Green pode aplicar-se a regiões mais gerais que são “limites” de domínios regulares, passando ao limite nos integrais em ambos os lados da igualdade dada no Teorema. Por exemplo o Teorema é válido se A é um intervalo aberto $]a, b[\times]c, d[$ (que não é um domínio regular porque a fronteira - a linha retangular - não é uma variedade). Isto pode ser visto aplicando o teorema a “retângulos com cantos arredondados” que aproximam A cada vez melhor e passando ao limite.*

Nota 41.4. *É frequente escrever-se*

$$\oint_{\partial A} P dx + Q dy$$

para o trabalho

$$\oint_{\partial A} (P, Q) \cdot d\vec{r}$$

Trata-se apenas de uma notação alternativa para o integral de linha de um campo vetorial.

Exemplo 41.5. *Calcular*

$$\oint_C (x^2 + y^2) dx + (x + y) dy$$

onde C é a fronteira do quadrado $[-1, 1] \times [-1, 1]$ percorrida no sentido horário.

Aplicando o Teorema de Green à região $A =]-1, 1[\times]-1, 1[$ limitada por C e notando que o sentido de $C = \partial A$ é o sentido oposto ao indicado no Teorema de Green temos

$$\begin{aligned} \oint_C (x^2 + y^2)dx + (x + y)dy &= - \int_{-1}^1 \int_{-1}^1 \frac{\partial}{\partial x} (x^2 + y^2) - \frac{\partial}{\partial y} (x + y) dx dy \\ &= - \int_{-1}^1 \int_{-1}^1 2x - 1 dx dy = -(-4) = 4. \end{aligned}$$

Observe-se que é muito mais simples calcular o trabalho de $(P, Q) = (x^2 + y^2, x + y)$ ao longo de C desta forma. Para fazer pela definição seria preciso parametrizar os quatro segmentos de reta que constituem C individualmente e calcular o trabalho de (P, Q) ao longo de cada um dos segmentos.

Exemplo 41.6. Calcular a área do círculo de raio 1 centrado na origem usando o Teorema de Green.

A área de uma região A do plano é dada por

$$\int \int_A 1 dx dy$$

Para a calcular usando o Teorema de Green achamos um campo vetorial (P, Q) no plano tal que

$$\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} = 1$$

Podemos por exemplo tomar $P(x, y) = 0$ e $Q(x, y) = x$. Temos então pelo Teorema de Green

$$\int \int_A 1 dx dy = \int_{\partial A} (0, x) \cdot d\vec{r}$$

em que A deve ser percorrida de forma a que A esteja à esquerda. Aplicando a fórmula acima ao círculo $A = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 = 1\}$ temos que ∂A é a circunferência de raio 1 percorrida no sentido direto. Uma parametrização que percorre ∂A nesse sentido é

$$g(\theta) = (\cos \theta, \sin \theta), \quad 0 \leq \theta \leq 2\pi$$

e portanto

$$\text{área}(A) = \int_0^{2\pi} (0, \cos \theta) \cdot (-\sin \theta, \cos \theta) d\theta = \int_0^{2\pi} \cos^2 \theta d\theta$$

Usando as fórmulas trigonométricas $\cos(2\theta) = \cos^2 \theta - \sin^2 \theta$ e $\sin^2 \theta = 1 - \cos^2 \theta$ obtém-se $\cos 2\theta = 2 \cos^2 \theta - 1 \Leftrightarrow \cos^2 \theta = \frac{1 + \cos(2\theta)}{2}$ e o valor para a área é

$$\int_0^{2\pi} \frac{1 + \cos(2\theta)}{2} d\theta = \pi.$$

Exemplo 41.7. Consideremos o campo $\vec{F}: \mathbb{R}^2 \setminus \{(0, 0)\} \rightarrow \mathbb{R}^2$ do exemplo 41.1, definido pela expressão

$$\vec{F}(x, y) = (P(x, y), Q(x, y)) = \left(-\frac{y}{x^2 + y^2}, \frac{x}{x^2 + y^2} \right)$$

Vamos agora calcular o trabalho realizado por \vec{F} ao longo da elipse

$$C = \{(x, y) \in \mathbb{R}^2: \frac{x^2}{4} + \frac{y^2}{9} = 1\}$$

percorrida no sentido direto.

Não seria fácil fazer as contas diretamente parametrizando a elipse. No entanto, podemos aplicar o Teorema de Green à região

$$A = \{(x, y) \in \mathbb{R}^2: x^2 + y^2 \geq 1 \text{ e } \frac{x^2}{4} + \frac{y^2}{9} \leq 1\}$$

limitada pela elipse e pela circunferência S de raio 1. Note-se que (P, Q) é fechado sse

$$\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} = 0$$

Uma vez que o campo \vec{F} é fechado, o Teorema de Green diz-nos que

$$\int \int_A 0 \, dx dy = \oint_{\partial A} \vec{F} \cdot d\vec{r}$$

A fronteira de A é formada pela elipse e pela circunferência, isto é,

$$\partial A = S \cup C$$

e de acordo com o Teorema de Green a componente C deve ser percorrida no sentido anti-horário enquanto que S deve ser percorrida no sentido horário. Assim temos

$$0 = \oint_{\partial A} \vec{F} \cdot d\vec{r} = \oint_S \vec{F} \cdot d\vec{r} + \oint_C \vec{F} \cdot d\vec{r}$$

sendo os sentidos de C e S os indicados acima. Conclui-se que

$$\oint_C \vec{F} \cdot d\vec{r} = - \oint_S \vec{F} \cdot d\vec{r} = -(-2\pi) = 2\pi$$

(o integral ao longo de S no sentido direto foi calculado no Exemplo 41.1).

O cálculo anterior exemplifica uma aplicação mais sofisticada do Teorema de Green: Se (P, Q) é fechado, o Teorema de Green pode ser usado para substituir o caminho de integração ao longo do qual queremos calcular o trabalho de (P, Q) por outro caminho mais conveniente.

Mais geralmente o argumento anterior permanece válido se substituirmos C por qualquer curva fechada (seccionalmente regular, sem auto-interseções) em $\mathbb{R}^2 \setminus \{(0, 0)\}$ que contenha a origem no seu interior. Obtemos

$$\oint_C \vec{F} \cdot d\vec{r} = \pm 2\pi$$

consoante o sentido em que C é percorrida é o direto ou o horário.

Note-se ainda que se C não contém a origem no seu interior, aplicando o Teorema de Green à região A limitada por C obtém-se

$$\oint_C \vec{F} \cdot d\vec{r} = 0$$

O Teorema de Green abre assim as portas ao cálculo do integral de um campo vetorial fechado em \mathbb{R}^2 ao longo de qualquer curva fechada contida no seu domínio, e portanto, à aplicação do Teorema 40.6(ii) como critério efetivo para decidir se um campo fechado é ou não conservativo.

42. HOMOTOPIA DE CAMINHOS

Exemplo 42.1. Seja $\vec{F}: \mathbb{R}^2 \setminus \{(0, 1)\} \rightarrow \mathbb{R}^2$ o campo vetorial definido por

$$\vec{F}(x, y) = (P(x, y), Q(x, y)) = \left(-\frac{y-1}{x^2 + (y-1)^2}, \frac{x}{x^2 + (y-1)^2} \right)$$

e C a fronteira do quadrado com vértices nos pontos $(3, 0)$, $(0, 3)$, $(-3, 0)$ e $(0, -3)$ percorrida no sentido positivo. Vamos calcular $\oint_C \vec{F} \cdot d\vec{r}$.

O campo vetorial \vec{F} é o ralo da banheira (Exemplo 41.1) com y substituído por $y-1$, ou seja, é um ralo da banheira centrado no ponto $(0, 1)$. Em particular é um campo fechado e é fácil calcular o trabalho de \vec{F} ao longo de circunferências centradas no ponto $(0, 1)$. Aplicando o Teorema de Green à região A limitada pela curva C e pela circunferência S de raio 1 centrada em $(0, 1)$, isto é, a

$$A = \{(x, y) \in \mathbb{R}^2 : x^2 + (y-1)^2 > 1, -3 < |x+y| < 3, -3 < |x-y| < 3\}$$

obtemos

$$0 = \int \int_A \frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} dx dy = \oint_C \vec{F} \cdot d\vec{r} + \oint_S \vec{F} \cdot d\vec{r}$$

E portanto

$$\oint_C \vec{F} \cdot d\vec{r} = \oint_S \vec{F} \cdot d\vec{r} = 2\pi$$

(o cálculo do trabalho sobre S é inteiramente análogo, e de facto, reduz-se, ao cálculo efectuado no Exemplo 41.1).

O mesmo raciocínio se aplicaria a qualquer curva fechada C_1 sem auto-interseções contendo o ponto $(0, 1)$ no seu interior. Teríamos

$$\oint_{C_1} \vec{F} \cdot d\vec{r} = \pm 2\pi$$

consoante C_1 seja percorrida no sentido direto ou no sentido horário.

Note-se ainda que se C_2 é uma curva fechada sem auto-interseções que não contém $(0, 1)$ no seu interior, então aplicando o Teorema de Green à região limitada pela curva vemos que $\oint_{C_2} \vec{F} \cdot d\vec{r} = 0$. Podemos assim usar o Teorema de Green para calcular o integral de \vec{F} ao longo de qualquer curva fechada contida no domínio $\mathbb{R}^2 \setminus \{(0, 1)\}$.

Dem. do Teorema 41.2. Suponhamos primeiro que a região A pode ser escrita simultaneamente como estando compreendida entre gráficos quer de funções de x quer de y , isto é, que A é da forma

$$A = \{(x, y) \in \mathbb{R}^2: a < x < b, f(x) < y < g(x)\}$$

e

$$A = \{(x, y) \in \mathbb{R}^2: c < y < d, h(y) < x < k(y)\}$$

com $f, g: [a, b] \rightarrow \mathbb{R}$ e $h, k: [c, d] \rightarrow \mathbb{R}$ seccionalmente C^1 . Vamos mostrar o Teorema de Green neste caso recorrendo ao Teorema de Fubini e ao Teorema Fundamental do Cálculo para funções reais de uma variável real. Temos

$$\begin{aligned} \int \int_A \frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} dx dy &= \int \int_A \frac{\partial Q}{\partial x} dx dy - \int \int_A \frac{\partial P}{\partial y} dx dy \\ &= \int_c^d \left(\int_{h(y)}^{k(y)} \frac{\partial Q}{\partial x} dx \right) dy - \int_a^b \left(\int_{f(x)}^{g(x)} \frac{\partial P}{\partial y} dy \right) dx \end{aligned}$$

Começamos por calcular a primeira parcela na última linha acima. Pelo Teorema Fundamental do Cálculo temos

$$(20) \quad \int_c^d \left(\int_{h(y)}^{k(y)} \frac{\partial Q}{\partial x} dx \right) dy = \int_c^d Q(k(y), y) - Q(h(y), y) dy = \int_c^d Q(k(y), y) dy - \int_c^d Q(h(y), y) dy$$

Observamos agora que a fronteira ∂A da região A é constituída por 4 troços parametrizados respetivamente por

$$\begin{aligned} g_1(y) &= (h(y), y), & c \leq y \leq d \\ g_2(x) &= (x, d), & h(d) \leq x \leq k(d) \\ g_3(y) &= (k(y), y), & c \leq y \leq d \\ g_4(x) &= (x, c), & h(c) \leq x \leq k(c) \end{aligned}$$

Os dois primeiros caminhos indicados percorrem a fronteira no sentido horário, enquanto que os últimos têm o sentido anti-horário. Uma vez que

$$g'_1(y) = (h'(y), 1), \quad g'_2(x) = (1, 0), \quad g'_3(y) = (k'(y), 1), \quad g'_4(x) = (1, 0)$$

vemos que o integral do campo $(0, Q)$ ao longo de ∂A percorrida no sentido anti-horário é dado por

$$\begin{aligned} \oint_{\partial A} (0, Q) \cdot d\vec{r} &= - \int_c^d (0, Q(h(y), y)) \cdot (h'(y), 1) dy - \int_{h(d)}^{k(d)} (0, Q(x, d)) \cdot (1, 0) dx \\ &\quad + \int_c^d (0, Q(k(y), y)) \cdot (k'(y), 1) dy + \int_{h(c)}^{k(c)} (0, Q(x, c)) \cdot (1, 0) dx \\ &= - \int_c^d Q(h(y), y) dy + \int_c^d Q(k(y), y) dy \end{aligned}$$

que é precisamente o termo direito em (20). Conclui-se assim que

$$\int \int_A \frac{\partial Q}{\partial x} dx dy = \oint_{\partial A} (0, Q) \cdot d\vec{r}$$

Cálculos inteiramente análogos mostram que

$$\int \int_A -\frac{\partial P}{\partial y} dx dy = \oint_{\partial A} (P, 0) \cdot d\vec{r}$$

o que conclui a demonstração do Teorema de Green neste caso particular. O caso geral do Teorema de Green reduz-se ao caso particular dividindo o domínio regular em regiões elementares do tipo considerado na demonstração acima. Pode mostrar-se que tal é sempre possível, e em qualquer caso, é fácil de verificar diretamente para qualquer A concreto ao qual queiramos aplicar o Teorema de Green. Ao dividir A em regiões elementares A_1, \dots, A_N , teremos

$$\begin{aligned} \int \int_A \frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} dx dy &= \sum_{i=1}^N \int \int_{A_i} \frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} dx dy \\ &= \sum_{i=1}^N \oint_{\partial A_i} P dx + Q dy \end{aligned}$$

As porções das fronteiras ∂A_i que sejam interiores a A pertencem à fronteira comum de duas regiões elementares adjacentes. Como tal aparecem duas vezes na soma acima, e são percorridas duas vezes em sentidos opostos. Isto tem como efeito que todos estes termos cancelam e portanto

$$\sum_{i=1}^N \oint_{\partial A_i} P dx + Q dy = \oint_{\partial A} P dx + Q dy$$

sendo ∂A percorrida no sentido indicado no enunciado. Isto completa a demonstração. \square

O Teorema de Green tem como consequência importante a possibilidade de deformar o caminho de integração de um campo vetorial fechado sem que isso afecte o valor do trabalho. Já vimos alguns exemplos acima para campos vetoriais no plano. Começamos por definir rigorosamente a ideia intuitiva de “deformação de caminhos”.

Definição 42.2. *Seja $D \subset \mathbb{R}^n$ um conjunto. Os caminhos $g_1, g_2: [a, b] \rightarrow D$ com o mesmo ponto inicial $A = g_1(a) = g_2(a)$ e final $B = g_1(b) = g_2(b)$ dizem-se homotópicos em D se existe uma função contínua*

$$H: [a, b] \times [0, 1] \rightarrow D$$

(dita uma homotopia) tal que

$$\begin{aligned} H(t, 0) &= g_1(t) && \text{para } a \leq t \leq b \\ H(t, 1) &= g_2(t) && \text{para } a \leq t \leq b \\ H(a, s) &= A && \text{para } 0 \leq s \leq 1 \\ H(b, s) &= B && \text{para } 0 \leq s \leq 1 \end{aligned}$$

A ideia da definição anterior é a seguinte: Podemos pensar no parâmetro s como o tempo. Ao fixar s obtemos um caminho

$$t \mapsto H(t, s) \quad a \leq t \leq b$$

com ponto inicial A e ponto final B . No “instante inicial” $s = 0$ este caminho é g_1 e, no “instante final” $s = 1$, é o caminho g_2 . H é uma “animação” que nos deforma g_1 em g_2 . A continuidade da função H garante a continuidade da deformação.

