Instituto Superior Técnico
Departamento de Matemática

Cálculo Diferencial e Integral II

Apontamentos das aulas teóricas

(Adaptado de um texto do Prof. Gustavo Granja)

Índice

1. Introdução 2
2. Noções topológicas em $\mathbb{R}^n$ 3
3. Sucessões em $\mathbb{R}^n$ 4
4. Limites de funções 6
5. Continuidade 7
6. Cálculo de limites 8
7. Teorema de Weierstrass 10
8. Derivada segundo um vetor 12
9. Diferenciabilidade 14
10. Funções de classe C^1 18
11. Regra de derivação da função composta 19
12. Aplicações geométricas da derivada 22
13. Lema de Schwarz 25
14. Fórmula de Taylor 29
15. Extremos 32
16. Teorema da Função Inversa 37
17. Teorema da Função Implícita 39
18. Método dos multiplicadores de Lagrange 49
19. Definição de integral 53
20. Teorema de Fubini 55
21. Propriedades do integral. Integrabilidade 63
22. Mudança de variáveis 68
23. Regra de Leibniz 80
24. Integrais de linha de campos escalares 83
25. Integrais de linha de campos vetoriais 86
26. Campos gradientes e campos fechados 89
27. Teorema Fundamental do Cálculo 94
28. Teorema de Green 98
29. Homotopia de caminhos 103
30. Conjuntos simplesmente conexos 105

Referências 108

1. Introdução

Este texto consiste num guião aproximado para as apresentações feitas nas aulas. Não se destina a substituir um livro de texto. Para um tratamento coerente dos assuntos desta cadeira recomendamos os livros [P1] e [P2]. Um excelente texto para os alunos mais interessados é a referência [S].

2. Noções topológicas em $\mathbb{R}^n$

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

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

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 vetor $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$ é o real não negativo dado pela expressão

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

e que o seu significado geométrico é o comprimento do vetor x. Tendo em conta o significado geométrico da adição de vetores, 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:*

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

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 $B_r(x) \cap A = \varnothing$ (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 interseção 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 $B_r(x)$ interseta A).*

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

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 $A = \{(x, y) \in \mathbb{R}^2 \colon x \geq 0\}$. Então $\operatorname{int} A = \{(x, y) \in \mathbb{R}^2 \colon x > 0\}$, $\operatorname{ext} A = \{(x, y) \in \mathbb{R}^2 \colon x < 0\}$, $\partial A = \{(x, y) \in \mathbb{R}^2 \colon x = 0\}$ *e* $\overline{A} = A$.

(b) Seja $A = \{(x,y) \in \mathbb{R}^2 \colon x^2 + y^2 = 1\}$. *Então* $\operatorname{int} A = \varnothing$, $\operatorname{ext} A = A^c = \{(x,y) \in \mathbb{R}^2 \colon x^2 + y^2 \neq 1\}$, $\partial A = A$, $\overline{A} = A$.
(c) Seja $A = \{(x,y,z) \in \mathbb{R}^3 \colon z > x^2 + y^2\}$. *Então* $\operatorname{int} A = A$, $\operatorname{ext} A = \{(x,y,z) \in \mathbb{R}^3 \colon z < x^2 + y^2\}$, $\partial A = \{(x,y,z) \in \mathbb{R}^3 \colon z = x^2 + y^2\}$, *e* $\overline{A} = \{(x,y,z) \in \mathbb{R}^3 \colon z \geq x^2 + y^2\}$.
(d) Seja $A = \{(x,y) \in \mathbb{R}^2 \colon x \in \mathbb{Q}\}$. *Então* $\operatorname{int} A = \operatorname{ext} A = \varnothing$ *e* $\partial A = \overline{A} = \mathbb{R}^2$.

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

(i) aberto *se* $A = \operatorname{int} A$,
(ii) fechado *se* $A = \overline{A}$ *(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 portanto 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 $\varnothing$.

3. Sucessões em $\mathbb{R}^n$

Uma sucessão em $\mathbb{R}^n$ é uma função $\mathbb{N} \to \mathbb{R}^n$ que a cada natural $k \in \mathbb{N}$ associa um vetor $x_k \in \mathbb{R}^n$. 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

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

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

Exemplo 3.1. *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 3.2. *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* $\lim_{k\to\infty} x_k = y$.

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

$$\lim x_k = y \quad \Leftrightarrow \|x_k - y\| \to 0.$$

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

Exemplo 3.3. *A sucessão do Exemplo 3.1 não converge.*

Proposição 3.4. *Seja (x_k) uma sucessão em $\mathbb{R}^n$. Então a sucessão x_k converge sse as suas sucessões coordenadas convergem e $\lim x_k = (y_1, \ldots, y_n) \in \mathbb{R}^n$ se e só se $\lim x_{i,k} = y_i$.*

Dem. Se $\lim x_k = y$ então $\|x_k - y\| \to 0$. Para cada $i = 1, \ldots, n$ 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 $|x_{i,k} - y_i| \to 0$ para cada i e portanto

$$\|x_k - y\| = \sqrt{(x_{1,k} - y_1)^2 + \ldots + (x_{n,k} - y_k)^2} \to 0$$

logo x_k converge para y. □

Exemplo 3.5.

(i) Vemos novamente que a sucessão do Exemplo 3.1 não converge uma vez que a sua segunda sucessão coordenada não converge.

(ii) Tem-se

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

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 3.6. *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 $\{x_k \colon k \in \mathbb{N}\} \subset \mathbb{R}^n$ é limitado.*

Proposição 3.7. *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 3.4 e 3.7, uma sucessão convergente é limitada: uma sucessão convergente tem sucessões coordenadas convergentes, portanto limitadas, e é portanto limitada. É também fácil deduzir este resultado diretamente das definições de sucessão convergente e limitada (exercício).

Exemplo 3.8.

(i) A sucessão $\left(\cos k, \operatorname{sen} k, \frac{k^2+k}{3k^2+5}\right)$ *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* $\left(0, \frac{1}{3}, 0\right)$ *e é portanto também limitada.*

4. Limites de funções

Chegamos agora a um dos conceitos fundamentais do Cálculo.

Definição 4.1. *Seja $A \subset \mathbb{R}^n$, $f\colon A \to \mathbb{R}^m$ uma função, e $a \in \overline{A}$ 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*

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

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:

$$\lim_{x\to a} f(x) = y \Leftrightarrow \forall_{\epsilon>0}\exists_{\delta>0}\ x \in A \text{ e } \|x-a\| < \delta \Rightarrow \|f(x)-y\| < \epsilon.$$

Exemplo 4.2. *Seja $A = \mathbb{R}^2 \setminus \{(0,0)\}$ e $f\colon A \to \mathbb{R}$ a função definida pela expressão*

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

Então

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

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 $(0,0) \in \overline{A}$. 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 $(x_k,y_k) = \left(\frac{1}{k},\frac{1}{k}\right)$ 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 $\lim_{(x,y)\to(0,0)} f(x,y)$ existir, terá de ser igual a $\sqrt{2}$. No entanto, calculando o limite de $f(x_k,y_k)$ com $(x_k,y_k) = \left(\frac{1}{k},0\right)$ obtemos

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

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

5. Continuidade

Definição 5.1. *Seja $A \subset \mathbb{R}^n$, $f\colon A \to \mathbb{R}^n$ uma função e $a \in A$. A função f diz-se* contínua em a *se $\lim_{x\to a} f(x) = f(a)$. 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 5.2.

(i) A função $f\colon \mathbb{R}^2 \to \mathbb{R}$ 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 $f\colon \mathbb{R}^2 \to \mathbb{R}$ 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 $\left(\frac{(-1)^k}{k}, 0\right)$ (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 5.3. *Seja $A \subset \mathbb{R}^n$ e $f\colon A \to \mathbb{R}^m$ uma função. Escrevendo*

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

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

Exemplo 5.4. *A função $f\colon \mathbb{R}^2 \to \mathbb{R}^2$ definida por $f(x,y) = (2x + e^y, \operatorname{sen}(xy))$, tem primeira função coordenada $f_1\colon \mathbb{R}^2 \to \mathbb{R}$ definida por $f_1(x,y) = 2x + e^y$ e segunda função coordenada $f_2\colon \mathbb{R}^2 \to \mathbb{R}$ definida por $f_2(x,y) = \operatorname{sen}(xy)$.*

Proposição 5.5 (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 $f\colon A \to \mathbb{R}^m$. Então f é contínua em $a \in A$ se e só se as funções coordenadas $f_1, \ldots, f_m\colon A \to \mathbb{R}$ são contínuas em a.

(iv) *Sejam* $A \subset \mathbb{R}^n$ *e* $f, g\colon A \to \mathbb{R}$ *funções contínuas em* $a \in A$. *Então* $f+g, f\cdot g\colon A \to \mathbb{R}$ *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* $A \subset \mathbb{R}^n$, $B \subset \mathbb{R}^m$, $f\colon A \to R^m$ *e* $g\colon B \to \mathbb{R}^p$. *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 $x_k \to a$, então $g(f(x_k)) \to g(f(a))$. Mas claramente, x_k pertence ao domínio A de f, logo, sendo f contínua em a, temos $f(x_k) \to f(a)$. Uma vez que x_k pertence ao domínio da composta, $f(x_k)$ pertence ao domínio de g. Como g é contínua em $f(a)$, $g(f(x_k)) \to g(f(a))$. □

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

Exemplo 5.6. *A função* f *do Exemplo 5.4 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,* $(x, y) \mapsto \operatorname{sen}(y)$ *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* $(x, y) \mapsto 2x + \operatorname{sen} y$, *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.*

6. 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 6.1.

(i) *Seja* $f\colon \mathbb{R}^2 \setminus \{(0,0)\} \to \mathbb{R}^2$ *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*

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

porque f *é contínua em* $(1, 1)$ *(como segue da Proposição 5.5).*

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 reta. Considerando uma reta com*

declive m que passa pela origem temos

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

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 reta $y=x$). Conclui-se que o limite de f na origem não existe.

(ii) Seja $f\colon \mathbb{R}^2 \to \mathbb{R}$ 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:

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

e uma vez que $f(x,0) = 0$, 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

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

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 $\lim_{x\to a} f(x) = b$ 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 6.2.

(i) Determinar $\lim_{(x,y)\to(0,0)} \frac{x^3+x^2y+y^3}{x^2+y^2}$ *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, $\frac{x^3+x^2y+y^3}{x^2+y^2} \to 0$ *quando* $(x,y) \to (0,0)$.

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

É imediato verificar que os limites direcionais 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*

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

Concluimos que

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

7. 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 7.1 (Teorema de Weierstrass)**.** *Seja* $A \subset \mathbb{R}^n$ *compacto e* $f\colon A \to \mathbb{R}$ *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 7.2 (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 interseta A e portanto v está distante de todos os termos da sucessão). □

Nota 7.3. *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 7.2 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 $\overline{A}$ implicam a Proposição 7.2.

Dem. do Teorema de Weierstrass 7.1. 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 $\ell = \inf\{f(x)\colon x \in A\} \in [-\infty, \infty[$. 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 $\{f(x)\colon x \in A\}$ é limitado inferiormente)

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

A Proposição 7.2 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

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

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. □

8. Derivada segundo um vetor

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

Definição 8.1. *Seja $A \subset \mathbb{R}^n$, $f\colon A \to \mathbb{R}$ e a um ponto interior a A. Dado $v \in \mathbb{R}^n$, define-se a* derivada de f em a segundo o vetor v *como sendo o limite*

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

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 vetor v é o seguinte (para $v \neq 0$): o vetor v determina uma reta com direção v que passa por a, que é descrita pela equação paramétrica $a + tv, t \in \mathbb{R}$. Restringindo a função f a essa reta 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 reta). A derivada segundo v é portanto uma medida da *taxa de crescimento da função f na direção de v* (ver no entanto a Nota 8.3 abaixo).

Exemplo 8.2. *Seja $f\colon \mathbb{R}^2 \to \mathbb{R}$ a função definida por $f(x, y) = x^2 + y^2$. Tomando por exemplo $a = (1, 1)$ e $v = (2, 3)$ temos*

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

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

Nota 8.3. *A derivada segundo um vetor v depende, não apenas da direção 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 vetor v é um dos vetores da base canónica de $\mathbb{R}^n$ é particularmente importante e recebe um nome específico.

Definição 8.4. *Seja $A \subset \mathbb{R}^n$, $f\colon A \to \mathbb{R}$ e $a \in \operatorname{int} A$. Para $i = 1, \ldots, n$, a* i-ésima derivada parcial de f *no ponto a é $D_{e_i} f(a)$, onde $e_i = (0, \ldots, 0, 1, 0, \ldots, 0)$ é o i-ésimo vetor 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

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

Isto significa que a i-ésima derivada parcial num ponto $(x_1, \ldots, x_n)$ 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 8.5.

(i) Se $f\colon \mathbb{R}^2 \to \mathbb{R}$ é definida por $f(x, y) = 2x + e^{xy}$ temos

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

(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.

A definição de derivada segundo um vetor e de derivada parcial para funções com valores vetoriais 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 $f\colon A \to \mathbb{R}^m$ é uma função, define-se a derivada de f segundo o vetor v no ponto a por

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

e a i-ésima derivada parcial por

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

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

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

tem-se

$$D_v f(a) = (D_v f_1(a), \ldots, D_v f_m(a)), \qquad \frac{\partial f}{\partial x_i}(a) = \left(\frac{\partial f_1}{\partial x_i}(a), \ldots, \frac{\partial f_m}{\partial x_i}(a)\right).$$

Exemplo 8.6. *Sendo $f\colon \mathbb{R}^2 \to \mathbb{R}^3$ a função definida por $f(x,y) = (x^2 + y, xe^{xy}, \operatorname{sen}(xy))$ tem-se*

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

9. Diferenciabilidade

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 reta tangente no ponto $(a, f(a))$. A derivada é então o declive dessa reta 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 reta 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 9.1. *Seja $A \subset \mathbb{R}^n$, $f\colon A \to \mathbb{R}^m$ e $a \in \operatorname{int} A$. A função f diz-se* diferenciável em *a se existe uma transformação linear $L\colon \mathbb{R}^n \to \mathbb{R}^m$ tal que*

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

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 9.2. *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 $\lim_{x\to a} L(x-a) = L(a-a) = L0 = 0$. Conclui-se então que $\lim_{x\to a} f(x) - f(a) = 0$, isto é, que f é contínua em a. □

Proposição 9.3. *Seja $A \subset \mathbb{R}^n$, $f\colon A \to \mathbb{R}^m$ e $a \in \operatorname{int} A$. Se f é diferenciável em a, então a derivada de f segundo o vetor 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

$$f(x) = f(a) + Df(a)(x-a) + \epsilon(x), \qquad \lim_{x\to a} \frac{\epsilon(x)}{\|x-a\|} = 0.$$

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

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

conclui-se que

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

□

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 vetor da base. Aplicando a Proposição 9.3 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 vetor $Df(a)e_i$, onde e_i é o i-ésimo vetor da base canónica de $\mathbb{R}^n$, ou seja, a i-ésima derivada parcial de f em a.

Definição 9.4. *Seja $A \subset \mathbb{R}^n$, $a \in \operatorname{int} A$ e $f\colon A \to \mathbb{R}^m$ 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 $f = (f_1, \ldots, f_m)$ 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 9.5. *Seja $A \subset \mathbb{R}^n$, $a \in \operatorname{int} A$ e $f\colon A \to \mathbb{R}^m$ 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 9.6. *Seja $f\colon \mathbb{R}^2 \to \mathbb{R}^3$ a função definida por $f(x,y) = (x^2+y, \operatorname{sen}(x+y), xe^{xy})$. A matriz Jacobiana de f num ponto (x,y) é dada por*

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

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

Proposição 9.7.

