FCUP
Dep. Matemática Pura

CURSO de ANÁLISE INFINITESIMAL

RESUMO das Aulas Teóricas

João Nuno Tavares

Dept. Matemática Pura, Faculdade de Ciências, Univ. Porto, 4050 Porto, Portugal[1]

[1]E-mail adress: jntavar@fc.up.pt

ÍNDICE

BIBLIOGRAFIA

1. J.E. Marsden, A.J. Tromba: “**Vector Calculus**”. W.H.Freeman and Company. New York, 1988.

2. T.M. Apostol: “**Calculus**, vol.2”. Xerox College Publishing International Textbook series, 1969.

3. R.E. Williamson, R.H. Crowell, H.F. Trotter. “**Calculus of Vector Functions**”. Prentice-Hall, Inc. Englewood Cliffs, 1989.

4. J.E. Marsden: “**Elementary Classical Analysis”**. W.H.Freeman and Company. New York, 1974.

5. C.H. Edwards, Jr.. “**Advanced Calculus of Several Variables**”. Academic Press, 1973.

6. M.R. Spiegel: “**Advanced Mathematics for Engineers and Scientists**”. Schaum's Outline Series. McGraw-Hill Book Company,1971.

7. M.R. Spiegel: “**Vector Analysis**”. Schaum's Outline Series. McGraw-Hill Book Company,1971.

Capítulo 1

Preliminares

1.1 $\mathbb{R}^n$. Produto interno e norma euclideana

Consideremos o conjunto $\mathbb{R}^n$:

$$\mathbb{R}^n \equiv \{\mathbf{x} = (x^1, ..., x^n) : \quad x^1, ..., x^n \in \mathbb{R}\} \tag{1.1.1}$$

munido da estrutura usual de espaço vectorial real de dimensão n. A **base canónica** de $\mathbb{R}^n$ é constituída pelos vectores:

$$\begin{aligned} \mathbf{e}_1 &= (1, 0, ..., 0) \\ \mathbf{e}_2 &= (0, 1, 0, ..., 0) \\ &\cdots \\ \mathbf{e}_n &= (0, 0, ..., 0, 1) \end{aligned} \tag{1.1.2}$$

e as coordenadas de um vector qualquer $\mathbf{x} \in \mathbb{R}^n$, relativamente a esta base, dizem-se as **coordenadas (cartesianas) usuais** do vector $\mathbf{x}$:

$$\begin{aligned} \mathbf{x} &= (x^1, x^2, \cdots, x^n) \\ &= x^1\,\mathbf{e}_1 + x^2\,\mathbf{e}_2 + \cdots + x^n\,\mathbf{e}_n \end{aligned} \tag{1.1.3}$$

Em $\mathbb{R}^2$ é comum utilizar a notação $\mathbf{i} = \mathbf{e}_1, \mathbf{j} = \mathbf{e}_2$, e anàlogamente em $\mathbb{R}^3$, a notação $\mathbf{i} = \mathbf{e}_1, \mathbf{j} = \mathbf{e}_2, \mathbf{k} = \mathbf{e}_3$, para os vectores das respectivas bases canónicas.

♣ **Definição 1.1** ... *Define-se o* **produto interno (euclideano)** *em* $\mathbb{R}^n$*, através da fórmula:*

$$\mathbf{x} \cdot \mathbf{y} \equiv \sum_{i=1}^{n} x^i y^i \tag{1.1.4}$$

$\forall \mathbf{x} = (x^1, ..., x^n), \mathbf{y} = (y^1, ..., y^n) \in \mathbb{R}^n$.

♣ **Proposição 1.1** ... *O produto interno (euclideano) verifica as propriedades seguintes:*

(é uma forma bilinear):

$$\begin{aligned}(\mathbf{x}+\mathbf{y})\cdot\mathbf{z} &= \mathbf{x}\cdot\mathbf{z}+\mathbf{y}\cdot\mathbf{z}\\ \mathbf{x}\cdot(\mathbf{y}+\mathbf{z}) &= \mathbf{x}\cdot\mathbf{y}+\mathbf{x}\cdot\mathbf{z}\\ \alpha\mathbf{x}\cdot\mathbf{y} &= \mathbf{x}\cdot\alpha\mathbf{y}=\alpha(\mathbf{x}\cdot\mathbf{y})\end{aligned} \tag{1.1.5}$$

(é uma forma simétrica):

$$\mathbf{x}\cdot\mathbf{y}=\mathbf{y}\cdot\mathbf{x} \tag{1.1.6}$$

(é não degenerada):

$$\mathbf{x}\cdot\mathbf{y}=0 \quad \forall\mathbf{y}\in\mathbb{R}^n \Rightarrow \mathbf{x}=\mathbf{O} \tag{1.1.7}$$

(é definida positiva)

$$\mathbf{x}\cdot\mathbf{x}\geq 0 \tag{1.1.8}$$

$\forall\mathbf{x},\mathbf{y},\mathbf{z}\in\mathbb{R}^n, \forall\alpha\in\mathbb{R}$.

Outra notação comum para o produto interno acima referido é $<\mathbf{x},\mathbf{y}>$.

♣ **Definição 1.2** ... *Define-se a* **norma euclideana** $\|\mathbf{x}\|$, *de um vector* $\mathbf{x}\in\mathbb{R}^n$, *através da fórmula:*

$$\begin{aligned}\|\mathbf{x}\| &\equiv \sqrt{\mathbf{x}\cdot\mathbf{x}}\\ &= \sqrt{\sum_{i=1}^{n}(x^i)^2}\end{aligned} \tag{1.1.9}$$

♣ **Proposição 1.2** ... *A norma euclideana verifica as propriedades seguintes:*

(é positiva e não degenerada):

$$\|\mathbf{x}\|\geq 0 \quad e \quad \|\mathbf{x}\|=0 \quad sse \quad \mathbf{x}=\mathbf{O} \tag{1.1.10}$$

(é homogénea (positiva)):

$$\|\alpha\mathbf{x}\|=|\alpha|\|\mathbf{x}\| \tag{1.1.11}$$

(desigualdade triangular):

$$\|\mathbf{x}+\mathbf{y}\|\leq\|\mathbf{x}\|+\|\mathbf{y}\| \tag{1.1.12}$$

$\forall\mathbf{x},\mathbf{y}\in\mathbb{R}^n, \forall\alpha\in\mathbb{R}$.

Todas as propriedades são de demonstração imediata com excepção da desigualdade triangular, que resulta imediatamente de uma outra importante desigualdade que passamos a enunciar:

♣ **Proposição 1.3** ... $\forall\mathbf{x},\mathbf{y}\in\mathbb{R}^n$ *é válida a seguinte* **desigualdade de Cauchy-Schwarz***:*

$$|\mathbf{x}\cdot\mathbf{y}|\leq\|\mathbf{x}\|\|\mathbf{y}\| \tag{1.1.13}$$

- Demonstração...

  Pondo $\mathbf{x} = (x^1, ..., x^n)$ e $\mathbf{y} = (y^1, ..., y^n)$, a desigualdade de Cauchy-Schwarz pode ser escrita na forma equivalente:

$$(\sum_{i=1}^{n} x^i y^i)^2 \leq (\sum_{i=1}^{n} (x^i)^2)(\sum_{i=1}^{n} (y^i)^2) \tag{1.1.14}$$

  que por sua vez resulta directamente da identidade:

$$(\sum_{i=1}^{n} x^i y^i)^2 = (\sum_{i=1}^{n} (x^i)^2)(\sum_{i=1}^{n} (y^i)^2) - \frac{1}{2}\sum_{j=1}^{n}\sum_{k=1}^{n} (x^j y^k - y^j x^k)^2 \tag{1.1.15}$$

  que pode ser verificada imediatamente,

♣.

Nota 1.1 ... *Qualquer função definida em* $\mathbb{R}^n$ *e com valores em* $\mathbb{R}$, *que verifique as três propriedades referidas na proposição anterior, diz-se uma* **norma** *em* $\mathbb{R}^n$. *Outros exemplos de normas em* $\mathbb{R}^n$ *são dados pelas fórmulas seguintes:*

$$\|\mathbf{x}\|_1 \equiv \sum_{i=1}^{n} |x^i| \tag{1.1.16}$$

$$\|\mathbf{x}\|_\infty \equiv \max_{1\leq i\leq n} |x^i| \tag{1.1.17}$$

Dados dois vectores não nulos $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, deduzimos da desigualdade de Cauchy-Schwarz que:

$$-1 \leq \frac{\mathbf{x}\cdot\mathbf{y}}{\|\mathbf{x}\|\|\mathbf{y}\|} \leq 1 \tag{1.1.18}$$

o que permite definir o **ângulo (não orientado)** $\phi \in [0, \pi]$, entre os referidos vectores não nulos $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, como sendo o único $\phi \in [0, \pi]$, tal que:

$$\cos\phi = \frac{\mathbf{x}\cdot\mathbf{y}}{\|\mathbf{x}\|\|\mathbf{y}\|} \in [-1, 1] \tag{1.1.19}$$

Portanto:

$$\mathbf{x}\cdot\mathbf{y} = \|\mathbf{x}\|\|\mathbf{y}\|\cos\phi \tag{1.1.20}$$

Dois vectores $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ dizem-se **ortogonais** se $\mathbf{x}\cdot\mathbf{y} = 0$.

Sejam $\mathbf{x}$ e $\mathbf{y}$ dois vectores em $\mathbb{R}^n$, com $\mathbf{x}$ não nulo. Então existe um único vector $\mathbf{u}$, na recta gerada por $\mathbf{x}$, e um único vector $\mathbf{v}$, ortogonal a $\mathbf{x}$, tais que $\mathbf{y} = \mathbf{u} + \mathbf{v}$. O vector $\mathbf{u}$ diz-se a **projecção ortogonal** de $\mathbf{y}$ sobre $\mathbf{x}$ e é dado por:

$$\mathbf{u} = \frac{\mathbf{y}\cdot\mathbf{x}}{\|\mathbf{x}\|^2}\mathbf{x} \tag{1.1.21}$$

Dado um vector não nulo $\mathbf{n} \neq \mathbf{O}$, o conjunto dos vectores $\mathbf{x} \in \mathbb{R}^n$ que são ortogonais a $\mathbf{n}$:

$$\{\mathbf{x} \in \mathbb{R}^n : \mathbf{x} \cdot \mathbf{n} = 0\} \tag{1.1.22}$$

formam um subespaço de dimensão $n-1$ em $\mathbb{R}^n$, que se diz o **hiperplano (vectorial) ortogonal a n**.

Dado um ponto arbitrário $\mathbf{p} \in \mathbb{R}^n$ e um vector não nulo $\mathbf{n} \neq \mathbf{O}$, o conjunto dos vectores $\mathbf{x} \in \mathbb{R}^n$ que verificam a equação:

$$(\mathbf{x} - \mathbf{p}) \cdot \mathbf{n} = 0 \tag{1.1.23}$$

formam um subespaço afim de dimensão $n - 1$ em $\mathbb{R}^n$, que se diz o **hiperplano (afim) que passa em p e é ortogonal a n**.

1.2 Produto vectorial e produto misto em $\mathbb{R}^3$

Definamos agora o chamado produto vectorial de dois vectores em $\mathbb{R}^3$:

♣ **Definição 1.3** ... *Dados dois vectores* $\mathbf{x} = (x^1, x^2, x^3), \mathbf{y} = (y^1, y^2, y^3) \in \mathbb{R}^3$*, define-se o* **produto vectorial** $\mathbf{x} \times \mathbf{y}$*, de* $\mathbf{x}$ *por* $\mathbf{y}$*, como sendo o seguinte vector de* $\mathbb{R}^3$*:*

$$\mathbf{x} \times \mathbf{y} \equiv (x^2y^3 - x^3y^2)\mathbf{i} + (x^3y^1 - x^1y^3)\mathbf{j} + (x^1y^2 - x^2y^1)\mathbf{k} \tag{1.2.1}$$

O produto vectorial $\mathbf{x} \times \mathbf{y}$, pode ser obtido desenvolvendo segundo a primeira linha, o determinante formal:

$$\mathbf{x} \times \mathbf{y} = \det \begin{bmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ x^1 & x^2 & x^3 \\ y^1 & y^2 & y^3 \end{bmatrix}$$

A seguinte proposição, cuja demonstração é simples, refere as propriedades mais importantes deste produto vectorial.

♣ **Proposição 1.4** ... *O produto vectorial verifica as propriedades seguintes:*

(é bilinear):

$$\begin{aligned} (\mathbf{x} + \mathbf{y}) \times \mathbf{z} &= \mathbf{x} \times \mathbf{z} + \mathbf{y} \times \mathbf{z} \\ \mathbf{x} \times (\mathbf{y} + \mathbf{z}) &= \mathbf{x} \times \mathbf{y} + \mathbf{x} \times \mathbf{z} \\ \alpha\mathbf{x} \times \mathbf{y} &= \mathbf{x} \times \alpha\mathbf{y} = \alpha(\mathbf{x} \times \mathbf{y}) \end{aligned} \tag{1.2.2}$$

(é antissimétrico):

$$\mathbf{x} \times \mathbf{y} = -\mathbf{y} \times \mathbf{x} \tag{1.2.3}$$

Além disso, se $\mathbf{x} \in \mathbb{R}^3$ *e* $\mathbf{y} \in \mathbb{R}^3$*, são ambos não nulos, então:*

(i)... $\mathbf{x} \times \mathbf{y}$ *é perpendicular a* $\mathbf{x}$ *e a* $\mathbf{y}$*, i.e.:*

$$(\mathbf{x} \times \mathbf{y}) \cdot \mathbf{x} = 0 = (\mathbf{x} \times \mathbf{y}) \cdot \mathbf{y} \tag{1.2.4}$$

Se $\mathbf{x}$ *e* $\mathbf{y}$ *são linearmente independentes,* $\mathbf{x} \times \mathbf{y}$ *é perpendicular ao plano gerado por* $\mathbf{x}$ *e* $\mathbf{y}$.

(ii)...

$$\|\mathbf{x} \times \mathbf{y}\| = \|\mathbf{x}\| \|\mathbf{y}\| \sin \phi \tag{1.2.5}$$

onde ϕ *é o ângulo entre* $\mathbf{x}$ *e* $\mathbf{y}$.

Portanto, $\|\mathbf{x} \times \mathbf{y}\|$ *é igual à área do paralelogramo cujos lados adjacentes são* $\mathbf{x}$ *e* $\mathbf{y}$.

(iii)... $\mathbf{x} \times \mathbf{y} = \mathbf{O} \Leftrightarrow \mathbf{x}$ *e* $\mathbf{y}$ *são linearmente dependentes.*

(iv)... O produto vectorial não é associativo. De facto:

$$(\mathbf{x} \times \mathbf{y}) \times \mathbf{z} = (\mathbf{x} \cdot \mathbf{z})\mathbf{y} - (\mathbf{y} \cdot \mathbf{z})\mathbf{x} \tag{1.2.6}$$

enquanto que:

$$\mathbf{x} \times (\mathbf{y} \times \mathbf{z}) = (\mathbf{x} \cdot \mathbf{z})\mathbf{y} - (\mathbf{x} \cdot \mathbf{y})\mathbf{z} \tag{1.2.7}$$

Em particular, se consideramos o paralelogramo de lados adjacentes $\mathbf{x} = (x^1, x^2, 0)$ e $\mathbf{y} = (y^1, y^2, 0)$, contido no subespaço $\mathbb{R}^2 \cong \mathbb{R}^2 \times \{\mathbf{O}\} \subset \mathbb{R}^3$, vemos que a respectiva área é dada por:

$$\begin{aligned} \|\mathbf{x} \times \mathbf{y}\| &= \|\det \begin{bmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ x^1 & x^2 & 0 \\ y^1 & y^2 & 0 \end{bmatrix} \| \\ &= |\det \begin{bmatrix} x^1 & x^2 \\ y^1 & y^2 \end{bmatrix} | \end{aligned} \tag{1.2.8}$$

Definamos agora, ainda em $\mathbb{R}^3$, o chamado produto misto.

♣ **Definição 1.4** ... *Dados três vectores* $\mathbf{x}, \mathbf{y}, \mathbf{z}$ *em* $\mathbb{R}^3$, *define-se o* **produto misto (ou triplo)** $[\mathbf{x}, \mathbf{y}, \mathbf{z}]$, *de* $\mathbf{x}, \mathbf{y}$ *e* $\mathbf{z}$ *(por esta ordem), através de:*

$$[\mathbf{x}, \mathbf{y}, \mathbf{z}] \equiv \mathbf{x} \cdot (\mathbf{y} \times \mathbf{z}) \tag{1.2.9}$$

É fácil ver que $[\mathbf{x}, \mathbf{y}, \mathbf{z}]$ é dado por:

$$[\mathbf{x}, \mathbf{y}, \mathbf{z}] = \det \begin{bmatrix} x^1 & x^2 & x^3 \\ y^1 & y^2 & y^3 \\ z^1 & z^2 & z^3 \end{bmatrix} \tag{1.2.10}$$

♣ **Proposição 1.5** ... *(i). São válidas as igualdades seguintes:*

$$\begin{aligned} [\mathbf{x}, \mathbf{y}, \mathbf{z}] &= [\mathbf{y}, \mathbf{z}, \mathbf{x}] = [\mathbf{z}, \mathbf{x}, \mathbf{y}] = -[\mathbf{y}, \mathbf{x}, \mathbf{z}] \\ &= -[\mathbf{x}, \mathbf{z}, \mathbf{y}] = -[\mathbf{z}, \mathbf{y}, \mathbf{x}] \end{aligned} \tag{1.2.11}$$

(ii). O volume $\mathbf{vol}(\mathbf{x}, \mathbf{y}, \mathbf{z})$, *do paralelipípedo de lados adjacentes* $\mathbf{x}, \mathbf{y}, \mathbf{z} \in \mathbb{R}^3$, *é igual ao módulo do produto misto (ver a figura (1.1)):*

$$\begin{aligned} \mathbf{vol}(\mathbf{x}, \mathbf{y}, \mathbf{z}) &= |[\mathbf{x}, \mathbf{y}, \mathbf{z}]| \\ &= |det \begin{bmatrix} x^1 & x^2 & x^3 \\ y^1 & y^2 & y^3 \\ z^1 & z^2 & z^3 \end{bmatrix} | \end{aligned} \tag{1.2.12}$$

Figure 1.1: $\mathbf{vol}(\mathbf{x},\mathbf{y},\mathbf{z}) = |[\mathbf{x},\mathbf{y},\mathbf{z}]|$

.