Exemplo 42.3. (i) Quaisquer dois caminhos $g_1, g_2: [a, b] \rightarrow \mathbb{R}^n$ com o mesmo ponto inicial e o mesmo ponto final são homotópicos (em \mathbb{R}^n). Intuitivamente isto é claro e é fácil neste caso escrever uma expressão para uma homotopia:

$$H(t, s) = g_1(t) + s(g_2(t) - g_1(t))$$

(ii) Os caminhos $g_1, g_2: [0, \pi] \rightarrow \mathbb{R}^2 \setminus \{(0, 0)\}$ definidos por

$$g_1(\theta) = (\cos \theta, \sin \theta) \quad g_2(\theta) = (\cos \theta, -\sin \theta)$$

não são homotópicos em $\mathbb{R}^2 \setminus \{(0, 0)\}$. Intuitivamente é fácil perceber que uma deformação de g_1 em g_2 teria que “passar por cima de $(0, 0)$ ” e isto pode de facto demonstrar-se.

(iii) O caminho $g_1: [0, 2\pi] \rightarrow \mathbb{R}^3 \setminus \{(0, 0, 0)\}$ definido por

$$g_1(\theta) = (\cos \theta, \sin \theta, 0)$$

é homotópico ao caminho constante $g_2(\theta) = (1, 0, 0)$. Pode definir-se uma homotopia começando por inclinar o plano da circunferência descrita por g_1 relativamente ao plano xy e depois contrair a circunferência ao longo desse plano.

Teorema 42.4. Seja $U \subset \mathbb{R}^n$ um aberto e $\vec{F}: U \rightarrow \mathbb{R}^n$ um campo vetorial fechado de classe C^1 . Se C_1 e C_2 são curvas parametrizadas por caminhos homotópicos então

$$\int_{C_1} \vec{F} \cdot d\vec{r} = \int_{C_2} \vec{F} \cdot d\vec{r}$$

43. CONJUNTOS SIMPLEMENTE CONEXOS

Dem. do Teorema 42.4. Temos a verificar que

$$(21) \quad \oint_{C_1} \vec{F} \cdot d\vec{r} - \oint_{C_2} \vec{F} \cdot d\vec{r} = 0$$

Seja $H: [a, b] \times [0, 1] \rightarrow U$ uma homotopia de classe C^2 entre os caminhos g_1 e g_2 que parametrizam C_1 e C_2 respetivamente. Como

$$g_1(t) = H(t, 0), \quad g_1'(t) = \frac{\partial H}{\partial t}(t, 0), \quad g_2(t) = H(t, 1), \quad g_2'(t) = \frac{\partial H}{\partial t}(t, 1)$$

pela definição de integral de caminho, a igualdade (21) é equivalente a

$$(22) \quad \int_a^b \vec{F}(H(t, 0)) \cdot \frac{\partial H}{\partial t}(t, 0) - \vec{F}(H(t, 1)) \cdot \frac{\partial H}{\partial t}(t, 1) dt = 0$$

O resultado segue da aplicação do Teorema de Green ao quadrado $[a, b] \times [0, 1]$ e ao campo vetorial $(P(t, s), Q(t, s))$ definido em $[a, b] \times [0, 1]$ pelas expressões:

$$P(t, s) = \vec{F}(H(t, s)) \cdot \frac{\partial H}{\partial t} = F_1(H(t, s)) \frac{\partial H_1}{\partial t}(t, s) + \cdots + F_n(H(t, s)) \frac{\partial H_n}{\partial t}(t, s)$$

e

$$Q(t, s) = \vec{F}(H(t, s)) \cdot \frac{\partial H}{\partial s} = F_1(H(t, s)) \frac{\partial H_1}{\partial s}(t, s) + \cdots + F_n(H(t, s)) \frac{\partial H_n}{\partial s}(t, s)$$

De facto,

$$\frac{\partial Q}{\partial t} = \sum_{i=1}^n \left(\left(\sum_{j=1}^n \frac{\partial F_i}{\partial x_j} \frac{\partial H_j}{\partial t} \right) \frac{\partial H_i}{\partial s} + \frac{\partial^2 H_i}{\partial t \partial s} \right)$$

e

$$\frac{\partial P}{\partial s} = \sum_{i=1}^n \left(\left(\sum_{j=1}^n \frac{\partial F_i}{\partial x_j} \frac{\partial H_j}{\partial s} \right) \frac{\partial H_i}{\partial t} + \frac{\partial^2 H_i}{\partial s \partial t} \right)$$

Uma vez que \vec{F} é fechado, $\frac{\partial F_i}{\partial x_j} = \frac{\partial F_j}{\partial x_i}$ e, pelo Lema de Schwarz, $\frac{\partial^2 H_i}{\partial t \partial s} = \frac{\partial^2 H_i}{\partial s \partial t}$ logo

$$\frac{\partial Q}{\partial t} - \frac{\partial P}{\partial s} = 0.$$

O Teorema de Green garante então que o trabalho realizado por (P, Q) ao longo da fronteira do quadrado $[a, b] \times [0, 1]$ no plano (s, t) percorrido no sentido anti-horário é nulo. A fronteira do quadrado é formada por quatro segmentos de reta, parametrizados por

$$\begin{aligned} g_1(t) &= (t, 0), & a \leq t \leq b \\ g_2(s) &= (b, s), & 0 \leq s \leq 1 \\ g_3(t) &= (t, 1), & a \leq t \leq b \\ g_4(s) &= (a, s), & 0 \leq s \leq 1 \end{aligned}$$

sendo que as duas últimas parametrizações têm a orientação contrária à requerida pelo Teorema de Green. Conclui-se que o trabalho de (P, Q) ao longo da fronteira do quadrado é dado por

$$\int_a^b P(t, 0) dt + \int_0^1 Q(b, s) ds - \int_a^b P(t, 1) dt - \int_0^1 Q(a, s) ds = 0$$

O primeiro e terceiro termos na soma anterior são nulos pois $\frac{\partial H}{\partial s}(a, s) = \frac{\partial H}{\partial s}(b, s) = 0$ (a homotopia é constante para $t = a, b$) e os restantes termos dão então a igualdade (22) que pretendíamos demonstrar. \square

Nota 43.1. *Assumi-se na demonstração anterior que a homotopia H (e portanto os caminhos que parametrizam C_1 e C_2) é de classe C^2 . Pode demonstrar-se que se dois caminhos seccionalmente regulares são homotópicos (através de uma homotopia contínua), então existe uma homotopia entre eles que é de classe C^2 em $[a, b] \times]0, 1[$ e o argumento na demonstração anterior garante então a validade do Teorema 42.4 na generalidade em que foi enunciado.*

Nota 43.2. Dois caminhos fechados $g_1, g_2: [a, b] \rightarrow U$ (recorde-se que isto significa que $g_1(a) = g_1(b)$ e $g_2(a) = g_2(b)$) dizem-se livremente homotópicos se existe uma função contínua $H: [a, b] \times [0, 1] \rightarrow U$ satisfazendo

$$H(t, 0) = g_1(t), \quad H(t, 1) = g_2(t), \quad H(a, s) = H(b, s) \text{ para todos os } t, s \in [0, 1].$$

(portanto os “caminhos intermédios” $t \mapsto H(t, s)$ são fechados para cada s mas o ponto inicial (e final) do caminho pode variar com s). É um exercício verificar que o argumento dado na demonstração anterior prova também que se \vec{F} é um campo fechado de classe C^1 e C_1, C_2 são curvas seccionalmente regulares parametrizadas por caminhos fechados livremente homotópicos, então

$$\oint_{C_1} \vec{F} \cdot d\vec{r} = \oint_{C_2} \vec{F} \cdot d\vec{r}$$

O Teorema 42.4 torna efetiva a caracterização de campos gradientes dada no Teorema 40.6(ii) desde que o campo \vec{F} seja fechado.

Exemplo 43.3. Determinar se o campo vetorial $\vec{F}: \mathbb{R}^3 \setminus \{(0, 0, z): z \in \mathbb{R}\} \rightarrow \mathbb{R}^3$ definido por

$$\vec{F}(x, y, z) = \left(\frac{2x}{x^2 + y^2} + z, \frac{2y}{x^2 + y^2}, x \right)$$

é um gradiente no seu domínio.

É fácil verificar que o campo é fechado:

$$\begin{aligned} \frac{\partial}{\partial y} \left(\frac{2x}{x^2 + y^2} + z \right) &= -\frac{4xy}{x^2 + y^2} = \frac{\partial}{\partial x} \left(\frac{2y}{x^2 + y^2} \right) \\ \frac{\partial}{\partial z} \left(\frac{2x}{x^2 + y^2} + z \right) &= 1 = \frac{\partial}{\partial x} (x) \\ \frac{\partial}{\partial z} \left(\frac{2y}{x^2 + y^2} \right) &= 0 = \frac{\partial}{\partial y} (x) \end{aligned}$$

Conclui-se do Teorema 42.4 que o trabalho de \vec{F} ao longo de caminhos fechados homotópicos é igual.

Intuitivamente é claro, e pode demonstrar-se, que qualquer caminho fechado contido no domínio de \vec{F} é, ou homotópico a um caminho constante, ou homotópico a uma circunferência horizontal com o centro no eixo dos zz percorrida um certo número de vezes.

No primeiro caso, o trabalho é 0 (porque o trabalho ao longo de um caminho constante é sempre 0) e para perceber o que se passa no segundo caso basta calcular o trabalho ao longo do caminho

$$g(t) = (\cos t, \sin t, 0), \quad 0 \leq t \leq 2\pi$$

Temos

$$\int_0^1 \vec{F}(g(t)) \cdot g'(t) dt = \int_0^1 (2 \cos t, 2 \sin t, \cos t) \cdot (-\sin t, \cos t, 0) dt = \int_0^1 0 dt = 0$$

Conclui-se que o integral de caminho de \vec{F} ao longo de qualquer caminho seccionalmente regular fechado é 0, e portanto, pelo Teorema 40.6(ii), o campo \vec{F} é um gradiente.

O argumento anterior mostra mais geralmente que se \vec{F} é um campo vetorial de classe C^1 definido no complementar do eixo dos zz , \vec{F} é um gradiente sse o integral de \vec{F} ao longo de uma circunferência horizontal centrada no eixo dos zz é 0.

O Teorema 42.4 permite-nos ainda identificar certos domínios em \mathbb{R}^n nos quais as condições de um campo vetorial ser fechado e de ser gradiente são equivalentes.

Definição 43.4. Um conjunto aberto conexo $U \subset \mathbb{R}^n$ diz-se simplesmente conexo se todo o caminho fechado em U é homotópico a um caminho constante.

Exemplo 43.5. (i) \mathbb{R}^n é simplesmente conexo para todo o n (pois como vimos no Exemplo 42.3(i) quaisquer dois caminhos em \mathbb{R}^n são homotópicos).

(ii) $\mathbb{R}^2 \setminus \{(0, 0)\}$ não é simplesmente conexo. Um caminho fechado que não é homotópico a um caminho constante é $g(t) = (\cos t, \sin t)$ com $0 \leq t \leq 2\pi$. Isto é claro intuitivamente (uma homotopia com um caminho constante “teria de passar por cima da origem”, o que não é permitido) e pode ser demonstrado - embora a demonstração seja surpreendentemente difícil.

Mais geralmente, um subconjunto aberto do plano $U \subset \mathbb{R}^2$ é simplesmente conexo se “não tem buracos”.

(iii) $\mathbb{R}^3 \setminus \{(0, 0, 0)\}$ é simplesmente conexo.

(iv) $U = \mathbb{R}^3 \setminus \{(x, y, 0) : x^2 + y^2 = 1\}$ não é simplesmente conexo. Qualquer caminho fechado que “dê uma volta” à circunferência que falta em U não pode ser deformado dentro de U num caminho constante.

Nota 43.6. Seja $U \subset \mathbb{R}^n$ um conjunto aberto conexo. É um bom exercício demonstrar que U é simplesmente conexo sse quaisquer dois caminhos $g_1, g_2: [a, b] \rightarrow U$ com os mesmos pontos inicial e final são homotópicos. Esta formulação alternativa justifica a terminologia “simplesmente conexo”. De facto, um caminho entre dois pontos $A, B \in U$ é uma maneira de mostrar que os pontos podem ser ligados em U . Dizer que um tal caminho é único a menos de deformação é dizer que “essencialmente” há apenas uma maneira de ligar A e B . Contraste-se com o caso de $U = \mathbb{R}^2 \setminus \{(0, 0)\}$, onde há “essencialmente” duas maneiras de ligar $(-1, 0)$ a $(1, 0)$ (o caminho pode passar por cima ou por baixo de $(0, 0)$). Uma tal região do plano, com “um buraco” diz-se duplamente conexa. Uma região plana com “dois buracos” diz-se triplamente conexa e assim sucessivamente.

Proposição 43.7. Seja $U \subset \mathbb{R}^n$ um aberto simplesmente conexo e $\vec{F}: U \rightarrow \mathbb{R}^n$ um campo vetorial de classe C^1 fechado. Então \vec{F} é um gradiente em U .

Dem. Qualquer caminho fechado em U é homotópico a um caminho constante logo temos

$$\oint_C \vec{F} \cdot d\vec{r} = 0$$

para toda a curva seccionalmente regular C contida em U . Segue-se do Teorema 40.6 que \vec{F} é um gradiente em U . \square

Exemplo 43.8. (i) O campo vetorial $\vec{F}: \mathbb{R}^3 \rightarrow \mathbb{R}^3$ definido por $\vec{F}(x, y, z) = (x^2 + y, x - y, z)$ é fechado. Uma vez que \mathbb{R}^3 é simplesmente conexo, conclui-se que \vec{F} é um gradiente.

(ii) Uma vez que $\mathbb{R}^3 \setminus \{(0, 0, 0)\}$ é simplesmente conexo, um campo vetorial de classe C^1 $\vec{F}: \mathbb{R}^3 \setminus \{(0, 0, 0)\}$ é gradiente sse é fechado.

43.9. Homotopia de caminhos no espaço dos referenciais. Seja R a variedade dos referenciais ortonormados em \mathbb{R}^3 descrita no Exemplo 33.1, que pode ser também vista como a variedade formada pelas rotações em \mathbb{R}^3 . Nesse exemplo vimos que podemos visualizar o espaço R identificando os pontos antipodais na fronteira da bola de raio π centrada na origem

$$\{(x, y, z) \in \mathbb{R}^3: x^2 + y^2 + z^2 \leq \pi^2\}$$

(um ponto v da bola corresponde à rotação de um ângulo $\|v\|$ em torno do eixo orientado determinado por v).

Um caminho em R pode ser visualizado usando um cinto. Apoiando um referencial no cinto e percorrendo-o a velocidade uniforme (imaginem um pequeno carro com um referencial dentro a percorrer o cinto) obtemos um caminho em R - a orientação do cinto em cada ponto determina o referencial no instante correspondente. Se a orientação no espaço do início e final do cinto é a mesma, trata-se de um caminho fechado. Uma homotopia de caminhos visualiza-se através de um movimento qualquer do cinto em que as extremidades se mantenham com a orientação fixa.