(1) Se $f\colon \mathbb{R}^n \to \mathbb{R}^m$ 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 $f\colon \mathbb{R}^n \to \mathbb{R}^m$ 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 $x \to a$. 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 $x \to a$, 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 9.8. *Seja $A \subset \mathbb{R}^n$, $a \in \operatorname{int} A$ e $f\colon A \to \mathbb{R}^m$. Então f é diferenciável em a sse cada uma das suas funções coordenadas $f_i\colon A \to \mathbb{R}$ com $i = 1, \ldots, m$ é diferenciável em a.*

Dem. Tendo em conta o Corolário 9.5, 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 $x \to a$. Uma vez que o limite se calcula coordenada a coordenada, este limite é nulo sse para cada $i = 1, \ldots, m$ tivermos

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

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 9.9. *Verificar se a função* $f \colon \mathbb{R}^2 \to \mathbb{R}$ *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* $\begin{bmatrix} \frac{\partial f}{\partial x}(0,0) & \frac{\partial f}{\partial y}(0,0) \end{bmatrix}$. *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*

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

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

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

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*

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

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 $\begin{bmatrix} 0 & 0 \end{bmatrix}$.

10. Funções de classe C^1

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

Teorema 10.2. *Seja $A \subset \mathbb{R}^n$, $f\colon A \to \mathbb{R}^m$ e $a \in \operatorname{int} A$. Se f é de classe C^1 em a, então f é diferenciável em a.*

Exemplo 10.3. *A função do Exemplo 9.9 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).*

Dem. do Teorema 10.2. Pela Proposição 9.8 (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.

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

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

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)

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

por definição de derivada no ponto a da função $g(x) = f(x,b)$ (derivada esta que é exactamente 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 $(x,y)\to(a,b)$, o que conclui a demonstração. □

11. Regra de derivação da função composta

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

Teorema 11.1. *Sejam $A\subset\mathbb{R}^n$, $f\colon A\to\mathbb{R}^m$, $a\in\operatorname{int}A$, $B\subset\mathbb{R}^m$ tal que $f(a)\in\operatorname{int}B$, e $g\colon B\to\mathbb{R}^p$. 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.

Exemplo 11.2. *Sejam $f\colon\mathbb{R}^2\to\mathbb{R}^2$ e $g\colon\mathbb{R}^2\to\mathbb{R}^3$ 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*

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

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

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

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 11.1 e da regra para a multiplicação de matrizes.

Corolário 11.3 (Regra da cadeia)**.** *Nas condições do Teorema 11.1, escrevendo $y = f(x)$ e $z = g(y)$, temos para cada $1 \leq i \leq p$ e $1 \leq j \leq n$*

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

ou, equivalentemente,

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

Exemplo 11.4. *Sendo $g\colon \mathbb{R}^3 \to \mathbb{R}$ uma função diferenciável, vamos determinar uma expressão para a derivadas parcial em ordem a x da função $g\left(xy + x^2, y^3x, e^{xy}\right)$ em termos das derivadas parciais de g:*

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

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

O Teorema 11.1 tem como consequência que as operações algébricas habituais preservam a diferenciabilidade.

Proposição 11.5. *Seja $A \subset \mathbb{R}^n$. Se $f, g\colon A \to \mathbb{R}$ são diferenciáveis em $a \in \operatorname{int} A$, 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 $f/g\colon A \to \mathbb{R}$ é a composta das funções $k\colon \{(x,y) \in \mathbb{R}^2 \colon y \neq 0\} \to \mathbb{R}$ definida por $k(x,y) = x/y$ e $h\colon A \to \mathbb{R}^2$ 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 11.1 garante então que f/g é diferenciável em a. □

Dem. do Teorema 11.1. Precisamos de mostrar que o quociente

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

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

$$Df(a)(x-a) = f(x) - f(a) - \epsilon_a(x) \tag{4}$$

e, por definição de derivada, temos

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

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 $f(x) \to f(a)$ quando $x \to a$, logo

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

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

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

A primeira parcela é claramente limitada (certamente as componentes do vetor 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. □

12. Aplicações geométricas da derivada

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

Definição 12.1. *Seja $A \subset \mathbb{R}^n$, $f\colon A \to \mathbb{R}$ e $a \in \operatorname{int} A$. O* gradiente de f em a *é o vetor*

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

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

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

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

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

(onde $\langle \cdot, \cdot \rangle$ denota o produto interno usual de vetores de $\mathbb{R}^n$). Recorde-se que se v e w são vetores 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 direção que o gradiente. Além disso se v é um vetor unitário com a mesma direção 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 vetor que aponta na direção em que o valor de f cresce mais rapidamente a partir de a. O seu comprimento é a taxa de crescimento de f nessa direção.

Exemplo 12.2. *Consideremos a função $f\colon \mathbb{R}^2 \to \mathbb{R}$ 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 direção radial, e que a taxa de crescimento aumenta à medida que nos afastamos da origem.

Definição 12.3. *Um* conjunto de nível *de uma função escalar $f\colon A \to \mathbb{R}$ (com $A \subset \mathbb{R}^n$) é um conjunto da forma*

$$\{(x_1, \ldots, x_n) \in A\colon f(x_1, \ldots, x_n) = c\}$$

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

Exemplo 12.4. *Para a função $f(x,y) = x^2 + y^2$ do exemplo anterior, os conjuntos do nível são*

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

No exemplo anterior $\nabla f(x,y)$ é um vetor perpendicular à linha de nível que passa por (x,y). Intuitivamente é claro que isto deve acontecer. Caso contrário, a derivada segundo um vetor 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.

Definição 12.5. *Um* caminho *em $\mathbb{R}^n$ é uma função contínua $\gamma \colon I \to \mathbb{R}^n$ 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

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

Intuitivamente, à medida que t se aproxima de t_0, a direção da razão incremental aproxima-se da direção 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 vetor tangente à curva descrita por γ* no ponto $\gamma(t_0)$. Pensando em $\gamma(t)$ como o vetor 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 retas tangentes a curvas assim como dos planos ortogonais a estas.

Exemplo 12.6. *Seja $\gamma \colon \mathbb{R} \to \mathbb{R}^3$ o caminho definido por*

$$\gamma(t) = (\cos t, \operatorname{sen} t, t).$$

γ descreve uma hélice em torno do eixo dos zz, que se projeta na circunferência de raio 1 no plano xy. Temos $\gamma'(t) = (-\operatorname{sen} t, \cos t, 1)$. A reta 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 vetor tangente $(0,1,1)$. É portanto dado pela equação

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

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 $f\colon \mathbb{R}^n \to \mathbb{R}$ uma função diferenciável e $\gamma\colon I \to \mathbb{R}^n$ um caminho diferenciável com imagem contida no conjunto de nível c de f,

$$f^{-1}(c) = \{(x_1, \ldots, x_n) \in \mathbb{R}^n \colon f(x_1, \ldots, x_n) = c\}.$$

Então a função composta

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

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

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

obtemos

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

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.

Uma aplicação da ortogonalidade do gradiente aos conjuntos de nível é a determinação de planos tangentes e retas normais a superfícies.

Exemplo 12.7. *Seja $C = \{(x, y, z) \in \mathbb{R}^3 \colon x^2 + y^2 = z^2\}$. Determinar o plano tangente e a reta 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 interseção 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 $f\colon \mathbb{R}^3 \to \mathbb{R}$ definida por $f(x, y, z) = x^2+y^2-z^2$. Um vetor 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 vetor normal a C em $(1, 1, \sqrt{2})$ é $(2, 2, -2\sqrt{2})$. Segue-se que a reta 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 reta $z = r$ no plano vertical definido 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 vetor perpendicular a C em $(1,1,\sqrt{2})$ e assim recuperar uma equação para a reta perpendicular nesse ponto.

13. Lema de Schwarz

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

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

e

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

(e analogamente para derivadas parciais destas funções).

Exemplo 13.1. *Seja $f\colon \mathbb{R}^2\to\mathbb{R}$ a função definida por $f(x,y)=xe^{xy^2}$. Então*

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

Definição 13.2. *Seja f uma função escalar de n variáveis reais. Uma* derivada parcial de ordem k de f *é uma expressão da forma*

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

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

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

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 13.3. *Seja $f\colon \mathbb{R}^2 \to \mathbb{R}$ a função definida por $f(x,y) = x^4y^2 + xy^2$. Temos*

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

e

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

Temos ainda, por exemplo,

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

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 13.4. *Seja $A \subset \mathbb{R}^n$, $a \in \operatorname{int} A$ e $f\colon A \to \mathbb{R}$ 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 de centro 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 o seno, a exponencial, etc.) definem funções de classe C^k no seu domínio. Por exemplo a função do Exemplo 13.3 é de classe C^∞.

Teorema 13.5 (Lema de Schwarz)**.** *Seja $A \subset \mathbb{R}^n$, $a \in \operatorname{int} A$, e $f\colon A \to \mathbb{R}$ 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 13.6. *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 $f\colon \mathbb{R}^3 \to \mathbb{R}$ 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

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

e, de forma inteiramente análoga, vemos que

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

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*

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

Ambos estes limites iterados são iguais aos limites de sucessões

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

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,

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

De facto, com essa escolha de k_n teremos

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

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

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

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

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

pelo que

$$\Delta_{a,b}(h,k)=\phi'(t_0)=\left(\frac{\partial f}{\partial x}(a+t_0h,b+k)-\frac{\partial f}{\partial x}(a+t_0h,b)\right)h.$$

Uma vez que a função

$$y\mapsto\frac{\partial f}{\partial x}(a+t_0h,y)$$

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

$$\frac{\partial f}{\partial x}(a+t_0h,b+k)-\frac{\partial f}{\partial x}(a+t_0h,b)=\frac{\partial^2 f}{\partial y\partial x}(a+t_0h,b+t_1k)k,$$

e portanto

$$\frac{\Delta_{a,b}(h,k)}{hk}=\frac{\left(\frac{\partial^2 f}{\partial y\partial x}(a+t_0h,b+t_1k)k\right)h}{hk}=\frac{\partial^2 f}{\partial y\partial x}(a+t_0h,b+t_1k).$$

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

$$\lim_{(h,k)\to(0,0)}\frac{\Delta_{a,b}(h,k)}{hk}=\lim_{(h,k)\to(0,0)}\frac{\partial^2 f}{\partial y\partial x}(a+t_0h,b+t_1k)=\frac{\partial^2 f}{\partial y\partial x}(a,b),$$

o que conclui a demonstração. □

Nota 13.7. *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. 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 $f\colon I \to \mathbb{R}$ é uma função $(k+1)$ vezes diferenciável em I, temos a fórmula de Taylor de ordem k com resto de Lagrange:

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

onde ξ é um ponto entre a e $a+h$. Cada termo não nulo na expressão acima domina or 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 $f\colon U \to \mathbb{R}$ e um vetor $h \in \mathbb{R}^n$ 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 reta que passa por a que tem a direção do vetor 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 $\phi'(t)$:

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

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

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

Esta expressão pode ser escrita na forma matricial

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

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* $n \times n$

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

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 vetor 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, $f\colon U \to \mathbb{R}$ 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

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

Dem. Seja $h \in \mathbb{R}^n$ tal que o segmento de reta 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 $K = \{x \in \mathbb{R}^n \colon \|x-a\| \le r\}$ 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 $\|h\| \le r$

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

Isto conclui a demonstração. □

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

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

e

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

Como

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

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

$$f(x,y) = 1 + x + 3y + \frac{1}{2}(x^2 + 6xy + 9y^2) + R_2(x,y)\ ,$$

onde

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

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*

$$e^u = 1 + u + \frac{u^2}{2} + \cdots$$

obtemos

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

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 $f\colon A \to \mathbb{R}$ com $A \subset \mathbb{R}^n$, dizemos que $a \in A$ é um *ponto de máximo* (relativo) de f se existir 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 $f\colon I \to \mathbb{R}$ é diferenciável, para que $a \in \operatorname{int} I$ 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 $f\colon A \to \mathbb{R}$ uma função diferenciável. Se $a \in \operatorname{int} A$ 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 vetor $\nabla f(a)$ 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. □

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* $T = \{(x,y) \in \mathbb{R}^2 \colon x \geq 0, y \geq 0, x+y \leq 1\}$ *e* $f \colon T \to \mathbb{R}$ *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* $(x,y) = \left(\frac{1}{3}, \frac{1}{3}\right)$*. 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, respetivamente (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 $f \colon \mathbb{R}^2 \to \mathbb{R}$ 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 respetivamente 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* $f \colon U \to \mathbb{R}$ *uma função de classe* C^3*. Seja* a *um ponto de estacionariedade de* f *e* $q(h) = h^T Hf(a) h$ *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 respetivamente*

(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.*

Dem. da Proposição 15.5. Para verificar as afirmações (iii)-(v), considerem-se as funções $\phi(t) = f(a+th)$ onde $h \in \mathbb{R}^n$ é um vetor 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 $S^t S = \mathrm{Id}$). As colunas de S são vetores próprios de $Hf(a)$. Sendo $\lambda_1, \ldots, \lambda_n$ os valores próprios de $Hf(a)$ e escrevendo $\Lambda = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$ 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

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

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. □

Exemplo 15.7. *Determinar e classificar os pontos de estacionariedade das seguintes funções.*

(i) $f\colon \mathbb{R}^2 \to \mathbb{R}$ *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* $\det(A - \lambda \operatorname{Id}) = 0$. *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* $\operatorname{tr} A = \lambda_1 + \lambda_2$ *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* $\operatorname{tr} A$. *No exemplo que estamos a tratar temos* $\det Hf(-1,1) = 3 > 0$ *e* $\operatorname{tr} Hf(-1,1) = 4$ *donde concluímos novamente que ambos os valores próprios de* $Hf(-1,1)$ *são positivos.*

(ii) $f\colon \mathbb{R}^3 \to \mathbb{R}$ *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

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

que são

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

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

(iii) $f\colon \mathbb{R}^2 \to \mathbb{R}$ *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 arbitráriamente 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.*

Por vezes há formas de obter informação a respeito da matriz Hessiana sem calcular valores próprios, o que é útil na prática. A proposição seguinte apresenta alguns critérios.

Proposição 15.8. *Seja* $U \subset \mathbb{R}^n$ *um aberto e* $f\colon U \to \mathbb{R}$ *uma função de classe* C^3. *Seja* a *um ponto de estacionariedade de* f *e* $q(h) = h^T Hf(a) h$ *a forma quadrática definida pela matriz Hessiana de* f *no ponto* a. *Então*

(i) Se q *for definida positiva tem-se* $\frac{\partial^2 f}{\partial x_i^2} > 0$ *para qualquer* $i = 1, \ldots, n$.

(ii) Se q *for definida negativa tem-se* $\frac{\partial^2 f}{\partial x_i^2} < 0$ *para qualquer* $i = 1, \ldots, n$.