- Demonstração...

  (i). Deduz-se de (1.2.10) aplicando as propriedades conhecidas sobre determinantes.

  (ii). O volume de um paralelipípedo é igual ao produto da área da base pela sua altura. A base é o paralelogramo de lados adjacentes $\mathbf{x}$ e $\mathbf{y}$, e por isso, a sua área é $\|\mathbf{x}\times\mathbf{y}\|$. A altura é igual à norma da projecção de $\mathbf{z}$ sobre um vector perpendicular à base. Mas $\mathbf{x}\times\mathbf{y}$ é perpendicular à base, e atendendo a (1.1.21), a projecção de $\mathbf{z}$ sobre $\mathbf{x}\times\mathbf{y}$, é igual a:

$$\frac{\mathbf{z}\cdot\mathbf{x}\times\mathbf{y}}{\|\mathbf{x}\times\mathbf{y}\|^2}\mathbf{x}\times\mathbf{y} \tag{1.2.13}$$

  donde se deduz fàcilmente o resultado,

♣.

Podemos ainda generalizar a noção de volume de um paralelipípedo, através da seguinte definição:

♣ **Definição 1.5** ... *Dados n vectores* $\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_n$ *em* $\mathbb{R}^n$, *define-se o* **volume** $n-$**dimensional**, $\mathbf{vol}\,(\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_n)$, *do paralelipípedo* $n-$*dimensional de lados adjacentes* $\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_n$, *através de:*

$$\mathbf{vol}\,(\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_n) \equiv |det \begin{bmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \\ \vdots \\ \mathbf{x}_n \end{bmatrix}| \tag{1.2.14}$$

onde

$$\begin{bmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \\ \vdots \\ \mathbf{x}_n \end{bmatrix}$$

designa a matriz quadrada $n \times n$, *cujas linhas são as coordenadas dos vectores* $\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_n$.

Portanto em $\mathbb{R}^2$ a área do paralelogramo de lados adjacentes $\mathbf{x}_1, \mathbf{x}_2$, é igual a $\mathbf{vol}(\mathbf{x}_1, \mathbf{x}_2)$ (ver (1.2.8)), em $\mathbb{R}^3$ o volume do paralelipípedo de lados adjacentes $\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3$, é igual a $\mathbf{vol}(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3)$ (ver (1.2.12)), etc....

Consideremos agora uma aplicação linear $T : \mathbb{R}^2 \to \mathbb{R}^2$. A imagem do quadrado $\mathcal{Q}$, de lados adjacentes $\mathbf{i}, \mathbf{j}$:

$$\mathcal{Q} = \{\alpha\mathbf{i} + \beta\mathbf{j} : 0 \leq \alpha, \beta \leq 1\}$$

é o paralelogramo $T(\mathcal{Q})$, de lados adjacentes $T(\mathbf{i})$ e $T(\mathbf{j})$. Pondo $T(\mathbf{i}) = a\mathbf{i} + b\mathbf{j} = (a, b)$ e $T(\mathbf{j}) = c\mathbf{i} + d\mathbf{j} = (c, d)$, sabemos por (1.2.8) (ver também (1.2.14)), que a área deste paralelogramo é igual a:

$$\begin{aligned} \text{área}\, T(\mathcal{Q}) = \mathbf{vol}(T(\mathbf{i}), T(\mathbf{j})) &= |\det \begin{bmatrix} a & b \\ c & d \end{bmatrix}| \\ &= |\det T| \end{aligned} \tag{1.2.15}$$

uma vez que o determinante de uma matriz é igual ao determinante da sua transposta. Portanto:

$$\text{área}\, T(\mathcal{Q}) = |\det T| \tag{1.2.16}$$

Mais geralmente, se $\mathcal{R}$ é um paralelogramo de lados adjacentes $\mathbf{u}$ e $\mathbf{v}$, então a imagem $T(\mathcal{R})$ é um paralelogramo de lados adjacentes $T(\mathbf{u})$ e $T(\mathbf{v})$, e é fácil provar que a área desta imagem é igual a (ver (1.2.12)):

$$\begin{aligned} \text{área}\, T(\mathcal{R}) = \mathbf{vol}(T(\mathbf{u}), T(\mathbf{v})) &= |\det \begin{bmatrix} T(\mathbf{u}) \\ T(\mathbf{v}) \end{bmatrix}| \\ &= |\det T|\,(\text{área}\,(\mathcal{R})) \end{aligned} \tag{1.2.17}$$

isto é:

$$\frac{\text{área}\, T(\mathcal{R})}{\text{área}\,(\mathcal{R})} = |\det T| \tag{1.2.18}$$

Um raciocínio análogo permite concluír que se $\mathcal{P}$ é um paralelipípedo em $\mathbb{R}^3$, e $T : \mathbb{R}^3 \to \mathbb{R}^3$ é uma transformação linear, então:

$$\frac{\mathbf{vol}\, T(\mathcal{P})}{\mathbf{vol}\,(\mathcal{P})} = |\det T| \tag{1.2.19}$$

sendo também válida uma fórmula análoga para paralelipípedos n-dimensionais e transformações lineares em $\mathbb{R}^n$.

Capítulo 2

Campos escalares

2.1 Conjuntos de nível. Gráficos. Exemplos

2.1.1 Conjuntos de nível. Gráficos

♣ <u>**Definição**</u> **2.1** ... *Dada uma função $f : \mathcal{D} \subseteq \mathbb{R}^n \to \mathbb{R}$ (um campo escalar), definida num subconjunto $\mathcal{D} \subseteq \mathbb{R}^n$, define-se, para cada $c \in \mathbb{R}$, o* **conjunto de nível** *c, notado por N_c, através de:*

$$N_c \equiv \{\mathbf{x} \in \mathcal{D} :\ f(\mathbf{x}) = c\} \tag{2.1.1}$$

onde $c \in \mathbb{R}$.

<u>**Exemplos**</u> ...

(i)... Os conjuntos de nível de $f(\mathbf{x}) = \|\mathbf{x} - \mathbf{a}\|$, $\mathbf{x} \in \mathbb{R}^3$, onde $\mathbf{a}$ é um vector fixo em $\mathbb{R}^3$, são esferas concêntricas de centro $\mathbf{a}$, e o próprio ponto $\mathbf{a}$.

(ii)... Os conjuntos de nível de $f(\mathbf{x}) = \|\mathbf{x} - \mathbf{a}\| + \|\mathbf{x} - \mathbf{b}\|$, $\mathbf{x} \in \mathbb{R}^2$, onde $\mathbf{a}, \mathbf{b}$ são dois vectores fixos em $\mathbb{R}^2$, são elipses com focos $\mathbf{a}$ e $\mathbf{b}$, e também o segmento que une $\mathbf{a}$ a $\mathbf{b}$.

♣ <u>**Definição**</u> **2.2** ... *Dada uma função $f : \mathcal{D} \subseteq \mathbb{R}^n \to \mathbb{R}$ (um campo escalar), definida num subconjunto $\mathcal{D} \subseteq \mathbb{R}^n$, define-se o* **gráfico** *de f, $\mathbf{gr}\, f$, como sendo o subconjunto de $\mathbb{R}^{n+1}$, dado por:*

$$\mathbf{gr}\, f \equiv \{(\mathbf{x}, z) \in \mathbb{R}^n \times \mathbb{R} = \mathbb{R}^{n+1} :\ z = f(\mathbf{x}), \quad \mathbf{x} \in \mathcal{D}\} \tag{2.1.2}$$

Notemos que $\mathbf{gr}\, f$, onde $f : \mathcal{D} \subseteq \mathbb{R}^n \to \mathbb{R}$, é igual ao conjunto de nível 0, da função:

$$F : \mathcal{D} \times \mathbb{R} \subseteq \mathbb{R}^{n+1} \to \mathbb{R}$$

definida por $F(\mathbf{x}, z) = f(\mathbf{x}) - z, \quad (\mathbf{x}, z) \in \mathcal{D} \times \mathbb{R}$.

Por outro lado, dada uma função $f : \mathcal{D} \subseteq \mathbb{R}^n \to \mathbb{R}$, as intersecções do gráfico $\mathbf{gr}\, f \subset \mathbb{R}^{n+1}$, com os hiperplanos "horizontais" $z \equiv c$, $c \in \mathbb{R}$, projectam-se sobre os conjuntos de nível $N_c \subset \mathbb{R}^n = \mathbb{R}^n \times \{\mathbf{O}\}$.

<u>**Exemplo**</u> ...

Para $f(x,y) = x^2 + y^2$, $\mathbf{x} = (x,y) \in \mathbb{R}^2$, tem-se que:

$$\mathbf{gr}\, f \equiv \{(x,y,z) \in \mathbb{R}^3 : z = x^2 + y^2, \quad \mathbf{x} = (x,y) \in \mathbb{R}^2\}$$

é um "parabolóide de revolução" em $\mathbb{R}^3$. Para cada $c > 0$, a intersecção de $\mathbf{gr}\, f$ com o plano $z \equiv c$, projecta-se na circunferência centrada na origem e de raio $\sqrt{c}$, contida no plano (x,y), que é o conjunto de nível c de f (ver figura 2.1).

Figure 2.1: Parabolóide de revolução

.

<u>**Exemplo**</u> ...

O gráfico de $f(x,y) = x^2 - y^2$ é a superfície de uma "sela" em $\mathbb{R}^3$, representada na figura 2.2. Aí se mostram algumas curvas de nível de f, no plano (x,y), obtidas por projecção de intersecções de $\mathbf{gr}\, f$ com planos $z \equiv c$.

Figure 2.2: Sela. Gráfico de $f(x,y) = x^2 - y^2$

.

<u>**Exemplo**</u> ...

Na figura 2.3 representam-se alguns conjuntos de nível da função $f(x,y) = x^2 + y^2 - z^2$

<u>**Exemplo**</u> ...

Os conjuntos de nível de uma forma linear $\phi : \mathbb{R}^n \to \mathbb{R}$, não nula, dada por:

$$\phi(\mathbf{x}) = \mathbf{x} \cdot \mathbf{a}$$

Figure 2.3: Conjuntos de nível de $f(x, y) = x^2 + y^2 - z^2$

.

com $\mathbf{a} \neq \mathbf{O}$ fixo, são os conjuntos:

$$N_c = \{\mathbf{x} \in \mathbb{R}^n : \mathbf{x} \cdot \mathbf{a} = c\}$$

e constituem portanto, uma família (a um parâmetro $c \in \mathbb{R}$) de hiperplanos paralelos entre si, e todos ortogonais a $\mathbf{a}$.

Exemplo ...

O gráfico de uma forma afim do tipo (4.1.4):

$$T(\mathbf{x}) = \mathbf{x} \cdot \mathbf{a} + b$$

onde $\mathbf{a} \in \mathbb{R}^n$ e $b \in \mathbb{R}$ são fixos, é o subconjunto de $\mathbb{R}^{n+1}$:

$$\mathbf{gr}\, T = \{(\mathbf{x}, z) \in \mathbb{R}^{n+1} : z = \mathbf{x} \cdot \mathbf{a} + b, \quad \mathbf{x} \in \mathbb{R}^n\} \tag{2.1.3}$$

Se definirmos a função $F : \mathbb{R}^{n+1} \to \mathbb{R}$, através de:

$$\begin{aligned} F(\mathbf{x}, z) &= T(\mathbf{x}) - z = (\mathbf{x} \cdot \mathbf{a}) - z + b \\ &= (\mathbf{x}, z) \cdot (\mathbf{a}, -1) + b \end{aligned} \tag{2.1.4}$$

vemos que $\mathbf{gr}\, T$ é igual ao conjunto de nível 0, N_0 da forma afim em $\mathbb{R}^{n+1}$, dada por (2.1.4), e é portanto um hiperplano (afim) em $\mathbb{R}^{n+1}$ perpendicular ao vector $(\mathbf{a}, -1) \in \mathbb{R}^{n+1}$.

Exemplo ...

Seja $T(x, y) = 2x - 5y + 3$, $(x, y) \in \mathbb{R}^2$. Os conjuntos de nível de T, são uma família de rectas paralelas em $\mathbb{R}^2$, todas perpendiculares ao vector $(2, -5)$. Por outro lado:

$$\mathbf{gr}\, T = \{(x, y, z) \in \mathbb{R}^3 : z = T(x, y) = 2x - 3y + 3\}$$

que é igual ao conjunto de nível 0 de:

$$F(x, y, z) = T(x, y) - z = 2x - 5y - z + 3$$

que é um plano (afim) em $\mathbb{R}^3$ perpendicular a $(2, -5, -1)$.

Exemplo ...

As figuras 2.4, 2.5, 2.6 e 2.7, mostram alguns conjuntos (superfícies) de nível, de formas quadráticas definidas em $\mathbb{R}^3$, em forma diagonal.

Figure 2.4: Elipsóide

.

Figure 2.5: Hiperbolóide de uma folha

.

2.2 Limites. Continuidade

2.2.1 Noções elementares de topologia em $\mathbf{R}^n$

Recordemos que definimos a **distância** (euclideana) $d(\mathbf{x},\mathbf{y})$, entre dois pontos $\mathbf{x},\mathbf{y}\in\mathbb{R}^n$, através de:

$$\begin{aligned} d(\mathbf{x},\mathbf{y}) &\equiv \|\mathbf{x}-\mathbf{y}\| \\ &= \sqrt{(x^1-y^1)^2+...+(x^n-y^n)^2} \end{aligned} \tag{2.2.1}$$

É fácil provar que d satisfaz as propriedades seguintes:

(i)...

$$d(\mathbf{x},\mathbf{y})\geq 0 \qquad e \qquad d(\mathbf{x},\mathbf{y})=0\Leftrightarrow \mathbf{x}=\mathbf{y}$$

(ii)...

$$d(\mathbf{x},\mathbf{y})=d(\mathbf{y},\mathbf{x}) \qquad \forall \mathbf{x},\mathbf{y}\in\mathbb{R}^n$$

(iii)...(desigualdade triangular)

$$d(\mathbf{x},\mathbf{y})\leq d(\mathbf{x},\mathbf{z})+d(\mathbf{z},\mathbf{y}) \qquad \forall \mathbf{x},\mathbf{y},\mathbf{z}\in\mathbb{R}^n$$

Figure 2.6: Cone elíptico

.

Figure 2.7: Hiperbolóide de duas folhas

.

Dado um ponto $\mathbf{a} \in \mathbb{R}^n$, e um número positivo $\epsilon > 0$, o conjunto:

$$B(\mathbf{a}, \epsilon) \equiv \{\mathbf{x} \in \mathbb{R}^n : d(\mathbf{x}, \mathbf{a}) = \|\mathbf{x} - \mathbf{a}\| < \epsilon\} \tag{2.2.2}$$

diz-se a **bola aberta centrada em a e de raio** ϵ.

Um subconjunto $\mathcal{U} \subseteq \mathbb{R}^n$ diz-se uma **vizinhança** de um ponto $\mathbf{p} \in \mathbb{R}^n$, se $\mathcal{U}$ contem alguma bola aberta centrada em $\mathbf{p}$ e de raio $\epsilon > 0$.

♣ **Definição 2.3** ...

(i)... Seja $\mathcal{U}$ um subconjunto de $\mathbb{R}^n$. Um ponto $\mathbf{a} \in \mathcal{U}$, diz-se um **ponto interior** *de $\mathcal{U}$, se existir uma bola aberta centrada em $\mathbf{a}$ e de raio ϵ, completamente contida em $\mathcal{U}$:*

$$\mathbf{a} \in \mathcal{U}, \textit{ é ponto interior de } \mathcal{U} \textit{ sse } \exists \epsilon > 0 : B(\mathbf{a}, \epsilon) \subset \mathcal{U}$$

(ii)... Um subconjunto $\mathcal{U} \subseteq \mathbb{R}^n$, diz-se **aberto** *se todo o ponto de $\mathcal{U}$ é ponto interior de $\mathcal{U}$.*

Exemplo ...

(i)... A bola aberta $B(\mathbf{a}, \epsilon)$, centrada em $\mathbf{a}$ e de raio ϵ, é um conjunto aberto de $\mathbb{R}^n$.

(ii)... $\mathcal{U} = \{(x, y) \in \mathbb{R}^2 : y > 0\}$ é um aberto de $\mathbb{R}^2$.

♣ **Definição 2.4** ... *Um subconjunto* $\mathcal{U} \subseteq \mathbb{R}^n$*, diz-se* **fechado** *se o seu complementar* $\mathcal{U}^C \equiv \mathbb{R}^n - \mathcal{U}$ *é aberto.*

♣ **Definição 2.5** ... *Seja* $\mathcal{U}$ *um subconjunto de* $\mathbb{R}^n$*. Um ponto* $\mathbf{a} \in \mathbb{R}^n$*, diz-se um* **ponto limite***, ou um* **ponto de acumulação** *de* $\mathcal{U}$*, se toda a bola aberta centrada em* $\mathbf{a}$ *e de raio* ϵ*, contem pelo menos um ponto de* $\mathcal{U}$*, diferente de* $\mathbf{a}$*:*

$$\mathbf{a} \in \mathcal{U}, \text{ é ponto limite de } \mathcal{U} \text{ sse } \forall \epsilon > 0 : B(\mathbf{a}, \epsilon) - \{\mathbf{a}\} \cap \mathcal{U} \neq \emptyset$$

Um ponto $\mathbf{a} \in \mathcal{U}$ *que não é ponto limite de* $\mathcal{U}$*, diz-se um* **ponto isolado** *de* $\mathcal{U}$*.*

♣ **Proposição 2.1** ... *(i). Seja* $\mathcal{U}$ *um subconjunto de* $\mathbb{R}^n$*. Um ponto* $\mathbf{a} \in \mathbb{R}^n$*, é ponto limite de* $\mathcal{U}$*, se e só se existe uma sucessão de pontos* $\mathbf{u}_k \in \mathcal{U}$*, tal que* $\lim_{k\to\infty} \mathbf{u}_k = \mathbf{a}$ *e* $\mathbf{u}_k \neq \mathbf{a} \ \ \forall k$*.*

(ii). Um subconjunto $\mathcal{U} \subseteq \mathbb{R}^n$*, é fechado, se e só se contém todos os seus pontos limite.*

- Demonstração...