É possível demonstrar que o caminho fechado em R determinado por

$$g(t) = (t, 0, 0), \quad -\pi \leq t \leq \pi$$

(este caminho é fechado porque os pontos antipodais $(-\pi, 0, 0)$ e $(\pi, 0, 0)$ são identificados em R) não é homotópico a um caminho constante. Note-se que este caminho associa a t a rotação de um ângulo t em torno do eixo dos xx . Logo, ao longo do caminho, o primeiro vetor do referencial mantém-se fixo e os restantes dois vetores giram uma volta completa em torno do primeiro. Pode ser visualizado com um cinto começando por esticá-lo e torcendo depois o cinto uma volta inteira em torno do seu eixo.

É fácil ver usando o nosso modelo que, se percorrermos este caminho fechado duas vezes, obtemos um caminho homotópico a um caminho constante (o caminho g está portanto “à volta de um buraco de ordem 2”!): Para visualizar esta homotopia imagine-se uma deformação em que cada percurso do caminho se começa a afastar em direções opostas até se transformar na concatenação dos caminhos

$$g_1(t) = (\pi \cos \frac{t}{2}, \pi \sin \frac{t}{2}, 0), \quad g_2(t) = (\pi \cos \frac{t}{2}, -\pi \sin \frac{t}{2}, 0) \quad \text{com } -\pi \leq t \leq \pi$$

Mas o segundo caminho identifica-se (devido à identificação dos pontos antipodais na superfície da esfera) com

$$g_3(t) = (-\pi \cos \frac{t}{2}, \pi \sin \frac{t}{2}, 0), \quad -\pi \leq t \leq \pi$$

que é exactamente $g_1(t)$ percorrido no sentido oposto. A concatenação de um caminho com ele próprio percorrido no sentido oposto é sempre homotópica a um caminho constante.

É possível visualizar esta homotopia com um cinto ou rodando o nosso braço. Ver os seguintes vídeos:

<https://www.youtube.com/watch?v=CHAFKtCgTM>

<https://www.youtube.com/watch?v=JDJKfs3HqRg>

44. FLUXOS

Vamos agora definir o integral de superfície de um campo vetorial em \mathbb{R}^3 e, mais geralmente, o integral de um campo vetorial em \mathbb{R}^n sobre uma variedade- $(n-1)$ $S \subset \mathbb{R}^n$. O fluxo de \vec{F} através de S é definido como o integral sobre S do campo escalar determinado pela componente de \vec{F} perpendicular a S . Se

$$\vec{n}: S \rightarrow \mathbb{R}^n$$

é um campo vetorial contínuo que associa a cada ponto de S um vetor perpendicular a S unitário, o fluxo na direção de \vec{n} é portanto definido como o seguinte integral de campo escalar

$$(23) \quad \int_S \vec{F} \cdot \vec{n}$$

Em \mathbb{R}^3 , o fluxo tem a seguinte interpretação física que é bom manter presente. Suponhamos que \vec{F} é a *densidade de fluxo* de um fluido (um líquido ou um gás) em movimento. Por definição,

$$\vec{F} = \rho \vec{v}$$

onde ρ é um campo escalar que dá a densidade de massa do fluido e \vec{v} é o campo (vetorial) das velocidades do fluido que a cada ponto do espaço associa a velocidade instantânea de um “elemento de fluido” nesse ponto. Então para S uma superfície

$$\int_S \vec{F} \cdot \vec{n}$$

calcula a *quantidade de massa de fluido que atravessa a superfície S por unidade de tempo na direção de \vec{n}* . Note-se que esta quantidade pode ser negativa, o que significa que, no cômputo geral, há mais matéria a atravessar S no sentido oposto a \vec{n} do que no sentido de \vec{n} . Se, como é habitual, ρ e \vec{v} dependerem do tempo, isto é, se forem campos vetoriais dependentes do tempo, o fluxo calcula a taxa instantânea a que a matéria atravessa S .

Seja $U \subset \mathbb{R}^2$ um aberto e $g: U \rightarrow S$ uma parametrização da superfície S . Escrevendo $g = g(s, t)$, temos que os vectores $\frac{\partial g}{\partial s}$ e $\frac{\partial g}{\partial t}$ formam uma base para o espaço tangente a S no ponto $g(s, t)$. Consequentemente, o produto externo

$$\frac{\partial g}{\partial s} \times \frac{\partial g}{\partial t}$$

é um vetor perpendicular a S em $g(s, t)$ e conclui-se que os dois (assumindo que U é conexo) campos vetoriais normais unitários a S na imagem da parametrização g são dados

pela expressão

$$\vec{n}(g(s, t)) = \pm \frac{\frac{\partial g}{\partial s} \times \frac{\partial g}{\partial t}}{\left\| \frac{\partial g}{\partial s} \times \frac{\partial g}{\partial t} \right\|}$$

Da expressão anterior, e da definição de integral de um campo escalar, deduz-se a seguinte fórmula que podemos usar para calcular o fluxo de um campo vetorial \vec{F} usando uma parametrização. Sendo $A \subset g(U) \subset S$, temos

$$\begin{aligned} \int \int_A \vec{F} \cdot \vec{n} &= \int \int_{g^{-1}(A)} \vec{F}(g(s, t)) \cdot \vec{n}(g(s, t)) \left\| \frac{\partial g}{\partial s} \times \frac{\partial g}{\partial t} \right\| ds dt \\ (24) \qquad &= \pm \int \int_{g^{-1}(A)} \vec{F}(g(s, t)) \cdot \frac{\frac{\partial g}{\partial s} \times \frac{\partial g}{\partial t}}{\left\| \frac{\partial g}{\partial s} \times \frac{\partial g}{\partial t} \right\|} \left\| \frac{\partial g}{\partial s} \times \frac{\partial g}{\partial t} \right\| ds dt \\ &= \pm \int \int_{g^{-1}(A)} \vec{F}(g(s, t)) \cdot \frac{\partial g}{\partial s} \times \frac{\partial g}{\partial t} ds dt \end{aligned}$$

sendo o sinal $+$ se o vetor normal $\frac{\partial g}{\partial s} \times \frac{\partial g}{\partial t}$ tem o mesmo sentido que \vec{n} , e $-$ caso contrário.

Nota 44.1. *Recorda-se que o produto externo de dois vetores $v, w \in \mathbb{R}^3$ se pode calcular através do determinante formal (em que $\mathbf{i}, \mathbf{j}, \mathbf{k}$ denotam os vetores da base canónica de \mathbb{R}^3)*

$$\begin{aligned} v \times w &= \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ v_1 & v_2 & v_3 \\ w_1 & w_2 & w_3 \end{vmatrix} = (v_2 w_3 - v_3 w_2) \mathbf{i} + (v_3 w_1 - v_1 w_3) \mathbf{j} + (v_1 w_2 - v_2 w_1) \mathbf{k} \\ &= (v_2 w_3 - v_3 w_2, v_3 w_1 - v_1 w_3, v_1 w_2 - v_2 w_1) \end{aligned}$$

Nota 44.2. *A definição de fluxo pressupõe a existência de um campo vetorial unitário normal à superfície, contínuo,*

$$\vec{n}: S \rightarrow \mathbb{R}^3$$

que se designa por orientação da superfície e que corresponde a uma “escolha de lado” da superfície. É interessante observar que uma tal escolha não é sempre possível. A banda de Möbius é o exemplo mais simples de uma superfície não orientável e um exemplo mais interessante é a garrafa de Klein que não pode ser entendida enquanto subconjunto de \mathbb{R}^3 . Para imaginar a garrafa de Klein como subconjunto de \mathbb{R}^4 a partir dos modelos tridimensionais apresentados no URL anterior, imagine-se que imediatamente “antes” da superfície se auto-intersestar em \mathbb{R}^3 saímos de \mathbb{R}^3 numa direção perpendicular, regressando apenas após a auto-interseção (da mesma forma que a auto-interseção de uma figura 8 no plano pode ser evitada em \mathbb{R}^3 deslocando uma das linhas do 8 na direção zz perto da auto-interseção).

Note-se que a banda de Möbius é naturalmente um subconjunto do espaço das rotações R do Exemplo 33.1. De facto, usando o modelo para R dado pela bola sólida de raio π com pontos antipodais da fronteira identificados, é imediato verificar que (para $0 < \epsilon < \pi$) a imagem em R do subconjunto

$$\{(x, y, 0) \in \mathbb{R}^3 : x^2 + y^2 \leq \pi^2, |y| < \epsilon\}$$

é uma banda de Möbius (de largura 2ϵ).

O bordo de uma banda de Möbius é uma curva fechada homotópica ao caminho obtido percorrendo a circunferência “central” da banda de Möbius duas vezes. A imagem do disco $\{(x, y, 0) \in \mathbb{R}^3: x^2 + y^2 \leq \pi^2\}$ em R obtém-se colando um disco ao bordo da banda de Möbius descrita no parágrafo anterior. É este disco que dá a homotopia (descrita no final da secção anterior) entre o bordo da banda de Möbius e um caminho constante.

Mais geralmente define-se o produto externo de $(n - 1)$ vetores $v_1, \dots, v_{n-1} \in \mathbb{R}^n$ pela fórmula

$$v_1 \times \cdots \times v_{n-1} = \begin{vmatrix} \mathbf{e}_1 & \cdots & \mathbf{e}_n \\ v_{1,1} & \cdots & v_{1,n} \\ \vdots & & \vdots \\ v_{n-1,1} & \cdots & v_{n-1,n} \end{vmatrix}$$

(onde \mathbf{e}_i denota o i -ésimo vetor da base canónica de \mathbb{R}^n e escrevemos $v_{i,j}$ para a componente j do vetor v_i). O significado geométrico deste produto externo é análogo ao do produto externo em \mathbb{R}^3 . Trata-se de um vetor que tem por magnitude o volume $(n - 1)$ -dimensional do paralelepípedo $(n - 1)$ -dimensional com arestas v_1, \dots, v_{n-1} e que, quando este volume não é nulo, é perpendicular ao plano gerado por v_1, \dots, v_{n-1} e forma juntamente com estes vetores uma base de \mathbb{R}^n orientada positivamente.

Conclui-se que se S é uma variedade- $(n - 1)$ em \mathbb{R}^n , $\vec{n}: S \rightarrow \mathbb{R}^n$ é uma normal unitária contínua, e $g: U \rightarrow S$ é uma parametrização com $g = g(t_1, \dots, t_{n-1})$, então para $A \subset g(U)$ temos

$$\int_A \vec{F} \cdot \vec{n} = \pm \int \cdots \int_{g^{-1}(A)} \vec{F}(g(t_1, \dots, t_{n-1})) \cdot \left(\frac{\partial g}{\partial t_1} \times \cdots \times \frac{\partial g}{\partial t_{n-1}} \right) dt_1 \cdots dt_{n-1}$$

sendo o sinal $+$ se o vetor perpendicular $\frac{\partial g}{\partial t_1} \times \cdots \times \frac{\partial g}{\partial t_{n-1}}$ tem o mesmo sentido que a normal \vec{n} dada, e $-$ caso contrário.

Exemplo 44.3. Calcular o fluxo de $\vec{F}(x, y, z) = (x, y, z)$ através da porção de parabolóide definido por

$$S = \{(x, y, z) \in \mathbb{R}^3: z = x^2 + y^2 - 1, z \leq 0\}$$

no sentido da normal que aponta para o eixo dos zz .

Em coordenadas cilíndricas, a superfície S é descrita pela equação $z = \rho^2 - 1$, e $z \leq 0 \Leftrightarrow \rho \leq 1$ logo uma parametrização para S em termos de duas das coordenadas cilíndricas é dada por

$$g(\rho, \theta) = (\rho \cos \theta, \rho \sin \theta, \rho^2 - 1), \quad 0 < \rho < 1, 0 < \theta < 2\pi$$

Temos

$$\frac{\partial g}{\partial \rho} \times \frac{\partial g}{\partial \theta} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \cos \theta & \sin \theta & 2\rho \\ -\rho \sin \theta & \rho \cos \theta & 0 \end{vmatrix} = (-2\rho^2 \cos \theta, -2\rho^2 \sin \theta, \rho)$$

Uma vez que a terceira componente deste vetor é positiva (ou seja, este vetor aponta para cima), trata-se de um vector normal a S que aponta para o “interior” do parabolóide, e

portanto, no sentido do eixo dos zz . Assim $\frac{\partial g}{\partial \rho} \times \frac{\partial g}{\partial \theta}$ tem o mesmo sentido que \vec{n} e o fluxo é dado por

$$\begin{aligned} \int \int_S \vec{F}(x, y, z) \cdot \vec{n} &= + \int \int_{g^{-1}(S)} \vec{F}(g(\rho, \theta)) \cdot \frac{\partial g}{\partial \rho} \times \frac{\partial g}{\partial \theta} d\rho d\theta \\ &= \int_0^1 \int_0^{2\pi} (\rho \cos \theta, \rho \sin \theta, \rho^2 - 1) \cdot (-2\rho^2 \cos \theta, -2\rho^2 \sin \theta, \rho) d\theta d\rho \\ &= \int_0^1 \int_0^{2\pi} -2\rho^3 + \rho^3 - \rho d\theta d\rho \\ &= 2\pi \left(-\frac{1}{2} - \frac{1}{4} \right) = -\frac{3\pi}{2} \end{aligned}$$

45. O TEOREMA DA DIVERGÊNCIA

É por vezes possível calcular fluxos de campos vetoriais sem aplicar a fórmula geral (24), recorrendo apenas à definição (23). Por exemplo, se é claro que o campo vetorial \vec{F} é tangente a S , então o campo escalar $\vec{F} \cdot \vec{n}$ é identicamente nulo, e portanto o fluxo é nulo. Mais geralmente se virmos diretamente que $\vec{F} \cdot \vec{n}$ é constante, então o fluxo pode ser calculado multiplicando a constante pela área da superfície (que pode ser já conhecida), etc...

Exemplo 45.1. Calcular o fluxo de $\vec{F}(x, y, z) = (x(x^2 + y^2 - 1), y(x^2 + y^2 - 1), z)$ através da fronteira de

$$V = \{(x, y, z) \in \mathbb{R}^3 : 0 \leq z \leq 1, x^2 + y^2 \leq 1\}$$

no sentido da normal exterior a V .

A fronteira de V pode ser escrita como uma união de três superfícies: a superfície cilíndrica

$$S_1 = \{(x, y, z) \in \mathbb{R}^3 : x^2 + y^2 = 1, 0 \leq z \leq 1\}$$

e os círculos

$$S_2 = \{(x, y, 0) \in \mathbb{R}^3 : x^2 + y^2 \leq 1\} \quad e \quad S_3 = \{(x, y, 1) \in \mathbb{R}^3 : x^2 + y^2 \leq 1\}$$

É fácil escrever expressões para as normais unitárias a S_1, S_2 e S_3 que apontam para o exterior de V . Temos respetivamente

$$\vec{n}_1(x, y, z) = (x, y, 0) \quad \text{para } (x, y, z) \in S_1$$

$\vec{n}_2(x, y, 0) = (0, 0, -1)$ para $(x, y, 0) \in S_2$ e $\vec{n}_3(x, y, 1) = (0, 0, 1)$ para $(x, y, 1) \in S_3$.
Portanto

$$\begin{aligned} \int \int_{\partial V} \vec{F} \cdot \vec{n}^{ext} &= \int \int_{S_1} \vec{F} \cdot (x, y, 0) + \int \int_{S_2} \vec{F} \cdot (0, 0, -1) + \int \int_{S_3} \vec{F} \cdot (0, 0, 1) \\ &= \int \int_{S_1} (x^2 + y^2)(x^2 + y^2 - 1) + \int \int_{S_2} -z + \int \int_{S_3} z \\ &= \int \int_{S_1} 0 + \int \int_{S_2} 0 + \int \int_{S_3} 1 \\ &= \text{area}(S_3) = \pi \end{aligned}$$

onde, na penúltima passagem, usamos que $x^2 + y^2 - 1 = 0$ em S_1 e que $z = 0$ em S_2 .