(iii) *Se existirem i e j tais que $\frac{\partial^2 f}{\partial x_i^2} < 0$ e $\frac{\partial^2 f}{\partial x_j^2} > 0$ a forma quadrática q é indefinida e consequentemente a é um ponto de sela.*

(iv) *Se a é um ponto de mínimo, tem-se $\frac{\partial^2 f}{\partial x_i^2} \geq 0$ para qualquer $i = 1, \ldots, n$.*

(v) *Se a é um ponto de máximo, tem-se $\frac{\partial^2 f}{\partial x_i^2} \leq 0$ para qualquer $i = 1, \ldots, n$.*

Proof. Todas as alíneas resultam da Proposição 15.5 ao observarmos que, sendo e_i o i-ésimo vetor da base canónica de $\mathbb{R}^n$, se tem

$$q(e_i) = e_i^T Hf(a) e_i = \frac{\partial^2 f}{\partial x_i^2} \, .$$

□

Nota 15.9. *Um critério computacionalmente eficiente para determinar se q é definida positiva resulta de aplicar o método da eliminação de Gauss: q é definida positiva se e só se todos os pivots de $Hf(a)$ forem positivos. Para outros critérios deste tipo ver o livro "Álgebra Linear como Introdução à Matemática Aplicada" de L. Magalhães, 1992, Texto Editora.*

16. Teorema da Função Inversa

Vamos agora explicar um dos teoremas fundamentais do Cálculo Diferencial em $\mathbb{R}^n$, 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 $f\colon X \to Y$, 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 $g\colon f(X) \to X$ (onde $f(X) = \{f(x)\colon x \in X\} \subset Y$ é a imagem de f) tal que

(i) $g \circ f = \mathrm{Id}_X$, isto é, $g(f(x)) = x$ para todo o $x \in X$,

(ii) $f \circ g = \mathrm{Id}_{f(X)}$, isto é, $f(g(y)) = y$ para todo o $y \in f(X)$.

Além disso a *função inversa* $g\colon f(X) \to X$, que se denota habitualmente por f^{-1}, é dada por

$$f^{-1}(y) = \text{único } x \text{ tal que } f(x) = y \, .$$

Exemplo 16.1. *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 16.2 (Teorema da Função Inversa)**.** *Seja $U \subset \mathbb{R}^n$ um aberto, x um ponto de U, e $f\colon U \to \mathbb{R}^n$ uma função de classe C^1. Se $\det Df(x) \neq 0$, então existe um aberto V contendo x tal que*

(i) *A restrição de f a V é injectiva;*

(ii) *$f(V)$ é aberto;*

(iii) *A função inversa $f^{-1}\colon f(V) \to V$ 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 16.3. *Recorde-se de Álgebra Linear, que a condição* $\det Df(x) \neq 0$ *é equivalente a dizer que a derivada* $Df(x)\colon \mathbb{R}^n \to \mathbb{R}^n$ *é invertível. Assim, em termos gerais, o Teorema 16.2 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 subtileza 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 16.4. *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*

$$g \circ f = \mathrm{Id}_V \Leftrightarrow g(f(x)) = x \text{ para todo o } x \in V.$$

Tendo em conta que a derivada da identidade é a identidade (em todos os pontos) temos, pela regra de derivação da função composta:

$$Dg(f(x))Df(x) = \mathrm{Id},$$

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 16.2 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 vejam a demonstração em [S].

Se $f\colon I \to \mathbb{R}$ 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 16.5 (A exponencial complexa)**.** *Seja* $f\colon \mathbb{R}^2 \to \mathbb{R}^2$ *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* $f(0, 0) = f(0, 2\pi) = (1, 0)$*.*

É fácil entender geometricamente o efeito da transformação f*: cada reta horizontal com ordenada* y *é enviada na semi-reta 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 16.6. *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.*

Exemplo 16.7. *Mostrar que a função* $f\colon \mathbb{R}^2 \to \mathbb{R}^2$ *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*

$$Df^{-1}(2, 5) = Df(1, 2)^{-1} = \begin{bmatrix} 2 & 1 \\ 1 & 4 \end{bmatrix}^{-1} = \frac{1}{7} \begin{bmatrix} 4 & -1 \\ -1 & 2 \end{bmatrix}.$$

17. 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). Comecemos 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

$$e^y + xy - 2 = 0 \tag{6}$$

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 (6) acima temos

$$\frac{\partial}{\partial x}(e^y + xy - 2) = y, \qquad \frac{\partial}{\partial y}(e^y + xy - 2) = e^y + x,$$

logo a equação linear que melhor a aproxima perto de $(0, \log 2)$ é

$$\log(2)x + 2(y - \log 2) = 0, \tag{7}$$

que pode ser resolvida em ordem a y. O Teorema da Função Implícita garante que a equação (6) 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 (6) e chama-se a *função implícita*.

Teorema 17.1 (Teorema da função implícita para funções de $\mathbb{R}^2$ para $\mathbb{R}$)**.** *Seja $U \subset \mathbb{R}^2$ um conjunto aberto, $f\colon U \to \mathbb{R}$ 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 $g\colon V \to W$ de classe C^1 tais que as condições seguintes são equivalentes:*

- $(x, y) \in V \times W$ *e* $f(x, y) = 0$*;*

- $y = g(x)$.

Nota 17.2. *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 reta 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:

$$\{(x, y) \in V \times W \colon f(x, y) = 0\} = \{(x, g(x)) \colon x \in V\}.$$

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 vetor 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 17.1 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 $g'(x)$:

$$\frac{d}{dx}\left(f(x, g(x))\right) = 0 \Leftrightarrow \frac{\partial f}{\partial x}(x, g(x)) + \frac{\partial f}{\partial y}(x, g(x))g'(x) = 0 \Leftrightarrow g'(x) = -\frac{\frac{\partial f}{\partial x}(x, g(x))}{\frac{\partial f}{\partial y}(x, g(x))}.$$

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 17.3. *Voltando ao exemplo (6), 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

$$g'(0) = -\frac{\log 2}{2}.$$

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 da Função Implícita.

Exemplo 17.4. *Seja $F\colon \mathbb{R}^2 \to \mathbb{R}$ 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 a y explicitamente:

$$x^2 + y^2 = 1 \Leftrightarrow y = \pm\sqrt{1 - x^2}.$$

Vemos então 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 17.1.

Este exemplo é esclarecedor se não nos lembrarmos qual é a derivada parcial que não se deve anular no enunciado do Teorema 17.1.

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.

Vejamos 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 vetor 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 17.5 (Teorema da Função Implícita)**.** *Seja $U \subset \mathbb{R}^n$ um aberto e $F\colon U \to \mathbb{R}^m$ uma função de classe C^1. Escrevemos os vetores de $\mathbb{R}^n$ na forma (x, y) com x um vetor de $\mathbb{R}^{n-m}$ formado por quaisquer $(n - m)$ coordenadas (não necessariamente as primeiras) e y o vetor 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 $g\colon V \to W$ de classe C^1 tais que $V \times W \subset U$ e as condições seguintes são equivalentes:

- $(x, y) \in V \times W$ *e* $F(x, y) = 0$*;*
- $y = g(x)$*.*

Nota 17.6. *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*

$$\{(x, g(x))\colon x \in V\}$$

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 17.7. *A condição* $\det \frac{\partial F}{\partial y} \neq 0$ *no Teorema 17.5 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 17.8. *As derivadas da função implícita $g(x)$ cuja existência o Teorema 17.5 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 17.9. *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 $F\colon \mathbb{R}^3 \to \mathbb{R}^2$ a função definida por

$$F(x, y, z) = (x^2 - y^2 + z^2 - 1, 3x + y^2 + z - 3).$$

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 17.5 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

$$\begin{cases} 2g_1'(-1) - 2g_2'(-1) = 2 \\ 3g_1'(-1) + 2g_2'(-1) = -1 \end{cases} \Leftrightarrow \begin{cases} g_1'(-1) = \frac{1}{5} \\ g_2'(-1) = -\frac{4}{5} \end{cases}$$

Dem. do Teorema 17.5. Vamos demonstrar o Teorema 17.5 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 17.5, seja $f \colon U \to \mathbb{R}^n$ a função definida por

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

A matriz Jacobiana de f pode ser escrita como uma matriz por blocos

$$Df(x, y) = \begin{bmatrix} \mathrm{Id} & 0 \\ \frac{\partial F}{\partial x} & \frac{\partial F}{\partial y} \end{bmatrix},$$

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 $f^{-1}(a,b) = (a, h(a,b))$ para alguma função h de classe C^1, logo a última equação, $(x,y) = f^{-1}(x,0)$, é equivalente a

$$\begin{cases} x = a \\ y = h(a,0) \end{cases} \quad \Leftrightarrow \quad y = h(x,0).$$

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 17.5 se pode tomar qualquer aberto contendo y_0 tal que $\{x_0\} \times \overline{W} \subset A$ e que então $V = g^{-1}(W)$. □

Nota 17.10. *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.*

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

(8)
$$\frac{\partial F}{\partial x}(x,g(x)) \cdot \mathrm{Id} + \frac{\partial F}{\partial y}(x,g(x)) Dg(x) = 0 \Leftrightarrow Dg(x) = -\left(\frac{\partial F}{\partial y}(x,g(x))\right)^{-1} \frac{\partial F}{\partial x}(x,g(x)).$$

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:

$$\frac{\partial y}{\partial x_i} = \frac{\partial g}{\partial x_i} = -\frac{\frac{\partial F}{\partial x_i}(x,g(x))}{\frac{\partial F}{\partial y}(x,g(x))}. \tag{9}$$

Exemplo 17.11. *Determinar se o subconjunto de $\mathbb{R}^3$ definido pelas condições*

$$\begin{cases} x + yz = 1 \\ x^3 z + x - y = 0 \end{cases}$$

é o gráfico de uma função de uma variável numa vizinhança do ponto $(1,1,0)$.

A função $F(x,y,z) = (x + yz - 1, x^3 z + x - y = 0)$ é de classe C^1 e $F(1,1,0) = (0,0)$. Temos

$$DF(x,y,z) = \begin{bmatrix} 1 & z & y \\ 3x^2 z + 1 & -1 & x^3 \end{bmatrix},$$

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* (8) *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 17.12. *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* (9) *temos*

$$\frac{\partial y}{\partial z} = \frac{\partial g}{\partial z} = -\frac{\frac{\partial F}{\partial z}}{\frac{\partial F}{\partial y}} = -\frac{ye^{yz}}{x + ze^{yz}}.$$

Portanto (recordando que y *é uma função de* x *e* z*),*

$$\frac{\partial^2 y}{\partial x \partial z} = \frac{\partial}{\partial x}\left(-\frac{ye^{yz}}{x + ze^{yz}}\right) = -\frac{\left(\frac{\partial y}{\partial x}e^{yz} + yz\frac{\partial y}{\partial x}e^{yz}\right)(x + ze^{yz}) - ye^{yz}\left(1 + z^2\frac{\partial y}{\partial x}e^{yz}\right)}{(x + ze^{yz})^2}.$$

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 17.13. *O Teorema 17.5 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 17.5 se verificam, é uma consequência da fórmula* (8) *que a função* g *é também de classe* C^k*. Só é necessário*

verificar que a expressões que se obtêm derivando sucessivamente a fórmula (como no exemplo anterior) são contínuas. A demonstração detalhada deixa-se como exercício.

Vimos na secção 12 que o gradiente de um campo escalar diferenciável $f : \mathbb{R}^n \to \mathbb{R}$ é perpendicular aos seus conjuntos de nível, no sentido em que $\nabla f(a)$ é perpendicular a todos os vetores tangentes a $f^{-1}(c)$ no ponto a, onde $c = f(a)$. Recorde-se que estes vetores tangentes são simplesmente os vetores velocidade $\gamma'(0)$ de caminhos $\gamma : I \to f^{-1}(c)$ tais que $\gamma(0) = a$. Podemos colocar a questão inversa: dado um vetor v ortogonal a $\nabla f(a)$, será que ele é tangente a $f^{-1}(c)$? A resposta é afirmativa sob certas hipóteses, e é uma consequência do Teorema da Função Implícita:

Proposição 17.14. *Seja $f : \mathbb{R}^n \to \mathbb{R}$ um campo escalar de classe C^1. Se $\nabla f(a) \neq 0$ então qualquer vetor ortogonal a $\nabla f(a)$ é tangente a $f^{-1}(c)$ no ponto a, onde $c = f(a)$*

Dem. Suponhamos sem perda de generalidade que $\frac{\partial f}{\partial x_1}(a) \neq 0$. Então existe uma função g de classe C^1 tal que as soluções da equação $f(x) = c$ são dadas por $x_1 = g(x_2, \ldots, x_n)$ numa vizinhança do ponto a. Dado um vetor $v \in \mathbb{R}^n$ ortogonal a $\nabla f(a)$, considere-se o caminho

$$\gamma(t) = (g(a_2 + v_2 t, \ldots, a_n + v_n t), a_2 + v_2 t, \ldots, a_n + v_n t).$$

Por construção, $\gamma(t) \in f^{-1}(c)$ para t suficientemente pequeno, $\gamma(0) = a$ e

$$\gamma'(0) = (\gamma_1'(0), v_2, \ldots, v_n).$$

Como $\gamma'(0)$ é necessariamente ortogonal a $\nabla f(a)$, e v também, temos

$$0 = \langle \nabla f(a), \gamma'(0) - v \rangle = \frac{\partial f}{\partial x_1}(a)\,(\gamma_1'(0) - v_1) \quad \Rightarrow \quad \gamma_1'(0) = v_1.$$

Portanto $v = \gamma'(0)$ é tangente a $f^{-1}(c)$. □

Este resultado pode ser facilmente generalizado:

Teorema 17.15. *Seja $F : \mathbb{R}^n \to \mathbb{R}^m$ uma função de classe C^1 e $c = F(a)$. Se $DF(a)$ tem característica máxima então os vetores tangentes a $F^{-1}(c)$ no ponto a são precisamente os vetores no núcleo de $DF(a)$.*

Dem. A demonstração é completamente análoga à da proposição anterior, uma vez que por um lado

$$\gamma(t) \in F^{-1}(c) \quad \Leftrightarrow \quad F(\gamma(t)) = c \quad \Rightarrow \quad DF(\gamma(0))\gamma'(0) = 0,$$

e por outro podemos supor sem perda de generalidade que

$$\det \frac{\partial F}{\partial (x_1, \ldots, x_m)}(a) \neq 0.$$

□

18. Método dos multiplicadores de Lagrange

Frequentemente é necessário optimizar quantidades sujeitas a certas restrições. Consideremos os seguintes exemplos.

Exemplo 18.1. *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 reta com equação $y = -2 + x$.