  (i)... Suponhamos que $\mathbf{a}$ é ponto limite de $\mathcal{U}$. Aplicando a definição, obtemos para cada $k \in \mathbf{N}$, um ponto $\mathbf{u}_k \neq \mathbf{a}$, em $\mathcal{U}$, e tal que $0 < \|u_k - \mathbf{a}\| < \frac{1}{k}$. Esta última condição implica que $\lim_{k\to\infty} \mathbf{u}_k = \mathbf{a}$, como é evidente. O recíproco é imediato.

  (ii)... Suponhamos que $\mathcal{U}$ é fechado, e que $\mathbf{a}$ é um ponto limite de $\mathcal{U}$. Se $\mathbf{a} \in \mathcal{U}^C$, então existe uma bola aberta centrada em $\mathbf{a}$, totalmente contida em $\mathcal{U}^C$ (uma vez que $\mathcal{U}^C$ é aberto), e que portanto não contém pontos de $\mathcal{U}$, o que contraria a definição de ponto limite.

  Recìprocamente, suponhamos que $\mathcal{U}$ contém todos os seus pontos limite. Se $\mathbf{b} \in \mathcal{U}^C$, então $\mathbf{b}$ não é ponto limite, e existe portanto uma bola aberta centrada em $\mathbf{a}$, que não contém pontos de $\mathcal{U}$. Logo $\mathcal{U}^C$ é aberto, e portanto $\mathcal{U}$ é fechado,

♣.

♣ **Definição 2.6** ...

(i)... Seja $\mathcal{U}$ *um subconjunto de* $\mathbb{R}^n$*. Um ponto* $\mathbf{a} \in \mathbb{R}^n$*, diz-se um* **ponto fronteira** *de* $\mathcal{U}$*, se toda a bola aberta centrada em* $\mathbf{a}$ *e de raio* ϵ*, contem pontos de* $\mathcal{U}$ *e do complementar de* $\mathcal{U}$*:*

$\mathbf{a} \in \mathbb{R}^n$*, é ponto fronteira de* $\mathcal{U}$ *sse:*

$$\forall \epsilon > 0 : B(\mathbf{a}, \epsilon) \cap \mathcal{U} \neq \emptyset \qquad e \qquad B(\mathbf{a}, \epsilon) \cap \mathcal{U}^{\mathcal{C}} \neq \emptyset$$

(ii)... O conjunto de todos os pontos fronteira de $\mathcal{U}$*, diz-se a* **fronteira** *de* $\mathcal{U}$*.*

♣ **Definição 2.7** ...

(i)... Um subconjunto $\mathcal{U} \subset \mathbb{R}^n$*, diz-se* **limitado** *se* $\mathcal{U}$ *está contido numa bola de raio suficientemente grande:*

$$\mathcal{U} \subset \mathbb{R}^n \ \text{ é limitado sse } \ \exists \mathbf{a} \in \mathbb{R}^n, \exists R > 0 : \mathcal{U} \subset B(\mathbf{a}, R)$$

(ii)... Um subconjunto $\mathcal{U} \subset \mathbb{R}^n$*, diz-se* **compacto***, se* $\mathcal{U}$ *é limitado e fechado.*

2.2.2 Limites e continuidade

♣ **Definição 2.8** ... *Seja $f : \mathcal{U} \subseteq \mathbb{R}^n \to \mathbb{R}$, um campo escalar, $\mathbf{p} \in \mathbb{R}^n$ um ponto limite de $\mathcal{U}$, e $l \in \mathbb{R}$ um número real.*

Diz-se que o **limite de f é l, quando $\mathbf{x}$ converge para $\mathbf{p}$**, *e nota-se por $\lim_{\mathbf{x}\to\mathbf{p}} f(\mathbf{x}) = l$, se se verifica a seguinte condição:*

$$\forall \epsilon > 0 \quad \exists \delta > 0 : \forall \mathbf{x} \in \mathcal{U},\, 0 < \|\mathbf{x} - \mathbf{p}\| < \delta \Rightarrow |f(\mathbf{x}) - l| < \epsilon \tag{2.2.3}$$

ou mais sucintamente se:

$$\lim_{\|\mathbf{x}-\mathbf{p}\|\to 0} |f(\mathbf{x}) - l| = 0 \tag{2.2.4}$$

♣ **Definição 2.9** ... *Seja $f : \mathcal{U} \subseteq \mathbb{R}^n \to \mathbb{R}$, um campo escalar, e $\mathbf{p} \in \mathcal{U}$ um ponto limite de $\mathcal{U}$.*

Diz-se que f é **contínua em $\mathbf{p}$**, *se se verifica a seguinte condição:*

$$\lim_{\mathbf{x}\to\mathbf{p}} f(\mathbf{x}) = f(\mathbf{p}) \tag{2.2.5}$$

f diz-se contínua em $\mathcal{U}$, se todo o ponto de $\mathcal{U}$ é ponto limite de $\mathcal{U}$, e se f é contínua em todo o ponto de $\mathcal{U}$.

Exemplo ...

Consideremos o campo escalar definido por:

$$f(x,y) = \begin{cases} 0 & \text{se } (x,y) = (0,0) \\ \frac{2x^3-y^3}{x^2+y^2} & \text{se } (x,y) \neq (0,0) \end{cases}$$

Como:

$$\begin{aligned} |2x^3 - y^3| &\leq 2|x|^3 + |y|^3 = 2|x||x|^2 + |y||y|^2 \\ &\leq 2\|\mathbf{x}\||x|^2 + 2\|\mathbf{x}\||y|^2 = 2\|\mathbf{x}\|(|x|^2 + |y|^2) = 2\|\mathbf{x}\|^3 \end{aligned}$$

onde $\mathbf{x} = (x,y)$ e usamos as desigualdades $|x| \leq \|\mathbf{x}\|$, $|y| \leq \|\mathbf{x}\|$. Portanto:

$$\frac{2x^3 - y^3}{x^2 + y^2} \leq \frac{2\|\mathbf{x}\|^3}{\|\mathbf{x}\|^2} = \|\mathbf{x}\|$$

o que demonstra imediatamente que:

$$\lim_{\mathbf{x}=(x,y)\to(0,0)} \frac{2x^3 - y^3}{x^2 + y^2} = 0 = f(0,0)$$

isto é, f é contínua em $\mathbf{O} = (0,0)$.

Exemplo ...

Consideremos o campo escalar definido por:

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

A restrição de f à recta de equação $y = \frac{1}{2}x$, é dada por (com excepção da origem $\mathbf{O} = (0,0)$):

$$f(x, y = \frac{1}{2}x) \equiv \frac{2}{5}$$

Como $f(0,0) = 0$, f não pode ser contínua em $\mathbf{O} = (0,0)$. Aliás o limite $\lim_{\mathbf{x}=(x,y)\to(0,0)} f$ não existe sequer.

2.3 Diferenciabilidade

2.3.1 Derivadas direccionais e derivadas parciais

Comecemos com um exemplo para motivar as definições que daremos em breve. Consideremos uma função $f : \mathbb{R}^2 \to \mathbb{R}$. Como já vimos, o respectivo gráfico $\mathbf{gr}\, f$ é o subconjunto de $\mathbb{R}^3$:

$$S \equiv \mathbf{gr}\, f = \{(x, y, z) \in \mathbb{R}^3 :\ z = f(x, y)\}$$

que representa uma "superfície" em $\mathbb{R}^3$, situada "sobre" o plano (x, y) (ver a figura 2.8).

Figure 2.8: A "superfície" $S \equiv \mathbf{gr}\, f = \{(x, y, z) \in \mathbf{R}^3 :\ z = f(x, y)\}$

.

Consideremos um vector não nulo $\mathbf{v} \neq \mathbf{O}$ no plano $\mathbb{R}^2$, das coordenadas (x, y), e um ponto qualquer $\mathbf{p}$ também nesse plano (e no domínio de f).

A recta que passa em $\mathbf{p}$ e é paralela a $\mathbf{v}$, consiste dos pontos de $\mathbb{R}^2$, da forma:

$$\{\mathbf{p} + t\mathbf{v} :\ t \in \mathbb{R}\}$$

e a intersecção do plano vertical (paralelo ao eixo dos zz), que contém esta recta, com $S = \mathbf{gr}\, f$, é uma curva análoga ao gráfico da função ϕ, real de variável real, definida por (ver a figura 2.8):

$$\phi(t) \equiv f(\mathbf{p} + t\mathbf{v}) \quad t \in \mathbb{R} \tag{2.3.1}$$

Esta curva está contida na "superfície" $S \equiv \mathbf{gr}\, f$. Por isso uma medida da "suavidade" dessa "superfície", no ponto $(\mathbf{p}, f(\mathbf{p}))$, e na direcção do vector $\mathbf{v}$, é dada pela existência da derivada $\phi'(0)$. Se esta derivada existe, ela representa a variação instantânea da restrição da função f, à recta acima descrita. Isto motiva a seguinte definição:

♣ **Definição 2.10** ... *Seja* $f : \mathcal{U} \subseteq \mathbb{R}^n$ *um campo escalar definido num subconjunto* $\mathcal{U}$ *de* $\mathbb{R}^n$, $\mathbf{p}$ *um ponto interior de* $\mathcal{U}$, *e* $\mathbf{v} \neq \mathbf{O}$ *um vector de* $\mathbb{R}^n$.

Define-se a **derivada direccional de f, em $\mathbf{p}$, na direcção de $\mathbf{v}$**, *notada por* $D_{\mathbf{v}} f(\mathbf{p})$, *através de:*

$$\boxed{D_{\mathbf{v}} f(\mathbf{p}) = \lim_{t\to 0} \frac{f(\mathbf{p}+t\mathbf{v})-f(\mathbf{p})}{t}} \tag{2.3.2}$$

Portanto $D_{\mathbf{v}} f(\mathbf{p}) = \phi'(0)$, onde ϕ é definida por (2.3.1).

De especial interesse é o caso em que $\mathbf{v} = \mathbf{e}_i$, onde $\mathbf{e}_i, i = 1, ..., n$, é a base canónica de $\mathbb{R}^n$. Neste caso, $D_{\mathbf{e}_i} f(\mathbf{p})$ diz-se a **i-derivada parcial de f em $\mathbf{p}$**, e nota-se por $\partial_i f(\mathbf{p})$, ou por $\frac{\partial f}{\partial x^i}(\mathbf{p})$:

$$\frac{\partial f}{\partial x^i}(\mathbf{p}) = \partial_i f(\mathbf{p}) = D_{\mathbf{e}_i} f(\mathbf{p})$$

Se $\mathbf{x} = (x^1, x^2, ..., x^n)$, então:

$$\begin{aligned}\frac{\partial f}{\partial x^i}(\mathbf{x}) = \partial_i f(\mathbf{x}) &= \lim_{t\to 0} \frac{f(\mathbf{x}+t\mathbf{e}_i) - f(\mathbf{x})}{t} \\ &= \lim_{t\to 0} \frac{f(x^1, ..., x^i + t, ..., x^n) - f(x^1, x^2, ..., x^n)}{t}\end{aligned} \tag{2.3.3}$$

o que significa que para calcular $\frac{\partial f}{\partial x^i}(\mathbf{x})$, devemos derivar a função f, considerando-a apenas como função de uma única variável real x^i, mantendo as outras variáveis fixas.

A noção de derivada direccional é manifestamente insuficiente. De facto, pode acontecer que uma função admita num ponto, uma derivada direccional na direcção de um qualquer vector, sem que por isso seja necessàriamente contínua nesse ponto.

Exemplo ...

Consideremos o campo escalar definido por:

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

Seja $\mathbf{v} = (a, b)$ um qualquer vector de $\mathbb{R}^2$. Temos então que, para $t \neq 0$:

$$\begin{aligned}\frac{f(\mathbf{O}+t\mathbf{v}) - f(\mathbf{O})}{t} &= \frac{f((0,0)+t(a,b)) - f((0,0))}{t} \\ &= \frac{f(ta, tb)}{t} = \frac{ab^2}{a^2+t^2b^4}\end{aligned}$$

e portanto:

$$D_{\mathbf{v}} f(\mathbf{O}) = D_{(a,b)} f((0,0)) = \begin{cases} 0 & \text{se } a = 0 \\ \frac{b^2}{a} & \text{se } a \neq 0 \end{cases}$$

Isto é, $D_{\mathbf{v}} f(\mathbf{O})$ existe para todo o $\mathbf{v} \in \mathbb{R}^2$. Por outro lado, f toma o valor constante e igual a $1/2$, quando restrita à parábola $x = y^2$ (excepto na origem), e por isso não é contínua em $\mathbf{O}$, já que $f(\mathbf{O}) = 0$.

Se não se impõe qualquer hipótese de continuidade sobre as derivadas direccionais pode acontecer que não haja qualquer ligação entre as derivadas direccionais num certo ponto, segundo os diversos vectores. Note que no exemplo anterior a aplicação $\mathbf{v} \to D_{\mathbf{v}} f(\mathbf{p})$ não é contínua.

Notemos que, se $D_{\mathbf{v}} f(\mathbf{p})$ existe, também existe $D_{\lambda\mathbf{v}} f(\mathbf{p})$, $\forall \lambda \in \mathbb{R}$, e:

$$D_{\lambda\mathbf{v}} f(\mathbf{p}) = \lambda\, D_{\mathbf{v}} f(\mathbf{p})$$

No entanto, não é verdade que, para $\mathbf{p}$ fixo, a aplicação:

$$\mathbf{v} \in \mathbb{R}^n \mapsto D_{\mathbf{v}} f(\mathbf{p})$$

seja linear, como mostra o exemplo anterior, com $\mathbf{p} = \mathbf{O}$.

O defeito da derivada direccional $D_{\mathbf{v}} f(\mathbf{p})$, reside no facto de apenas considerar o comportamento de f, ao longo das rectas que passam em $\mathbf{p}$, enquanto que uma boa noção de derivada, deve reflectir o comportamento global de f, em toda uma vizinhança de $\mathbf{p}$.

Por todos estes motivos, somos conduzidos à noção de diferencial, que a seguir trataremos.

2.3.2 Diferencial. Gradiente

Consideremos de novo, uma função $f : \mathbb{R}^2 \to \mathbb{R}$. Como já vimos, o respectivo gráfico $\mathbf{gr}\, f$ é o subconjunto de $\mathbb{R}^3$:

$$S \equiv \mathbf{gr}\, f = \{(x, y, z) \in \mathbb{R}^3 : z = f(x, y)\}$$

que representa uma "superfície" em $\mathbb{R}^3$, situada "sobre" o plano (x, y) (ver a figura 2.9).

Figure 2.9: A "superfície" $S \equiv \mathbf{gr}\, f$, e o plano tangente

.

Uma medida da suavidade desta "superfície", sobre uma vizinhança de um ponto $\mathbf{p} \in \mathbb{R}^2$, é dada pela existência de um plano tangente, que passe no ponto $(\mathbf{p}, f(\mathbf{p})) \in S$, e que seja uma aproximação óptima de S, numa vizinhança de $\mathbf{p}$.

Um tal plano, se existir, pode ser representado como o gráfico de uma forma afim $T : \mathbb{R}^2 \to \mathbb{R}$, tal que $T(\mathbf{p}) = f(\mathbf{p})$. Se $\mathbf{x} \in \mathbb{R}^2$, está próximo de $\mathbf{p}$, então a diferença:

$$f(\mathbf{x}) - T(\mathbf{x})$$

representa o "desvio" entre o valor exacto $f(\mathbf{x})$, avaliado em $S = \mathbf{gr}\, f$, e o "valor aproximado" $T(\mathbf{x})$, avaliado no plano $\mathbf{gr}\, T$.

Quando este desvio converge mais ràpidamente para 0, do que $h = \|\mathbf{x}-\mathbf{p}\|$, diz-se que $\mathbf{f}$ é diferenciável em $\mathbf{p}$. Mais formalmente:

♣ **Definição 2.11** ... *Seja $f : \mathcal{U} \to \mathbb{R}$ um campo escalar, definido num subconjunto aberto $\mathcal{U} \subseteq \mathbb{R}^n$. Fixemos um qualquer ponto $\mathbf{p} \in \mathcal{U}$.*

Diz-se que $\mathbf{f}$ *é* **diferenciável (ou derivável) em p**, *se existe uma função afim $T_{\mathbf{p}} : \mathbb{R}^n \to \mathbb{R}$ (que depende de f e de $\mathbf{p}$), tal que:*

$$T_{\mathbf{p}}(\mathbf{p}) = f(\mathbf{p}) \tag{2.3.4}$$

e que satisfaz a condição:

$$\lim_{\|\mathbf{x}-\mathbf{p}\|\to 0} \frac{|f(\mathbf{x}) - T_{\mathbf{p}}(\mathbf{x})|}{\|\mathbf{x}-\mathbf{p}\|} = 0 \tag{2.3.5}$$

Neste caso, a função afim $T_{\mathbf{p}}$ diz-se a **aproximação afim óptima** *de* $\mathbf{f}$ *em* $\mathbf{p}$, *e o seu gráfico $\mathbf{gr}\, T_{\mathbf{p}}$, diz-se o* **hiperplano tangente** *a $S = \mathbf{gr}\, f$, no ponto $(\mathbf{p}, f(\mathbf{p}))$.*

A parte linear de $T_{\mathbf{p}}$, diz-se a **diferencial** *de* $\mathbf{f}$ *em* $\mathbf{p}$, *e nota-se por $df_{\mathbf{p}}$. Portanto a diferencial $df_{\mathbf{p}}$, é uma forma linear:*

$$df_{\mathbf{p}} : \mathbf{h} \in \mathbb{R}^n \mapsto df_{\mathbf{p}}(\mathbf{h}) \in \mathbb{R}$$

$\mathbf{f}$ *diz-se* **diferenciável** *em $\mathcal{U}$, se o é em todo o ponto $\mathbf{p} \in \mathcal{U}$, e neste caso diz-se que $S = \mathbf{gr}\, f$ é uma* **hipersuperfície** *(ou uma* **variedade diferenciável** *de dimensão n) em $\mathbb{R}^{n+1}$, de equação $z = f(\mathbf{x})$, $\mathbf{x} \in \mathcal{U}$.*

É fácil ver que se existe uma função afim $T_{\mathbf{p}}$, que satisfaz (2.3.5), então ela é única, e portanto a diferencial $df_{\mathbf{p}}$, está univocamente determinada.