Definição 45.2. Seja $U \subset \mathbb{R}^n$ um aberto e $\vec{F}: U \rightarrow \mathbb{R}^n$ um campo vetorial de classe C^1 . A divergência de \vec{F} é o campo escalar $\text{div } \vec{F}: U \rightarrow \mathbb{R}$ definido pela expressão

$$\text{div } \vec{F} = \frac{\partial F_1}{\partial x_1} + \dots + \frac{\partial F_n}{\partial x_n}$$

Usa-se também a notação

$$\nabla \cdot \vec{F}$$

para a divergência de \vec{F} .

Note-se que a divergência é o traço da matriz Jacobiana da função \vec{F} . Deve ser encarada como “um tipo de derivada” de \vec{F} . Trata-se da derivada relevante para o “Teorema Fundamental do Cálculo para fluxos” que iremos agora ver.

Exemplo 45.3. Se $\vec{F}: \mathbb{R}^3 \rightarrow \mathbb{R}^3$ é o campo vetorial definido por $\vec{F}(x, y, z) = (x^3 + yz, xy - z^2, e^x + \text{sen } z)$, a divergência de \vec{F} é

$$\text{div } \vec{F} = \frac{\partial}{\partial x} (x^3 + yz) + \frac{\partial}{\partial y} (xy - z^2) + \frac{\partial}{\partial z} (e^x + \text{sen } z) = 3x^2 + x + \cos z$$

Definição 45.4. Um aberto $V \subset \mathbb{R}^n$ diz-se um domínio regular se V é limitado, $V = \text{int}(\bar{V})$ e $\partial V = \text{front } V$ é uma variedade- $(n - 1)$.

A ideia da definição anterior é que um domínio regular é uma região aberta de \mathbb{R}^n limitada por uma variedade- $(n - 1)$.

Exemplo 45.5. 1. $V = \{(x, y, z) \in \mathbb{R}^3 : x^2 + y^2 + z^2 < 1\}$ é um domínio regular em \mathbb{R}^3 .

Mais geralmente, para todo o $n \geq 1$, $V = \{(x_1, \dots, x_n) \in \mathbb{R}^n : x_1^2 + \dots + x_n^2 \leq 1\}$ é um domínio regular em \mathbb{R}^n . A fronteira é $\partial V = \{(x_1, \dots, x_n) \in \mathbb{R}^n : x_1^2 + \dots + x_n^2 = 1\}$.

2. $V = \{(x, y, z) \in \mathbb{R}^3 : z > x^2 + y^2\}$ não é um domínio regular porque não é limitado.

3. $V = \{(x, y, z) \in \mathbb{R}^3 : \sqrt{x^2 + y^2} < z < 1\}$ não é um domínio regular porque $\text{front } V$ não é uma variedade (o vértice e a aresta do cone são pontos onde $\text{front } V$ não tem plano tangente).

4. A condição $V = \text{int}(\bar{V})$ destina-se a excluir conjuntos como $\{(x, y) \in \mathbb{R}^2 : x^2 + y^2 < 4 \text{ e } x^2 + y^2 \neq 1\}$ que satisfazem todas as restantes condições.

O seguinte Teorema deve-se a Gauss e é um dos Teoremas Fundamentais do Cálculo Integral. É também conhecido por Teorema de Gauss.

Teorema 45.6 (Teorema da Divergência). *Seja $U \subset \mathbb{R}^n$ um aberto, $\vec{F}: U \rightarrow \mathbb{R}^n$ um campo vetorial de classe C^1 e V um domínio regular com $\bar{V} \subset U$. Então*

$$\int_V \operatorname{div} f = \int_{\partial V} \vec{F} \cdot \vec{n}^{ext}$$

Note-se a forma geral do enunciado anterior que é comum a todos os Teoremas Fundamentais do Cálculo: o integral de uma certa derivada (neste caso a divergência) numa região é igual ao integral da função que se está a derivar sobre as “extremidades” da região (que neste caso formam uma variedade- $(n - 1)$).

Nota 45.7. *O Teorema da Divergência aplica-se mais geralmente a domínios V que podem ser “aproximados por domínios regulares”. A generalização obtém-se passando ao limite nos integrais sobre os volumes e as suas fronteiras dos domínios regulares que aproximam V . Pode, por exemplo, aplicar-se o Teorema da Divergência quando V é um cilindro ou um cone sólido (limitado) ou um cubo.*

Observe-se que no caso em que V é um (hiper)cubo, o Teorema da Divergência é uma consequência imediata do Teorema Fundamental do Cálculo de Cálculo 1. Consideremos por exemplo o caso em que $V = [0, 1] \times [0, 1] \times [0, 1] \subset \mathbb{R}^3$. O campo vetorial $\vec{F}(x, y, z) = (F_1(x, y, z), F_2(x, y, z), F_3(x, y, z))$ pode ser escrito como a soma

$$\vec{F}(x, y, z) = (F_1(x, y, z), 0, 0) + (0, F_2(x, y, z), 0) + (0, 0, F_3(x, y, z))$$

e claramente o Teorema é válido para \vec{F} se fôr válido para cada uma das três parcelas. A verificação do Teorema para cada uma das parcelas é inteiramente análoga. Os cálculos relevantes no caso da primeira parcela são

$$\operatorname{div}(F_1, 0, 0) = \frac{\partial F_1}{\partial x}$$

Portanto

$$\begin{aligned} \int_{[0,1] \times [0,1] \times [0,1]} \operatorname{div} F &= \int_0^1 \left(\int_0^1 \left(\int_0^1 \frac{\partial F_1}{\partial x} dx \right) dy \right) dz \\ &= \int_0^1 \int_0^1 F_1(1, y, z) - F_1(0, y, z) dy dz \\ &= \int_{\{1\} \times [0,1] \times [0,1]} (F_1, 0, 0) \cdot (1, 0, 0) + \int_{\{0\} \times [0,1] \times [0,1]} (F_1, 0, 0) \cdot (-1, 0, 0) \end{aligned}$$

A última expressão é exactamente igual a

$$\int_{\partial V} (F_1, 0, 0) \cdot \vec{n}^{ext}$$

uma vez que $(F_1, 0, 0)$ é tangente às faces do cubo contidas nos planos $y = 0$, $y = 1$, $z = 0$ e $z = 1$ (pelo que a contribuição dessas faces para o fluxo é nula) e as normais unitárias

exteriores às faces contidas nos planos $x = 1$ e $x = 0$ são respectivamente $\vec{n} = (1, 0, 0)$ e $\vec{n} = (-1, 0, 0)$.

O Teorema da Divergência permite interpretar $\text{div } \vec{F}$ como uma *medida da intensidade de criação de fluxo* de \vec{F} num ponto. De facto, sendo x um ponto no domínio de \vec{F} e aplicando o Teorema 45.6 à bola de raio ϵ centrada em x obtemos

$$\int \int \int_{B_\epsilon(x)} \text{div } F = \int \int_{\partial B_\epsilon(x)} \vec{F} \cdot \vec{n}^{ext}$$

Uma vez que \vec{F} é de classe C^1 , a função $\text{div } \vec{F}$ é contínua e portanto o lado esquerdo da igualdade é, para ϵ suficientemente pequeno, aproximadamente $\text{div } F(x) \cdot \text{vol}(B_\epsilon(x))$. Passando ao limite quando $\epsilon \rightarrow 0$ obtém-se (exercício)

$$\text{div } F = \lim_{\epsilon \rightarrow 0} \frac{1}{\text{vol}(B_\epsilon(x))} \int \int_{\partial B_\epsilon(x)} \vec{F} \cdot \vec{n}^{ext}$$

Por exemplo, se $\vec{F} = \rho \vec{v}$ é a densidade de fluxo de um fluido, quando a divergência de \vec{F} é positiva em x , o fluxo para fora de uma pequena bola centrada em x é positivo. Isto significa que está a “sair fluido de x ” - há uma torneira (infinitesimal) aberta em x . Por outro lado se $\text{div } F(x) < 0$, está a desaparecer fluido em x - há um sumidouro ou ralo infinitesimal em x .

Exemplo 45.8. O campo vetorial $\vec{F}(x, y, z) = (x, 0, 0)$ tem divergência constante igual a 1. É um bom exercício esboçar o aspeto deste campo vetorial em \mathbb{R}^3 e perceber intuitivamente a forma como está a ser criado fluxo de \vec{F} em cada ponto de \mathbb{R}^3 .

Dever ser claro que o argumento dado acima proporciona uma interpretação para o significado da divergência de um campo vetorial em \mathbb{R}^n para qualquer n , e não apenas para $n = 3$.

Exemplo 45.9. Seja $V = \{(x, y, z) \in \mathbb{R}^3 : x^2 + y^2 + z^2 \leq 1\}$ e $\vec{F}(x, y, z) = (x, y, z)$. Verifiquemos diretamente a fórmula dada pelo Teorema 45.6 neste caso:

Temos $\text{div } \vec{F} = 1 + 1 + 1 = 3$ logo

$$\int \int \int_V \text{div } \vec{F} = 3 \text{vol}(V) = 3 \cdot \frac{4\pi}{3} = 4\pi.$$

Por outro lado, temos $\partial V = \{(x, y, z) \in \mathbb{R}^3 : x^2 + y^2 + z^2 = 1\}$ e a normal unitária exterior a V é dada pela expressão

$$\vec{n}(x, y, z) = (x, y, z) \quad \text{para } (x, y, z) \in \partial V$$

logo

$$\int \int_{\partial V} \vec{F} \cdot \vec{n}^{ext} = \int \int_{\partial V} x^2 + y^2 + z^2 = \int \int_{\partial V} 1 = \text{área}(\partial V) = 4\pi.$$

46. O TEOREMA DA DIVERGÊNCIA (CONT.)

Exemplo 46.1. Usar o Teorema da Divergência para calcular o fluxo de $\vec{F}(x, y, z) = (0, 0, 1)$ ao longo de

$$S = \{(x, y, z) \in \mathbb{R}^3 : z = 1 - x^2 - y^2, z \geq 0\}$$

no sentido da normal \vec{n} com terceira componente positiva.

O parabolóide S não limita nenhuma região do espaço, pelo que o Teorema da Divergência não pode ser aplicado diretamente para calcular o fluxo através de S . Podemos no entanto “fechar” S juntando-lhe a “tampa”

$$T = \{(x, y, 0) \in \mathbb{R}^3 : x^2 + y^2 \leq 1\}$$

Juntamente com S , T limita a região do espaço

$$V = \{(x, y, z) \in \mathbb{R}^3 : 0 \leq z \leq 1 - x^2 - y^2\}$$

e aplicando o Teorema da Divergência a V obtemos

$$\int \int \int_V \operatorname{div} \vec{F} = \int \int_{\partial V} \vec{F} \cdot \vec{n}^{ext} = \int \int_S \vec{F} \cdot \vec{n}^{ext} + \int \int_T \vec{F} \cdot \vec{n}^{ext}$$

A normal \vec{n} ao parabolóide dada no enunciado é a normal exterior a V , e a normal exterior a tampa T é dada pela expressão $\vec{n}^{ext} = (0, 0, -1)$. Como $\operatorname{div} \vec{F} = 0$ temos

$$\begin{aligned} 0 &= \int \int_S \vec{F} \cdot \vec{n} + \int \int_T \vec{F} \cdot (0, 0, -1) \\ \int \int_S \vec{F} \cdot \vec{n} &= - \int \int_T -1 \\ \int \int_S \vec{F} \cdot \vec{n} &= \text{área}(T) = \pi \end{aligned}$$

Note-se que o argumento acima mostra mais geralmente que o fluxo através de qualquer superfície S que seja fechada pela tampa T (isto é, cujo bordo seja a circunferência $\{(x, y, 0) \in \mathbb{R}^3 : x^2 + y^2 = 1\}$) seria o mesmo (desde que S fosse orientada convenientemente).

Se pensarmos no fluxo em termos físicos como o fluxo de um fluido, podemos interpretar fisicamente a equação dada pelo Teorema da Divergência

$$\int \int_S \vec{F} \cdot \vec{n} = \int \int \int_V \operatorname{div} \vec{F} + \int \int_T \vec{F} \cdot (0, 0, 1)$$

em termos de conservação de massa da seguinte forma: “o que sai” por S é igual ao que “é criado” em V + “o que entra” por T . Com um pouco de prática podemos usar este procedimento para escrever diretamente uma fórmula para o fluxo através de uma superfície que não limite uma região do espaço usando o Teorema da Divergência.

Exemplo 46.2. Calcular o fluxo do campo

$$\vec{F}(x, y, z) = \left(x - \frac{yz}{y^2 + 1}, y + \frac{x^2}{x^2 + 1}, z \right)$$

ao longo da superfície do cilindro

$$S = \{(x, y, z) \in \mathbb{R}^3 : x^2 + y^2 = 1, 0 \leq z \leq 1\}$$

no sentido da normal \vec{n} que aponta para fora do cilindro.

Vamos aplicar o Teorema da Divergência a

$$V = \{(x, y, z) \in \mathbb{R}^3 : x^2 + y^2 \leq 1, 0 \leq z \leq 1\}$$

A fronteira de V é dada por

$$\partial V = S \cup T_0 \cup T_1$$

com

$$T_0 = \{(x, y, 0) \in \mathbb{R}^3 : x^2 + y^2 \leq 1\} \quad e \quad T_1 = \{(x, y, 1) \in \mathbb{R}^3 : x^2 + y^2 \leq 1\}$$

O Teorema da Divergência diz-nos que

$$\int \int \int_V \operatorname{div} \vec{F} = \int \int_S \vec{F} \cdot \vec{n} + \int \int_{T_0} \vec{F} \cdot (0, 0, -1) + \int \int_{T_1} \vec{F} \cdot (0, 0, 1)$$

Como $\operatorname{div} \vec{F} = 1 + 1 + 1 = 3$ obtemos

$$\begin{aligned} 3 \operatorname{vol}(V) &= \int \int_S \vec{F} \cdot \vec{n} + \int \int_{T_0} -z + \int \int_{T_1} z \\ 3\pi &= \int \int_S \vec{F} \cdot \vec{n} + \int \int_{T_0} 0 + \int \int_{T_1} 1 \\ \int \int_S \vec{F} \cdot \vec{n} &= 3\pi - \operatorname{área}(T_1) = 3\pi - \pi = 2\pi. \end{aligned}$$

Dem. do Teorema 45.6. Vamos escrever a demonstração no caso em que $n = 3$ mas o caso geral é inteiramente análogo. Podemos escrever o campo vetorial $\vec{F} = (F_1, F_2, F_3)$ como a seguinte soma

$$\vec{F} = (F_1, 0, 0) + (0, F_2, 0) + (0, 0, F_3)$$

Claramente, basta verificar a fórmula dada no Teorema da Divergência para cada uma das três parcelas e a verificação para cada uma delas é inteiramente análoga. Consideremos portanto apenas o caso de um campo vetorial da forma $\vec{F} = (0, 0, F_3)$.