Exemplo 18.2. *Achar as dimensões que uma caixa com base retangular 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 18.3 (Método dos multiplicadores de Lagrange)**.** *Seja $U \subset \mathbb{R}^n$ um aberto e $f\colon U \to \mathbb{R}$ uma função de classe C^1. Seja $F\colon U \to \mathbb{R}^m$ uma função de classe C^1, $S = \{x \in U\colon F(x) = 0\}$ 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 $\lambda_1, \ldots, \lambda_m \in \mathbb{R}$ (chamados multiplicadores de Lagrange*) tais que a função*

$$g(x) = f(x) + \lambda_1 F_1(x) + \ldots + \lambda_m F_m(x)$$

(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 $F_1 = \ldots = F_m = 0$ 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 $x_1, \ldots, x_n$ e $\lambda_1, \ldots, \lambda_m$. Genericamente esperamos obter um conjunto finito de soluções, que são os "pontos de estacionaridade" 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 18.4. *Há um critério de classificação destes "pontos de estacionaridade" envolvendo derivadas de segunda ordem de f (e de F) mas não o iremos estudar por este ser pouco prático.*

Dem. da Proposição 18.3. Seja $\gamma\colon\ ]-\epsilon,\epsilon[\ \to S$ um caminho de classe C^1 com $\gamma(0)=x$. 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,

$$\frac{d}{dt}\left(f(\gamma(t))\right) = \langle \nabla f(\gamma(t)), \gamma'(t)\rangle$$

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 vetor tangente a S arbitrário. Pelo Teorema 17.15, os vetores tangentes a S no ponto x são precisamente os vetores no núcleo de $DF(x)$, e portanto $\nabla f(x)$ tem que pertencer ao espaço das linhas de $DF(x)$, ou seja, tem que ser uma combinação linear dos vetores $\nabla F_1(x),\ldots,\nabla F_m(x)$. Logo, têm que existir $\lambda_1,\ldots,\lambda_m\in\mathbb{R}$ tais que

$$\nabla f(x) + \lambda_1\nabla F_1(x) + \ldots + \lambda_m \nabla F_m(x) = 0$$

Isto conclui a demonstração. □

Nota 18.5. *Como observámos na demonstração anterior, a condição $\nabla g(x)=0$ na Proposição 18.3 é equivalente à condição de $\nabla f(x)$ ser ortogonal a S. 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 projeção de $\nabla f(x)$ no plano tangente a x dá uma direção em S ao longo da qual f cresce. Na direção oposta f decresce, logo f não tem um extremo em x.*

Exemplo 18.6. *Vamos aplicar a Proposição 18.3 para resolver o problema enunciado no Exemplo 18.1. 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*

$$\langle (x_0,y_0), \tfrac{1}{\sqrt{2}}(-1,1)\rangle + \sqrt{2}$$

(a primeira parcela é a projeção do vetor (x_0,y_0) na direção 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

$$f(x,y) = -\tfrac{1}{\sqrt{2}}x + \tfrac{1}{\sqrt{2}}y + \sqrt{2}$$

Vamos portanto considerar a função auxiliar

$$g(x,y) = -\tfrac{1}{\sqrt{2}}x + \tfrac{1}{\sqrt{2}}y + \sqrt{2} + \lambda(y-x^2)$$

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

$$C = \{(x, y) \in \mathbb{R}^2 \colon y = x^2, f(x, y) \leq 50\}$$

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 reta que passa por P e é ortogonal à reta de equação $y = -2 + x$ com esta última. A equação paramétrica da reta ortogonal é

$$(\tfrac{1}{2}, \tfrac{1}{4}) + t(1, -1) \Leftrightarrow \begin{cases} x = \frac{1}{2} + t \\ y = \frac{1}{4} - t \end{cases}$$

e substituindo em $y = -2 + x$ obtemos

$$\tfrac{1}{4} - t = -2 + \tfrac{1}{2} + t \Leftrightarrow 2t = \tfrac{7}{4} \Leftrightarrow t = \tfrac{7}{8},$$

pelo que

$$Q = (\tfrac{1}{2}, \tfrac{1}{4}) + \tfrac{7}{8}(1, -1) = (\tfrac{11}{8}, -\tfrac{5}{8}).$$

Nota 18.7. *O exemplo anterior podia resolver-se alternativamente minimizando a função real de variável real $f(x, x^2)$ onde*

$$f(x, y) = -\tfrac{1}{\sqrt{2}}x + \tfrac{1}{\sqrt{2}}y + \sqrt{2}$$

é a função distância à reta (e na realidade este método seria mais eficiente).

Exemplo 18.8. *Para resolver o problema enunciado no Exemplo 18.2 devemos considerar a função*

$$g(x, y, z) = xyz + \lambda(xy + 2xz + 2yz - 6)$$

e resolver o sistema

$$\text{(10)} \qquad \begin{cases} \frac{\partial g}{\partial x} = 0 \\ \frac{\partial g}{\partial y} = 0 \\ \frac{\partial g}{\partial z} = 0 \\ xy + 2xz + 2yz = 6 \end{cases} \Leftrightarrow \begin{cases} yz + \lambda(y + 2z) = 0 \\ xz + \lambda(x + 2z) = 0 \\ xy + 2\lambda(x + y) = 0 \\ xy + 2xz + 2yz = 6 \end{cases}$$

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* (10) *implica que* $xy=0$ *o que é impossível pela última equação em* (10). *Conclui-se que* $z+\lambda\neq 0$ *e portanto*

$$x=y=-\frac{2\lambda z}{z+\lambda}.$$

Substituindo na terceira equação em (10) *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 (10) *mostra que* $x=0$ *é impossível. Portanto* $x=y=-4\lambda$. *Substituindo numa das primeiras equações em* (10) *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* (10) *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

$$xy+2xz+2yz=6$$

obtém-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 18.9. *O método dos multiplicadores de Lagrange pode também ser utilizado para achar extremos de funções em conjuntos de* $\mathbb{R}^n$ *limitados por conjuntos de nível. Por exemplo, imagine-se que queremos achar os extremos de uma função* f *de classe* C^1 *no conjunto* $B=\{(x,y,z)\in\mathbb{R}^3\colon x^2+y^2+z^2\leq 1\}$. *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 18.3. 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 respetivamente.*

Os problemas de multiplicadores de Lagrange são habitualmente chamados de problemas de *extremos condicionados*, pois consistem em procurar pontos de extremo de uma função *condicionados* por determinadas restrições.

19. Definição de integral

Recorde-se de Cálculo 1 que se $I = [a, b] \subset \mathbb{R}$ é um intervalo e $f\colon I \to \mathbb{R}$ 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 $f > 0$ e negativamente quando $f < 0$. O significado do integral de uma função escalar de várias variáveis é análogo.

Definição 19.1. *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 $f\colon I \to \mathbb{R}$ 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 19.2. *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* $s\colon I \to \mathbb{R}$ *diz-se uma* função em escada *se é limitada e constante no interior de cada um dos intervalos de alguma partição de* I.

Definição 19.3. *O* volume n-dimensional *(ou simplesmente volume) do intervalo* $I = [a_1, b_1] \times \cdots \times [a_n, b_n]$ *é definido por*

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

(ou escrevemos simplesmente $\mathrm{vol}(I)$ *quando não houver ambiguidade).*

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 19.4. *Seja $I \subset \mathbb{R}^n$ um intervalo limitado e $s\colon I \to \mathbb{R}$ 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 $\operatorname{int} I_j$. Define-se o* integral de s em I *por*

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

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 $s_j \operatorname{vol}(I_j)$. 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 19.5. *Seja $s\colon [0,1] \times [0,2] \to \mathbb{R}$ 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

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

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

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

Em rigor, é necessário verificar que a Definição 19.4 faz sentido. De facto, se $s\colon I \to \mathbb{R}$ 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 19.6. *Seja $I \subset \mathbb{R}^n$ um intervalo limitado e $f\colon I \to \mathbb{R}$ uma função limitada. O* integral inferior *de f é definido por*

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

O integral superior *de f é definido por*

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

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.

20. Teorema de Fubini

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 20.1 (Teorema de Fubini - versão preliminar para funções de duas variáveis)**.** *Seja $I = [a,b] \times [c,d] \subset \mathbb{R}^2$ um intervalo e $f\colon I \to \mathbb{R}$ uma função integrável tal que para cada $x \in [a,b]$, a função $f_x\colon [c,d] \to \mathbb{R}$ 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 intersetando a região debaixo do gráfico de f com os planos verticais paralelos ao plano yz com abcissa x.

Exemplo 20.2. *Seja* $f\colon [0,1]\times[0,1] \to \mathbb{R}$ *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*

$$A(x) = x + \left[\frac{y^2}{2} \right]_{y=0}^{y=1} = x + \frac{1}{2}$$

e

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

É frequente usar a notação

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

para o integral de uma função $f\colon I \to \mathbb{R}$ definida num intervalo limitado $I \subset \mathbb{R}^n$, a que se chama um *integral múltiplo*. Assim, um integral de uma função de duas variáveis costuma escrever-se

$$\iint_I f(x,y)dxdy$$

e o de uma função de três variáveis por

$$\iiint_I f(x,y,z)dxdydz.$$

O cálculo de um tal integral reduz-se, 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 $f\colon [a,b]\times[c,d] \to \mathbb{R}$ (tal que para cada $x \in [a,b]$ a função $f_x\colon [c,d] \to \mathbb{R}$ definida por $f_x(y) = f(x,y)$ é integrável), temos

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

No integral dentro de parêntesis, 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 20.3. *Seja $f\colon \mathbb{R}^2 \to \mathbb{R}$ a função definida por $f(x,y) = x^2y$. Assumindo a integrabilidade de f, calcular*

$$\iint_{[0,1]\times[0,1]} f(x,y)dxdy.$$

Pelo Teorema de Fubini, este integral é dado por

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

Vejamos agora o enunciado geral do Teorema de Fubini.

Teorema 20.4 (Teorema de Fubini)**.** *Sejam $I \subset \mathbb{R}^n$ e $J \subset \mathbb{R}^m$ intervalos limitados e $f\colon I \times J \to \mathbb{R}$ uma função integrável. Para cada $x \in I$ seja $f_x\colon J \to \mathbb{R}$ a função definida por $f_x(y) = f(x,y)$ e para cada $y \in J$ seja $f_y\colon I \to \mathbb{R}$ a função definida por $f_y(x) = f(x,y)$. Então as funções de I para $\mathbb{R}$ definidas por*

$$x \mapsto \underline{\int}_J f_x \qquad e \qquad x \mapsto \overline{\int}_J f_x$$

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

$$y \mapsto \underline{\int}_I f_y \qquad e \qquad y \mapsto \overline{\int}_I f_y$$

são integráveis e

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

Nota 20.5. *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

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

Aplicando várias vezes o Teorema 20.4, qualquer integral múltiplo sobre um rectângulo se reduz ao cálculo iterado de integrais de funções de uma variável. Por exemplo

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

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

Definição 20.6. *Seja $A \subset \mathbb{R}^n$ um conjunto limitado, e $f\colon A \to \mathbb{R}$ uma função limitada. Define-se*

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

onde I é um intervalo limitado de $\mathbb{R}^n$ que contém A e $\tilde{f}\colon I \to \mathbb{R}$ 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.

Vejamos 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 20.7.

(1) Seja $T = \{(x,y) \in \mathbb{R}^2\colon x \geq 0, y \geq 0, x+y \leq 1\}$. Calcular $\iint_T xy\, dxdy$.

De acordo com a Definição 20.6, temos que integrar a função $\tilde{f}\colon [0,1]\times[0,1] \to \mathbb{R}$ 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 $\tilde{f}_x\colon [0,1] \to \mathbb{R}$ 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

$$\begin{aligned}\iint_T xy\ dxdy &= \int_0^1 \left(\int_0^1 \tilde{f}_x(y)dy\right) dx \\ &= \int_0^1 \left(\int_0^{1-x} xydy + \int_{1-x}^1 0\, dy\right) dx \\ &= \int_0^1 \left(\int_0^{1-x} xydy\right) dx \\ &= \int_0^1 \left[x\frac{y^2}{2}\right]_{y=0}^{y=1-x} dx \\ &= \int_0^1 \frac{1}{2}x(1-x)^2 dx = \frac{1}{24}.\end{aligned}$$

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 abcissa* x *sobre a qual temos que somar a função integranda é a fatia* $0 \leq y \leq 1-x$*).*

(2) Calcular $\iint_S xdxdy$*, onde* $S = \{(x,y) \in \mathbb{R}^2 \colon x \geq 0,\ 2x^2 \leq y \leq 1+x^2\}$.

As parábolas $y = 1+x^2$ *e* $y = 2x^2$ *intersetam-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*

$$\begin{aligned}\iint_S xdxdy &= \int_0^1 \left(\int_{2x^2}^{1+x^2} xdy\right) dx \\ &= \int_0^1 x\left(1+x^2-2x^2\right) dx = \int_0^1 (x-x^3)dx = \frac{1}{4}.\end{aligned}$$

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 diretamente 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*

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

(3) Calcular

$$\iiint_T 1 \, dxdydz,$$

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

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 definem* T *obtemos uma expressão para a fatia:*

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

O Teorema de Fubini diz que

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

onde na segunda igualdade usámos 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* $(0, 1 - z)$*).*

Os integrais múltiplos têm diversas aplicações.

Definição 20.8. *O* volume n-dimensional *de* $A \subset \mathbb{R}^n$ *é definido como* $\mathrm{vol}_n(A) = \int_A 1$. *(Escrevemos apenas* $\mathrm{vol}(A)$ *quando não houver ambiguidade.)*

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

$$\overline{f} = \frac{\int_A f}{\mathrm{vol}(A)}$$

Um exemplo particularmente útil é quando f é uma das funções coordenadas x_i. O número $\overline{x}_i$ 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 $(\overline{x}, \overline{y}, \overline{z})$, onde

$$\overline{x} = \frac{\int_A x \, dxdydz}{\int_A 1 \, dxdydz}, \quad \overline{y} = \frac{\int_A y \, dxdydz}{\int_A 1 \, dxdydz}, \quad \overline{z} = \frac{\int_A z \, dxdydz}{\int_A 1 \, dxdydz}.$$

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 $\overline{z} = 0$.

Se $f\colon A \to \mathbb{R}$ 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:

$$\text{Massa/Carga} = \int_A f \ .$$

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

$$\vec{r}_{CM} = \frac{\sum_i m_i \vec{r}_i}{\sum_i m_i}, \qquad I_L = \sum_i m_i d_i^2,$$

onde $\vec{r}_i$ é a posição da partícula i, m_i é a respetiva massa, e d_i é a distância dessa partícula ao eixo L. A massa total é o somatório $\sum_i m_i$. Portanto o centro de massa é uma média *ponderada* das posições das partículas.

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* $f\colon A \to \mathbb{R}$ (por unidade de área ou volume respetivamente). Por exemplo a coordenada x do centro de massa é dada por

$$x_{CM} = \frac{\int_A x\, f(x,y,z)\ dxdydz}{\int_A f(x,y,z)\ dxdydz}$$

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

$$I_L = \int_A f(x,y,z)\, d_L(x,y,z)^2\ dxdydz$$

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

Exemplo 20.9. *Escrever um integral triplo que calcula o volume do sólido*

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

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:

$$\begin{aligned}
\text{vol}(V) &= \iiint_T 1 dxdydz \\
&= \iint_{\{(x,y)\colon x\geq 0, y\geq 0, x+y\leq 1\}} \left(\int_0^{1+x^2+y^2} 1 dz \right) dxdy \\
&= \int_0^1 \left(\int_0^{1-x} \left(\int_0^{1+x^2+y^2} 1 dz \right) dy \right) dx.
\end{aligned}$$

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 interseta 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 a seguinte expressão para o volume da porção de V compreendida entre os planos $z = \frac{3}{2}$ e $z = 2$:

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

Analogamente, o volume da porção contida entre os planos $z = 1$ e $z = \frac{3}{2}$ é dado por:

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

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

Dem. do Teorema 20.4. 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 $f_x \colon J \to \mathbb{R}$ é ela própria uma função em escada, e é imediato verificar que as funções $I \to \mathbb{R}$ definidas por

$$x \mapsto \underline{\int_J} f_x \qquad \text{e} \qquad x \mapsto \overline{\int_J} f_x$$

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 $x \mapsto \underline{\int}_J f_x$ 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 $g\colon K \to \mathbb{R}$ é integrável sse para cada $\epsilon > 0$ existem funções em escada $s, t\colon K \to \mathbb{R}$ 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 se e só se 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 $s, t\colon I \times J \to \mathbb{R}$ 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

$$\underline{\int}_J s(x,y)dy \leq \underline{\int}_J f(x,y)dy \leq \underline{\int}_J t(x,y)dy$$

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

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

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

$$x \mapsto \underline{\int}_J f(x,y)dy$$

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. □

21. 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 21.1 (Propriedades do integral)**.** *Seja $I \subset \mathbb{R}^n$ um intervalo limitado e $f, g\colon I \to \mathbb{R}$ 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 $\left|\int_I f\right| \leq \int_I |f|$.