Como $df_{\mathbf{p}}$ é a parte linear de $T_{\mathbf{p}}$, $T_{\mathbf{p}}$ é da forma:

$$T_{\mathbf{p}}(\mathbf{x}) = df_{\mathbf{p}}(\mathbf{x}) + c$$

Pondo $\mathbf{x} = \mathbf{p} + \mathbf{h}$, com $h = \|\mathbf{h}\| = \|\mathbf{x}-\mathbf{p}\|$, e atendendo a que $T_{\mathbf{p}}(\mathbf{p}) = f(\mathbf{p})$, podemos escrever que:

$$T_{\mathbf{p}}(\mathbf{p}+\mathbf{h}) = df_{\mathbf{p}}(\mathbf{h}) + f(\mathbf{p}) \tag{2.3.6}$$

e o limite (2.3.5), pode então ser escrito na forma:

$$\lim_{h=\|\mathbf{h}\|\to 0} \frac{|f(\mathbf{p}+\mathbf{h}) - f(\mathbf{p}) - df_{\mathbf{p}}(\mathbf{h})|}{\|\mathbf{h}\|} = 0 \tag{2.3.7}$$

ou ainda na forma:

$$\boxed{f(\mathbf{p}+\mathbf{h}) = f(\mathbf{p}) + df_{\mathbf{p}}(\mathbf{h}) + o(\|\mathbf{h}\|)} \tag{2.3.8}$$

onde $\lim_{\|\mathbf{h}\|\to 0} \frac{o(\|\mathbf{h}\|)}{\|h\|} = 0$. Esta última fórmula diz-se a **fórmula de Taylor de primeira ordem** para f, em $\mathbf{p}$.

Podemos portanto dar a seguinte definição alternativa de diferenciabilidade de um campo escalar:

♣ **Definição 2.12** ... *Seja $f : \mathcal{U} \to \mathbb{R}$ um campo escalar, definido num subconjunto aberto $\mathcal{U} \subseteq \mathbb{R}^n$. Fixemos um qualquer ponto $\mathbf{p} \in \mathcal{U}$.*

Diz-se que f é **diferenciável (ou derivável) em p**, *se existe uma aplicação linear:*

$$df_{\mathbf{p}} : \mathbb{R}^n \to \mathbb{R}$$

(que depende de f e de $\mathbf{p}$), tal que:

$$\boxed{\lim_{\|\mathbf{h}\|\to 0} \frac{|f(\mathbf{p}+\mathbf{h}) - f(\mathbf{p}) - df_{\mathbf{p}}(\mathbf{h})|}{\|\mathbf{h}\|} = 0} \tag{2.3.9}$$

Esta aplicação linear $df_{\mathbf{p}} : \mathbb{R}^n \to \mathbb{R}$ (que é única), diz-se a **diferencial** *de f em $\mathbf{p}$.*

f diz-se **diferenciável** *em $\mathcal{U}$, se o é em todo o ponto $\mathbf{p} \in \mathcal{U}$, e neste caso diz-se que $S = \mathbf{gr}\, f$ é uma* **variedade diferenciável** *de dimensão n, em $\mathbb{R}^{n+1}$, de equação $z = f(\mathbf{x})$, $\mathbf{x} \in \mathcal{U}$.*

Exemplo ...

Se $f = L : \mathbb{R}^n \to \mathbb{R}$ é uma aplicação linear, então L é diferenciável em todo o ponto $\mathbf{p} \in \mathbb{R}^n$ e:

$$dL_{\mathbf{p}} = L \qquad \forall \mathbf{p} \in \mathbb{R}^n \tag{2.3.10}$$

Com efeito, o numerador em (2.3.7) é neste caso igual a (pondo $df_{\mathbf{p}} = dL_{\mathbf{p}} = L$):

$$|f(\mathbf{p}+\mathbf{h}) - f(\mathbf{p}) - df_{\mathbf{p}}(\mathbf{h})| = |L(\mathbf{p}+\mathbf{h}) - L(\mathbf{p}) - L(\mathbf{h})| = 0$$

uma vez que estamos a supôr que L é linear (e portanto, $L(\mathbf{p}+\mathbf{h}) = L(\mathbf{p}) + L(\mathbf{h})$).

Exemplo ...

Seja $f(\mathbf{x}) = \mathbf{S}(\mathbf{x}) \cdot \mathbf{x}$, $\mathbf{x} \in \mathbb{R}^n$ uma forma quadrática, onde $\mathbf{S} : \mathbb{R}^n \to \mathbb{R}^n$ é uma aplicação linear simétrica.

Então f é diferenciável em todo o ponto $\mathbf{x} \in \mathbb{R}^n$, e a diferencial $df_{\mathbf{x}}$, é dada por:

$$df_{\mathbf{x}}(\mathbf{h}) = \mathbf{S}(\mathbf{x}) \cdot \mathbf{h} + \mathbf{S}(\mathbf{h}) \cdot \mathbf{x} \tag{2.3.11}$$

Com efeito, substituindo (2.3.11) no numerador de (2.3.7), obtemos (com $\mathbf{p} = \mathbf{x}$):

$$\begin{aligned}
|f(\mathbf{x}+\mathbf{h}) - f(\mathbf{x}) - df_{\mathbf{x}}(\mathbf{h})| &= |\mathbf{S}(\mathbf{x}+\mathbf{h}) \cdot (\mathbf{x}+\mathbf{h}) - \mathbf{S}(\mathbf{x}) \cdot \mathbf{x} - \mathbf{S}\mathbf{x} \cdot \mathbf{h} - \mathbf{S}\mathbf{h} \cdot \mathbf{x}| \\
&= |\mathbf{S}(\mathbf{x}) \cdot \mathbf{x} + \mathbf{S}(\mathbf{x}) \cdot \mathbf{h} + \mathbf{S}(\mathbf{h}) \cdot \mathbf{x} + \mathbf{S}(\mathbf{h}) \cdot \mathbf{h} - \\
&\qquad \mathbf{S}(\mathbf{x}) \cdot \mathbf{x} - (\mathbf{h}) - \mathbf{S}(\mathbf{x}) \cdot \mathbf{h} - \mathbf{S}(\mathbf{h}) \cdot \mathbf{x}| \\
&= |\mathbf{S}(\mathbf{h}) \cdot \mathbf{h}|
\end{aligned}$$

Resta agora provar que:

$$\lim_{\mathbf{h}\to\mathbf{O}} \frac{|\mathbf{S}(\mathbf{h}) \cdot \mathbf{h}|}{\|\mathbf{h}\|} = 0$$

Diagonalizando f numa base ortonormal de $\mathbb{R}^n$, obtemos:

$$f(\mathbf{h}) = \lambda_1 (h^1)^2 + \lambda_2 (h^2)^2 + \cdots + \lambda_n (h^n)^2$$

onde $h^1, ..., h^n$ são as coordenadas de $\mathbf{h}$ na base referida. Daqui se deduz que:

$$\begin{aligned} |f(\mathbf{h})| = |\mathbf{S}(\mathbf{h}) \cdot \mathbf{h}| &= |\lambda_1(h^1)^2 + \lambda_2(h^2)^2 + \cdots + \lambda_n(h^n)^2| \\ &\leq (\max_{1\leq i\leq n} |\lambda_i|)((h^1)^2 + \cdots + (h^n)^2) \\ &= M\,\|\mathbf{h}\|^2 \end{aligned}$$

onde $M = \max_{1\leq i\leq n} |\lambda_i|$. Portanto (se $\mathbf{h} \neq \mathbf{O}$):

$$\frac{|\mathbf{S}(\mathbf{h}) \cdot \mathbf{h}|}{\|\mathbf{h}\|} \leq \frac{M\,\|\mathbf{h}\|^2}{\|\mathbf{h}\|} = M\,\|\mathbf{h}\|$$

o que prova o que se pretendia,

♣.

Vejamos agora algumas propriedades da diferencial. Aplicando (2.3.8), fàcilmente se deduz a seguinte proposição:

♣ **Proposição 2.2** ... *Se f é diferenciável em $\mathbf{p}$, então f é contínua em $\mathbf{p}$.*

♣ **Proposição 2.3** ... *Suponhamos que f é diferenciável em $\mathbf{p}$, com diferencial $df_{\mathbf{p}} \in (\mathbb{R}^n)^*$. Então a derivada direccional $D_{\mathbf{v}}f(\mathbf{p})$ existe, para todo o vector $\mathbf{v} \in \mathbb{R}^n$, e tem-se que:*

$$D_{\mathbf{v}}f(\mathbf{p}) = df_{\mathbf{p}}(\mathbf{v}) \qquad \forall \mathbf{v} \in \mathbb{R}^n \tag{2.3.12}$$

- Demonstração...

Se $\mathbf{v} = \mathbf{O}$, então $D_{\mathbf{O}}f(\mathbf{p}) = 0 = df_{\mathbf{p}}(\mathbf{O})$. Podemos por isso supôr que $\mathbf{v} \neq \mathbf{O}$. Façamos $\mathbf{h} = t\mathbf{v}$ na fórmula de Taylor (2.3.8), com $t \neq 0$, para obter:

$$f(\mathbf{p} + t\mathbf{v}) - f(\mathbf{p}) = df_{\mathbf{p}}(t\mathbf{v}) + o(|t|\|\mathbf{v}\|) = t\,df_{\mathbf{p}}(\mathbf{v}) + o(|t|\|\mathbf{v}\|)$$

Dividamos por t, e façamos $t \to 0$, para obter (2.3.12).

Para provar (2.3.13) basta utilizar a linearidade de $df_{\mathbf{p}}$, a fórmula (2.3.12) e a definição de derivada parcial,

♣.

Se $\mathbf{v} = \sum_{i=1}^n v^i \mathbf{e}_i = (v^1, ..., v^n)$, então:

$$\boxed{df_{\mathbf{p}}(\mathbf{v}) = \sum_{i=1}^n v^i \partial_i f(\mathbf{p}) = \sum_{i=1}^n v^i \frac{\partial f}{\partial x^i}(\mathbf{p})} \tag{2.3.13}$$

Como as funções coordenadas $x^i : \mathbb{R}^n \to \mathbb{R}$, são lineares sabemos que $dx^i = x^i$, i.e., $dx^i(\mathbf{v}) = v^i$, $\forall \mathbf{v} \in \mathbb{R}^n$. Daí que a fórmula (2.3.13) se escreva usualmente na forma:

$$\boxed{df_{\mathbf{p}}(\mathbf{v}) = \sum_{i=1}^n \frac{\partial f}{\partial x^i}(\mathbf{p})dx^i(\mathbf{v})} \tag{2.3.14}$$

ou mais sucintamente:

$$\boxed{df = \sum_{i=1}^n \frac{\partial f}{\partial x^i}\, dx^i} \tag{2.3.15}$$

Como $df_{\mathbf{p}}$ é uma forma linear em $\mathbb{R}^n$, sabemos pela proposição 2.8.1, que existe um único vector, que notamos por $\nabla f(\mathbf{p}) \in \mathbb{R}^n$, tal que:

$$\boxed{df_{\mathbf{p}}(\mathbf{v}) = \nabla f(\mathbf{p}) \cdot \mathbf{v} \quad \forall \mathbf{v} \in \mathbb{R}^n} \tag{2.3.16}$$

♣ **Definição 2.13** ... ... *Suponhamos que f é diferenciável em $\mathbf{p}$, com diferencial $df_{\mathbf{p}} \in (\mathbb{R}^n)^*$.*

Ao único vector $\nabla f(\mathbf{p}) \in \mathbb{R}^n$, que verifica (2.3.16), chama-se **gradiente de f em $\mathbf{p}$**.

Se f é diferenciável num aberto $\mathcal{U} \subseteq \mathbb{R}^n$, então o campo de vectores:

$$\mathbf{p} \in \mathcal{U} \mapsto \nabla f(\mathbf{p}) \in \mathbb{R}^n$$

diz-se o **campo de gradientes** *do campo escalar f, em $\mathcal{U}$.*

A fórmula (2.3.13):

$$df_{\mathbf{p}}(\mathbf{v}) = \sum_{i=1}^n v^i \partial_i f(\mathbf{p}) = \sum_{i=1}^n v^i \frac{\partial f}{\partial x^i}(\mathbf{p})$$

mostra que as componentes do vector gradiente $\nabla f(\mathbf{p})$, na base canónica de $\mathbb{R}^n$, são:

$$\nabla f(\mathbf{p}) = \Big(\frac{\partial f}{\partial x^1}(\mathbf{p}), \frac{\partial f}{\partial x^2}(\mathbf{p}), ..., \frac{\partial f}{\partial x^n}(\mathbf{p})\Big) \tag{2.3.17}$$

Portanto:

$$df_{\mathbf{p}}(\mathbf{v}) = \nabla f(\mathbf{p}) \cdot \mathbf{v} \qquad \forall \mathbf{v} \in \mathbb{R}^n$$

Notando as coordenadas cartesianas usuais em $\mathbb{R}^{n+1} = \mathbb{R}^n \times \mathbb{R}$, por $(\mathbf{x} = (x^1, \cdots, x^n), x^{n+1})$, podemos escrever a equação cartesiana do hiperplano tangente à hipersuperfície $\mathbf{gr}\, f$, no ponto $(\mathbf{p}, f(\mathbf{p}))$, numa das formas seguintes:

$$\begin{aligned} x^{n+1} &= df_{\mathbf{p}}(\mathbf{x} - \mathbf{p}) + f(\mathbf{p}) \\ &= \nabla f(\mathbf{p}) \cdot (\mathbf{x} - \mathbf{p}) + f(\mathbf{p}) \\ &= \frac{\partial f}{\partial x^1}(\mathbf{p})\,(x^1 - p^1) + \cdots + \frac{\partial f}{\partial x^n}(\mathbf{p})\,(x^n - p^n) + f(\mathbf{p}) \end{aligned} \tag{2.3.18}$$

onde $\mathbf{p} = (p^1, ..., p^n)$.

De facto a aproximação afim óptima de f em $\mathbf{p}$, é dada por (pondo $\mathbf{p} + \mathbf{h} = \mathbf{x}$ em (2.3.6)):

$$T_{\mathbf{p}}(\mathbf{x}) = df_{\mathbf{p}}(\mathbf{x} - \mathbf{p}) + f(\mathbf{p})$$

e o hiperplano tangente à hipersuperfície $\mathbf{gr}\, f$, no ponto $(\mathbf{p}, f(\mathbf{p}))$, é o gráfico desta aplicação $T_{\mathbf{p}}$.

Exemplo ...

Para uma função diferenciável de duas variáveis, $f : \mathcal{U} \subseteq \mathbb{R}^2 \to \mathbb{R}$, o plano tangente ao gráfico de f, no ponto $(a, b, f(a, b))$, $\mathbf{p} = (a, b) \in \mathcal{U}$, tem por equação cartesiana:

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

Assim por exemplo, se $f(x, y) = x^3 + y^5$, então $\frac{\partial f}{\partial x} = 3x^2$ e $\frac{\partial f}{\partial y} = 5y^4$. Portanto, por exemplo no ponto $\mathbf{p} = (a, b) = (-1, 3)$, onde f vale $f(\mathbf{p}) = 242$, temos que:

$$\frac{\partial f}{\partial x}(-1,3) = 3 \qquad \frac{\partial f}{\partial y}(-1,3) = 405$$

e o plano tangente ao $\mathbf{gr}\, f$, no ponto de coordenadas $(-1, 3, 242)$, é o plano em $\mathbb{R}^3$, de equação cartesiana:

$$z = \frac{\partial f}{\partial x}(-1,3)\,(x-a) + \frac{\partial f}{\partial y}(-1,3)\,(y-b) + f(-1,3)$$

isto é:

$$z = 3(x+1) + 405(y-3) + 242$$

Vejamos qual o significado do vector gradiente $\nabla f(\mathbf{p})$. Seja $f : \mathcal{U} \subset \mathbb{R}^n \to \mathbb{R}$ um campo escalar, definido e diferenciável no aberto $\mathcal{U}$, e $\mathbf{p} \in \mathcal{U}$ um ponto onde $\nabla f(\mathbf{p}) \neq \mathbf{O}$. Suponhamos que se pretende determinar a direcção, definida por um vector unitário $\mathbf{u}$, segundo a qual f cresce mais ràpidamente, em $\mathbf{p}$.