Suponhamos primeiro que a região V é da forma especial

$$V = \{(x, y, z) \in \mathbb{R}^3 : a(x, y) \leq z \leq b(x, y), (x, y) \in D\}$$

onde $D \subset \mathbb{R}^2$ é um domínio no plano e $a(x, y), b(x, y)$ são funções de classe C^1 (isto é, V é a região do espaço compreendida entre os gráficos de duas funções de x, y sobre um conjunto do plano). Então

$$\begin{aligned} \int \int \int_V \operatorname{div} \vec{F} &= \int \int_D \left(\int_{a(x,y)}^{b(x,y)} \frac{\partial F_3}{\partial z} dz \right) dx dy \\ &= \int \int_D F_3(x, y, b(x, y)) - F_3(x, y, a(x, y)) dx dy \end{aligned}$$

Por outro lado podemos escrever

$$\partial V = S_1 \cup C \cup S_2$$

com

$$S_1 = \{(x, y, a(x, y)) : (x, y) \in D\} \quad S_2 = \{(x, y, b(x, y)) : (x, y) \in D\}$$

$$C = \{(x, y, z) : a(x, y) \leq z \leq b(x, y) \text{ e } (x, y) \in \partial D\}$$

C é a “parte vertical” de ∂V , sobre a fronteira da projeção D de V no plano xy . Uma vez que $(0, 0, F_3)$ é tangente a C , a componente do fluxo de \vec{F} ao longo de C é nula. A função $g_1 : D \rightarrow \mathbb{R}^3$ definida por $g_1(x, y) = (x, y, a(x, y))$ é uma parametrização de S_1 e temos

$$\frac{\partial g_1}{\partial x} \times \frac{\partial g_1}{\partial y} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ 1 & 0 & \frac{\partial a}{\partial x} \\ 0 & 1 & \frac{\partial a}{\partial y} \end{vmatrix} = \left(-\frac{\partial a}{\partial x}, -\frac{\partial a}{\partial y}, 1 \right)$$

Uma vez que S_1 é a “parte de baixo” de ∂V e $\frac{\partial g_1}{\partial x} \times \frac{\partial g_1}{\partial y}$ tem terceira componente positiva, vemos que este último vetor aponta no sentido oposto à normal exterior a V em S_1 . A componente do fluxo para o exterior de V ao longo de S_1 é portanto

$$\begin{aligned} \int \int_{S_1} \vec{F} \cdot \vec{n}^{ext} &= - \int \int_D (0, 0, F_3(x, y, a(x, y))) \cdot \left(-\frac{\partial a}{\partial x}, -\frac{\partial a}{\partial y}, 1 \right) dx dy \\ &= - \int \int_D F_3(x, y, a(x, y)) dx dy \end{aligned}$$

De forma inteiramente análoga vemos que a componente do fluxo ao longo de S_2 é

$$\begin{aligned} \int \int_{S_2} \vec{F} \cdot \vec{n}^{ext} &= + \int \int_D (0, 0, F_3(x, y, b(x, y))) \cdot \left(-\frac{\partial b}{\partial x}, -\frac{\partial b}{\partial y}, 1 \right) dx dy \\ &= \int \int_D F_3(x, y, b(x, y)) dx dy \end{aligned}$$

e portanto

$$\begin{aligned} \int \int_{\partial V} \vec{F} \cdot \vec{n}^{ext} &= \int \int_{S_1} \vec{F} \cdot \vec{n}^{ext} + \int \int_{S_2} \vec{F} \cdot \vec{n}^{ext} + \int \int_C \vec{F} \cdot \vec{n}^{ext} \\ &= - \int \int_D F_3(x, y, a(x, y)) dx dy + \int \int_D F_3(x, y, b(x, y)) dx dy + 0 \end{aligned}$$

o que conclui a demonstração do Teorema da Divergência neste caso particular. Se V é um domínio regular qualquer, pode provar-se (e é fácil verificar diretamente para um qualquer V concreto dado) que V pode ser subdividido em regiões elementares (isto é, compreendidas entre o gráfico de funções de quaisquer duas das variáveis). Os fluxos através de componentes da fronteira destas regiões elementares que sejam interiores a V cancelam aos pares e portanto, somando as fórmulas dadas pelo Teorema da Divergência aplicado às regiões elementares, obtém-se o Teorema em completa generalidade. \square

47. O TEOREMA DA DIVERGÊNCIA (CONC.)

47.1. **A relação do Teorema da Divergência com o Teorema de Green.** Seja $U \subset \mathbb{R}^2$ um aberto, $(P, Q): U \rightarrow \mathbb{R}^2$ um campo vetorial de Classe C^1 e $V \subset U$ um domínio regular. O Teorema de Green afirma que

$$\int \int_D \frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} dx dy = \oint_{\partial V} (P, Q) \cdot d\vec{r}$$

Claramente,

$$\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} = \text{div}(Q, -P)$$

Note-se agora que rodando $(a, b) \in \mathbb{R}^2$ 90° para a direita obtém-se o vetor $(b, -a)$ (pensando em (a, b) como o número complexo $a + bi$, temos $(-i)(a + bi) = b - ai$). Se $\vec{t}: \partial V \rightarrow \mathbb{R}^2$ é um vector tangente unitário a ∂V com o sentido indicado no enunciado do Teorema de Green (isto é, tal que V se encontra à esquerda de ∂V quando esta é percorrida com esta orientação) então a normal \vec{n}^{ext} exterior a V obtém-se de \vec{t} rodando 90° para a direita. Logo

$$\oint_{\partial V} (P, Q) \cdot d\vec{r} = \int_{\partial V} (P, Q) \cdot \vec{t} = \int_{\partial V} (Q, -P) \cdot \vec{n}^{ext}$$

pelo que a fórmula dada pelo Teorema de Green é exatamente a fórmula dada pelo Teorema da Divergência aplicado ao campo vetorial $(Q, -P)$. O Teorema de Green é portanto equivalente ao Teorema da Divergência em \mathbb{R}^2 .

47.2. **As leis de Gauss do Eletromagnetismo.** Seja $\vec{E}: U \rightarrow \mathbb{R}^3$ o campo elétrico numa região do espaço. A Lei de Gauss, uma das leis fundamentais do Eletromagnetismo, afirma que o fluxo do campo elétrico através de uma superfície S que limita uma região $V \subset U$ do espaço é proporcional à carga elétrica contida em V . Se $\rho: U \rightarrow \mathbb{R}^3$ for a densidade de carga, esta lei é descrita pela fórmula

$$\int \int_S \vec{E} \cdot \vec{n}^{ext} = \frac{1}{\epsilon} \int \int \int_V \rho$$

onde ϵ é a permissividade elétrica do meio (suposto homogéneo) e V é o volume limitado por S . Pelo Teorema da Divergência, a equação anterior é equivalente a

$$\int \int \int_V \text{div } \vec{E} = \frac{1}{\epsilon} \int \int \int_V \rho$$

e uma vez que esta equação é válida para todo o $V \subset U$, ela é por sua vez equivalente à seguinte “forma diferencial” da equação de Gauss

$$(25) \quad \text{div } \vec{E} = \frac{1}{\epsilon} \rho$$

De facto, a forma integral obtém-se integrando (25) e, reciprocamente, podemos recuperar o valor de uma função contínua $f: U \rightarrow \mathbb{R}$ a partir dos integrais de f sobre bolas contidas

em U através da fórmula

$$f(x) = \lim_{\epsilon \rightarrow 0} \frac{1}{\text{vol } B_\epsilon(x)} \int \int \int_{B_\epsilon(x)} f$$

onde x é um ponto qualquer de U (exercício). Aplicando esta fórmula à forma integral da equação de Gauss obtém-se (25).

A lei de Gauss afirma que as cargas positivas são fontes do campo elétrico, enquanto que as cargas negativas são sumidouros.

Há também uma lei análoga para o campo magnético $\vec{B}: U \rightarrow \mathbb{R}^3$:

$$\text{div } \vec{B} = 0$$

que traduz a inexistência de “cargas magnéticas”.

47.3. A equação da continuidade. Seja $\vec{F} = \rho\vec{v}$ a densidade de fluxo de um fluido, onde $\rho: U \times \mathbb{R} \rightarrow \mathbb{R}$ e $\vec{v}: U \times \mathbb{R} \rightarrow \mathbb{R}^3$ são campos dependentes do tempo. A lei de conservação da massa numa região $V \subset U$ é descrita pela equação

$$\int \int_{\partial V} \rho\vec{v} \cdot \vec{n}^{ext} = -\frac{d}{dt} \left(\int \int \int_V \rho \right)$$

pois, do lado esquerdo da igualdade temos a quantidade de massa que atravessa a fronteira de V por unidade de tempo, enquanto que do lado direito está a diminuição da massa no interior de V por unidade de tempo.

Aplicando o Teorema da Divergência e a regra de Leibniz obtemos a equação equivalente

$$\int \int \int_V \text{div}(\rho\vec{v}) = \int \int \int_V -\frac{\partial \rho}{\partial t}$$

Como a equação anterior vale para todo o $V \subset U$, ela é equivalente (como explicado acima) à equação diferencial

$$\text{div}(\rho\vec{v}) + \frac{\partial \rho}{\partial t} = 0$$

que se designa por *equação da continuidade*. Trata-se de uma equação diferencial que traduz a lei da conservação de massa durante o escoamento de um fluido.

47.4. O Teorema de Stokes. Seja $S \subset \mathbb{R}^3$ uma superfície (isto é, uma variedade-2) e $\vec{n}: S \rightarrow \mathbb{R}^3$ um campo vetorial normal unitário contínuo. O Teorema de Stokes, o último teorema fundamental do cálculo que iremos estudar, afirma que dado um campo vetorial \vec{F} de classe C^1 definido num aberto que contenha S temos

$$(26) \quad \int \int_S \text{rot } \vec{F} \cdot \vec{n} = \oint_{\partial S} \vec{F} \cdot d\vec{r}$$

onde ∂S denota o *bordo de* S (informalmente, a linha que limita S) e $\text{rot } \vec{F} = \nabla \times \vec{F}$ é uma certa derivada de \vec{F} , chamada o *rotacional* de \vec{F} , definida pela expressão

$$(27) \quad \text{rot } \vec{F} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ F_1 & F_2 & F_3 \end{vmatrix} = \left(\frac{\partial F_3}{\partial y} - \frac{\partial F_2}{\partial z}, \frac{\partial F_1}{\partial z} - \frac{\partial F_3}{\partial x}, \frac{\partial F_2}{\partial x} - \frac{\partial F_1}{\partial y} \right)$$

e que corresponde à parte anti-simétrica da matriz Jacobiana de \vec{F} . Usa-se também a notação

$$\nabla \times \vec{F}$$

para o rotacional de \vec{F} .

Exemplo 47.5. *Seja $\vec{F}(x, y, z) = (x + yz, x, y^2 + z)$. Então*

$$\nabla \times \vec{F} = \text{rot } \vec{F} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ x + yz & x & y^2 + z \end{vmatrix} = (2y, y, 1 - z)$$

Note-se a forma geral da equação (26), comum a todos os teoremas fundamentais do cálculo: o integral de uma certa derivada de \vec{F} sobre S é igual ao integral de \vec{F} sobre a “extremidade” de S , ∂S .

Para que (26) tenha um significado preciso, é necessário explicar o sentido em que ∂S é percorrido (senão o termo direito da igualdade só estaria definido a menos de um sinal). A regra é a seguinte:

(28)

supondo que estamos em pé sobre a superfície S do lado para o qual aponta a normal, junto a ∂S , temos que percorrer o bordo ∂S de forma a que S esteja à nossa esquerda. Alternativamente, podemos achar o sentido com a **regra da mão direita**: apoiando a mão direita na superfície com o polegar a apontar na direção da normal dada e os restantes dedos junto a ∂S , o bordo ∂S deve ser percorrido no sentido em que apontam os dedos.

Exemplo 47.6. *Se $S \subset \mathbb{R}^3$ fôr a superfície determinada por uns calções em \mathbb{R}^3 , orientada com a normal \vec{n} exterior aos calções, o bordo de S é formado por três circunferências. A circunferência de cima (a cintura) deve ser percorrida no sentido horário (quando visto de cima), enquanto que as circunferências de baixo (as componentes do bordo correspondentes às pernas) devem ser percorridas no sentido anti-horário (vistas também de cima).*

Nota 47.7. *Um domínio regular $D \subset \mathbb{R}^2$ pode ser visto como uma superfície contida no plano xy em \mathbb{R}^3 e um campo vetorial (P, Q) em \mathbb{R}^2 pode ser promovido ao campo vectorial em \mathbb{R}^3 definido por*

$$\vec{F}(x, y, z) = (P(x, y), Q(x, y), 0)$$

Temos

$$\text{rot } \vec{F} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ P(x, y) & Q(x, y) & 0 \end{vmatrix} = \left(0, 0, \frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} \right)$$

Considerando a normal unitária $\vec{n} = (0, 0, 1)$ a $D \subset \mathbb{R}^3$ e aplicando a fórmula de Stokes (26) obtemos

$$\begin{aligned} \int \int_D \operatorname{rot} \vec{F} \cdot (0, 0, 1) &= \oint_{\partial D} \vec{F} \cdot d\vec{r} \\ \int \int_D \frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} dx dy &= \oint_{\partial D} \vec{F} \cdot d\vec{r} \end{aligned}$$

e o sentido atribuído a ∂D pela regra da mão direita é o do Teorema de Green. Conclui-se portanto que o Teorema de Green pode ser visto como um caso particular do Teorema de Stokes. Ou, alternativamente, o Teorema de Stokes pode ser encarado como uma generalização do Teorema de Green ao caso de um pedaço de superfície curva. Deste último ponto de vista, o sentido atribuído ao bordo pelo Teorema de Stokes é “o mesmo que no Teorema de Green”.

Uma outra característica que o Teorema de Stokes tem em comum com o Teorema de Green é o facto de se aplicar apenas a uma dimensão (neste caso a \mathbb{R}^n com $n = 3$). Existe na realidade um Teorema de Stokes geral que contém como casos particulares todos os Teoremas Fundamentais do Cálculo que estudámos mas nesse Teorema já não se pode encarar as funções integrandas como campos escalares ou vetoriais - trata-se antes de campos tensoriais que, nos casos particulares que estudámos, se podem identificar com campos escalares ou vetoriais. A referência [S] contém o enunciado geral do Teorema de Stokes e a sua demonstração.

48. O TEOREMA DE STOKES

Começemos por enunciar formalmente o Teorema de Stokes.