Sejam $A, B \subset \mathbb{R}^n$ conjuntos disjuntos e $f\colon A \cup B \to \mathbb{R}$ 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 $s_1, s_2, t_1, t_2\colon I \to \mathbb{R}$ em escada tais que $s_1 \leq f \leq t_1$, $s_2 \leq g \leq t_2$,

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

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 respetivamente.

A demonstração das restantes propriedades deixa-se como exercício. □

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 21.2. *Seja $I \subset \mathbb{R}^n$ um intervalo compacto e $f\colon I \to \mathbb{R}$ uma função contínua. Então f é integrável em I.*

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

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

e tais que

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

Isto garante a integrabilidade de f.

Para cada $n \geq 0$, seja P_n a partição de $I = [a_1, b_1] \times \cdots \times [a_n, b_n]$ 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

$$s_n(x) = \min\{f(y)\colon y \in I_j \text{ onde } I_j \text{ é um subintervalo contendo } x\},$$

e, analogamente,

$$t_n(x) = \max\{f(y)\colon y \in I_j \text{ onde } I_j \text{ é um subintervalo contendo } x\}.$$

Assim, no interior do subintervalo I_j da partição P_n, as funções s_n (respetivamente t_n) tomam como valores o mínimo (respetivamente 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

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

Vamos agora ver que

$$\lim_{n\to\infty} \max\{t_n(x) - s_n(x)\colon x \in I\} = 0 \tag{11}$$

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

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

Suponhamos que (11) é falso. Então existe $l > 0$, uma sucessão crescente de naturais $n_1 \leq n_2 \leq \ldots \leq n_k \leq \ldots$ 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

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

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. □

O critério de integrabilidade dado pelo Teorema 21.2 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 estendendo 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 21.2. 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 21.3. *Um conjunto $A \subset \mathbb{R}^n$ tem* conteúdo nulo *se para todo o $\epsilon > 0$ existe uma família finita $I_1, \ldots, I_N$ de intervalos limitados de $\mathbb{R}^n$ tal que*

(i) $A \subset I_1 \cup \ldots \cup I_N$;
(ii) $\sum_{k=1}^{N} \operatorname{vol}(I_k) < \epsilon$.

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

Nota 21.4. *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 21.5 (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 $f\colon I \to \mathbb{R}$ é uma função integrável, então o gráfico de f, isto é o conjunto $\{(x, f(x))\colon x \in I\} \subset \mathbb{R}^{n+1}$, tem conteúdo nulo.

Dem. A demonstração das duas primeiras afirmações fica como exercício. Dado $\epsilon > 0$, sejam $s, t\colon I \to \mathbb{R}$ 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 $\cup_i \operatorname{int} I_i$ está contido em

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

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 $\sum_k \operatorname{vol}(J_k) < \frac{\epsilon}{2}$. Como f é limitada, existe $R > 0$ tal que $|f(x)| \leq R$ para todo o $x \in I$. Sendo $F = I \setminus \cup_i \operatorname{int} I_i$ 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 $\sum_k \operatorname{vol}(M_k) < \frac{\epsilon}{4R}$. Podemos então tomar $J_k = M_k \times [-R, R]$ □

Exemplo 21.6.

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

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

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

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

que são contínuas e portanto integráveis (pelo Teorema 21.2).

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 21.5(i) que se $\operatorname{int} A \neq \varnothing$, então A não tem conteúdo nulo.

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

Teorema 21.7. *Seja* $I \subset \mathbb{R}^n$ *um intervalo e* $f \colon I \to \mathbb{R}$ *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 21.2 e Proposição 21.5(iii): Dado $\epsilon > 0$, precisamos de explicar como achar funções em escada $s, t \colon I \to \mathbb{R}$ 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 interseta 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 21.2 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 $\frac{\epsilon}{2}$. □

Exemplo 21.8. *Seja* $S = \{(x, y) \in \mathbb{R}^2 \colon x^2 + y^2 \leq 1\}$*. Qualquer função contínua* $f \colon S \to \mathbb{R}$ *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* $\{(x, y) \in \mathbb{R}^2 \colon x^2 + y^2 = 1\}$ *(a fronteira de* S*) que tem conteúdo nulo pelo Exemplo 21.6.*

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

Nota 21.9. *O* critério de integrabilidade de Lebesgue *estabelece que uma função limitada* $f \colon I \to \mathbb{R}$ *(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 $\sum_{k=1}^{\infty} \mathrm{vol}(I_k) < \epsilon$. *Recomenda-se aos alunos interessados que consultem* [S] *para uma demonstração deste resultado.*

22. Mudança de variáveis

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 mudanças 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 22.1. *Sendo* $S = \{(x, y) \in \mathbb{R}^2 \colon x^2 + y^2 \leq 1\}$ *vamos calcular o integral*

$$\iint_S \sqrt{x^2 + y^2}dxdy$$

fazendo a mudança de variável definida pela expressão

$$\begin{cases} x = r\cos\theta \\ y = r\,\mathrm{sen}\,\theta \end{cases}$$

(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*, quando escrita em termos das coordenadas* (r, θ)*, é simplesmente dada por* $0 \leq r \leq 1$ *e* $0 \leq \theta < 2\pi$*. Substituindo as expressões para* x *e* y *em função de* (r, θ)*, vemos que a função integranda escrita nas novas variáveis é a função*

$$\sqrt{r^2\cos^2\theta + r^2\,\mathrm{sen}^2\,\theta} = r.$$

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

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

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!

Definição 22.2. *Uma* transformação de coordenadas *(ou mudança de coordenadas, ou mudança de variável) em $\mathbb{R}^n$ é uma aplicação $\varphi\colon U \to \mathbb{R}^n$, com U um subconjunto 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 22.3. *Veremos mais tarde que a definição anterior é equivalente à seguinte, que é mais natural mas menos prática: $\varphi\colon U \to \mathbb{R}^n$ é uma transformação de coordenadas se é de classe C^1, injectiva, $V = \varphi(U)$ é aberto e a função inversa $\varphi^{-1}\colon V \to U$ é de classe C^1.*

Uma mudança de variável $\varphi\colon U \to V = \varphi(U)$ 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 22.4.

(1) **Coordenadas polares.** *Trata-se da transformação de coordenadas definida por $\varphi\colon\ ]0,+\infty[\times]0,2\pi[\to \mathbb{R}^2$, onde*

$$\varphi(r,\theta) = (r\cos\theta, r\,\mathrm{sen}\,\theta).$$

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

$$\begin{cases} x = r\cos\theta \\ y = r\,\mathrm{sen}\,\theta \end{cases}$$

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 vetor (x, y) faz com o semi-eixo positivo dos xx.

Claramente, φ é de classe C^1, e a interpretação geométrica de (r, θ) torna 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:

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

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

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

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

$$1 \leq (r\cos\theta)^2 + (r\,\mathrm{sen}\,\theta)^2 < 4 \Leftrightarrow 1 \leq r^2 < 4 \Leftrightarrow 1 \leq r < 2,$$

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

$$0 < r\,\mathrm{sen}\,\theta \leq r\cos\theta \Leftrightarrow 0 < \mathrm{sen}\,\theta \leq \cos\theta \Leftrightarrow 0 < \theta \leq \tfrac{\pi}{4}.$$

(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*

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

dada por

$$\varphi(r, \theta, z) = (r\cos\theta, r\,\mathrm{sen}\,\theta, z)$$

ou, em termos de coordenadas, por

$$\begin{cases} x = r\cos\theta \\ y = r\,\mathrm{sen}\,\theta \\ z = z \end{cases}.$$

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 $\mathbb{R}^3 \setminus \{(x, 0, z) \colon x \geq 0, z \in \mathbb{R}\}$. O Jacobiano da transformação φ é dado por

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

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

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

fica em coordenadas cílindricas

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

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

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

limitada pelo eixo dos zz, a parábola $z = r^2$ e a reta $z = 2 - r$. Uma vez que esta interseção 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, 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 dada pela função*

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

definida por

$$\varphi(\rho, \theta, \phi) = (\rho\,\mathrm{sen}\,\phi\cos\theta, \rho\,\mathrm{sen}\,\phi\,\mathrm{sen}\,\theta, \rho\cos\phi),$$

ou, em coordenadas,

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

Geometricamente, ρ corresponde à distância do ponto (x,y,z) à origem, ϕ é o ângulo (no intervalo $]0,\pi[$) entre o vetor (x,y,z) e o semi-eixo positivo dos zz, e θ é o ângulo da projeção do vetor (x,y,z) no plano xy com o semieixo 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, bem como o facto de que a sua imagem consiste no conjunto $\mathbb{R}^3 \setminus \{(x,0,z)\colon x \geq 0, z \in \mathbb{R}\}$. Quanto ao Jacobiano, temos

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

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

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

transforma-se no intervalo

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

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 22.5. *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 22.6 (Teorema de mudança de variáveis)**.** *Sejam U, V abertos limitados de $\mathbb{R}^n$ e $\varphi\colon U \to V$ uma transformação de coordenadas. Seja $A \subset V$ um conjunto e $f\colon A \to \mathbb{R}$ 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

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

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
- $\left| \det \frac{\partial(y_1, \ldots, y_n)}{\partial(x_1, \ldots, x_n)} \right| dx$ é dy escrito em termos das variáveis x.

A demonstração do Teorema 22.6 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.

Nota 22.7. *Um exercício interessante consiste em verificar que a transformação de coordenadas esféricas pode ser obtida compondo duas transformações de coordenadas cilíndricas. Com efeito, seja φ a transformação de coordenadas cilíndricas vista no Exemplo 22.4:*

$$\varphi(r, \theta, z) = (r \cos\theta, r \operatorname{sen}\theta, z) \ ,$$

em que $(r, \theta, z) \in]0, +\infty[\times [0, 2\pi[\times \mathbb{R}$. Podemos agora usar uma nova transformação de coordenadas cilíndricas para substituir z e r por $\rho = \sqrt{z^2 + r^2}$ e pelo ângulo ϕ que é formado pelo vetor (z, r) com o eixo das abcissas no plano zr:

$$\begin{cases} z &= \rho \cos\phi \\ r &= \rho \operatorname{sen}\phi \\ \theta &= \theta \end{cases}$$

Então deverá ter-se $\rho \in]0, +\infty[$ e, uma vez que $r > 0$, ter-se-á $\phi \in]0, \pi[$. A variável θ mantém o domínio $]0, 2\pi[$. Assim a equação $\psi(\rho, \theta, \phi) = (r, \theta, z)$ define uma transformação de coordenadas

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

cujo Jacobiano satisfaz a condição $|\det D(\rho, \theta, \phi)\psi| = \rho$. A transformação de coordenadas esféricas vista no Exemplo 22.4 coincide com a composição $\psi \circ \varphi$. O factor de conversão de áreas a usar quando se utilizam as coordenadas esféricas num integral é, por isso,

$$\begin{aligned} |\det D(\psi \circ \varphi)(\rho, \theta, \phi)| &= \left|\det\big(D\psi(r, \theta, z) D\varphi(\rho, \theta, \phi)\big)\right| \\ &= |\det D\psi(r, \theta, z) \det D\varphi(\rho, \theta, \phi)| \\ &= r\rho = \rho^2 \operatorname{sen}\phi \ . \end{aligned}$$

Vejamos alguns exemplos de aplicação do Teorema 22.6

Exemplo 22.8. *Calcular o momento de inércia em relação ao eixo dos xx da placa*

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

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

$$I_x = \iint_S y^2 f(x, y)\ dxdy = \iint_S y^2 dxdy\,.$$

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

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

Exemplo 22.9. *Escrever uma expressão para*

$$\iint_S f(x, y) dxdy$$

em coordenadas polares, onde

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

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 reta $x = 1$ 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

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

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

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

Exemplo 22.10. *Calcular o centróide do sólido*

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

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 cilindricas em torno do eixo dos yy:

$$\begin{cases} x = r\cos\theta \\ y = y \\ z = r\,\mathrm{sen}\,\theta \end{cases}\,.$$

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 semireta $y = \sqrt{3}r$ e o arco de circunferência $y = \sqrt{1-r^2}$. A trigonometria elementar mostra que estas linhas se intersetam 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 $\overline{x} = \overline{z} = 0$. Resta-nos portanto calcular

$$\overline{y} = \frac{\iiint_V y dx dy dz}{\iiint_V 1 dx dy dz}$$

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

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

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

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

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

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

Portanto nestas coordenadas esféricas, V é descrito por

$$0 \leq \rho \leq 1, \qquad 0 \leq \phi \leq \tfrac{\pi}{6}$$

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

$$\begin{aligned} \iiint_V y \; dxdydz &= \int_0^{2\pi}\int_0^1\int_0^{\frac{\pi}{6}} \rho\cos\phi\rho^2 \operatorname{sen}\phi d\phi d\rho d\theta \\ &= 2\pi \int_0^1 \rho^3 d\rho \int_0^{\frac{\pi}{6}} \tfrac{1}{2}\operatorname{sen} 2\phi d\phi \\ &= 2\pi \cdot \frac{1}{4} \cdot \frac{1-\frac{1}{2}}{4} = \frac{\pi}{16}. \end{aligned}$$

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

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

Exemplo 22.11. *Sendo $S = \{(x,y) \in \mathbb{R}^2 \colon 1 \leq xy \leq 2, x \leq y \leq 4x, x > 0, y > 0\}$, calcular*

$$\iint_S \operatorname{sen}(xy) dxdy$$

usando uma mudança de coordenadas apropriada.

A região de integração pode escrever-se na forma

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

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

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

pelo que $\varphi^{-1}(u,v) = \left(\sqrt{\frac{u}{v}}, \sqrt{uv}\right)$. 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

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

pelo que o Teorema 22.6 implica que

$$\iint_S \operatorname{sen}(xy)dxdy = \int_1^4 \left(\int_1^2 \operatorname{sen}(u) \left| \det \frac{\partial(x,y)}{\partial(u,v)} \right| du \right) dv.$$

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,

$$\left| \det \frac{\partial(x,y)}{\partial(u,v)} \right| = \frac{1}{\left| \det \frac{\partial(u,v)}{\partial(x,y)} \right|}.$$

Vimos já que

$$\left| \det \frac{\partial(u,v)}{\partial(x,y)} \right| = \frac{2y}{x} = 2v,$$

logo

$$\iint_S \operatorname{sen}(xy)dxdy = \int_1^4 \left(\int_1^2 \operatorname{sen}(u) \frac{1}{2v} du \right) dv = (\cos 2 - \cos 1) \log 2.$$

Vamos agora dar uma breve explicação da fórmula de mudança de variáveis (Teorema 22.6). 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 22.12. *O significado preciso do parágrafo anterior é o seguinte. Se escrevermos*

$$[v_1, \ldots, v_n] = \{t_1 v_1 + \ldots + t_n v_n \in \mathbb{R}^n \colon t_1, \ldots, t_n \in [0,1]\}$$

para o paralelepípedo n-dimensional cujas arestas são os vetores $v_1, \ldots, v_n \in \mathbb{R}^n$*, esperaríamos que qualquer noção razoável de volume n-dimensional tivesse as seguintes propriedades:*

(i) Sendo $\{e_1, \ldots, e_n\}$ *a base canónica de* $\mathbb{R}^n$*,* $\operatorname{vol}([e_1, \ldots, e_n]) = 1$*;*
(ii) $\operatorname{vol}([v_1, \ldots, v_i, \ldots, v_j, \ldots, v_n]) = \operatorname{vol}([v_1, \ldots, v_j, \ldots, v_i, \ldots, v_n])$*;*
(iii) Dado $\alpha \in \mathbb{R}$*, tem-se* $\operatorname{vol}([\alpha v_1, \ldots, v_n]) = |\alpha| \operatorname{vol}([v_1, \ldots, v_n])$*;*
(iv) $\operatorname{vol}([v_1, \ldots, v_i + \alpha v_j, \ldots, v_j, \ldots, v_n]) = \operatorname{vol}([v_1, \ldots, v_i, \ldots, v_j, \ldots, v_n])$*.*

(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* $\mathbb{R}^n \times \cdots \mathbb{R}^n \to \mathbb{R}_0^+$ *que a um n-tuplo* $(v_1, \ldots, v_n)$ *de vetores 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 $\varphi \colon U \to V$ uma transformação de coordenadas em $\mathbb{R}^n$, $A \subset V$ e $f \colon A \to \mathbb{R}$ uma função contínua. Suponhamos que podemos aproximar o conjunto $\varphi^{-1}(A)$ por uma união de intervalos $I_1, \ldots, I_n$ 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

$$\sum_{i=1}^{n} \operatorname{vol}(\varphi(I_i)) f(y_i), \tag{12}$$

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

$$\operatorname{vol}(L_{x_i}(I_i)) = |\det D\varphi(x_i)| \operatorname{vol}(I_i)$$

Obtemos assim a segunte aproximação para (12):

$$\sum_{i=1}^{n} \operatorname{vol}(I_i) |\det D\varphi(x_i)| f(\varphi(x_i)).$$

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 22.6. Novamente sugere-se aos alunos interessados que consultem [S] para uma demonstração do Teorema.