Se designamos por θ, o ângulo entre $\mathbf{u}$ e $\nabla f(\mathbf{p})$, então:

$$df_{\mathbf{p}}(\mathbf{u}) = \nabla f(\mathbf{p}) \cdot \mathbf{u} = \|\nabla f(\mathbf{p})\| \cos\theta$$

Como $\cos\theta$ atinge o seu valor máximo, igual a $+1$, quando $\theta = 0$, isto é quando $\nabla f(\mathbf{p})$ e $\mathbf{u}$, são colineares, concluímos que $\|\nabla f(\mathbf{p})\|$ é o valor máximo de $\mathbf{u} \mapsto df_{\mathbf{p}}(\mathbf{u})$, quando $\mathbf{u}$ é um vector unitário. Além disso, esse valor máximo é atingido em $\mathbf{u} = \frac{\nabla f(\mathbf{p})}{\|\nabla f(\mathbf{p})\|}$. Fica assim demonstrada a seguinte proposição;

♣ **Proposição 2.4** ... *Seja $f : \mathcal{U} \subset \mathbb{R}^n \to \mathbb{R}$ um campo escalar, definido e diferenciável no aberto $\mathcal{U}$, e $\mathbf{p} \in \mathcal{U}$ um ponto onde $\nabla f(\mathbf{p}) \neq \mathbf{O}$.*

Então o valor máximo da função :

$$\mathbf{u} \mapsto df_{\mathbf{p}}(\mathbf{u}) = \nabla f(\mathbf{p}) \cdot \mathbf{u}$$

(onde $\|\mathbf{u}\| = 1$, e $\mathbf{p}$ está fixo), é igual a $\|\nabla f(\mathbf{p})\|$, e é atingido em $\mathbf{u} = \frac{\nabla f(\mathbf{p})}{\|\nabla f(\mathbf{p})\|}$.

♣ **Definição 2.14** ... *Um campo escalar $f : \mathcal{U} \subset \mathbb{R}^n \to \mathbb{R}$, definido no aberto $\mathcal{U}$, diz-se de classe C^1 em $\mathcal{U}$, se todas as derivadas parciais $\frac{\partial f}{dx^i}(\mathbf{p})$ existem e são contínuas em $\mathcal{U}$.*

♣ **Proposição 2.5** ... *Se $f : \mathcal{U} \subset \mathbb{R}^n \to \mathbb{R}$, é de classe C^1 em $\mathcal{U}$, então f é diferenciável em todo o ponto de $\mathcal{U}$.*

- (*)Demonstração...

É claro que o candidato natural para $df_{\mathbf{p}}$, é a aplicação $\mathbf{h} \mapsto \nabla f(\mathbf{p}) \cdot \mathbf{h}$. Portanto devemos provar que:

$$\lim_{h=\|\mathbf{h}\|\to 0} \frac{|f(\mathbf{p}+\mathbf{h}) - f(\mathbf{p}) - \nabla f(\mathbf{p}) \cdot \mathbf{h})|}{\|\mathbf{h}\|} = 0 \tag{2.3.19}$$

Façamos $\mathbf{p} = (p^1, ..., p^n)$, $\mathbf{h} = (h^1, ..., h^n)$, e para cada $k = 1, 2, ..., n$, $\mathbf{h}_k = (h^1, ..., h^k, 0, ..., 0)$ (e ainda $\mathbf{h}_0 = \mathbf{O}$). Portanto, $\mathbf{p} + \mathbf{h}_0 = \mathbf{p}$ e $\mathbf{p} + \mathbf{h}_n = \mathbf{p} + \mathbf{h}$.

Temos então que:

$$f(\mathbf{p}+\mathbf{h}) - f(\mathbf{p}) = \sum_{k=1}^{n} ((f(\mathbf{p}+\mathbf{h}_k) - f(\mathbf{p}+\mathbf{h}_{k-1}))$$

já que os termos sucessivos se cancelam, deixando apenas o primeiro e o último. Mas $\mathbf{p} + \mathbf{h}_k$ e $\mathbf{p} + \mathbf{h}_{k-1}$ diferem apenas na componente k, e por isso, podemos aplicar o teorema do valor médio do cálculo de uma variável, para obter:

$$f(\mathbf{p}+\mathbf{h}_k) - f(\mathbf{p}+\mathbf{h}_{k-1}) = \frac{\partial f}{\partial x^k}(\mathbf{z}_k)\, h^k$$

onde $\mathbf{z}_k$ é um ponto no segmento de recta que une $\mathbf{p} + \mathbf{h}_{k-1}$ a $\mathbf{p} + \mathbf{h}_k$. Somando, vem que:

$$f(\mathbf{p}+\mathbf{h}) - f(\mathbf{p}) = \sum_{k=1}^{n} \frac{\partial f}{\partial x^k}(\mathbf{z}_k)\, h^k$$

Como:

$$\nabla f(\mathbf{p}) \cdot \mathbf{h} = \sum_{k=1}^{n} \frac{\partial f}{\partial x^k}(\mathbf{p})\, h^k$$

vem que:

$$\begin{aligned} |f(\mathbf{p}+\mathbf{h}) - f(\mathbf{p}) - \nabla f(\mathbf{p}) \cdot \mathbf{h})| &= |\sum_{k=1}^{n} \Big(\frac{\partial f}{\partial x^k}(\mathbf{z}_k) - \frac{\partial f}{\partial x^k}(\mathbf{p})\Big)\, h^k \,| \\ &\leq \sum_{k=1}^{n} |\, \frac{\partial f}{\partial x^k}(\mathbf{z}_k) - \frac{\partial f}{\partial x^k}(\mathbf{p}) \,|\; |h^k| \end{aligned} \tag{2.3.20}$$

onde utilizamos a desigualdade triangular e o facto de que:

$$|h^k| \leq \|\mathbf{h}\| \qquad k = 1, ..., n$$

Finalmente, como as derivadas parciais são por hipótese contínuas em $\mathbf{p}$ e como $\mathbf{h}_k \to \mathbf{O}$ quando $\mathbf{h} \to \mathbf{O}$, obtemos o limite (2.3.19),

♣.

Para que fique clara a relação entre os diversos conceitos anteriormente definidos, consideremos as condições seguintes, sobre um campo escalar $f : \mathcal{U} \subset \mathbb{R}^n \to \mathbb{R}$, definido num aberto $\mathcal{U}$:

(**A**)... f admite derivadas parciais $\frac{\partial f}{\partial x^i}(\mathbf{p})$ $(i = 1, ..., n)$, num ponto $\mathbf{p} \in \mathcal{U}$, contínuas em $\mathbf{p}$.

(**B**)... f é diferenciável em $\mathbf{p}$.

(**C**)... f admite em $\mathbf{p}$, derivadas direccionais segundo um qualquer vector $\mathbf{v}$.

(**D**)... f admite derivadas parciais $\frac{\partial f}{\partial x^i}(\mathbf{p})$ $(i = 1, ..., n)$, em $\mathbf{p}$.

(**E**)... f é contínua em $\mathbf{p}$.

(**F**)... f é contínua em $\mathbf{p}$, relativamente a cada uma das suas variáveis.

Temos então o seguinte quadro de implicações:

$$\begin{array}{ccccccc} & & & & (\mathbf{C}) & \Longrightarrow & (\mathbf{D}) \\ (\mathbf{A}) & \Longrightarrow & (\mathbf{B}) & & & & \Downarrow \\ & & & & (\mathbf{E}) & \Longrightarrow & (\mathbf{F}) \end{array}$$

No entanto, nenhuma destas implicações é uma equivalência, como mostram os exemplos seguintes.

Exemplo ...

Consideremos o campo escalar definido por:

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

f não é contínua em $(0, 0)$. Existem as derivadas parciais em $(0, 0)$, e são iguais a: $\frac{\partial f}{\partial x}(0, 0) = 0$ e $\frac{\partial f}{\partial y}(0, 0) = 0$.

Mas f não é diferenciável em $(0, 0)$ (por nem sequer ser contínua nasse ponto).

Exemplo ...

Consideremos o campo escalar definido por:

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

Neste caso, f é contínua em $(0, 0)$, $\frac{\partial f}{\partial x}(0, 0)$ não existe, e $\frac{\partial f}{\partial y}(0, 0) = 0$.

Mas f não é diferenciável em $(0, 0)$ (porque $\frac{\partial f}{\partial x}(0, 0)$ não existe).

Exemplo ...

Consideremos o campo escalar definido por:

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

Neste caso, f é contínua em $(0,0)$, as derivadas parciais em $(0,0)$, existem e são iguais a: $\frac{\partial f}{\partial x}(0,0) = 0$ e $\frac{\partial f}{\partial y}(0,0) = 0$.

f admite derivadas direccionais em $(0,0)$, segundo um qualquer vector $\mathbf{v} = (a,b) \in \mathbb{R}^2$.

No entanto, f não é diferenciável em $(0,0)$!

Exemplo ...

Consideremos o campo escalar definido por:

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

Neste caso, f é contínua em $(0,0)$, as derivadas parciais em $(0,0)$, existem e são iguais a: $\frac{\partial f}{\partial x}(0,0) = 0$ e $\frac{\partial f}{\partial y}(0,0) = 0$.

Mas as derivadas parciais $\frac{\partial f}{\partial x}$ e $\frac{\partial f}{\partial y}$ não são contínuas em $(0,0)$.

No entanto, f é diferenciável em $(0,0)$!

Exemplo ...

Consideremos o campo escalar definido por:

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

Neste caso, f é contínua em $(0,0)$, mas f não admite derivadas direccionais segundo um qualquer vector não nulo $\mathbf{v} \in \mathbb{R}^2$!

Evidentemente que f não é diferenciável em $(0,0)$.

Exemplo ...

Consideremos o campo escalar definido por:

$$f(x,y) = \begin{cases} 0 & \text{se } (x,y) = (0,0) \\ \frac{xy^3+x^3y}{x^4+y^4} & \text{se } (x,y) \neq (0,0) \end{cases}$$

Neste caso, f é contínua em $(0,0)$, relativamente a cada uma das variáveis x e y, mas f não é contínua em $(0,0)$.

As derivadas parciais em $(0,0)$, existem e são iguais a: $\frac{\partial f}{\partial x}(0,0) = 0$ e $\frac{\partial f}{\partial y}(0,0) = 0$.

Não existe a derivada direccional $D_{\mathbf{v}}f(0,0)$, sempre que $\mathbf{v} = (a,b)$ satisfaz a condição "$a \neq 0$ ou $b \neq 0$".

Portanto, f não é diferenciável em $(0,0)$.

2.4 Regra da Cadeia I. Hipersuperfícies regulares e Hiperplanos tangentes

♣ **Proposição 2.6** (**Regra da Cadeia I**) ... *Seja $f : \mathcal{U} \subset \mathbb{R}^n \to \mathbb{R}$, um campo escalar, definido num aberto $\mathcal{U}$, e $\alpha : \mathbf{I} \subseteq \mathbb{R} \to \mathcal{U}$ uma curva em $\mathcal{U}$.*

Se a curva α é derivável em $t \in \mathbf{I}$, e se f é diferenciável em $\alpha(t) \in \mathcal{U}$, então $f \circ \alpha : \mathbf{I} \to \mathbb{R}$ é derivável em t e:

$$\begin{aligned} (f \circ \alpha)'(t) &= df_{\alpha(t)}(\dot{\alpha}(t)) \\ &= \nabla f(\alpha(t)) \cdot \dot{\alpha}(t) \end{aligned} \tag{2.4.1}$$

- Demonstração...

  Pondo $\mathbf{p} = \alpha(t)$, existe uma bola aberta $B(\mathbf{p}, r)$ totalmente contida em $\mathcal{U}$, e $\alpha(t+\epsilon)$ estará em $B(\mathbf{p}, r) \subset \mathcal{U}$, desde que $\epsilon \neq 0$ seja suficientemente pequeno.

  Seja $\mathbf{h} = \alpha(t+\epsilon) - \alpha(t)$, de tal forma que $\alpha(t+\epsilon) = \alpha(t) + \mathbf{h} = \mathbf{p} + \mathbf{h}$. Temos então que, pela fórmula de Taylor (2.3.8):

$$\begin{aligned} (f \circ \alpha)(t+\epsilon) - (f \circ \alpha)(t) &= f(\mathbf{p} + \mathbf{h}) - f(\mathbf{p}) \\ &= \nabla f(\mathbf{p}) \cdot \mathbf{h} + o(\|\mathbf{h}\|) \end{aligned}$$

  onde $\lim_{\|\mathbf{h}\| \to 0} \frac{o(\|\mathbf{h}\|)}{\|\mathbf{h}\|} = 0$.

  Portanto, como $\mathbf{h} = \alpha(t+\epsilon) - \alpha(t)$, obtemos:

$$\frac{(f \circ \alpha)(t+\epsilon) - (f \circ \alpha)(t)}{\epsilon} = \nabla f(\mathbf{p}) \cdot \frac{\alpha(t+\epsilon) - \alpha(t)}{\epsilon} + \frac{o(\|\mathbf{h}\|)}{\epsilon}$$

  Finalmente, fazendo $\epsilon \to 0$ (o que implica que $\|\mathbf{h}\| \to 0$), obtemos:

$$\begin{aligned} (f \circ \alpha)'(t) &= \nabla f(\alpha(t)) \cdot \dot{\alpha}(t) \\ &= df_{\alpha(t)}(\dot{\alpha}(t)) \end{aligned}$$

  como se pretendia,

♣.

Pondo $\alpha(t) = (x^1(t), \cdots, x^n(t))$, é usual escrever a regra da cadeia I (2.4.1), na forma:

$$\boxed{\frac{d}{dt} f(x^1(t), \cdots, x^n(t)) = \frac{\partial f}{\partial x^1}(\alpha(t)) \frac{dx^1}{dt} + \cdots + \frac{\partial f}{\partial x^n}(\alpha(t)) \frac{dx^n}{dt}} \tag{2.4.2}$$

A regra da cadeia (2.4.1), pode ser aplicada para calcular a diferencial de uma função diferenciável, como se mostra nos exemplos seguintes.

Exemplo ...

Seja $f(\mathbf{x}) = \mathbf{S}\mathbf{x} \cdot \mathbf{x}$ uma forma quadrática. Para calcular $df_{\mathbf{x}}$, podemos utilizar (2.4.1) (supondo já sabido que f é de facto diferenciável). Assim, consideremos uma curva $\alpha : \mathbf{I} \to \mathbb{R}^n$, derivável tal que $\alpha(0) = \mathbf{x}$ e $\dot{\alpha}(o) = \mathbf{v}$. Então, por (2.4.1), temos que:

$$\begin{aligned} df_{\mathbf{x}}(\mathbf{v}) &= (f \circ \alpha)'(0) \\ &= \frac{d}{dt}\Big|_{t=0} \mathbf{S}\alpha(t) \cdot \alpha(t) \\ &= \mathbf{S}\dot{\alpha}(0) \cdot \alpha(0) + \mathbf{S}\alpha(0) \cdot \dot{\alpha}(0) \\ &= \mathbf{S}\mathbf{x} \cdot \mathbf{v} + \mathbf{S}\mathbf{v} \cdot \mathbf{x} \\ &= \mathbf{S}\mathbf{x} \cdot \mathbf{v} + \mathbf{v} \cdot \mathbf{S}\mathbf{x} \qquad \text{já que } \mathbf{S} \text{ é simétrica} \\ &= 2\,\mathbf{S}\mathbf{x} \cdot \mathbf{v} \end{aligned} \tag{2.4.3}$$

Em particular, deduzimos que $\nabla f(\mathbf{x}) = 2\,\mathbf{S}\mathbf{x}$.

Exemplo ...

Seja $f(\mathbf{x}) = [\mathbf{x}, \mathbf{a}, \mathbf{b}] = \mathbf{x} \cdot \mathbf{a} \times \mathbf{b}$ um campo escalar definido em $\mathbb{R}^3$ (onde $\mathbf{a}, \mathbf{b}$ são vectores fixos em $\mathbb{R}^3$).

Uma vez mais (supondo já sabido que f é de facto diferenciável), consideremos uma curva $\alpha : \mathbf{I} \to \mathbb{R}^n$, derivável tal que $\alpha(0) = \mathbf{x}$ e $\dot{\alpha}(o) = \mathbf{v}$. Então, por (2.4.1), temos que:

$$
\begin{aligned}
df_{\mathbf{x}}(\mathbf{v}) &= (f \circ \alpha)'(0) \\
&= \frac{d}{dt}\,|_{t=0}\; \alpha(t) \cdot \mathbf{a} \times \mathbf{b} \\
&= \mathbf{v} \cdot \mathbf{a} \times \mathbf{b} \\
&= [\mathbf{v}, \mathbf{a}, \mathbf{b}]
\end{aligned}
$$

Em particular, deduzimos que $\nabla f(\mathbf{p}) \equiv \mathbf{a} \times \mathbf{b}$.

Exemplo ...

Seja $f(\mathbf{x}) = \|\mathbf{x}\|^{2-n}$, $\mathbf{x} \in \mathbb{R}^n - \{\mathbf{O}\}$, onde $n \geq 3$.

Então, f é diferenciável em $\mathbb{R}^n - \{\mathbf{O}\}$, e se $\alpha : \mathbf{I} \to \mathbb{R}^n - \{\mathbf{O}\}$ é uma curva , derivável tal que $\alpha(0) = \mathbf{x}$ e $\dot{\alpha}(o) = \mathbf{v}$, temos que:

$$
\begin{aligned}
df_{\mathbf{x}}(\mathbf{v}) &= (f \circ \alpha)'(0) \\
&= \frac{d}{dt}\,|_{t=0}\; \|\alpha(t)\|^{2-n} \\
&= \frac{d}{dt}\,|_{t=0}\; [\alpha(t) \cdot \alpha(t)]^{\frac{2-n}{2}} \\
&= \frac{2-n}{2}[\alpha(0) \cdot \alpha(0)]^{\frac{-n}{2}}\, 2(\alpha(0) \cdot \dot{\alpha}(0)) \\
&= (2-n)\|\mathbf{x}\|^{-n}\mathbf{x} \cdot \mathbf{v}
\end{aligned}
$$

Em particular vemos que $\nabla f(\mathbf{x}) = (2-n)\|\mathbf{x}\|^{-n}\mathbf{x}$.

♣ **Definição 2.15** ... *Um subconjunto $S \subset \mathbb{R}^n$ diz-se uma* **hipersuperfície regular**, *em $\mathbb{R}^n$, se S é o conjunto de nível* 0*:*

$$S = f^{-1}(0)$$

de uma função:

$$f : \mathcal{U} \subseteq \mathbb{R}^n \to \mathbb{R}$$

de classe C^1, tal que:

$$\nabla f(\mathbf{p}) \neq \mathbf{O} \qquad \forall \mathbf{p} \in S = f^{-1}(0)$$

Designemos por $\mathbf{I} \subseteq \mathbb{R}$ um qualquer intervalo aberto em $\mathbb{R}$, que contem 0.