Definição 48.1. *Seja $M \subset \mathbb{R}^3$ uma variedade-2. Um subconjunto limitado $S \subset M$ diz-se um domínio regular em M se existe uma função $f: \mathbb{R}^3 \rightarrow \mathbb{R}$ de classe C^1 tal que*

$$S = \{(x, y, z) \in M: f(x) < 0\} \quad e \quad (x \in M \text{ e } f(x) = 0) \Rightarrow \nabla f(x) \notin (T_x M)^\perp$$

O bordo de um domínio regular S é

$$\partial S = \{x \in M: f(x) = 0\}$$

Note-se que em consequência da definição anterior, ∂S é uma variedade de dimensão 1. De facto, se a equação $F(x, y, z) = 0$ define a variedade M perto de $x_0 \in \partial S$, o vector $\nabla F(x_0)$ gera $(T_{x_0} M)^\perp$, pelo que a condição imposta a $\nabla f(x_0)$ é precisamente que a matriz Jacobiana da função $(x, y, z) \mapsto (F(x, y, z), f(x, y, z))$ tenha característica máxima nos pontos em que $(F, f) = (0, 0)$, ou seja, nos pontos de ∂S .

Exemplo 48.2. *Seja $M = \{(x, y, z) \in \mathbb{R}^3: x^2 + y^2 + z^2 = 1\}$, o subconjunto*

$$S = \{(x, y, z) \in M: z > 0\}$$

é um domínio regular. De facto $f(x, y, z) = -z$ e $\nabla f = (0, 0, -1)$ não pertence ao espaço ortogonal a M nos pontos do equador (que são os pontos onde f se anula).

Nota 48.3. A ideia de domínio regular numa superfície é a de um subconjunto limitado por uma curva na superfície (e que é limitado em \mathbb{R}^3). A curva $f = 0$ na definição separa o “interior da curva” definido pela condição $f < 0$ do “exterior” definido pela equação $f > 0$.

É possível demonstrar que o conceito de domínio regular em \mathbb{R}^n usado no Teorema da Divergência pode ser definido de forma análoga: um aberto limitado $V \subset \mathbb{R}^n$ é um domínio regular sse existe uma função $f: \mathbb{R}^n \rightarrow \mathbb{R}$ de classe C^1 tal que

$$V = \{x \in \mathbb{R}^n : f(x) < 0\} \quad e \quad (f(x) = 0 \text{ e } x \in M) \Rightarrow \nabla f(x) \neq 0.$$

Teorema 48.4 (Teorema de Stokes). Seja $U \subset \mathbb{R}^3$ um aberto, $\vec{F}: U \rightarrow \mathbb{R}^3$ um campo vetorial de classe C^1 , $M \subset U$ uma variedade-2, $\vec{n}: M \rightarrow \mathbb{R}^3$ um campo vetorial unitário normal a M e $S \subset M$ um domínio regular. Então

$$\int \int_S \text{rot } \vec{F} \cdot \vec{n} = \oint_{\partial S} \vec{F} \cdot d\vec{r}$$

sendo o bordo ∂S percorrido no sentido dado pela regra da mão direita (28).

Exemplo 48.5. Vamos verificar a fórmula dada pelo Teorema 48.4 para o campo vetorial

$$\vec{F}(x, y, z) = (y, z, x).$$

e o pedaço de parabolóide

$$S = \{(x, y, z) \in \mathbb{R}^3 : z = 1 - x^2 - y^2, z > 0\}$$

orientado com a normal \vec{n} que tem a terceira componente positiva.

Temos

$$\text{rot } \vec{F} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ y & z & x \end{vmatrix} = (-1, -1, -1)$$

Uma parametrização de S é dada por

$$g(x, y) = (x, y, 1 - x^2 - y^2), \quad \text{com } x^2 + y^2 < 1$$

e

$$\frac{\partial g}{\partial x} \times \frac{\partial g}{\partial y} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ 1 & 0 & -2x \\ 0 & 1 & -2y \end{vmatrix} = (2x, 2y, 1)$$

Este campo vetorial tem terceira componente positiva logo está de acordo com a orientação escolhida para S . Temos portanto

$$\begin{aligned} \int \int_S \text{rot } \vec{F} \cdot \vec{n} &= \int \int_{\{(x,y): x^2+y^2 < 1\}} (-1, -1, -1) \cdot (2x, 2y, 1) dx dy \\ &= \int \int_{\{(x,y): x^2+y^2 < 1\}} -2x - 2y - 1 dx dy \end{aligned}$$

Por simetria, os integrais das funções x e y no círculo anulam-se pelo que o último integral acima é igual a $-\text{área}(S) = -\pi$.

Por outro lado, o bordo de S é a curva

$$\partial S = \{(x, y, 0) : x^2 + y^2 = 1\}$$

e uma parametrização para ∂S é

$$g(\theta) = (\cos \theta, \sin \theta, 0), \quad 0 < \theta < 2\pi$$

O sentido dado a ∂S pela regra (28) é o sentido anti-horário quando ∂S é olhado desde um ponto acima do plano xy , logo a parametrização g percorre ∂S no sentido do Teorema de Stokes. Substituindo na definição de trabalho temos

$$\begin{aligned} \oint_{\partial S} \vec{F} \cdot d\vec{r} &= \int_0^{2\pi} \vec{F}(\cos \theta, \sin \theta, 0) \cdot (-\sin \theta, \cos \theta, 0) d\theta \\ &= \int_0^{2\pi} (\sin \theta, 0, \cos \theta) \cdot (-\sin \theta, \cos \theta, 0) d\theta \\ &= \int_0^{2\pi} -\sin^2 \theta d\theta = -\pi \end{aligned}$$

conforme afirma o Teorema de Stokes.

Exemplo 48.6. Calcular o trabalho de $\vec{F}(x, y, z) = (y, z, x)$ ao longo da curva

$$C = \{(x, y, z) \in \mathbb{R}^3 : x^2 + y^2 + z^2 = a^2, x + y + z = 0\}$$

(onde $a > 0$) percorrida num sentido que parece o horário quando visto de muito acima do plano xy .

Escolhendo uma superfície S que tenha C como bordo podemos usar a fórmula (26) para substituir o cálculo do trabalho pelo cálculo de um fluxo através de S . A superfície mais simples que podemos tomar é a interseção da esfera de raio a com o plano que corta C na superfície esférica:

$$S = \{(x, y, z) \in \mathbb{R}^3 : x^2 + y^2 + z^2 \leq a^2, x + y + z = 0\}$$

A normal a S é perpendicular ao plano definido pela equação e portanto é dada pela expressão

$$\vec{n} = \pm \frac{1}{\sqrt{3}}(1, 1, 1)$$

A normal que é compatível com a orientação de C de acordo com a regra (28) é a normal que aponta para baixo do plano, isto é $\vec{n} = -\frac{1}{\sqrt{3}}(1, 1, 1)$. Vimos no exemplo anterior que

$$\text{rot } \vec{F}(x, y, z) = (-1, -1, -1)$$

portanto

$$\oint_C \vec{F} \cdot d\vec{r} = \int \int_S \text{rot } \vec{F} \cdot \vec{n} = \int \int_S \sqrt{3} = \sqrt{3} \text{área}(S) = \sqrt{3}\pi a^2$$

(como S é a interseção da esfera com um plano que passa pela origem, trata-se de um círculo de raio igual ao da esfera).

49. O TEOREMA DE STOKES (CONT.)

O Teorema de Stokes 48.4 permite atribuir um significado ao rotacional de um campo vetorial \vec{F} . De facto, seja $U \subset \mathbb{R}^3$ um aberto e $\vec{F}: U \rightarrow \mathbb{R}^3$ um campo vetorial de classe C^1 . Aplicando a fórmula (26) a um pequeno disco $D_\epsilon(x)$ centrado em $x \in U$ e contido num plano perpendicular ao vetor \vec{n} obtemos

$$\int \int_{D_\epsilon(x)} \text{rot } \vec{F} \cdot \vec{n} = \int_{\partial D_\epsilon(x)} \vec{F} \cdot d\vec{r}$$

Para ϵ suficientemente pequeno, $\text{rot } \vec{F}$ é praticamente constante ao longo do disco e o integral do lado esquerdo é aproximadamente igual a

$$\text{área}(D_\epsilon(x)) \left(\text{rot } \vec{F}(x) \cdot \vec{n} \right)$$

Conclui-se que

$$\text{rot } \vec{F}(x) \cdot \vec{n} = \lim_{\epsilon \rightarrow 0} \frac{1}{\text{área}(D_\epsilon(x))} \oint_{\partial D_\epsilon(x)} \vec{F} \cdot d\vec{r}$$

Do lado direito da igualdade anterior temos uma medida da circulação (ou trabalho) de \vec{F} ao longo de uma circunferência perpendicular a \vec{n} . Assim, vemos que $\text{rot } \vec{F}(x)$ é um vetor que aponta na direção perpendicular ao plano que passa por x no qual a circulação de \vec{F} (em torno de x) é máxima e que a magnitude de $\text{rot } \vec{F}(x)$ é uma medida da intensidade dessa circulação. Em termos concretos, se \vec{F} for a densidade de fluxo de ar (isto é a força do vento!) o rotacional diz-nos como devemos colocar um moínho de vento de forma a que o moínho gire o mais rápido possível.

Exemplo 49.1. Considere-se o campo vetorial $\vec{F}(x, y, z) = (-y, x, 0)$. Este campo é horizontal e “roda” em torno do eixo dos zz . Calculando o rotacional obtemos

$$\text{rot } \vec{F} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ -y & x & 0 \end{vmatrix} = (0, 0, 2)$$

Exemplo 49.2. Seja $\vec{F}(x, y, z) = (-y, x, x + z)$. Calcular o fluxo de $\text{rot } \vec{F}$ através da superfície

$$S = \{(x, y, z) \in \mathbb{R}^3: x^2 + y^2 < 2x, z + x = 0\}$$

no sentido da normal que tem terceira componente positiva.

De acordo com o Teorema de Stokes, o fluxo em questão é o trabalho de \vec{F} ao longo do bordo

$$\partial S = \{(x, y, z) \in \mathbb{R}^3: x^2 + y^2 = 2x, z + x = 0\}$$

que é a elipse de interseção do cilindro com equação

$$x^2 - 2x + y^2 = 0 \Leftrightarrow (x - 1)^2 + y^2 = 1$$

com o plano de equação $x + z = 0$.

A orientação de ∂S que é compatível com a normal \vec{n} de acordo com a regra (28) é o sentido direto quando ∂S é vista de muito acima do plano xy . Uma parametrização para ∂S é dada por

$$g(\theta) = (1 + \cos \theta, \sin \theta, -(1 + \cos \theta)), \quad 0 < \theta < 2\pi$$

e esta parametrização induz em ∂S a orientação correta logo

$$\begin{aligned} \int \int_S \operatorname{rot} \vec{F} \cdot d\vec{r} &= \oint_{\partial S} \vec{F} \cdot d\vec{r} \\ &= \int_0^{2\pi} \vec{F}(1 + \cos \theta, \sin \theta, -1 - \cos \theta) \cdot (-\sin \theta, \cos \theta, \sin \theta) d\theta \\ &= \int_0^{2\pi} (-\sin \theta, 1 + \cos \theta, 0) \cdot (-\sin \theta, \cos \theta, \sin \theta) d\theta \\ &= \int_0^{2\pi} \sin^2 \theta + \cos^2 \theta + \cos \theta d\theta = 2\pi. \end{aligned}$$

Em certas situações, o Teorema de Stokes pode também ser aplicado para calcular trabalhos deformando o caminho de integração tal como o Teorema de Green. De facto, a fórmula (27) que define o rotacional mostra que

$$\operatorname{rot} \vec{F} = 0 \Leftrightarrow \vec{F} \text{ é um campo fechado.}$$

Por essa razão, os campos fechados em \mathbb{R}^3 dizem-se também *irrotacionais*. Se aplicarmos o Teorema de Stokes a um campo irrotacional, ele diz-nos que

$$\int_{\partial S} \vec{F} \cdot d\vec{r} = 0$$

para toda a superfície orientada S contida no domínio de \vec{F} . Por exemplo, aplicando esta fórmula a um “par de calções” S , ela diz-nos que o trabalho ao longo da cintura dos calções tem de ser igual ao trabalho ao longo dos bordos das pernas quando estes são percorridos no mesmo sentido (todos no horário ou todos no anti-horário) quando visto de muito acima dos calções). Isto também pode ser deduzido do Teorema 42.4 deformando a cintura num caminho que percorre os bordos das duas pernas unidos por um caminho que é percorrido duas vezes em sentidos inversos. Vejamos um exemplo mais concreto.

Exemplo 49.3. Seja $U = \mathbb{R}^3 \setminus \{(x, y, z) \in \mathbb{R}^3 : x^2 + y^2 = 0\}$ e $\vec{F}: U \rightarrow \mathbb{R}^3$ o campo vetorial definido por

$$\vec{F}(x, y, z) = \left(-\frac{y}{x^2 + y^2} + z, \frac{x}{x^2 + y^2} + y^2, x \right)$$

Vamos calcular o trabalho realizado por \vec{F} ao longo da curva

$$C = \{(x, y, z) \in \mathbb{R}^3 : x^2 + y^2 = 1, z = \sin(xy)\}$$

no sentido que parece o horário visto de $(0, 0, 100)$.

A expressão para \vec{F} mostra que será fácil de calcular o trabalho de \vec{F} ao longo de uma circunferência horizontal centrada no eixo dos zz . Aplicando o Teorema de Stokes à superfície

$$S = \{(x, y, z) \in \mathbb{R}^3 : x^2 + y^2 = 1, -2 \leq z \leq \text{sen}(xy)\}$$

orientada com a normal que aponta para fora do cilindro, temos $\partial S = C \cup D$ com

$$D = \{(x, y, -2) \in \mathbb{R}^3 : x^2 + y^2 = 1\}$$

As orientações de C, D são respetivamente, a que foi dada, e a que corresponde ao sentido anti-horário visto de $(0, 0, 100)$. Obtemos assim

$$0 = \oint_C \vec{F} \cdot d\vec{r} + \oint_D \vec{F} \cdot d\vec{r}$$

Uma parametrização para D é dada por

$$g(\theta) = (\cos \theta, \text{sen } \theta, -2), 0 < \theta < 2\pi$$

que tem o sentido dado pelo Teorema de Stokes. Logo

$$\oint_D \vec{F} \cdot d\vec{r} = \int_0^{2\pi} \pi(-2 - \text{sen } \theta, \cos \theta + \text{sen}^2 \theta, \cos \theta) \cdot (-\text{sen } \theta, \cos \theta, 0) d\theta = 2\pi$$

pelo que o trabalho pretendido é

$$\oint_C \vec{F} \cdot d\vec{r} = -2\pi.$$

50. POTENCIAIS VETOR

A Fórmula do Teorema de Stokes (26) permite calcular fluxos de campos que sejam rotacionais, da mesma forma que o Teorema Fundamental do Cálculo para integrais de linha permite calcular o trabalho de campos gradientes. Por essa razão é importante saber quando é que um campo vetorial \vec{F} é o rotacional de outro. De facto, quando isso acontece o fluxo de \vec{F} através de uma superfície orientada pode ser calculado através de um integral de linha.

A título de comparação lembremos primeiro as propriedades de campos vectoriais em \mathbb{R}^3 relevantes para o cálculo de integrais de linha usando o Teorema Fundamental do Cálculo:

$$\vec{F} = \nabla\varphi \quad \Rightarrow \quad \text{rot } \vec{F} = 0 \quad (\text{gradiente implica fechado})$$

Por outro lado

$$\text{rot } \vec{F} = 0 \quad \Rightarrow \quad \vec{F} = \nabla\varphi \text{ num domínio simplesmente conexo.}$$

Definição 50.1. Seja $U \subset \mathbb{R}^3$ um aberto e $\vec{F}: U \rightarrow \mathbb{R}^3$ um campo vetorial. O campo \vec{F} diz-se um rotacional se existe $\vec{A}: U \rightarrow \mathbb{R}^3$ tal que $\vec{F} = \text{rot } \vec{A}$. Nesse caso $\text{rot } \vec{A}$ diz-se um potencial vetor para \vec{F} . O campo \vec{F} diz-se solenoidal se $\text{div } \vec{F} = 0$.