Para terminar a nossa discussão do Teorema de Mudança de Variáveis vamos fazer um cálculo útil e algo surpreendente.

Exemplo 22.13 (Cálculo do volume da bola de raio 1 em $\mathbb{R}^n$)**.** *Seja*

$$B_n = \{(x_1, \ldots, x_n) \in \mathbb{R}^n \colon x_1^2 + \ldots + x_n^2 \leq 1\}$$

a bola de raio 1 *em* $\mathbb{R}^n$*. Por definição o volume n-dimensional é dado por*

$$\mathrm{vol}(B_n) = \int_{B_n} 1.$$

A projeção de B_n *no plano* x_1x_2 *é o disco unitário* $\{(x_1, x_2) \in \mathbb{R}^2 \colon x_1^2 + x_2^2 \leq 1\}$*, e o conjunto dos pontos de* B_n *que têm primeiras duas coordenadas* (x_1, x_2) *neste disco é*

$$\{(x_1, x_2, x_3, \ldots, x_n) \in \mathbb{R}^n \colon x_3^2 + \ldots + x_n^2 \leq 1 - x_1^2 - x_2^2\}.$$

Logo, o Teorema de Fubini garante que

$$\int_{B_n} 1 = \int_{\{(x_1,x_2)\in\mathbb{R}^2 \colon x_1^2+x_2^2\leq 1\}} \left(\int_{(x_3,\ldots,x_n)\in\mathbb{R}^{n-2} \colon x_3^2+\ldots+x_n^2\leq 1-x_1^2-x_2^2\}} 1 dx_3 \cdots dx_n \right) dx_1 dx_2$$

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 dado* $\lambda \in \mathbb{R}^+$ *e* $A \subset \mathbb{R}^k$ *se definir* $\lambda A = \{\lambda x \colon x \in A\}$ *então* $\mathrm{vol}(\lambda A) = \lambda^k \mathrm{vol}(A)$*. Daqui se conclui que*

$$\mathrm{vol}(B_n) = \int_{\{(x_1,x_2)\in\mathbb{R}^2 \colon x_1^2+x_2^2\leq 1\}} \mathrm{vol}(B_{n-2}) \left(1 - x_1^2 - x_2^2\right)^{\frac{n}{2}-1} dx_1 dx_2.$$

Mudando para coordenadas polares

$$\begin{cases} x_1 = r\cos\theta \\ x_2 = r\,\mathrm{sen}\,\theta \end{cases},$$

este integral transforma-se em

$$\mathrm{vol}(B_n) = \int_0^{2\pi} \int_0^1 \mathrm{vol}(B_{n-2}) r(1-r^2)^{\frac{n}{2}-1} dr d\theta = 2\pi\,\mathrm{vol}(B_{n-2}) \left[\frac{(1-r^2)^{\frac{n}{2}}}{-2\left(\frac{n}{2}\right)} \right]_{r=0}^{r=1} = \frac{2\pi}{n}\,\mathrm{vol}(B_{n-2}).$$

A fórmula anterior permite calcular os volumes de B_n *por recorrência a partir dos volumes* $\mathrm{vol}(B_1) = 2$ *e* $\mathrm{vol}(B_2) = \pi$*. Para os primeiros valores de n obtemos*

n	1	2	3	4	5	6
$\mathrm{vol}(B_n)$	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:

$$\mathrm{vol}(B_n) = \begin{cases} \frac{\pi^{n/2}}{(n/2)!} & \text{se } n \text{ é par,} \\ \\ \frac{2^{(n+1)/2}\pi^{(n-1)/2}}{n!!} & \text{se } n \text{ é ímpar,} \end{cases}$$

onde $n!! = n(n-2)\cdots 3 \cdot 1$. *Note-se que*

$$\lim_{n\to\infty} \mathrm{vol}(B_n) = 0.$$

Isto quer dizer, 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* $\sqrt{n}$.

23. 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 23.1. *Seja* $A \subset \mathbb{R}^n$ *um compacto tal que* ∂A *tem conteúdo nulo, e* $U \subset \mathbb{R}^m$ *um aberto. Seja* $f\colon A\times U \to \mathbb{R}$ *uma função contínua e* $\varphi\colon U \to \mathbb{R}$ *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 21.7 garante que as funções contínuas são integráveis em A.

Exemplo 23.2. *Seja* $\varphi\colon \mathbb{R}\to\mathbb{R}$ *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 23.3. *Seja $\varphi\colon \mathbb{R} \to \mathbb{R}$ a função definida por*

$$\varphi(t) = \iint_{\{(x,y)\in\mathbb{R}^2\,:\,x^2+y^2\leq t^2\}} e^{t(x^2+y^2)}dxdy$$

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 diretamente 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,

$$\varphi'(t) = \frac{\partial\psi}{\partial u}(t,t) + \frac{\partial\psi}{\partial v}(t,t). \tag{13}$$

A derivada parcial de ψ em ordem a u existe pelo Teorema Fundamental do Cálculo:

$$\frac{\partial\psi}{\partial u}(u,v) = 2\pi u e^{vu^2}.$$

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 (13)*,*

$$\varphi'(t) = 2\pi t e^{t^3} + 2\pi \int_0^t r^3 e^{tr^2} dr.$$

A demonstração da regra de Leibniz é uma consequência simples do seguinte resultado fundamental que está também na base da demonstração dos Teoremas 21.2 e 21.7.

Teorema 23.4 (Teorema de Heine-Cantor)**.** *Seja $A \subset \mathbb{R}^n$ compacto e $f\colon A \to \mathbb{R}^m$ contínua. Então f é* uniformemente contínua em A*, isto é, para todo o $\epsilon > 0$ 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 o $\delta > 0$ 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_i} 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_i} converge também para a. Mas f é contínua em a e portanto temos

$$\lim_{i\to\infty} \|f(x_{k_i}) - f(y_{k_i})\| = \|f(a) - f(a)\| = 0 \ ,$$

o que contradiz que a desigualdade $\|f(x_{k_i}) - f(y_{k_i})\| \geq \epsilon$ seja válida para todo o i. □

Nota 23.5. *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 $f\colon \mathbb{R} \to \mathbb{R}$ 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 23.1.

(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

$$|\varphi(t) - \varphi(s)| = \left| \int_A f(x,t)dx - \int_A f(x,s)dx \right| \leq \int_A |f(x,t) - f(x,s)|dx \leq \int_A \epsilon = \epsilon \,\mathrm{vol}(A),$$

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 vetor da base canónica de $\mathbb{R}^m$,

$$\begin{aligned} \frac{\partial \varphi}{\partial t_i}(t) &= \lim_{h\to 0} \frac{\varphi(t + he_i) - \varphi(t)}{h} \\ &= \lim_{h\to 0} \int_A \frac{f(x, t + he_i) - f(x,t)}{h} dx. \end{aligned}$$

Para provar a regra de Leibniz precisamos de comparar o integral 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

$$\left|\int_A \frac{f(x,t+he_i)-f(x,t)}{h}dx-\int_A \frac{\partial f}{\partial t_i}(x,t)dx\right|\leq \epsilon\,\mathrm{vol}(A),$$

e portanto

$$\lim_{h\to 0}\int_A \frac{f(x,t+he_i)-f(x,t)}{h}dx=\int_A \frac{\partial f}{\partial t_i}(x,t)dx,$$

conforme queríamos demonstrar.

□

24. Integrais de linha de campos escalares

Definição 24.1. *Uma* curva $C\subset\mathbb{R}^n$ *é a imagem de um caminho, ou seja, de uma função contínua* $g: I\to\mathbb{R}^n$*, onde* $I\subset\mathbb{R}$ *é um intervalo. Se o caminho é injetivo então diz-se uma* parametrização *de* C*. A curva diz-se* regular simples *se admite uma parametrização de classe* C^1 *cuja derivada não se anula.*

Definição 24.2. *Seja* $C\subset\mathbb{R}^n$ *uma curva regular simples parametrizada por caminho de classe* C^1 $g\colon[t_0,t_1]\to\mathbb{R}^n$*. Define-se*

$$\int_C f=\int_{t_0}^{t_1} f(g(t))\|g'(t)\|dt\,. \tag{14}$$

Nota 24.3. *Note-se que o integral só está definido para curvas regulares simples compactas. Esta definição pode ser estendida a curvas* seccionalmente regulares *compactas, ou seja, curvas que podem ser decompostas num número finito de curvas regulares simples compactas. É frequente usar-se a notação*

$$\int_C f\ ds$$

para o integral de um campo escalar numa curva (ds representa o "comprimento de arco").

O integral de campos escalares em curvas tem todas as aplicações habituais dos integrais. Pode ser usado para calcular comprimentos, 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 comprimento de uma curva C é

$$l(C) = \int_C 1\,,$$

a coordenada i do centróide é dada por

$$\overline{x_i} = \frac{\int_C x}{l(C)}\,,$$

o momento de inércia em torno do eixo L é dado por

$$I_L = \int_C f d_L^2\,,$$

onde f é a densidade de massa e d_L é a distância ao eixo L, etc...

Exemplo 24.4. *Calcular a massa do arco de parábola*

$$C = \{(x,y) \in \mathbb{R}^2 \colon y = x^2, 0 \leq x \leq 1\}$$

sabendo que a densidade de massa é dada por $f(x,y) = x$.

Uma parametrização para C *é dada por* $g\colon [0,1] \to \mathbb{R}^2$ *definida por* $g(x) = (x, x^2)$. *Temos* $g'(x) = (1, 2x)$ *e* $\|g'(x)\| = \sqrt{1+4x^2}$.

Logo, a massa é

$$\int_C x = \int_0^1 x\sqrt{1+4x^2}dx = \left[\frac{(1+4x^2)^{\frac{3}{2}}}{8\cdot\frac{3}{2}}\right]_0^1 = \frac{5\sqrt{5}-1}{12}\,.$$

Exemplo 24.5. *Calcular o comprimento da curva parametrizada por* $g(\theta) = (\cos\theta, \operatorname{sen}\theta, \theta)$ *com* $0 \leq \theta \leq 4\pi$.

Temos $g'(\theta) = (-\operatorname{sen}\theta, \cos\theta, 1)$ *e* $\|g'(\theta)\| = \sqrt{\cos^2\theta + \operatorname{sen}^2\theta + 1} = \sqrt{2}$. *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* $2\pi\sqrt{2}$.

Nota 24.6. *O comprimento de uma linha* C *é dado por* $\int_C 1$. *Sendo* $g\colon [t_0, t_1] \to C$ *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.*

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 (14) dependesse da parametrização g não estaríamos verdadeiramente a definir um valor associado apenas a C e f.

Proposição 24.7 (Propriedades do integral de campos escalares em curvas)**.** *Seja $C \subset \mathbb{R}^n$ uma curva regular simples compacta, e $f, k \colon C \to \mathbb{R}$ funções. Então:*

(i) (Linearidade) Para cada $\alpha, \beta \in \mathbb{R}$, temos $\int_C \alpha f + \beta k = \alpha \int_C f + \beta \int_C k$.
(ii) (Monotonia) Se $f \leq k$, então $\int_C f \leq \int_C k$.
(iii) (Desigualdade triangular) $\left|\int_C f\right| \leq \int_C |f|$.
(iv) (Aditividade) Se $C = C_1 \cup C_2$ e $C_1 \cap C_2$ é finito então $\int_C f = \int_{C_1} f + \int_{C_2} f$.
(v) (Invariância por reparametrização) Se $g : [t_0, t_1] \to \mathbb{R}^n$ e $h \colon [s_0, s_1] \to \mathbb{R}^n$ são duas parametrizações de classe C^1 de C, então

$$\int_{t_0}^{t_1} f(g(t))\|g'(t)\|dt = \int_{s_0}^{s_1} f(h(s))\|h'(s)\|ds\,.$$

Dem. da Proposição 24.7(v). As parametrizações g e h são funções injectivas. A função que "traduz" do parâmetro t para o parâmetro s é a função $\phi \colon [t_0, t_1] \to [s_0, s_1]$ definida por

$$s = \phi(t) = h^{-1}(g(t))$$

Claramente ϕ é uma funcão 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 numa curva, 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 $\mathbb{R}$. Vamos adiar a justificação que ϕ é C^1 e fazer a mudança de variável $s = \phi(t)$ no integral

$$\int_{s_0}^{s_1} f(h(s))\|h'(s)\|ds\,.$$

Como $\phi^{-1}([s_0, s_1]) = g^{-1}(h([s_0, s_1])) = g^{-1}(C) = [t_0, t_1]$ e $f(h(\phi(t))) = f(h(h^{-1}(g(t)))) = f(g(t))$, obtemos

$$\int_{[t_0,t_1]} f(g(t))\|h'(\phi(t))\||\phi'(t)|dt\,. \tag{15}$$

Por outro lado, uma vez que $h \circ \phi = g$, a regra de derivação da função composta diz-nos que

$$g'(t) = h'(\phi(t))\phi'(t)\,.$$

Substituindo na equação (15) obtemos a expressão pretendida.

Resta-nos justificar que a função $\phi = h^{-1} \circ g$ é de classe C^1. Para tal notamos que dado $t^* \in [t_0, t_1]$ e definindo $s^* = \phi(t^*)$ podemos assumir sem perda de generalidade

que $h_1'(s^*) \neq 0$. A função $F(t,s) = h_1(s) - g_1(t)$ é claramente de classe C^1 e satisfaz $F(t^*, s^*) = 0$ e

$$\frac{\partial F}{\partial s}(t^*, s^*) = h_1'(s^*) \neq 0\,.$$

Desta forma, podemos escrever as soluções da equação $F(t,s) = 0$ numa vizinhança do ponto (t^*, s^*) na forma $s = \phi(t)$, onde a função ϕ é de classe C^1 e coincide necessariamente com $h^{-1} \circ g$. □

25. Integrais de linha de campos vetoriais

Um *campo vetorial* em $\mathbb{R}^n$ é uma função $\vec{F}\colon U \to \mathbb{R}^n$ com $U \subset \mathbb{R}^n$. Uma maneira de visualizar um campo vetorial (quando $n = 2$ ou 3) é desenhar para alguns pontos $x \in U$ um vetor $\vec{F}(x)$ com origem em x.