♣ **Proposição 2.7** ... *Seja* $S = f^{-1}(0) \subset \mathbb{R}^n$*, uma hipersuperfície regular, em* $\mathbb{R}^n$*, onde:*

$$f : \mathcal{U} \subseteq \mathbb{R}^n \to \mathbb{R}$$

é uma função de classe C^1*, tal que* $\nabla f(\mathbf{p}) \neq \mathbf{O}, \quad \forall \mathbf{p} \in S$.

Seja $\mathbf{p} \in S$*. Então para toda a curva diferenciável* $\alpha : \mathbf{I} \to \mathbb{R}^n$*, tal que* $\alpha(\mathbf{I}) \subset S$ *e* $\alpha(0) = \mathbf{p}$*, o respectivo vector tangente* $\dot{\alpha}(0)$ *é ortogonal ao vector* $\nabla f(\mathbf{p})$ *(ver a figura 2.10).*

- Demonstração...

  Por hipótese, $f(\alpha(t)) \equiv 0$. Portanto pela regra da cadeia:

  $$0 = df_{\mathbf{p}}(\dot{\alpha}(0)) = \nabla f(\mathbf{p}) \cdot \dot{\alpha}(0)$$

  o que demonstra o resultado,

♣.

Figure 2.10: Hiperplano tangente

.

♣ **Definição 2.16** ... *Seja* $S \subset \mathbb{R}^n$ *uma* **hipersuperfície regular***, em* $\mathbb{R}^n$ *(isto é,* $S = f^{-1}(0)$*, com* $\nabla f(\mathbf{p}) \neq \mathbf{O}, \quad \forall \mathbf{p} \in S$*, onde* $f : \mathcal{U} \subseteq \mathbb{R}^n \to \mathbb{R}$ *é uma função de classe* C^1*). Seja* $\mathbf{p} \in S$.

(i)... À linha recta:

$$\{\mathbf{p} + t\,\nabla f(\mathbf{p}) : \ t \in \mathbb{R}\} \tag{2.4.4}$$

dá-se o nome de **linha normal** *a* S *em* $\mathbf{p}$.

(ii)... O conjunto dos vectores $\mathbf{v} \in \mathbb{R}^n$ *tais que:*

$$(\mathbf{v} - \mathbf{p}) \cdot \nabla f(\mathbf{p}) = 0 \tag{2.4.5}$$

diz-se o **hiperplano tangente** *a* S *em* $\mathbf{p}$ *(ver a figura 2.10).*

(ii)... O conjunto dos vectores $\mathbf{v} \in \mathbb{R}^n$ *tais que:*

$$\mathbf{v} \cdot \nabla f(\mathbf{p}) = 0 \tag{2.4.6}$$

diz-se o **espaço tangente** *a* S *em* $\mathbf{p}$*, e nota-se por:*

$$T_{\mathbf{p}}S \equiv \{\mathbf{v} \in \mathbb{R}^n : \mathbf{v} \cdot \nabla f(\mathbf{p}) = 0\} \tag{2.4.7}$$

Notemos que o espaço tangente $T_{\mathbf{p}}S$, é um subespaço vectorial de $\mathbb{R}^n$, de dimensão $n-1$, e que o hiperplano tangente é o transladado deste subespaço, que passa em $\mathbf{p}$.

A proposição anterior mostra que o vector velocidade $\dot{\alpha}(0)$, de uma qualquer curva diferenciável $\alpha : \mathbf{I} \to \mathbb{R}^n$, tal que $\alpha(\mathbf{I}) \subset S$ e $\alpha(0) = \mathbf{p}$, é ortogonal ao vector $\nabla f(\mathbf{p})$, e portanto é um vector que pertence ao espaço tangente $T_{\mathbf{p}}S$. Como veremos posteriormente, qualquer vector $\mathbf{v} \in T_{\mathbf{p}}S$, é deste tipo, isto é, $\mathbf{v}$ é vector velocidade de uma curva diferenciável $\alpha : \mathbf{I} \to \mathbb{R}^n$, tal que $\alpha(\mathbf{I}) \subset S$ e $\alpha(0) = \mathbf{p}$.

Exemplo ...

Seja $f : \mathcal{U} \subset \mathbb{R}^n$ uma função de classe C^1, e seja $S = \mathbf{gr}\, f \subset \mathbb{R}^{n+1}$ o respectivo gráfico em $\mathbb{R}^{n+1}$.

$S = \mathbf{gr}\, f$ é uma hipersuperfície em $\mathbb{R}^{n+1}$, de acordo com a definição anterior. De facto, se definirmos:

$$F : \mathcal{U} \times \mathbb{R} \subseteq \mathbb{R}^{n+1} \to \mathbb{R}$$

através de:

$$F(\mathbf{x}, z) = f(\mathbf{x}) - z \qquad (\mathbf{x}, z) \in \mathcal{U} \times \mathbb{R} \subseteq \mathbb{R}^{n+1}$$

então $\mathbf{gr}\, f = F^{-1}(0)$, e além disso:

$$\nabla F(\mathbf{p}, f(\mathbf{p})) = (\nabla f(\mathbf{p}), -1) \neq \mathbf{O} \qquad \forall (\mathbf{p}, f(\mathbf{p})) \in S$$

O hiperplano tangente a $S = \mathbf{gr}\, f$ no ponto $(\mathbf{p}, f(\mathbf{p}))$, é o hiperplano afim de equação:

$$\begin{aligned} ((\mathbf{w}, z) - (\mathbf{p}, f(\mathbf{p})) \cdot \nabla F(\mathbf{p}, f(\mathbf{p})) &= ((\mathbf{w}, z) - (\mathbf{p}, f(\mathbf{p})) \cdot (\nabla f(\mathbf{p}), -1) \\ &= (\mathbf{w} - \mathbf{p}, z - f(\mathbf{p})) \cdot (\nabla f(\mathbf{p}), -1) \\ &= (\mathbf{w} - \mathbf{p}) \cdot \nabla f(\mathbf{p}) = z - f(\mathbf{p}) \end{aligned}$$

ou, pondo $\mathbf{w} = (w^1, \cdots, w^n)$ e $\mathbf{p} = (p^1, \cdots, p^n)$:

$$(w^1 - p^1)\frac{\partial f}{\partial x^1}(\mathbf{p}) + \cdots + (w^n - p^n)\frac{\partial f}{\partial x^n}(\mathbf{p}) = z - f(\mathbf{p})$$

o que está de acordo com (2.3.18).

O espaço tangente $T_{\mathbf{p}}\mathbf{gr}\, f$, é constituído pelos vectores $\mathbf{v} = (v^1, \cdots, v^{n+1}) \in \mathbb{R}^{n+1}$, tais que $\mathbf{v} \cdot \nabla F(\mathbf{p}, f(\mathbf{p})) = 0$, isto é:

$$T_{\mathbf{p}}\mathbf{gr}\, f = \{\mathbf{v} = (v^1, \cdots, v^{n+1}) : v^1 \frac{\partial f}{\partial x^1}(\mathbf{p}) + \cdots + v^n \frac{\partial f}{\partial v^n}(\mathbf{p}) - v^{n+1} = 0\}$$

Exemplo ...

O Hiperbolóide de uma folha S, dado por (ver a figura 2.5):

$$x^2 + 4y^2 - 3z^2 = 1$$

é uma superfície regular em $\mathbb{R}^3$, já que $S = f^{-1}(0)$ onde:

$$f(x, y, z) = x^2 + 4y^2 - 3z^2 - 1$$

e $\nabla f(x, y, z) = (2x, 8y, -6z)$ que se anula apenas em $(0, 0, 0)$, ponto que não pertence a S.

O plano tangente a S, no ponto $(-3, 1, -2) \in S$, é dado por:

$$(x + 3)(-6) + (y - 1)(8) + (z + 2)(12) = 0$$

ou $-6x + 8y + 12z = 2$. O espaço tangente nesse ponto, é dado por:

$$T_{(-3,1,-2)}S = \{(x, y, z) : \, -6x + 8y + 12z = 0\}$$

♣ **Proposição 2.8** ... (**Teorema do valor médio**) ... *Seja $f : \mathcal{U} \subset \mathbb{R}^n \to \mathbb{R}$, um campo escalar, definido e diferenciável num aberto $\mathcal{U}$.*

Sejam $\mathbf{p}, \mathbf{q} \in \mathcal{U}$, dois pontos em $\mathcal{U}$, e suponhamos que o segmento de recta que une $\mathbf{p}$ a $\mathbf{q}$, está contido em $\mathcal{U}$:

$$\{\mathbf{p} + t(\mathbf{q} - \mathbf{p}) : \, 0 \leq t \leq 1\} \subset \mathcal{U}$$

Então existe um ponto $\mathbf{u} = \mathbf{p} + t_o(\mathbf{q} - \mathbf{p})$, $0 < t_o < 1$, nesse segmento, tal que:

$$\begin{aligned} f(\mathbf{q}) - f(\mathbf{p}) &= \nabla f(\mathbf{u}) \cdot (\mathbf{q} - \mathbf{p}) \\ &= \nabla f(\mathbf{p} + t_o(\mathbf{q} - \mathbf{p})) \cdot (\mathbf{q} - \mathbf{p}) \end{aligned} \tag{2.4.8}$$

• Demonstração...

Seja $\alpha(t) = \mathbf{p} + t(\mathbf{q} - \mathbf{p})$, e consideremos a função real de variável real (contínua e diferenciável):

$$\phi(t) \equiv (f \circ \alpha)(t) \qquad 0 \leq t \leq 1 \tag{2.4.9}$$

Aplicando a regra da cadeia obtemos:

$$\begin{aligned} \phi'(t) = (f \circ \alpha)'(t) &= \nabla f(\alpha(t)) \cdot (\dot{\alpha}(t)) \\ &= \nabla f(\alpha(t)) \cdot (\mathbf{q} - \mathbf{p}) \end{aligned} \tag{2.4.10}$$

Portanto, existe $t_o : \, 0 < t_o < 1$ tal que $\phi(1) - \phi(0) = \phi'(t_o)(1 - 0)$, isto é:

$$f(\mathbf{q}) - f(\mathbf{q}) = \nabla f(\mathbf{p} + t_o(\mathbf{q} - \mathbf{p})) \cdot (\mathbf{q} - \mathbf{p})$$

como se pretendia,

♣.

Pondo $\mathbf{h} = \mathbf{q} - \mathbf{p}$, de tal forma que $\mathbf{q} = \mathbf{p} + \mathbf{h}$, podemos reescrever (2.4.8), na forma equivalente:

$$f(\mathbf{p}+\mathbf{h}) = f(\mathbf{p}) + \nabla f(\mathbf{p}+t_o\,\mathbf{h}) \cdot \mathbf{h} \tag{2.4.11}$$

com $0 < t_o < 1$.

♣ **Corolário 2.1** ... *Seja $f : \mathcal{U} \subset \mathbb{R}^n \to \mathbb{R}$, um campo escalar, definido e diferenciável num aberto $\mathcal{U}$, e suponhamos que $\mathcal{U}$ é poligonalmente conexo (isto é, dois quaiquer pontos de $\mathcal{U}$ podem ser unidos por uma linha poligonal, contida em $\mathcal{U}$).*

Então, se $df_{\mathbf{x}} = 0, \quad \forall \mathbf{x} \in \mathcal{U}$, f é constante.

2.5 Derivadas parciais de ordem superior. Teorema de Schwartz. Fórmula de Taylor (de ordem dois)

Seja $f : \mathcal{U} \subset \mathbb{R}^n \to \mathbb{R}$, um campo escalar, definido e diferenciável num aberto $\mathcal{U}$. Existem portanto cada uma das derivadas parciais $\frac{\partial f}{\partial x^i}$, $i = 1, \cdots, n$, e cada uma delas define um novo campo escalar em $\mathcal{U}$, através de:

$$\mathbf{x} \in \mathcal{U} \mapsto \frac{\partial f}{\partial x^i}(\mathbf{x})$$

À derivada parcial em ordem a x^j, $j = 1, \cdots, n$, deste novo campo escalar:

$$\frac{\partial}{\partial x^j}\Big(\frac{\partial f}{\partial x^i}\Big)(\mathbf{x})$$

(se existir), chama-se a segunda derivada (i, j), de f em $\mathbf{x}$, e nota-se por:

$$\frac{\partial^2 f}{\partial x^j \partial x^i}(\mathbf{x}) \equiv \frac{\partial}{\partial x^j}\Big(\frac{\partial f}{\partial x^i}\Big)(\mathbf{x})$$

As derivadas de ordem superior definem-se anàlogamente.

♣ **Definição 2.17** ... *Um campo escalar $f : \mathcal{U} \subset \mathbb{R}^n \to \mathbb{R}$, definido num aberto $\mathcal{U}$, diz-se de* **classe** C^k *em $\mathcal{U}$, se existirem e forem contínuas em $\mathcal{U}$, todas as derivadas parciais até à ordem k (inclusivé).*

Aplicando essencialmente o teorema do valor médio, é possível demonstrar a seguinte proposição:

♣ **Proposição 2.9** ... (**Teorema de Schwartz**) ... *Seja* $f : \mathcal{U} \subset \mathbb{R}^n \to \mathbb{R}$, *um campo escalar, de classe* C^2 *num aberto* $\mathcal{U}$.

Então, $\forall i, j = 1, \cdots, n$ *tem-se que:*

$$\frac{\partial^2 f}{\partial x^j \partial x^i}(\mathbf{x}) = \frac{\partial^2 f}{\partial x^i \partial x^j}(\mathbf{x}) \qquad \forall \mathbf{x} \in \mathcal{U} \tag{2.5.1}$$

Em particular, para uma função $f : \mathcal{U} \subset \mathbb{R}^n \to \mathbb{R}$, de classe C^2 num aberto $\mathcal{U}$, podemos definir a seguinte matriz:

$$\mathbf{H}_f(\mathbf{x}) = \begin{bmatrix} \frac{\partial^2 f}{\partial (x^1)^2}(\mathbf{x}) & \frac{\partial^2 f}{\partial x^1 \partial x^2}(\mathbf{x}) & \cdots & \frac{\partial^2 f}{\partial x^1 \partial x^n}(\mathbf{x}) \\ \frac{\partial^2 f}{\partial x^2 \partial x^1}(\mathbf{x}) & \frac{\partial^2 f}{\partial (x^2)^2}(\mathbf{x}) & \cdots & \frac{\partial^2 f}{\partial x^2 \partial x^n}(\mathbf{x}) \\ \vdots & \cdots & \cdots & \vdots \\ \vdots & \cdots & \cdots & \vdots \\ \frac{\partial^2 f}{\partial x^n \partial x^1}(\mathbf{x}) & \cdots & \cdots & \frac{\partial^2 f}{\partial (x^n)^2}(\mathbf{x}) \end{bmatrix} \tag{2.5.2}$$

matriz que é simétrica pelo teorema de Schwartz, e que se diz a **matriz Hessiana** de f em $\mathbf{x}$.

À forma quadrática associada a $\mathbf{H}_f(\mathbf{x})$, chama-se o **Hessiano de** f **em** $\mathbf{x}$, e nota-se por $\mathbf{hess}_f(\mathbf{x})$:

$$\begin{aligned} (\mathbf{hess}_f(\mathbf{x}))(\mathbf{h}) &= \mathbf{H}_f(\mathbf{x})\mathbf{h} \cdot \mathbf{h} \\ &= \mathbf{h}^{tr}\, \mathbf{H}_f(\mathbf{x})\, \mathbf{h} \qquad \text{em notação matricial} \\ &= \sum_{i,j=1}^{n} \frac{\partial^2 f}{\partial x^i \partial x^j}(\mathbf{x})\, h^i\, h^j \end{aligned} \tag{2.5.3}$$

onde $\mathbf{h} = (h^1, \cdots, h^n) \in \mathbb{R}^n$.

Exemplo ...

Para uma função de duas variáveis $f : \mathbb{R}^2 \to \mathbb{R}$, o Hessiano tem o aspecto seguinte:

$$\begin{aligned} (\mathbf{hess}_f(\mathbf{x}))(\mathbf{h}) &= \mathbf{h}^{tr}\, \mathbf{H}_f(\mathbf{x})\, \mathbf{h} \qquad \text{em notação matricial} \\ &= [u \quad v] \begin{bmatrix} \frac{\partial^2 f}{\partial x^2}(\mathbf{x}) & \frac{\partial^2 f}{\partial x \partial y}(\mathbf{x}) \\ \frac{\partial^2 f}{\partial y \partial x}(\mathbf{x}) & \frac{\partial^2 f}{\partial y^2}(\mathbf{x}) \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix} \\ &= \frac{\partial^2 f}{\partial x^2}(\mathbf{x})\, u^2 + 2\frac{\partial^2 f}{\partial x \partial y}(\mathbf{x})\, uv + \frac{\partial^2 f}{\partial y^2}(\mathbf{x})\, v^2 \end{aligned}$$

onde se pôs $\mathbf{h} = (u, v)$.

Suponhamos agora que f é uma função de classe C^2 numa bola aberta $B(\mathbf{p}, r)$, de centro $\mathbf{p} \in \mathbb{R}^n$ e raio $r > 0$.

Se $\mathbf{h} \in \mathbb{R}^n$ é um vector de norma $\|\mathbf{h}\| < r$, então $\mathbf{p} + \mathbf{h} \in B(\mathbf{p}, r)$, e o segmento de recta:

$$\alpha(t) = \mathbf{p} + t\,\mathbf{h} \qquad 0 \leq t \leq 1$$

que une $\mathbf{p}$ a $\mathbf{p}+\mathbf{h}$ está contido em $B(\mathbf{p}, r)$.

Consideremos a função real de variável real (também de classe C^2):

$$\begin{aligned} \phi(t) &\equiv (f \circ \alpha)(t) \\ &= f(\mathbf{p} + t\,\mathbf{h}) \qquad 0 \leq t \leq 1 \end{aligned} \tag{2.5.4}$$

Pela fórmula de Taylor para funções reais de variável real, podemos escrever que:

$$\begin{aligned} f(\mathbf{p}+\mathbf{h}) - f(\mathbf{p}) &= \phi(1) - \phi(0) \\ &= \phi'(0) + \frac{1}{2!}\phi''(c) \end{aligned} \tag{2.5.5}$$

para algum $c: \ 0 < c < 1$.