Proposição 50.2. Seja $U \subset \mathbb{R}^3$ um aberto e $\vec{F}: U \rightarrow \mathbb{R}^3$ um campo vetorial de classe C^1 . Se \vec{F} é o rotacional de um campo vetorial de classe C^2 , então \vec{F} é solenoidal, isto é,

$$\vec{F} = \text{rot } \vec{A} \quad \Rightarrow \quad \text{div } \vec{F} = 0$$

Dem. Se $\vec{F} = \text{rot } A$, então

$$(F_1, F_2, F_3) = \left(\frac{\partial A_3}{\partial y} - \frac{\partial A_2}{\partial z}, \frac{\partial A_1}{\partial z} - \frac{\partial A_3}{\partial x}, \frac{\partial A_2}{\partial x} - \frac{\partial A_1}{\partial y} \right)$$

logo

$$\text{div } \vec{F} = \frac{\partial^2 A_3}{\partial x \partial y} - \frac{\partial^2 A_2}{\partial x \partial z} + \frac{\partial^2 A_1}{\partial y \partial z} - \frac{\partial^2 A_3}{\partial y \partial x} + \frac{\partial^2 A_2}{\partial z \partial x} - \frac{\partial^2 A_1}{\partial z \partial y}$$

é 0 pelo Lema de Schwarz. \square

Nota 50.3. *É um bom exercício verificar que a condição de \vec{A} ser C^2 na Proposição anterior é desnecessária. Se \vec{F} é um campo vetorial de classe C^1 que admite um potencial vetor \vec{A} de classe C^1 então $\text{div } F = 0$ (use a interpretação da divergência como uma medida de criação de fluxo).*

Exemplo 50.4. *O campo vetorial $\vec{F}(x, y, z) = (x + y^2, z, -x)$ não é solenoidal:*

$$\nabla \cdot \vec{F} = 1 + 0 + 0 \neq 0$$

logo não é um rotacional.

Analogamente ao que acontece relativamente à existência de potenciais escalares, se a “forma” do domínio do campo for sujeita a certas restrições, um campo solenoidal é necessariamente um rotacional.

Definição 50.5. *Um conjunto $A \subset \mathbb{R}^n$ diz-se em estrela se existe $P \in A$ tal que, para todo o $x \in A$, o segmento de reta que une P a x está contido em A .*

Em termos concretos, um conjunto A diz-se em estrela se há um “ponto central” P em A que “vê todos os outros”.

Exemplo 50.6. *1. \mathbb{R}^n é um conjunto em estrela. Qualquer ponto $P \in \mathbb{R}^n$ verifica a condição da definição. Isto é verdade mais geralmente para qualquer conjunto convexo que é, por definição, um conjunto com a seguinte propriedade mais restritiva: Para todo o $x, y \in A$, o segmento de reta que une x a y está contido em A . As bolas, regiões limitadas por planos, etc... são conjuntos convexos e portanto em estrela.*

2. $\mathbb{R}^2 \setminus \{(0, 0)\}$ não é um conjunto em estrela. Dado qualquer P neste conjunto, os pontos que, do ponto de vista de P , estão por detrás da origem, são “invisíveis” (o segmento de reta que os une a P não está contido em $\mathbb{R}^2 \setminus \{(0, 0)\}$).

3. $\mathbb{R}^2 \setminus \{(x, 0) : x \geq 0\}$ é um conjunto em estrela. Os possíveis “centros” P são os pontos no semi-eixo negativo dos xx .

Proposição 50.7. *Seja $U \subset \mathbb{R}^3$ um aberto em estrela e $\vec{F}: U \rightarrow \mathbb{R}^3$ um campo vetorial solenoidal. Então \vec{F} é um rotacional em U .*

Nota 50.8. *Prova-se mais geralmente que um campo solenoidal é um rotacional se o domínio U satisfaz a seguinte condição: Se S é uma variedade-2 compacta contida em U , a região de \mathbb{R}^3 limitada por S está inteiramente contida em U . Por exemplo,*

$$U = \mathbb{R}^3 \setminus \{(0, 0, z) : z \in \mathbb{R}\}$$

é uma região do espaço com esta propriedade, que não é um conjunto em estrela. Qualquer campo solenoidal neste domínio é um rotacional.

Tal já não é verdade para

$$U = \mathbb{R}^3 \setminus \{(0, 0, 0)\}$$

(considere-se uma superfície esférica cujo interior contenha a origem) ou para

$$U = \mathbb{R}^3 \setminus \{(x, y, 0) \in \mathbb{R}^3 : x^2 + y^2 = 1\}$$

(considere-se um toro à volta da circunferência que foi retirada).

Exemplo 50.9. Seja $\vec{F}(x, y, z) = (x + y^2, -y + z^2, xy)$. Este campo vetorial é de classe C^1 e

$$\operatorname{div} \vec{F} = 1 - 1 + 0 = 0$$

Uma vez que o domínio do campo é \mathbb{R}^3 , que é um conjunto em estrela, conclui-se que \vec{F} é um rotacional.

50.10. Cálculo de um potencial vetor. Suponhamos que \vec{F} é um rotacional. Para achar um potencial vetor \vec{A} , temos que resolver o sistema

$$\operatorname{rot} \vec{A} = \vec{F} \Leftrightarrow \begin{cases} \frac{\partial A_3}{\partial y} - \frac{\partial A_2}{\partial z} = F_1 \\ \frac{\partial A_1}{\partial z} - \frac{\partial A_3}{\partial x} = F_2 \\ \frac{\partial A_2}{\partial x} - \frac{\partial A_1}{\partial y} = F_3 \end{cases}$$

que é um sistema do mesmo género do que resolvemos para achar um potencial escalar, mas mais complicado. Quando existe, a solução do sistema para a determinação de um potencial escalar é única a menos de uma constante (em cada componente conexa do domínio). Para o sistema acima temos muito maior liberdade na escolha de uma solução (quando esta existe). De facto a equação

$$\operatorname{rot} \nabla \varphi = 0$$

mostra que se \vec{A} é uma solução do sistema, o mesmo é verdade para $\vec{A}' = \vec{A} + \nabla \varphi$ (mais geralmente, duas soluções do sistema diferem num campo *fechado* arbitrário).

Podemos usar esta liberdade na escolha de uma solução para tornar o sistema mais simples. De facto, é sempre possível achar uma solução satisfazendo a restrição adicional que uma das componentes de \vec{A} é identicamente nula. A imposição desta restrição torna o sistema bastante mais simples.

Para ver que é possível anular uma das componentes do potencial vetor, vejamos por exemplo como obter, a partir de um potencial vetor arbitrário \vec{A} para \vec{F} , um novo potencial vetor \vec{A}' tal que $A'_1 = 0$. Escrevendo

$$\vec{A}' = \vec{A} + \nabla \varphi$$

queremos que

$$0 = A'_1 = A_1 + \frac{\partial \varphi}{\partial x} = 0$$

Mas para isto basta tomar

$$\varphi(x, y, z) = - \int A_1(x, y, z) dx$$

Ou seja,

$$\vec{A}' = \vec{A} - \nabla \left(\int A_1(x, y, z) dx \right)$$

é um novo potencial vetor para \vec{F} cuja primeira componente é identicamente nula.

Exemplo 50.11. Calculemos um potencial vetor para o campo \vec{F} do Exemplo 50.9. Temos que resolver o sistema

$$\text{rot } \vec{A} = \vec{F} \Leftrightarrow \begin{cases} \frac{\partial A_3}{\partial y} - \frac{\partial A_2}{\partial z} = x + y^2 \\ \frac{\partial A_1}{\partial z} - \frac{\partial A_3}{\partial x} = -y + z^2 \\ \frac{\partial A_2}{\partial x} - \frac{\partial A_1}{\partial y} = xy \end{cases}$$

Para simplificar o sistema vamos assumir (vimos acima que podemos fazer isto) que $A_3 = 0$. O sistema fica então

$$\begin{cases} -\frac{\partial A_2}{\partial z} = x + y^2 \\ \frac{\partial A_1}{\partial z} = -y + z^2 \\ \frac{\partial A_2}{\partial x} - \frac{\partial A_1}{\partial y} = xy \end{cases}$$

As primeiras duas equações dizem-nos que

$$A_2(x, y, z) = -xz - y^2z + \alpha(x, y) \quad A_1(x, y, z) = -yz + \frac{z^3}{3} + \beta(x, y)$$

para algumas funções α, β . Substituindo na terceira equação ficamos com

$$\begin{aligned} \frac{\partial}{\partial x} (-xz - y^2z + \alpha(x, y)) - \frac{\partial}{\partial y} \left(-yz + \frac{z^3}{3} + \beta(x, y) \right) &= xy \\ -z + \frac{\partial \alpha}{\partial x} - \left(-z + \frac{\partial \beta}{\partial y} \right) &= xy \\ \frac{\partial \alpha}{\partial x} - \frac{\partial \beta}{\partial y} &= xy \end{aligned}$$

Podemos por exemplo tomar $\beta(x, y) = 0$ e $\alpha(x, y) = \frac{x^2y}{2}$. Conclui-se que

$$\vec{A}(x, y, z) = \left(-yz + \frac{z^3}{3}, -xz - y^2z + \frac{x^2y}{2}, 0 \right)$$

é um potencial vetor para \vec{F} (o que se pode verificar facilmente substituindo no sistema inicial).

51. O TEOREMA DE STOKES (CONC.)

Exemplo 51.1. Seja $\vec{F}: \mathbb{R}^3 \rightarrow \mathbb{R}^3$ o campo vetorial definido por

$$\vec{F}(x, y, z) = (y + z, xz, x^2 + y)$$

e

$$S = \{(x, y, z) \in \mathbb{R}^3: x^2 + y^2 + z^2 = 1, z > 0\}$$

Vamos calcular o fluxo de \vec{F} através de S no sentido da normal com terceira componente positiva recorrendo ao Teorema de Stokes.

Temos $\text{div } \vec{F} = 0 + 0 + 0 = 0$. Como \mathbb{R}^3 é um conjunto em estrela, existe um potencial vetor \vec{A} para \vec{F} que podemos calcular resolvendo o sistema

$$\text{rot } \vec{A} = \vec{F} \Leftrightarrow \begin{cases} \frac{\partial A_3}{\partial y} - \frac{\partial A_2}{\partial z} = y + z \\ \frac{\partial A_1}{\partial z} - \frac{\partial A_3}{\partial x} = xz \\ \frac{\partial A_2}{\partial x} - \frac{\partial A_1}{\partial y} = x^2 + y \end{cases}$$

Uma vez achado o potencial vetor, o Teorema de Stokes diz-nos que

$$\int \int_S \vec{F} \cdot \vec{n} = \int \int_S \text{rot } \vec{A} \cdot \vec{n} = \oint_{\partial S} \vec{A} \cdot d\vec{r}$$

O bordo de S é a circunferência horizontal

$$\partial S = \{(x, y, 0) \in \mathbb{R}^3: x^2 + y^2 = 1\}$$

Para calcular o trabalho de \vec{A} ao longo de S , a terceira componente de A é irrelevante (vai ser multiplicada por 0 na fórmula para o trabalho) por isso, em princípio, as contas serão mais simples se escolhermos uma das componentes A_1 ou A_2 iguais a 0. Vamos por exemplo tomar $A_1 = 0$. O sistema acima fica então

$$\begin{cases} \frac{\partial A_3}{\partial y} - \frac{\partial A_2}{\partial z} = y + z \\ -\frac{\partial A_3}{\partial x} = xz \\ \frac{\partial A_2}{\partial x} = x^2 + y \end{cases} \Leftrightarrow \begin{cases} \frac{\partial A_3}{\partial y} - \frac{\partial A_2}{\partial z} = y + z \\ A_3(x, y, z) = -\frac{x^2 z}{2} + \alpha(y, z) \\ A_2(x, y, z) = \frac{x^3}{3} + xy + \beta(y, z) \end{cases}$$

e substituindo na primeira equação obtém-se

$$\begin{aligned} \frac{\partial}{\partial y} \left(-\frac{x^2 z}{2} + \alpha(y, z) \right) - \frac{\partial}{\partial z} \left(\frac{x^3}{3} + xy + \beta(y, z) \right) &= y + z \\ \frac{\partial \alpha}{\partial y} - \frac{\partial \beta}{\partial z} &= y + z \end{aligned}$$

Podemos por exemplo tomar $\beta = 0$ (convém não complicar a componente A_2) e $\alpha(y, z) = \frac{y^2}{2} + yz$. Conclui-se que

$$\vec{A}(x, y, z) = \left(0, \frac{x^3}{3} + xy, -\frac{x^2 z}{2} + \frac{y^2}{2} + yz \right)$$

é um potencial vetor para \vec{F} . O sentido determinado pela normal \vec{n} dada no bordo ∂S é o sentido anti-horário quando visto de muito acima do plano xy . A parametrização

$$g(\theta) = (\cos \theta, \sin \theta, 0), \quad 0 \leq \theta \leq 2\pi$$

percorre ∂S neste sentido. O Teorema de Stokes diz então que

$$\begin{aligned} \int \int_S \vec{F} \cdot \vec{n} &= \oint_{\partial S} \vec{A} \cdot d\vec{r} \\ &= \int_0^{2\pi} \vec{A}(\cos \theta, \sin \theta, 0) \cdot (-\sin \theta, \cos \theta, 0) d\theta \\ &= \int_0^{2\pi} \left(0, \frac{\cos^3 \theta}{3} + \sin \theta \cos \theta, \frac{\sin^2 \theta}{2} \right) \cdot (-\sin \theta, \cos \theta, 0) d\theta \\ &= \int_0^{2\pi} \frac{\cos^4 \theta}{3} + \sin \theta \cos^2 \theta d\theta \\ &= \int_0^{2\pi} \frac{1}{3} \cos^2 \theta (1 - \sin^2 \theta) d\theta + \int_0^{2\pi} \sin \theta \cos^2 \theta d\theta \\ &= \frac{1}{3} \int_0^{2\pi} \cos^2 \theta - \frac{1}{4} \sin^2(2\theta) d\theta + 0 \\ &= \frac{1}{3} \left(\pi - \frac{\pi}{4} \right) = \frac{\pi}{4} \end{aligned}$$

Vejamos agora uma versão mais precisa da Proposição 50.7.

Proposição 51.2. *Seja $U \subset \mathbb{R}^3$ um aberto em estrela e $\vec{F}: U \rightarrow \mathbb{R}^3$ um campo vetorial solenoidal de classe C^1 . Seja $P \in U$ tal que para todo o $t \in [0, 1]$, o ponto $P + t(x - P) \in U$. Então*

$$\vec{A}(x) = \int_0^1 \vec{F}(P + t(x - P)) \times (t(x - P)) dt$$

é um potencial vetor para \vec{F} em U .