Os campos vetoriais 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 vetorial 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 vetor $\vec{F}(x)$. 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 vetor tangente à curva em x, unitário, que aponta na direção em que estamos a percorrer a curva, a componente tangente de $\vec{F}(x)$ é

$$\vec{F}(x) \cdot \vec{t}(x)$$

Seja $g\colon [t_0, t_1] \to C$ 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 vetorial.

Definição 25.1. *Seja C uma curva regular simples compacta em $\mathbb{R}^n$, e suponhamos que foi escolhido um sentido no qual C deve ser percorrida. Se $g\colon [t_0, t_1] \to C$ é uma parametrização de classe C^1, definimos*

$$\int_C \vec{F} \cdot d\vec{r} = \pm \int_{t_0}^{t_1} \vec{F}(g(t)) \cdot g'(t) dt \tag{16}$$

em que o sinal é $+$ se a parametrização percorre C no sentido pretendido e $-$ caso contrário.

Nota 25.2. *Para que a expressão no termo direito* (16) *faça sentido, não é necessário que g seja uma parametrização de uma curva regular simples. Basta que $g\colon [t_0,t_1]\to\mathbb{R}^n$ 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 curva regular simples (pode auto intersetar-se e conter pontos nos quais não há reta tangente). Define-se o* integral de linha de um campo vetorial $\vec{F}$ 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 25.3. *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 25.4. *Sendo $C=\{(x,y)\in\mathbb{R}^2\colon x^2+y^2=1\}$ 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 $g(t)=(\cos t,\operatorname{sen} t)$ com $0\le t\le 2\pi$. Temos $g'(t)=(-\operatorname{sen} t,\cos t)$. A parametrização percorre C no sentido contrário ao pretendido portanto

$$\begin{aligned}\int_C (x-y,x^2)\cdot d\vec{r} &= -\int_0^{2\pi}(\cos t-\operatorname{sen} t,\cos^2 t)\cdot(-\operatorname{sen} t,\cos t)dt\\ &= -\int_0^{2\pi} -\operatorname{sen} t\cos t+\operatorname{sen}^2 t+\cos^3 t\ dt=-\pi.\end{aligned}$$

Exemplo 25.5. *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

$$g(t)=(0,1,0)+t\,((1,2,3)-(0,1,0))=(t,1+t,3t),\qquad \text{com } 0\le t\le 1\,.$$

Temos

$$g'(t)=(1,1,3)\,,$$

e portanto

$$\int_C \vec{F}\cdot d\vec{r}=\int_0^1(2,3t,1)\cdot(1,1,3)dt=\int_0^1 2+3t+3dt=5+\tfrac{3}{2}=\tfrac{13}{2}\,.$$

O integral de linha de um campo vetorial tem as propriedades habituais de um integral.

Proposição 25.6 (Propriedades do integral de linha)**.** *Seja $U \subset \mathbb{R}^n$ um aberto, sejam $\vec{F}, \vec{G} \colon U \to \mathbb{R}^n$ campos vetoriais contínuos, e seja C a imagem de um caminho seccionalmente regular $g \colon [a,b] \to U$. Então:*

(i) (Linearidade) Dados $\alpha, \beta \in \mathbb{R}$ tem-se $\int_C \left(\alpha\vec{F} + \beta\vec{G}\right) \cdot d\vec{r} = \alpha \int_C \vec{F} \cdot d\vec{r} + \beta \int_C \vec{G} \cdot d\vec{r}$.

(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

$$\int_C \vec{F} \cdot d\vec{r} = \int_{C_1} \vec{F} \cdot d\vec{r} + \int_{C_2} \vec{F} \cdot \vec{r}$$

.

(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 $\phi \colon [c,d] \to [a,b]$ é uma função bijectiva, de classe C^1, cuja derivada não se anula, e $h \colon [c,d] \to U$ é o caminho definido por $h = g \circ \phi$, então

$$\int_a^b \vec{F}(g(t)) \cdot g'(t)dt = \pm \int_c^d F(h(s)) \cdot h'(s)ds$$

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. □

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 25.7. *Calcular o trabalho realizado por $\vec{F}(x,y) = (y + xy, x^2)$ ao longo da fronteira do triângulo $\{(x,y) \colon x, y \geq 0,\ x + y \leq 1\}$ percorrido no sentido anti-horário.*

Escrevendo $C_1 = \{(x,0) \in \mathbb{R}^2 \colon 0 \leq x \leq 1\}$, $C_2 = \{(x,y) \in \mathbb{R}^2 \colon x + y = 1, x \geq 0, y \geq 0\}$ e $C_3 = \{(0,y) \in \mathbb{R}^2 \colon 0 \leq y \leq 1\}$ 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 dado por

$$\begin{aligned}\int_C \vec{F} \cdot d\vec{r} &= \int_{C_2} \vec{F} \cdot d\vec{r} \\ &= \int_0^1 \vec{F}(1-t,t) \cdot (-1,1)dt = \int_0^1 (t + (1-t)t, (1-t)^2) \cdot (-1,1)dt \\ &= \int_0^1 -2t + t^2 + 1 - 2t + t^2 \; dt = -\frac{1}{3}\,.\end{aligned}$$

26. Campos gradientes e campos fechados

Vamos agora ver a primeira metade do Teorema Fundamental do Cálculo para integrais de linha de campos vetoriais. Embora fundamental, este resultado é uma consequência imediata da definição e da Regra de Barrow de Cálculo 1.

Teorema 26.1 (Regra de Barrow)**.** *Seja $\varphi\colon U \to \mathbb{R}$ uma função de classe C^1 e $\vec{F} = \nabla\varphi$. Se $g\colon [t_0, t_1] \to U$ 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

$$\nabla\varphi(g(t)) \cdot g'(t) = \frac{d}{dt}\left(\varphi(g(t))\right)\,.$$

Logo, o integral acima é

$$\int_{t_0}^{t_1} \frac{d}{dt}\left(\varphi(g(t))\right) dt = \varphi(g(t_1)) - \varphi(g(t_0)) = \varphi(B) - \varphi(A)\,,$$

onde na primeira igualdade usámos a regra de Barrow de Cálculo 1. □

Um campo vetorial $\vec{F}$ 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 26.1 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 26.1 diz-nos, em particular, que o integral de linha de um campo gradiente depende apenas das extremidades do caminho em questão. Um campo vetorial com esta propriedade diz-se *conservativo*.

Exemplo 26.2. *Um campo vetorial constante num aberto de $\mathbb{R}^n$ é conservativo. Consideremos por exemplo o campo vetorial $\vec{F}(x,y,z) = (1,3,-2)$. 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 secção.

Teorema 26.3. *Seja $U \subset \mathbb{R}^n$ um aberto e $\vec{F}\colon U \to \mathbb{R}^n$ um campo vetorial 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 vetorial $\vec{F}$ é conservativo, o Teorema 26.1 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 vetorial é ou não um gradiente.

Uma condição necessária para que um campo seja gradiente é-nos dada pelo Lema de Schwarz (Teorema 13.5). Por exemplo, suponhamos que $\vec{F}\colon \mathbb{R}^2 \to \mathbb{R}^2$ 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 (necessariamente de classe C^2) $\varphi\colon \mathbb{R}^2 \to \mathbb{R}$. 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 26.4. *Seja $U \subset \mathbb{R}^n$ um aberto. Um campo vetorial $\vec{F} = (F_1, \ldots, F_n)\colon U \to \mathbb{R}^n$ 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 26.5. *Seja $U \subset \mathbb{R}^n$ um aberto e $\vec{F}\colon U \to \mathbb{R}^n$ um campo vetorial 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 26.6. *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 conseguissemos 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 26.1 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 26.7. *Determinar quais dos seguintes campos vetoriais são conservativos e, em caso afirmativo, calcular um potencial escalar para o campo.*

(1) $\vec{F}\colon \mathbb{R}^3 \to \mathbb{R}^3$ *definido por* $\vec{F}(x,y,z) = (y+z, x, x+yz)$.

Vamos ver se $\vec{F}$ *é fechado:*

$$\frac{\partial F_1}{\partial y} = \frac{\partial}{\partial y}(y+z) = 1 = \frac{\partial}{\partial x}(x) = \frac{\partial F_2}{\partial x}\,;$$

$$\frac{\partial F_1}{\partial z} = \frac{\partial}{\partial z}(y+z) = 1 = \frac{\partial}{\partial x}(x+yz) = \frac{\partial F_3}{\partial x}\,;$$

$$\frac{\partial F_2}{\partial z} = \frac{\partial}{\partial z}(y+z) = 1 \neq z = \frac{\partial}{\partial y}(x+yz) = \frac{\partial F_3}{\partial y}\,.$$

Conclui-se que $\vec{F}$ *não é fechado e portanto não é conservativo.*

(2) $\vec{F}\colon \mathbb{R}^2 \to \mathbb{R}^2$ *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 se e só se é possível resolver o sistema de equações*

$$\begin{cases} \frac{\partial \varphi}{\partial x} = x^2 + y \\ \frac{\partial \varphi}{\partial y} = x - y \end{cases} \tag{17}$$

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 (17) *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,

$$A(y) = -\frac{y^2}{2} + C \qquad \text{com } C \in \mathbb{R}\,,$$

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.

Exemplo 26.8. *Mostrar que $\vec{F}\colon \mathbb{R}^3 \to \mathbb{R}^3$ 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 $g\colon [0, 2\pi] \to \mathbb{R}^3$ definido por $g(t) = (t\cos t, t\,\mathrm{sen}\, t, t)$.*

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

$$\frac{\partial}{\partial y}\left(x^2 + e^{xz} + A(y, z)\right) = z \Leftrightarrow \frac{\partial A}{\partial y} = z \Leftrightarrow A(y, z) = yz + B(z)\,.$$

Finalmente, substituindo

$$\varphi(x,y,z) = x^2 + e^{xz} + yz + B(z)$$

na última equação temos

$$\frac{\partial}{\partial z}\left(x^2 + e^{xz} + yz + B(z)\right) = y + xe^{xz} \Leftrightarrow xe^{xz} + y + B'(z) = y + xe^{xz} \Leftrightarrow B'(z) = 0\,.$$

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

$$\varphi(g(2\pi)) - \varphi(g(0)) = \varphi(2\pi,0,2\pi) - \varphi(0,0,0) = 4\pi^2 + e^{4\pi^2} - 1.$$

Exemplo 26.9 (Um campo radial é conservativo)**.** *Um* campo radial *é um campo*

$$\vec{F}\colon \mathbb{R}^n \setminus \{0\} \to \mathbb{R}^n$$

que no ponto $(x_1,\ldots,x_n)$ tem a direção da reta que une o ponto à origem (isto é a direção do vetor $(x_1,\ldots,x_n)$) e cuja magnitude depende apenas da distância à origem.

Escrevendo $\vec{e}_r$ para o versor radial *– um vetor unitário que aponta na direção oposta à da origem – temos a expressão*

$$\vec{e}_r = \frac{(x_1,\ldots,x_n)}{\|(x_1,\ldots,x_n)\|} = \left(\frac{x_1}{\sqrt{x_1^2+\ldots+x_n^2}},\ldots,\frac{x_n}{\sqrt{x_1^2+\ldots+x_n^2}}\right),$$

e escrevendo $r = \sqrt{x_1^2+\ldots+x_n^2}$ para a distância à origem, concluimos que um campo radial é um campo vetorial $\vec{F}\colon \mathbb{R}^n \setminus \{0\}$ que pode ser escrito na forma

$$\vec{F}(x_1,\ldots,x_n) = f(r)\vec{e}_r = f(r)\left(\frac{x_1}{r},\ldots,\frac{x_n}{r}\right).$$

Para verificar que um tal campo é conservativo podemos achar um potencial. Temos que resolver a equação

$$\nabla\phi = \vec{F}$$

para a incógnita $\phi\colon \mathbb{R}^n \setminus \{0\} \to \mathbb{R}$, ou seja o sistema

$$\begin{cases} \frac{\partial\phi}{\partial x_1} = f(r)\frac{x_1}{r} \\ \quad\vdots \\ \frac{\partial\phi}{\partial x_n} = f(r)\frac{x_n}{r} \end{cases}$$

Uma vez que

$$\frac{\partial r}{\partial x_i} = \frac{1}{2}\left(x_1^2+\ldots x_n^2\right)^{-\frac{1}{2}} 2x_i = \frac{x_i}{r}\,,$$

cada equação do sistema acima é da forma

$$\frac{\partial\phi}{\partial x_i} = f(r)\frac{\partial r}{\partial x_i}\,.$$

Sendo $V(r) = \int f(r)dr$ *uma primitiva da função* $f\colon]0, +\infty[\to \mathbb{R}$*, temos*

$$\frac{\partial}{\partial x_i}\left(V\left(\sqrt{x_1^2 + \ldots + x_n^2}\right)\right) = V'(r)\frac{\partial r}{\partial x_i} = f(r)\frac{x_i}{r}\,.$$

Conclui-se que

$$\phi(x_1, \ldots, x_n) = V\left(\sqrt{x_1^2 + \ldots + x_n^2}\right)$$

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 gravitação universal de Newton é dado pela expressão*

$$\vec{F}(x, y, z) = -\frac{GM}{r^2}\vec{e}_r$$

(com G *a constante de gravitação universal). Uma vez que*

$$\int -\frac{GM}{r^2}dr = \frac{GM}{r}\,,$$

temos a seguinte expressão para o potencial gravítico:

$$V(r) = \frac{GM}{r}.$$

27. Teorema Fundamental do Cálculo

Definição 27.1. *Seja* U *um subconjunto de* $\mathbb{R}^n$*. Um* caminho *em* U *é uma função contínua* $g\colon [a, b] \to U$ *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* $g\colon [a, b] \to A$ *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 27.2.

(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) $U = \mathbb{R}^2 \setminus \{(x, 0)\colon x \in \mathbb{R}\}$ *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 Valo Intermédio, 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 27.2(iii), as componentes conexas são $\{(x, y) \in \mathbb{R}^2\colon y > 0\}$ e $\{(x, y) \in \mathbb{R}^2\colon y < 0\}$.

Teorema 27.3 (Teorema Fundamental do Cálculo)**.** *Seja* $U \subset \mathbb{R}^n$ *um aberto e* $\vec{F}\colon U \to \mathbb{R}^n$ *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 $\varphi \colon U \to \mathbb{R}$ *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 qualquer de* V*, 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 $g \colon [a,b] \to U$ um caminho seccionalmente regular fechado e $A = g(a) = g(b)$. Seja $h \colon [a,b] \to \mathbb{R}^n$ o caminho constante definido por $h(t) = A$. Como $\vec{F}$ é conservativo temos
$$\int_a^b \vec{F}(g(t)) \cdot g'(t)dt = \int_a^b \vec{F}(h(t)) \cdot h'(t)dt = \int_a^b \vec{F}(h(t)) \cdot 0dt = 0\,.$$
- (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
$$\oint_C \vec{F} \cdot d\vec{r} = \int_{C_1} \vec{F} \cdot d\vec{r} - \int_{C_2} \vec{F} \cdot d\vec{r}.$$
Concluimos portanto que
$$\int_{C_1} \vec{F} \cdot d\vec{r} = \int_{C_2} \vec{F} \cdot d\vec{r},$$
ou seja, que $\vec{F}$ é conservativo.
- (iii)$\Rightarrow$(i): É uma consequência imediata da regra de Barrow (Teorema 26.1)
- (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
$$\frac{\partial\varphi}{\partial x_i}(x) = \lim_{h\to 0} \frac{\varphi(x + he_i) - \varphi(x)}{h}\,,$$

com e_i o i-ésimo vetor da base canónica de $\mathbb{R}^n$. Para h suficientemente pequeno, o segmento de reta 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 reta $[x, x + he_i]$. 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,

$$\lim_{h \to 0} \int_0^1 F_i(x + the_i) dt = F_i(x)\,,$$

ou seja,

$$\frac{\partial \varphi}{\partial x_i}(x) = F_i(x)\,,$$

o que conclui a demonstração.