Aplicando a regra da cadeia, e atendendo a que $\dot{\alpha}(t) \equiv \mathbf{h} = (h_1, ..., h_n)$, obtemos:

$$\begin{aligned} \phi'(t) = (f \circ \alpha)'(t) &= \nabla f(\alpha(t)) \cdot (\dot{\alpha}(t)) \\ &= \nabla f(\alpha(t)) \cdot \mathbf{h} \\ &= \sum_{i=1}^{n} \frac{\partial f}{\partial x^i}(\alpha(t))\, h_i \end{aligned} \tag{2.5.6}$$

e em particular:

$$\phi'(0) = \nabla f(\mathbf{p}) \cdot \mathbf{h}$$

Aplicando de novo a regra da cadeia a (2.5.6), obtemos:

$$\begin{aligned} \phi''(t) = (f \circ \alpha)''(t) &= \sum_{j=1}^{n} \frac{\partial}{\partial x^j}\Big(\sum_{i=1}^{n} \frac{\partial f}{\partial x^i}(\alpha(t))\, h_i\Big)\, h_j \\ &= \sum_{i,j=1}^{n} \frac{\partial^2 f}{\partial x^i \partial x^j}(\alpha(t))\, h_i h_j \\ &= \mathbf{hess}_f(\alpha(t))(\mathbf{h}) \\ &= \mathbf{hess}_f(\mathbf{p} + t\,\mathbf{h})\,(\mathbf{h}) \end{aligned} \tag{2.5.7}$$

atendendo a (2.5.3).

Em particular:

$$\phi''(c) = \mathbf{hess}_f(\mathbf{p} + c\,\mathbf{h})\,(\mathbf{h})$$

e portanto (2.5.5), fica na forma:

$$\begin{aligned} f(\mathbf{p}+\mathbf{h}) - f(\mathbf{p}) &= \phi'(0) + \frac{1}{2!}\phi''(c) \\ &= \nabla f(\mathbf{p}) \cdot \mathbf{h} + \frac{1}{2!}\mathbf{hess}_f(\mathbf{p} + c\,\mathbf{h})\,(\mathbf{h}) \end{aligned} \tag{2.5.8}$$

para algum $c: \ 0 < c < 1$.

Fica assim parcialmente demonstrada a seguinte proposição:

♣ <u>**Proposição**</u> **2.10** ... *(***Fórmula de Taylor de ordem** *2) ... Seja f um campo escalar de classe C^2 numa bola aberta $B(\mathbf{p}, r)$, de centro $\mathbf{p} \in \mathbb{R}^n$ e raio $r > 0$.*

Então, para todo o vector $\mathbf{h} \in \mathbb{R}^n$, tal que $\mathbf{p} + \mathbf{h} \in B(\mathbf{p}, r)$, é válida a seguinte **fórmula de Taylor de ordem 2, centrada em p***:*

$$\boxed{f(\mathbf{p}+\mathbf{h}) - f(\mathbf{p}) = \nabla f(\mathbf{p}) \cdot \mathbf{h} + \tfrac{1}{2!}\mathbf{hess}_f(\mathbf{p} + c\,\mathbf{h})\,(\mathbf{h})} \tag{2.5.9}$$

para algum $c:\ 0 < c < 1$.

Esta fórmula pode ainda ser escrita na forma:

$$\boxed{f(\mathbf{p}+\mathbf{h}) - f(\mathbf{p}) = \nabla f(\mathbf{p}) \cdot \mathbf{h} + \tfrac{1}{2!}\mathbf{hess}_f(\mathbf{p})\,(\mathbf{h}) + o(\|\mathbf{h}\|^2)} \tag{2.5.10}$$

onde:

$$\lim_{\mathbf{h}\to\mathbf{O}} \frac{o(\|\mathbf{h}\|^2)}{\|\mathbf{h}\|^2} = 0$$

- <u>Demonstração</u>...

Resta demonstrar (2.5.10). Para isso, definamos a função $o(\|\mathbf{h}\|^2)$, através de:

$$o(\|\mathbf{h}\|^2)(\mathbf{h}) = \begin{cases} 0 & \text{se } \mathbf{h} = \mathbf{O} \\ \frac{1}{2!}(\mathbf{hess}_f(\mathbf{p}+c\,\mathbf{h}))\,(\mathbf{h}) - \mathbf{hess}_f(\mathbf{p})\,(\mathbf{h})) & \text{se } \mathbf{h} \neq \mathbf{O} \end{cases}$$

Com esta definição a fórmula (2.5.9) fica na forma (2.5.10), e para completar a demonstração, resta provar que $\lim_{\mathbf{h}\to\mathbf{O}} \frac{o(\|\mathbf{h}\|^2)}{\|\mathbf{h}\|^2} = 0$.

Mas, para $\mathbf{h} \neq \mathbf{O}$, tem-se que:

$$\begin{aligned} |o(\|\mathbf{h}\|^2)(\mathbf{h})| &= |\frac{1}{2!}(\mathbf{hess}_f(\mathbf{p}+c\,\mathbf{h}))\,(\mathbf{h}) - \mathbf{hess}_f(\mathbf{p})\,(\mathbf{h}))| \\ &= \frac{1}{2!}|\sum_{i,j=1}^{n}\Big(\frac{\partial^2 f}{\partial x^i \partial x^j}(\mathbf{p}+c\,\mathbf{h}) - \frac{\partial^2 f}{\partial x^i \partial x^j}(\mathbf{p})\Big)\,h_i h_j| \\ &\leq \frac{1}{2!}\Big(\sum_{i,j=1}^{n}|\frac{\partial^2 f}{\partial x^i \partial x^j}(\mathbf{p}+c\,\mathbf{h}) - \frac{\partial^2 f}{\partial x^i \partial x^j}(\mathbf{p})|\Big)\|\mathbf{h}\|^2 \end{aligned}$$

Como cada derivada parcial de segunda ordem é contínua (f é por hipótese, de classe C^2), tem-se que:

$$\frac{\partial^2 f}{\partial x^i \partial x^j}(\mathbf{p}+c\,\mathbf{h}) \to \frac{\partial^2 f}{\partial x^i \partial x^j}(\mathbf{p})$$

quando $\mathbf{h} \to \mathbf{O}$, o que permite concluir que:

$$\lim_{\mathbf{h}\to\mathbf{O}} \frac{o(\|\mathbf{h}\|^2)}{\|\mathbf{h}\|^2} = 0$$

como se pretendia,

♣.

Em coordenadas cartesianas a fórmula de Taylor de ordem 2, centrada em $\mathbf{p}$ (2.5.10), escreve-se na forma:

$$\boxed{f(\mathbf{p}+\mathbf{h}) = f(\mathbf{p}) + \sum_{i=1}^{n} \frac{\partial f}{\partial x^i}(\mathbf{p})\, h^i + \frac{1}{2!}\sum_{i,j=1}^{n} \frac{\partial^2 f}{\partial x^i \partial x^j}(\mathbf{p})\, h^i h^j + o(\|\mathbf{h}\|^2)} \tag{2.5.11}$$

<u>**Exemplo**</u> ...

Calculemos a fórmula de Taylor de ordem 2, centrada em $\mathbf{p} = (0,0)$, da função $f(x,y) = e^x \cos y$.

Como:

$$f(0,0) = 1 \qquad \frac{\partial f}{\partial x}(0,0) = 1 \qquad \frac{\partial f}{\partial y}(0,0) = 0$$

e:

$$\frac{\partial^2 f}{\partial x^2}(0,0) = 1 \qquad \frac{\partial^2 f}{\partial x \partial y}(0,0) = 0 \qquad \frac{\partial^2 f}{\partial y^2}(0,0) = -1$$

temos que:

$$f(\mathbf{O}+\mathbf{h}) = f(u,v) = 1 + u + \frac{1}{2}u^2 - \frac{1}{2}v^2 + o(\|\mathbf{h}\|^2)$$

onde $\mathbf{h} = (u,v)$.

2.6 Extremos de campos escalares

♣ <u>**Definição**</u> **2.18** ... *Seja $f : \mathcal{U} \subseteq \mathbb{R}^n \to \mathbb{R}$, um campo escalar definido num aberto $\mathcal{U}$ de $\mathbb{R}^n$.*

(i)... Um ponto $\mathbf{x}_o \in \mathcal{U}$, diz-se um **mínimo local** *de f, se existe uma bola aberta $B(\mathbf{x}_o, r) \subset \mathcal{U}$, centrada em $\mathbf{x}_o$ e de raio $r > 0$, na qual:*

$$f(\mathbf{x}) \geq f(\mathbf{x}_o) \qquad \forall \mathbf{x} \in B(\mathbf{x}_o, r) \tag{2.6.1}$$

(ii)... Um ponto $\mathbf{x}_o \in \mathcal{U}$, diz-se um **máximo local** *de f, se existe uma bola aberta $B(\mathbf{x}_o, r) \subset \mathcal{U}$, centrada em $\mathbf{x}_o$ e de raio $r > 0$, na qual:*

$$f(\mathbf{x}) \leq f(\mathbf{x}_o) \qquad \forall \mathbf{x} \in B(\mathbf{x}_o, r) \tag{2.6.2}$$

(iii)... . Um ponto $\mathbf{x}_o \in \mathcal{U}$, diz-se um **extremo local** *de f, se $\mathbf{x}_o$ é um máximo ou um mínimo local de f.*

(iv)... Um ponto $\mathbf{x}_o \in \mathcal{U}$*, diz-se um* **ponto crítico** *de* f*, se* f *é diferenciável em* $\mathbf{x}_o$ *e:*

$$df_{\mathbf{x}_o} = 0 \qquad (\Leftrightarrow \quad \nabla f(\mathbf{x}_o) = \mathbf{O}) \tag{2.6.3}$$

Um ponto crítico $\mathbf{x}_o \in \mathcal{U}$*, diz-se um* **ponto crítico não degenerado** *de* f*, se* f *é de classe* C^2 *e se a matriz hessiana* $\mathbf{H}_f(\mathbf{x}_o)$*, é não degenerada (isto é, tem determinante não nulo).*

Quando f *é de classe* C^2*,* $\mathbf{x}_o \in \mathcal{U}$ *é ponto crítico, e* $det\,\mathbf{H}_f(\mathbf{x}_o) = 0$*, então* $\mathbf{x}_o$ *diz-se um* **ponto crítico degenerado** *de* f*.*

Um ponto crítico $\mathbf{x}_o \in \mathcal{U}$*, que não é extremo local de* f*, diz-se um* **ponto sela** *de* f*.*

♣ **Proposição 2.11** ... *Seja* $f : \mathcal{U} \subseteq \mathbb{R}^n \to \mathbb{R}$*, um campo escalar definido e diferenciável num aberto* $\mathcal{U}$ *de* $\mathbb{R}^n$*, e seja* $\mathbf{x}_o \in \mathcal{U}$*, um* **extremo local** *de* f*.*

Então $\mathbf{x}_o$*, é* **ponto crítico** *de* f*, isto é:* $df_{\mathbf{x}_o} = 0$ *(*$\Leftrightarrow \quad \nabla f(\mathbf{x}_o) = \mathbf{O}$*).*

- Demonstração...

  Suponhamos que $\mathbf{x}_o$ é um máximo local. Então, $\forall \mathbf{h} \in \mathbb{R}^n$, a função real de variável real $\phi(t) = f(\mathbf{x}_o + t\,\mathbf{h})$, tem um máximo local em $t = 0$. Portanto $\phi'(0) = 0$. Aplicando a regra da cadeia, obtemos então que:
  $$0 = \phi'(0) = df_{\mathbf{x}_o}(\mathbf{h}) = \nabla f(\mathbf{x}_o) \cdot \mathbf{h}$$
  $\forall \mathbf{h} \in \mathbb{R}^n$, isto é : $df_{\mathbf{x}_o} = 0$ ($\Leftrightarrow \quad \nabla f(\mathbf{x}_o) = \mathbf{O}$),

♣.

Notemos que o recíproco desta proposição é falso, isto é, um ponto crítico $\mathbf{x}_o$ (onde $df_{\mathbf{x}_o} = 0$), não é necessàriamente um extremo local de f.

Exemplo ...

A função $f(x, y) = x^2 - y^2$, $(x, y) \in \mathbb{R}^2$, tem um ponto crítico em $\mathbf{O} = (0, 0)$. No entanto, como $f(x, 0) = x^2 \geq 0 = f(0, 0)$ e $f(0, y) = -y^2 \leq 0 = f(0, 0)$, vemos que $\mathbf{O} = (0, 0)$ não pode ser extremo local de f.

Para calcular os extremos locais de f, devemos resolver a equação $\nabla f(\mathbf{x}) = \mathbf{O}$, isto é, calcular os pontos $\mathbf{x}$ tais que:
$$\frac{\partial f}{\partial x^i}(\mathbf{x}) = 0 \qquad i = 1, \cdots, n \tag{2.6.4}$$
A proposição anterior afirma que os candidatos a extremos locais de f, se encontram entre as soluções de (2.6.4).

Para decidir quais destas soluções são máximos ou mínimos locais de f, devemos dar um critério de segunda ordem, baseado na análise do Hessiano de f.

Suponhamos então que f é de classe C^2, e que $\mathbf{x}_o$ é um ponto crítico não degenerado de f (isto é, $\mathbf{x}_o$ é solução de (2.6.4) e $\det \mathbf{H}_f(\mathbf{x}_o) \neq 0$).

Como $\nabla f(\mathbf{x}_o) = \mathbf{O}$, a fórmula de Taylor de ordem 2, centrada em $\mathbf{x}_o$ (ver (2.5.10)), tem o aspecto seguinte:

$$f(\mathbf{x}_o + \mathbf{h}) - f(\mathbf{x}_o) = \frac{1}{2!}\mathbf{hess}_f(\mathbf{x}_o)\,(\mathbf{h}) + o(\|\mathbf{h}\|^2) \tag{2.6.5}$$

onde:

$$\lim_{\mathbf{h}\to\mathbf{O}} \frac{o(\|\mathbf{h}\|^2)}{\|\mathbf{h}\|^2} = 0 \tag{2.6.6}$$

Como $o(\|\mathbf{h}\|^2)$ é um "infinitésimo de ordem $\geq$ dois" (isto é, verifica a condição (2.6.6)), parece natural esperar que o sinal de $f(\mathbf{x}_o + \mathbf{h}) - f(\mathbf{x}_o)$, numa pequena vizinhança de $\mathbf{x}_o$, seja o mesmo do da forma quadrática $\mathbf{h} \mapsto \mathbf{hess}_f(\mathbf{x}_o)\,(\mathbf{h})$. Portanto, a natureza do ponto crítico $\mathbf{x}_o$ deverá ser determinada pelo sinal do Hessiano de f em $\mathbf{x}_o$.

De facto, é válida a seguinte proposição:

♣ **Proposição 2.12** ... *Seja $f : \mathcal{U} \subseteq \mathbb{R}^n \to \mathbb{R}$, um campo escalar de classe C^2 num aberto $\mathcal{U}$ de $\mathbb{R}^n$, e seja $\mathbf{x}_o \in \mathcal{U}$, um* **ponto crítico não degenerado** *de f.*

Consideremos o Hessiano de f em $\mathbf{x}_o$, $\mathbf{hess}_f(\mathbf{x}_o)$ (ver (2.5.3)):

$$\begin{aligned}(\mathbf{hess}_f(\mathbf{x}_o))(\mathbf{h}) &= \mathbf{h}^{tr}\,\mathbf{H}_f(\mathbf{x}_o)\,\mathbf{h} \quad \textit{em notação matricial}\\ &= \sum_{i,j=1}^{n} \frac{\partial^2 f}{\partial x^i \partial x^j}(\mathbf{x}_o)\,h^i\,h^j\end{aligned} \tag{2.6.7}$$

onde $\mathbf{h} = (h^1, \cdots, h^n)$, e $\mathbf{H}_f(\mathbf{x}_o)$ é a matriz Hessiana de f em $\mathbf{x}_o$ (ver (2.5.2)).

Então:

(i)... se $\mathbf{hess}_f(\mathbf{x}_o)$ *é* **definido positivo** *(isto é, todos os valores próprios de $\mathbf{H}_f(\mathbf{x}_o)$, são estritamente positivos), $\mathbf{x}_o$ é um* **mínimo local** *(estrito) de f.*

(ii)... se $\mathbf{hess}_f(\mathbf{x}_o)$ *é* **definido negativo** *(isto é, todos os valores próprios de $\mathbf{H}_f(\mathbf{x}_o)$, são estritamente negativos), $\mathbf{x}_o$ é um* **máximo local** *(estrito) de f.*

(iii)... se $\mathbf{hess}_f(\mathbf{x}_o)$ *é* **indefinido** *(isto é, todos os valores próprios são não nulos, mas existem valores próprios de $\mathbf{H}_f(\mathbf{x}_o)$, positivos e negativos), $\mathbf{x}_o$ é ponto sela de f.*

- Demonstração...

Sejam $\lambda_1, \cdots, \lambda_n$ os valores próprios de $\mathbf{H}_f(\mathbf{x}_o)$. Como sabemos, todos são reais. Designemos por $\lambda_{\min}$ e por $\lambda_{\max}$, respectivamente os valores próprios mínimo e máximo de $\mathbf{H}_f(\mathbf{x}_o)$. É fácil ver que:

$$\lambda_{\min}\,\|\mathbf{h}\|^2 \leq \mathbf{hess}_f(\mathbf{x}_o)(\mathbf{h}) \leq \lambda_{\max}\,\|\mathbf{h}\|^2 \qquad \forall \mathbf{h} \in \mathbb{R}^n \tag{2.6.8}$$

Deduzimos então desta dupla desigualdade e da fórmula de Taylor (2.6.5), que:

$$\begin{aligned}\frac{1}{2}\,\lambda_{\min}\,\|\mathbf{h}\|^2 + o(\|\mathbf{h}\|^2) &\leq f(\mathbf{x}_o + \mathbf{h}) - f(\mathbf{x}_o)\\ &\leq \frac{1}{2}\,\lambda_{\max}\,\|\mathbf{h}\|^2 + o(\|\mathbf{h}\|^2)\end{aligned} \tag{2.6.9}$$

(i)... Se $\mathbf{hess}_f(\mathbf{x}_o)$ é **definido positivo**, isto é, se todos os valores próprios de $\mathbf{H}_f(\mathbf{x}_o)$, são estritamente positivos, então $\lambda_{\min} > 0$.