Dem. Suponhamos para simplificar os cálculos que $P = 0$. Temos a verificar que $\text{rot } \vec{A} = \vec{F}$, ou seja que

$$\text{rot} \left(\int_0^1 \vec{F}(tx, ty, tz) \times (tx, ty, tz) dt \right) = \vec{F}(x, y, z)$$

Pela regra de Leibniz, a equação acima é equivalente a

$$\int_0^1 \text{rot} \left(\vec{F}(tx, ty, tz) \times (tx, ty, tz) \right) dt = \vec{F}(x, y, z)$$

Deixamos como exercício mostrar que a condição $\text{div } \vec{F} = 0$ implica que

$$\text{rot} \left(\vec{F}(tx, ty, tz) \times (tx, ty, tz) dt \right) = \frac{d}{dt} \left(t^2 \vec{F}(tx, ty, tz) \right)$$

O resultado é agora uma consequência imediata do Teorema Fundamental do Cálculo de Cálculo 1:

$$\begin{aligned} \int_0^1 \operatorname{rot} \left(\vec{F}(tx, ty, tz) \times (tx, ty, tz) \right) dt &= \int_0^1 \frac{d}{dt} \left(t^2 \vec{F}(tx, ty, tz) \right) dt \\ &= 1 \cdot \vec{F}(x, y, z) - 0 \cdot \vec{F}(0, 0, 0) = \vec{F}(x, y, z). \end{aligned}$$

□

Dem. do Teorema 48.4. Suponhamos primeiro que S é a imagem por uma parametrização de classe C^2 de um domínio regular no plano. Assim, seja $U \subset \mathbb{R}^2$ um aberto, $g: U \rightarrow M$ uma parametrização de classe C^2 , e D um domínio regular no plano tal que $\overline{D} \subset U$ e $g(D) = S$. Temos então $g(\partial D) = \partial S$. Temos a provar que

$$\int \int_S \operatorname{rot} \vec{F} \cdot \vec{n} = \oint_{\partial S} \vec{F} \cdot d\vec{r}$$

Podemos escrever

$$\vec{F} = (F_1, 0, 0) + (0, F_2, 0) + (0, 0, F_3)$$

e claramente basta mostrar a fórmula de Stokes para cada uma das parcelas. A demonstração para cada uma das parcelas é análoga e vamos fazê-la apenas para a primeira para simplificar os cálculos. Temos

$$\operatorname{rot}(F_1, 0, 0) = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ F_1 & 0 & 0 \end{vmatrix} = \left(0, \frac{\partial F_1}{\partial z}, -\frac{\partial F_1}{\partial y} \right)$$

Escrevendo $g(s, t) = (g_1(s, t), g_2(s, t), g_3(s, t))$ para a parametrização de M , temos

$$\frac{\partial g}{\partial s} \times \frac{\partial g}{\partial t} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \frac{\partial g_1}{\partial s} & \frac{\partial g_2}{\partial s} & \frac{\partial g_3}{\partial s} \\ \frac{\partial g_1}{\partial t} & \frac{\partial g_2}{\partial t} & \frac{\partial g_3}{\partial t} \end{vmatrix} = \left(\frac{\partial g_2}{\partial s} \frac{\partial g_3}{\partial t} - \frac{\partial g_3}{\partial s} \frac{\partial g_2}{\partial t}, \frac{\partial g_3}{\partial s} \frac{\partial g_1}{\partial t} - \frac{\partial g_1}{\partial s} \frac{\partial g_3}{\partial t}, \frac{\partial g_1}{\partial s} \frac{\partial g_2}{\partial t} - \frac{\partial g_2}{\partial s} \frac{\partial g_1}{\partial t} \right)$$

e substituindo na definição de fluxo obtemos a seguinte fórmula para o termo esquerdo na igualdade dada pelo Teorema de Stokes (no sentido da normal determinada pela parametrização g):

(29)

$$\begin{aligned} \int \int_S \operatorname{rot}(F_1, 0, 0) \cdot \vec{n} &= \int \int_D \left(0, \frac{\partial F_1}{\partial z}, -\frac{\partial F_1}{\partial y} \right) \cdot \frac{\partial g}{\partial s} \times \frac{\partial g}{\partial t} ds dt \\ &= \int \int_D \frac{\partial F_1}{\partial z} \left(\frac{\partial g_3}{\partial s} \frac{\partial g_1}{\partial t} - \frac{\partial g_1}{\partial s} \frac{\partial g_3}{\partial t} \right) - \frac{\partial F_1}{\partial y} \left(\frac{\partial g_1}{\partial s} \frac{\partial g_2}{\partial t} - \frac{\partial g_2}{\partial s} \frac{\partial g_1}{\partial t} \right) ds dt \end{aligned}$$

Por outro lado, sendo $h:]u_0, u_1[\rightarrow \mathbb{R}^2$ uma parametrização de ∂D que percorre o bordo ∂D no sentido anti-horário, temos que

$$u \mapsto g(h(u)), \quad u_0 < u < u_1$$

é uma parametrização de ∂S . É possível verificar que a orientação dada a ∂S por esta parametrização é compatível com a normal a S determinada pela parametrização g (isto é,

que as orientações induzidas em S e ∂S satisfazem a regra (28)). Substituindo na fórmula para o trabalho temos a seguinte expressão para o termo direito da igualdade no Teorema de Stokes

$$\oint_{\partial S} \vec{F} \cdot d\vec{r} = \int_{u_0}^{u_1} (F_1(g(h(u))), 0, 0) \frac{d}{du} (g(h(u))) du$$

Uma vez que

$$\frac{d}{du} (g(h(u))) = Dg(h(u))h'(u) = \begin{bmatrix} \frac{\partial g_1}{\partial s} & \frac{\partial g_1}{\partial t} \\ \frac{\partial g_2}{\partial s} & \frac{\partial g_2}{\partial t} \\ \frac{\partial g_3}{\partial s} & \frac{\partial g_3}{\partial t} \end{bmatrix} \begin{bmatrix} h'_1(u) \\ h'_2(u) \end{bmatrix}$$

temos

$$(30) \quad \begin{aligned} \oint_{\partial S} \vec{F} \cdot d\vec{r} &= \int_{u_0}^{u_1} F_1(g(h(u))) \left(\frac{\partial g_1}{\partial s}(h(u))h'_1(u) + \frac{\partial g_1}{\partial t}(h(u))h'_2(u) \right) du \\ &= \int_{u_0}^{u_1} \left(F_1 \circ g \frac{\partial g_1}{\partial s}, F_1 \circ g \frac{\partial g_1}{\partial t} \right) \cdot h'(u) du \end{aligned}$$

O último integral é o trabalho do campo

$$(P(s, t), Q(s, t)) = \left(F_1(g(s, t)) \frac{\partial g_1}{\partial s}(s, t), F_1(g(s, t)) \frac{\partial g_1}{\partial t}(s, t) \right)$$

ao longo do bordo ∂D . Temos

$$\frac{\partial Q}{\partial s} - \frac{\partial P}{\partial t} = \frac{\partial}{\partial s} \left((F_1 \circ g) \frac{\partial g_1}{\partial s} \right) - \frac{\partial}{\partial t} \left((F_1 \circ g) \frac{\partial g_1}{\partial t} \right) =$$

$$\left(\frac{\partial F_1}{\partial x} \frac{\partial g_1}{\partial s} + \frac{\partial F_1}{\partial y} \frac{\partial g_2}{\partial s} + \frac{\partial F_1}{\partial z} \frac{\partial g_3}{\partial s} \right) \frac{\partial g_1}{\partial t} + F_1 \frac{\partial^2 g_1}{\partial s \partial t} - \left(\frac{\partial F_1}{\partial x} \frac{\partial g_1}{\partial t} + \frac{\partial F_1}{\partial y} \frac{\partial g_2}{\partial t} + \frac{\partial F_1}{\partial z} \frac{\partial g_3}{\partial t} \right) \frac{\partial g_1}{\partial s} - F_1 \frac{\partial^2 g_1}{\partial t \partial s}$$

que (usando o Lema de Schwarz) se simplifica em

$$\left(\frac{\partial F_1}{\partial y} \frac{\partial g_2}{\partial s} + \frac{\partial F_1}{\partial z} \frac{\partial g_3}{\partial s} \right) \frac{\partial g_1}{\partial t} - \left(\frac{\partial F_1}{\partial y} \frac{\partial g_2}{\partial t} + \frac{\partial F_1}{\partial z} \frac{\partial g_3}{\partial t} \right) \frac{\partial g_1}{\partial s}$$

Esta expressão é exatamente a função integranda em (29), logo aplicando o Teorema de Green a (30), obtemos (29), o que conclui a demonstração do Teorema de Stokes no caso em que S é a imagem por uma parametrização de classe C^2 de um domínio regular no plano. O caso geral do Teorema deduz-se deste caso particular notando que qualquer domínio regular numa superfície pode ser dividido em regiões que são imagens de domínios regulares no plano (e o facto que parametrizações de classe C^1 podem ser aproximadas por parametrizações de classe C^2). \square

52. MAIS SOBRE AS LEIS DO ELETROMAGNETISMO. RESUMO DA MATÉRIA

52.1. **As leis de Faraday e de Ampère.** Falámos já das leis de Gauss do Eletromagnetismo que são

$$\operatorname{div} \vec{E} = \frac{\rho}{\epsilon} \quad \operatorname{div} \vec{B} = 0$$

As restantes leis de Maxwell do eletromagnetismo são a lei de Faraday e lei de Ampère que envolvem o rotacional dos campos eléctrico e magnético. A Lei de Faraday diz que

$$\nabla \times \vec{E} = -\frac{\partial \vec{B}}{\partial t}$$

Aplicando o Teorema de Stokes vemos que, para qualquer superfície S ,

$$\oint_{\partial S} \vec{E} \cdot d\vec{r} = -\frac{d}{dt} \left(\int \int_S \vec{B} \cdot \vec{n} \right)$$

Isto diz-nos que uma variação do fluxo magnético através de uma superfície gera trabalho do campo eléctrico no bordo da superfície. Este é o princípio utilizado para converter energia mecânica em eléctrica (nos alternadores dos carros, barragens, moínhos de vento, etc...): a energia mecânica é utilizada para fazer rodar ímans e portanto causar variação do campo magnético (e portanto do seu fluxo) com o tempo. Dispondo de forma conveniente um circuito eléctrico (sobre o bordo ∂S de uma superfície determinada em função da configuração dos ímans) ir-se-á gerar trabalho do campo eléctrico sobre o circuito, dando azo a uma corrente alterna.

A lei de Ampère afirma que

$$\nabla \times \vec{B} = \mu \left(\vec{J} + \epsilon \frac{\partial \vec{E}}{\partial t} \right)$$

onde \vec{J} é a densidade de corrente eléctrica e μ é uma constante determinada pelo meio chamada a permeabilidade magnética. Aplicando o Teorema de Stokes obtemos

$$\oint_{\partial S} \vec{B} \cdot d\vec{r} = \mu \int \int_S \left(\vec{J} + \epsilon \frac{\partial \vec{E}}{\partial t} \right) \cdot \vec{n}$$

Esta equação mostra que as correntes eléctricas (assim como a variação de fluxo do campo eléctrico) dão azo a um campo magnético. Por exemplo se tomarmos para S uma superfície que interseste um solenóide de tal forma que o bordo da superfície passe pelo interior do solenóide vemos que o trabalho do campo magnético ao longo do bordo da superfície será não nulo. Em particular a passagem de corrente num solenóide gera um campo magnético não nulo que, dentro do solenóide, tem aproximadamente a direcção do eixo do solenóide. É esta a origem da aplicação do termo solenoidal a um campo que, como o magnético, tenha divergência nula.

52.2. Resumo da matéria para o segundo teste.

(1) Teoremas da Função Inversa e Implícita

- $\det Df(x) \neq 0 \Rightarrow f$ é localmente invertível perto de x

$$D(f^{-1})(f(x)) = (Df(x))^{-1}$$

- $\det \frac{\partial F}{\partial y}(x_0, y_0) \neq 0 \Rightarrow$ perto de (x_0, y_0) temos $F(x, y) = 0 \Leftrightarrow y = g(x)$

$$Dg(x_0) = - \left(\frac{\partial F}{\partial y}(x_0, y_0) \right)^{-1} \frac{\partial F}{\partial x}(x_0, y_0)$$

(2) Variedades, espaços tangente e normal

- $M = \{x: F(x) = 0\}$ variedade se $F(x) = 0 \Rightarrow DF(x)$ tem característica máxima. $T_x M$ é o espaço das linhas de $DF(x)$, e $(T_x M)^\perp$ é o núcleo de $DF(x)$.

- Se $g: A \rightarrow M$ é uma parametrização, $T_{g(t)} M$ tem por base as colunas de $Dg(t)$.

(3) Multiplicadores de Lagrange: Sendo $M = \{x: F(x) = 0\}$, um extremo de f em M ocorre num ponto em que $\nabla f(x) \in (T_x M)^\perp \Leftrightarrow \nabla f(x) = \lambda_1 \nabla F_1(x) + \dots + \lambda_k \nabla F_k(x)$

(4) Integrais de campos escalares em variedades: Sendo $g: A \rightarrow M$ uma parametrização

$$\int_M f = \int_A f(g(t)) \sqrt{\det Dg(t)^T Dg(t)}$$

(5) Integrais de campos vetoriais

- Trabalho:

$$\int_C \vec{F} \cdot d\vec{r} = \pm \int_{t_0}^{t_1} \vec{F}(g(t)) \cdot g'(t) dt$$

- Fluxo:

$$\iint_S \vec{F} \cdot d\vec{r} = \pm \iint_A \vec{F}(g(s, t)) \cdot \left(\frac{\partial g}{\partial s} \times \frac{\partial g}{\partial t} \right) ds dt$$

(6) Teoremas Fundamentais do Cálculo

- Para integrais de linha:

$$\int_A^B \nabla \varphi \cdot d\vec{r} = \varphi(B) - \varphi(A)$$

Gradiente \Leftrightarrow Conservativo $\Leftrightarrow \oint_C \vec{F} \cdot d\vec{r} = 0$ para todo o C . Gradiente \Rightarrow Fechado. Fechado num conjunto simplesmente conexo \Rightarrow gradiente.

- Teorema de Green:

$$\iint_A \frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} dx dy = \int_{\partial A} (P, Q) \cdot d\vec{r}$$

- Teorema da Divergência:

$$\iiint_V \operatorname{div} \vec{F} = \iint_{\partial V} \vec{F} \cdot \vec{n}^{ext}$$

- Teorema de Stokes:

$$\int \int_S \operatorname{rot} \vec{F} \cdot \vec{n} = \oint_S \vec{F} \cdot d\vec{r}$$

Rotacional \Rightarrow Solenoidal (i.e. $\operatorname{div} \vec{F} = 0$). Solenoidal num conjunto em estrela \Rightarrow rotacional.

53. RESOLUÇÃO DE EXERCÍCIOS

54. RESOLUÇÃO DE EXERCÍCIOS

REFERENCES

- [DM] Departamento de Matemática do IST, *Exercícios de Análise Matemática I e II*, IST Press (2010).
[P1] G. Pires, *Cálculo Diferencial e Integral em \mathbb{R}^n* , IST Press (2014).
[P2] G. Pires, *Exercícios de Cálculo Integral em \mathbb{R}^n* , IST Press (2007).
[S] M. Spivak, *Calculus on Manifolds*, Perseus books (1971).