□

Exemplo 27.4 (O ralo da banheira)**.** *Seja* $\vec{F}\colon \mathbb{R}^2 \setminus \{(0,0)\} \to \mathbb{R}^2$ *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 escoa através do ralo de uma banheira, pelo que é frequente refererir-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 $g(\theta) = (\cos\theta, \operatorname{sen}\theta)$ com $0 \leq \theta \leq 2\pi$, obtemos

$$\begin{aligned}\oint_C \vec{F}\cdot d\vec{r} &= \int_0^{2\pi} \vec{F}(\cos\theta,\operatorname{sen}\theta)\cdot(-\operatorname{sen}\theta,\cos\theta)d\theta \\ &= \int_0^{2\pi} \left(\frac{-\operatorname{sen}\theta}{\cos^2\theta+\operatorname{sen}^2\theta}, \frac{\cos\theta}{\cos^2\theta+\operatorname{sen}^2\theta}\right)\cdot(-\operatorname{sen}\theta,\cos\theta)d\theta \\ &= \int_0^{2\pi} 1 d\theta = 2\pi\,.\end{aligned}$$

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 27.3(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 conexa do domínio podem ser diferentes). Os potenciais que encontrámos 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).

28. Teorema de Green

O seguinte Teorema Fundamental do Cálculo para integrais de linha no plano ajudar-nos-á a perceber melhor o exemplo anterior.

Teorema 28.1 (Teorema de Green)**.** *Seja $U \subset \mathbb{R}^2$ um aberto, e $(P,Q)\colon U \to \mathbb{R}^2$ um campo vetorial de classe C^1. Seja $A \subset \mathbb{R}^2$ um aberto limitado com $A = \operatorname{int} \overline{A}$, e $\overline{A} \subset U$, tal que $\partial A = \partial A$ é uma união de curvas regulares. Então*

$$\iint_A \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} \right) dxdy = \oint_{\partial A} (P,Q) \cdot d\vec{r}$$

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 de Green 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 28.2. *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 28.3. *É frequente escrever-se*

$$\oint_{\partial A} P dx + Q dy$$

para o integral de linha

$$\oint_{\partial A} (P,Q) \cdot d\vec{r}\,.$$

Trata-se apenas de uma notação alternativa para o integral de linha de um campo vetorial.

Exemplo 28.4. *Calcular*

$$\oint_C (x+y)dx + (x^2+y^2)dy$$

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*

$$\begin{aligned}\oint_C (x+y)dx + (x^2+y^2)dy &= -\int_{-1}^{1}\int_{-1}^{1}\left(\frac{\partial}{\partial x}\left(x^2+y^2\right) - \frac{\partial}{\partial y}(x+y)\right)dxdy \\ &= -\int_{-1}^{1}\int_{-1}^{1}(2x-1)\,dxdy = -(-4) = 4\,.\end{aligned}$$

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 28.5. *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*

$$\iint_A 1\;dxdy\,.$$

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*

$$\iint_A 1\;dxdy = \int_{\partial A}(0,x)\cdot d\vec{r},$$

onde ∂A *deve ser percorrida de forma a que* A *esteja à esquerda. Aplicando a fórmula acima ao círculo* $A = \{(x,y)\in\mathbb{R}^2 \colon x^2+y^2 = 1\}$*, temos que* ∂A *é a circunferência de raio* 1 *percorrida no sentido direto. Uma parametrização que percorre* ∂A *nesse sentido é*

$$g(\theta) = (\cos\theta, \operatorname{sen}\theta), \quad 0 \le \theta \le 2\pi\,,$$

e portanto

$$\text{área}(A) = \int_0^{2\pi}(0,\cos\theta)\cdot(-\operatorname{sen}\theta,\cos\theta)d\theta = \int_0^{2\pi}\cos^2\theta d\theta\,.$$

Usando as fórmulas trigonométricas $\cos(2\theta) = \cos^2\theta - \operatorname{sen}^2\theta$ *e* $\operatorname{sen}^2\theta = 1-\cos^2\theta$ *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 portanto*

$$\int_0^{2\pi}\frac{1+\cos(2\theta)}{2}d\theta = \pi\,.$$

Exemplo 28.6. *Consideremos o campo* $\vec{F}\colon \mathbb{R}^2\setminus\{(0,0)\} \to \mathbb{R}^2$ *do exemplo 27.4, 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*

$$C = \{(x, y) \in \mathbb{R}^2 \colon \frac{x^2}{4} + \frac{y^2}{9} = 1\}$$

percorrida no sentido direto.

Não seria fácil fazer as contas diretamente parametrizando a elipse. No entanto, podemos aplicar o Teorema de Green à região

$$A = \{(x, y) \in \mathbb{R}^2 \colon x^2 + y^2 \geq 1 \text{ e } \frac{x^2}{4} + \frac{y^2}{9} \leq 1\}$$

limitada pela elipse e pela circunferência S *de raio* 1. *Note-se que* (P, Q) *é fechado se e só se*

$$\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} = 0\,.$$

Uma vez que o campo $\vec{F}$ *é fechado, o Teorema de Green diz-nos que*

$$\iint_A 0 \; dxdy = \oint_{\partial A} \vec{F} \cdot d\vec{r}\,.$$

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*

$$\oint_C \vec{F} \cdot d\vec{r} = -\oint_S \vec{F} \cdot \vec{r} = -(-2\pi) = 2\pi$$

(o integral ao longo de S *no sentido direto foi calculado no Exemplo 27.4).*

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 27.3(ii) como critério efetivo para decidir se um campo fechado é ou não conservativo.*

Exemplo 28.7. *Seja $\vec{F}\colon \mathbb{R}^2 \setminus \{(0,1)\} \to \mathbb{R}^2$ 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 27.4) 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

$$A = \{(x,y) \in \mathbb{R}^2 \colon x^2 + (y-1)^2 > 1,\ |x+y| < 3,\ |x-y| < 3\}\,,$$

obtemos

$$0 = \iint_A \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right) dxdy = \oint_C \vec{F} \cdot d\vec{r} + \oint_S \vec{F} \cdot d\vec{r}\,,$$

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 27.4).

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 28.1. 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

$$A = \{(x,y) \in \mathbb{R}^2 \colon a < x < b,\ f(x) < y < g(x)\}$$

e

$$A = \{(x,y) \in \mathbb{R}^2 \colon c < y < d,\ h(y) < x < k(y)\}\,,$$

com $f, g \colon [a,b] \to \mathbb{R}$ e $h, k \colon [c,d] \to \mathbb{R}$ 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

$$\begin{aligned}\iint_A \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right) dxdy &= \iint_A \frac{\partial Q}{\partial x} dxdy - \iint_A \frac{\partial P}{\partial y} dxdy \\ &= \int_c^d \left(\int_{h(y)}^{k(y)} \frac{\partial Q}{\partial x} dx\right) dy - \int_a^b \left(\int_{f(x)}^{g(x)} \frac{\partial P}{\partial y} dy\right) dx\,.\end{aligned}$$

Começamos por calcular a primeira parcela na última linha acima. Pelo Teorema Fundamental do Cálculo temos

$$\begin{aligned}\int_c^d \left(\int_{h(y)}^{k(y)} \frac{\partial Q}{\partial x} dx\right) dy &= \int_c^d Q(k(y),y) - Q(h(y),y) dy \\ &= \int_c^d Q(k(y),y) dy - \int_c^d Q(h(y),y) dy\,.\end{aligned} \tag{18}$$

Observamos agora que a fronteira ∂A da região A é constituída por 4 troços parametrizados respetivamente por

$$\begin{aligned} g_1(y) &= (h(y), y)\,, & c \le y \le d\,,\\ g_2(x) &= (x, d)\,, & h(d) \le x \le k(d)\,,\\ g_3(y) &= (k(y), y)\,, & c \le y \le d\,,\\ g_4(x) &= (x, c)\,, & h(c) \le x \le k(c)\,. \end{aligned}$$

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 (18). Conclui-se assim que

$$\iint_A \frac{\partial Q}{\partial x} dxdy = \oint_{\partial A} (0,Q)\cdot d\vec{r}.$$

Cálculos inteiramente análogos mostram que

$$\iint_A -\frac{\partial P}{\partial y} dxdy = \oint_{\partial A} (P,0)\cdot d\vec{r},$$

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 $A_1, \ldots, A_N$, teremos

$$\begin{aligned}\iint_A \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right) dxdy &= \sum_{i=1}^N \iint_{A_i} \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right) dxdy \\ &= \sum_{i=1}^N \oint_{\partial A_i} Pdx + Qdy\,.\end{aligned}$$

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

$$\sum_{i=1}^N \oint_{\partial A_i} Pdx + Qdy = \oint_{\partial A} Pdx + Qdy\,,$$

sendo ∂A percorrida no sentido indicado no enunciado. Isto completa a demonstração. □

29. Homotopia de caminhos

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 29.1. *Seja $D \subset \mathbb{R}^n$ um conjunto. Os caminhos $g_1, g_2 \colon [a,b] \to D$ 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*

$$H \colon [a,b] \times [0,1] \to D$$

(dita uma homotopia*) tal que*

$$\begin{aligned} H(t,0) &= g_1(t) \quad \text{para } a \le t \le b\,, \\ H(t,1) &= g_2(t) \quad \text{para } a \le t \le b\,, \\ H(a,s) &= A \quad \text{para } 0 \le s \le 1\,, \\ H(b,s) &= B \quad \text{para } 0 \le s \le 1\,. \end{aligned}$$

A ideia da definição anterior é a seguinte: Podemos pensar no parâmetro s como o tempo. Ao fixar s obtemos um caminho

$$t \mapsto H(t,s) \quad a \le t \le b$$

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 29.2.

(i) Quaisquer dois caminhos $g_1, g_2 \colon [a,b] \to \mathbb{R}^n$ 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 $g_1, g_2 \colon [0,\pi] \to \mathbb{R}^2 \setminus \{(0,0)\}$ definidos por

$$g_1(\theta) = (\cos\theta, \operatorname{sen}\theta) \quad e \quad g_2(\theta) = (\cos\theta, -\operatorname{sen}\theta)$$

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 $g_1 \colon [0,2\pi] \to \mathbb{R}^3 \setminus \{(0,0,0)\}$ definido por

$$g_1(\theta) = (\cos\theta, \operatorname{sen}\theta, 0)$$

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 29.3. *Seja $U \subset \mathbb{R}^n$ um aberto e $\vec{F} \colon U \to \mathbb{R}^n$ um campo vetorial fechado de classe C^1. Se C_1 e C_2 são curvas parametrizadas por caminhos homotópicos então*

$$\int_{C_1} \vec{F} \cdot d\vec{r} = \int_{C_2} \vec{F} \cdot d\vec{r}\,.$$

30. Conjuntos simplesmente conexos

Dem. do Teorema 29.3. Temos a verificar que

$$\oint_{C_1} \vec{F} \cdot d\vec{r} - \oint_{C_2} \vec{F} \cdot d\vec{r} = 0\,. \tag{19}$$

Seja $H\colon [a,b] \times [0,1] \to U$ 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 linha, a igualdade (19) é equivalente a

$$\int_a^b \left(\vec{F}(H(t,0)) \cdot \frac{\partial H}{\partial t}(t,0) - \vec{F}(H(t,1)) \cdot \frac{\partial H}{\partial t}(t,1) \right) dt = 0\,. \tag{20}$$

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, $\frac{\partial^2 H_i}{\partial t \partial s} = \frac{\partial^2 H_i}{\partial s \partial t}$, pelo que

$$\frac{\partial Q}{\partial t} - \frac{\partial P}{\partial s} = 0.$$

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

$$\begin{aligned} g_1(t) &= (t,0)\,, & a \le t \le b\,,\\ g_2(s) &= (b,s)\,, & 0 \le s \le 1\\ g_3(t) &= (t,1)\,, & a \le t \le b\\ g_4(s) &= (a,s)\,, & 0 \le s \le 1\,, \end{aligned}$$

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 (20) que pretendíamos demonstrar. □

Nota 30.1. *Assumiu-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 29.3 na generalidade em que foi enunciado.*

Nota 30.2. *Dois caminhos* fechados $g_1, g_2 \colon [a,b] \to U$ *(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 $H \colon [a,b] \times [0,1] \to U$ satisfazendo*

$$H(t,0) = g_1(t), \quad H(t,1) = g_2(t), \quad H(a,s) = H(b,s)$$

para todos os $t \in [a,b]$ e $s \in [0,1]$ (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 29.3 torna efetiva a caracterização de campos gradientes dada no Teorema 27.3(ii) desde que o campo $\vec{F}$ seja fechado.

Exemplo 30.3. *Determinar se o campo vetorial $\vec{F} \colon \mathbb{R}^3 \setminus \{(0,0,z) \colon z \in \mathbb{R}\} \to \mathbb{R}^3$ 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 29.3 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 homotópico a um caminho constante ou 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). Para perceber o que se passa no segundo caso, basta calcular o trabalho ao longo do caminho

$$g(t) = (\cos t, \operatorname{sen} t, 0), \qquad 0 \leq t \leq 2\pi.$$

Temos

$$\int_0^1 \vec{F}(g(t)) \cdot g'(t)dt = \int_0^1 (2\cos t, 2\operatorname{sen} t, \cos t) \cdot (-\operatorname{sen} t, \cos t, 0)dt = \int_0^1 0 \; dt = 0,$$

pelo que o integral de linha de $\vec{F}$ ao longo de qualquer caminho seccionalmente regular fechado é 0, e portanto, pelo Teorema 27.3(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 se e só se o integral de $\vec{F}$ ao longo de uma circunferência horizontal centrada no eixo dos zz é 0.

O Teorema 29.3 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 30.4. *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 30.5.

(i) $\mathbb{R}^n$ *é simplesmente conexo para todo o n (pois como vimos no Exemplo 29.2(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 $g(t) = (\cos t, \operatorname{sen} t)$ 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) $U = \mathbb{R}^3 \setminus \{(x,y,0)\colon x^2+y^2=1\}$ *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 30.6. *Seja* $U \subset \mathbb{R}^n$ *um conjunto aberto conexo. É um bom exercício demonstrar que* U *é simplesmente conexo sse quaisquer dois caminhos* $g_1, g_2\colon [a,b] \to U$ *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 30.7. *Seja* $U \subset \mathbb{R}^n$ *um aberto simplesmente conexo e* $\vec{F}\colon U \to \mathbb{R}^n$ *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 27.3 que $\vec{F}$ é um gradiente em U. □

Exemplo 30.8.

(i) O campo vetorial $\vec{F}\colon \mathbb{R}^3 \to \mathbb{R}^3$ *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 $\vec{F}\colon \mathbb{R}^3 \setminus \{(0,0,0)\}$ *é gradiente se e só se é fechado.*

Referências

[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).