Como $\lim_{\mathbf{h}\to\mathbf{O}} \frac{o(\|\mathbf{h}\|^2)}{\|\mathbf{h}\|^2} = 0$, podemos encontrar um $\delta > 0$ tal que:

$$|o(\|\mathbf{h}\|^2)| < \frac{1}{4}\lambda_{\min}\,\|\mathbf{h}\|^2 \qquad \forall\mathbf{h} : \|\mathbf{h}\| < \delta$$

e de (2.6.9) deduzimos então que:

$$\begin{aligned} f(\mathbf{x}_o + \mathbf{h}) - f(\mathbf{x}_o) &\geq \frac{1}{2}\,\lambda_{\min}\,\|\mathbf{h}\|^2 + o(\|\mathbf{h}\|^2) \\ &\geq \frac{1}{2}\,\lambda_{\min}\,\|\mathbf{h}\|^2 - \frac{1}{4}\lambda_{\min}\,\|\mathbf{h}\|^2 \\ &= \frac{1}{4}\lambda_{\min}\,\|\mathbf{h}\|^2 \\ &> 0 \qquad \text{sempre que } 0 < \|\mathbf{h}\| < \delta \end{aligned}$$

o que implica que f tem um mínimo local (estrito) em $\mathbf{x}_o$.

(ii)... demonstra-se de forma análoga.

(iii)... Se $\mathbf{hess}_f(\mathbf{x}_o)$ é **indefinido**, isto é, se existem valores próprios de $\mathbf{H}_f(\mathbf{x}_o)$, positivos e negativos, mas todos não nulos, então $\lambda_{\min} < 0$ e $\lambda_{\max} > 0$.

Sejam $\mathbf{v}_{\max}$ e $\mathbf{v}_{\min}$, vectores próprios de norma 1, associados respectivamente aos valores próprios $\lambda_{\min} < 0$ e $\lambda_{\max} > 0$, de tal forma que:

$$\mathbf{hess}_f(\mathbf{x}_o)(\mathbf{v}_{\max}) = \mathbf{v}_{\max}^{tr}\mathbf{H}_f(\mathbf{x}_o)\mathbf{v}_{\max} = \mathbf{v}_{\max}^{tr}\lambda_{\max}\mathbf{v}_{\max} = \lambda_{\max}$$

e anàlogamente, $\mathbf{hess}_f(\mathbf{x}_o)(\mathbf{v}_{\min}) = \lambda_{\min}$.

Fazendo $\mathbf{h} = t\mathbf{v}_{\max}$ em (2.6.5), vem que:

$$f(\mathbf{x}_o + t\mathbf{v}_{\max}) - f(\mathbf{x}_o) = \frac{1}{2}t^2\,\lambda_{\max} + o(t^2)$$

e anàlogamente:

$$f(\mathbf{x}_o + t\mathbf{v}_{\min}) - f(\mathbf{x}_o) = \frac{1}{2}t^2\,\lambda_{\min} + o(t^2)$$

Como $\lambda_{\min} < 0$ e $\lambda_{\max} > 0$, é agora fácil ver que, para todo o t suficientemente pequeno:

$$f(\mathbf{x}_o + t\mathbf{v}_{\max}) - f(\mathbf{x}_o) > 0 \quad e \quad f(\mathbf{x}_o + t\mathbf{v}_{\max}) - f(\mathbf{x}_o) > 0$$

o que implica que $\mathbf{x}_o$ é um ponto sela de f,

♣.

Quando o ponto crítico é degenerado, os critérios anteriores não são conclusivos, como mostram os exemplos seguintes:

Exemplo ...

(i)... A função $f(x,y) = x^2 - y^3$, tem um ponto crítico no ponto $(0,0)$, que é degenerado uma vez que a matriz hessiana nesse ponto é a matriz (semi-definida positiva):

$$\mathbf{H}_f(0,0) = \begin{bmatrix} 2 & 0 \\ 0 & 0 \end{bmatrix}$$

Neste caso $(0,0)$ é ponto sela de f.

(ii)... A função $f(x,y) = x^3 - y^2$, tem um ponto crítico no ponto $(0,0)$, que é degenerado uma vez que a matriz hessiana nesse ponto é a matriz (semi-definida negativa):

$$\mathbf{H}_f(0,0) = \begin{bmatrix} 0 & 0 \\ 0 & -2 \end{bmatrix}$$

Neste caso $(0,0)$ é ponto sela de f.

(iii)... A função $f(x,y) = x^4 + y^4$, tem um ponto crítico no ponto $(0,0)$, que é degenerado uma vez que a matriz hessiana nesse ponto é a matriz nula. Neste caso $(0,0)$ é ponto mínimo local de f.

(iv)... A função $f(x,y) = -x^4 - y^4$, tem um ponto crítico no ponto $(0,0)$, que é degenerado uma vez que a matriz hessiuana nesse ponto é a matriz nula. Neste caso $(0,0)$ é ponto máximo local de f.

Na prática, são úteis os seguintes critérios para decidir se o hessiano é definido positivo ou negativo:

(i)... Para funções de duas variáveis:

Se a matriz hessiana de f em $\mathbf{x}_o$ é:

$$\mathbf{H}_f(\mathbf{x}_o) = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

então $\mathbf{hess}_f(\mathbf{x}_o)$ é definido positivo se e só se:

$$a > 0 \qquad e \qquad \det \begin{bmatrix} a & b \\ c & d \end{bmatrix} > 0$$

e é definido negativo sse:

$$a < 0 \qquad e \qquad \det \begin{bmatrix} a & b \\ c & d \end{bmatrix} > 0$$

(ii)... Para funções de três variáveis:

Se a matriz hessiana de f em $\mathbf{x}_o$ é:

$$\mathbf{H}_f(\mathbf{x}_o) = \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & k \end{bmatrix}$$

então $\mathbf{hess}_f(\mathbf{x}_o)$ é definido positivo se e só se:

$$a > 0 \qquad e \qquad \det \begin{bmatrix} a & b \\ d & e \end{bmatrix} > 0 \qquad e \qquad \det \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & k \end{bmatrix} > 0$$

e é definido negativo sse:

$$a < 0 \quad e \quad \det\begin{bmatrix} a & b \\ d & e \end{bmatrix} > 0 \quad e \quad \det\begin{bmatrix} a & b & c \\ d & e & f \\ g & h & k \end{bmatrix} < 0$$

2.7 Máximos e mínimos condicionados. Multiplicadores de Lagrange

♣ **Proposição 2.13** (**Método dos multiplicadores de Lagrange I**) ... *Seja $S = g^{-1}(0) \subset \mathbb{R}^n$, uma hipersuperfície regular, em $\mathbb{R}^n$, onde:*

$$g : \mathcal{U} \subseteq \mathbb{R}^n \to \mathbb{R}$$

é uma função de classe C^1, tal que $\nabla g(\mathbf{p}) \neq \mathbf{O}$, $\forall \mathbf{p} \in S$, e seja:

$$f : \mathcal{U} \subseteq \mathbb{R}^n \to \mathbb{R}$$

um campo escalar diferenciável em $\mathcal{U}$.

Se a restrição de f à hipersuperfície S, $f|_S$, tem um máximo ou um mínimo local num ponto $\mathbf{x}_o \in S$, então existe um número real λ tal que:

$$\nabla f(\mathbf{x}_o) = \lambda \, \nabla g(\mathbf{x}_o) \tag{2.7.1}$$

λ diz-se um **multiplicador de Lagrange**.

- Demonstração...

  Seja $\mathbf{v} \in T_{\mathbf{x}_o}S$, um qualquer vector do espaço tangente a S em $\mathbf{x}_o$. Como assinalamos antes, $\mathbf{v}$ é o vector velocidade de uma uma curva diferenciável $\alpha : \mathbf{I} \to \mathbb{R}^n$, tal que $\alpha(\mathbf{I}) \subset S$ e $\alpha(0) = \mathbf{x}_o$: $\mathbf{v} = \dot{\alpha}(0)$.

  É claro que $f \circ \alpha$ tem um extremo local em $t = 0$, e por isso:

  $$0 = (f \circ \alpha)'(0) = \nabla f(\alpha(0)) \cdot \dot{\alpha}(0) = \nabla f(\mathbf{x}_o) \cdot \mathbf{v}$$

  o que significa que $\nabla f(\mathbf{x}_o)$ é ortogonal a $T_{\mathbf{x}_o}S$ (uma vez que $\mathbf{v}$ é arbitrário).

  Mas, $T_{\mathbf{x}_o}S$ é o subespaço de $\mathbb{R}^n$ ortogonal a $\nabla g(\mathbf{x}_o)$, e portanto existe $\lambda \in \mathbb{R}$ tal que:

  $$\nabla f(\mathbf{x}_o) = \lambda \, \nabla g(\mathbf{x}_o)$$

♣.

Um ponto $\mathbf{x}_o \in S = g^{-1}(0)$, no qual a restrição de f a S, tem um extremo local, diz-se um **extremo condicionado** de f. Para calcular estes extremos condicionados, devemos portanto encontrar as soluções $\mathbf{x} = (x^1, ..., x^n), \lambda$, do sistema de equações:

$$\begin{aligned}
\frac{\partial f}{\partial x^1}(x^1, ..., x^n) - \lambda \frac{\partial g}{\partial x^1}(x^1, ..., x^n) &= 0 \\
\frac{\partial f}{\partial x^2}(x^1, ..., x^n) - \lambda \frac{\partial g}{\partial x^2}(x^1, ..., x^n) &= 0 \\
&\vdots \\
\frac{\partial f}{\partial x^n}(x^1, ..., x^n) - \lambda \frac{\partial g}{\partial x^n}(x^1, ..., x^n) &= 0 \\
g(x^1, ..., x^n) &= 0
\end{aligned} \tag{2.7.2}$$

Exemplo ...

Calcular o ponto pertencente à recta em $\mathbb{R}^2$, de equação $y - x - 1 = 0$, que está mais próximo da origem.

O problema pode ser reformulado, como o problema de encontrar o mínimo condicionado da função $f(x, y) = x^2 + y^2$, restrita à recta de equação $g(x, y) = y - x - 1 = 0$. O sistema (2.7.2), tem neste caso o aspecto seguinte:

$$\begin{aligned}
2x - \lambda(-1) &= 0 \\
2y - \lambda(+1) &= 0 \\
y - x - 1 = 0
\end{aligned}$$

cuja solução única é $x = -1/2, y = 1/2, \lambda = 1$. O ponto procurado é portanto o ponto $(-1/2, 1/2)$.

Exemplo ...

Calcule o volume máximo possível para uma caixa paralelipipédica, cuja área da respectiva superfície é igual a $10\, m^2$.

O problema é equivalente a calcular o máximo condicionado da função $f(x, y, z) = xyz$, quando restrita à superfície de $\mathbb{R}^3$ dada pela equação:

$$g(x, y, z) = 2(xy + xz + yz) - 10 = 0$$

O sistema (2.7.2) é neste caso:

$$\begin{aligned}
yz - \lambda(y + z) &= 0 \\
xz - \lambda(x + z) &= 0 \\
xy - \lambda(y + x) &= 0 \\
2(xy + xz + yz) - 10 &= 0
\end{aligned}$$

Se $x = 0$, então viria que $\lambda z = 0 = \lambda y$ e $yz = 5$, o que implica que $y \neq 0$, $z \neq 0$, e ainda $\lambda = 0$. Da primeira equação viria que $yz = 0$, o que é absurdo. Logo $x \neq 0$. Da mesma forma se conclui que $y \neq 0$, $z \neq 0$, $x + y \neq 0$, etc. Eliminando λ nas duas primeiras equações, dá que $x = y$, e anàlogamente se deduz que $y = z$.

Finalmente, substituindo estes valores na última equação, obtemos $3x^2 = 5$, ou $x = \sqrt{5/3}$, e portanto o volume máximo é igual a $xyz = (5/3)^{3/2}$.

A proposição anterior pode ser utilizada para estabelecer a seguinte propriedade extremal dos valores próprios de uma matriz simétrica (ou da forma quadrática associada):

♣ **Proposição 2.14** ... *Seja* $\mathbf{S} : \mathbb{R}^n \to \mathbb{R}^n$ *um endomorfismo simétrico de* $\mathbb{R}^n$, *e* $q : \mathbb{R}^n \to \mathbb{R}$ *a forma quadrática associada a* $\mathbf{S}$, *definida por* $q(\mathbf{x}) = \mathbf{x} \cdot \mathbf{S}(\mathbf{x})$.

A base ortonormada $[\mathbf{U}] = [\mathbf{u}_1\, \mathbf{u}_2 \dots \mathbf{u}_n]$, *de* $\mathbb{R}^n$, *constituída por vectores próprios de* $\mathbf{S}$ *(*$\mathbf{S}(\mathbf{u}_k) = \lambda_k\, \mathbf{u}_k$, $k = 1, \dots, n$*), e relativamente à qual a matriz de* $\mathbf{S}$ *é a matriz diagonal:*

$$[\mathbf{S}]_{\mathbf{U}} = \mathbf{diag}(\lambda_1, \lambda_2, \dots, \lambda_n)$$

pode ser escolhida de tal forma que, para cada $k = 1, \dots, n$, $\lambda_k = q(\mathbf{u}_k)$ *é o valor máximo de* q, *restrita à esfera unitária no subespaço de* $\mathbb{R}^n$, *perpendicular aos vectores* $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_{k-1}$*".*

- Demonstração...

  Com efeito, escolhamos $\mathbf{u}_1$ como sendo um máximo condicionado da restrição de q, à esfera $S_1 \equiv \{\mathbf{x} \in \mathbb{R}^n : \|\mathbf{x}\|^2 = 1\}$ (isto é sempre possível...). Consideremos o subespaço de $\mathbb{R}^n$, perpendicular a $\mathbf{u}_1$:

  $$V(\mathbf{u}_1) \equiv \{\mathbf{x} \in \mathbb{R}^n : \mathbf{x} \cdot \mathbf{u}_1 = 0\}$$

  e escolhamos $\mathbf{u}_2$ como sendo um máximo condicionado da restrição de q, à esfera $S_2 \equiv \{\mathbf{x} \in V(\mathbf{u}_1) : \|\mathbf{x}\|^2 = 1\}$ (isto é sempre possível...). Consideremos de seguida, o subespaço de $\mathbb{R}^n$, perpendicular a $\mathbf{u}_1$ e a $\mathbf{u}_2$:

  $$V(\mathbf{u}_1, \mathbf{u}_2) \equiv \{\mathbf{x} \in \mathbb{R}^n : \mathbf{x} \cdot \mathbf{u}_1 = 0 = \mathbf{x} \cdot \mathbf{u}_2\}$$

  e escolhamos $\mathbf{u}_3$ como sendo um máximo condicionado da restrição de q, à esfera $S_3 \equiv \{\mathbf{x} \in V(\mathbf{u}_1, \mathbf{u}_2) : \|\mathbf{x}\|^2 = 1\}$ (isto é sempre possível...).

  Procedendo sucessivamente desta forma, conseguimos n vectores $\mathbf{u}_1, \dots, \mathbf{u}_n$ que são evidentemente ortonormais. Resta provar que eles são vectores próprios de $\mathbf{S}$.

  Como por construção, q tem um máximo condicionado em $\mathbf{u}_1$, quando restrita à esfera S_1, existe um multiplicador de Lagrange λ_1, tal que:

  $$\nabla q(\mathbf{u}_1) = \lambda_1 \nabla g(\mathbf{u}_1) \tag{2.7.3}$$

  onde $g(\mathbf{x}) = \|\mathbf{x}\|^2 - 1$. Como vimos em (2.4.3), o gradiente de q é dado por $\nabla q(\mathbf{x}) = 2\mathbf{S}(\mathbf{x})$, e em particular $\nabla g(\mathbf{x}) = 2\mathbf{x}$. Portanto a condição (**??**) é equivalente a:

  $$\mathbf{S}(\mathbf{u}_1) = \lambda_1 \mathbf{u}_1$$

  o que significa exactamente que $\mathbf{u}_1$ é vector próprio associado ao valor próprio λ_1.

  O mesmo argumento pode ser utilizado sucessivamente, para concluir que $\mathbf{u}_k$ é vector próprio de $\mathbf{S}$.

  A forma quadrática associada a $\mathbf{S}$ pode então ser escrita na forma diagonal:

  $$q(\mathbf{x}) = q(y^1, \dots, y^n) = \lambda_1 (y^1)^2 + \lambda_2 (y^2)^2 + \dots + \lambda_n (y^n)^2 \tag{2.7.4}$$

  e é claro que $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n$,

♣.

2.8 Apêndice. Revisão de alguns conceitos de Álgebra Linear

2.8.1 Formas lineares e afins

Uma aplicação linear $\phi : \mathbb{R}^n \to \mathbb{R}$, é usualmente chamada uma **forma linear** (ou **funcional linear**) em $\mathbb{R}^n$. O conjunto de tais formas lineares tem uma estrutura natural de espaço vectorial real de dimensão n, que se diz o **espaço dual** de $\mathbb{R}^n$ e se representa por $(\mathbb{R}^n)^*$.

♣ **Proposição 2.15** ... *Uma função $\phi : \mathbb{R}^n \to \mathbb{R}$ é uma forma linear em $\mathbb{R}^n$, se e só se ϕ é do tipo:*

$$\phi(\mathbf{x}) = \mathbf{x} \cdot \mathbf{a} \qquad \forall \mathbf{x} \in \mathbb{R}^n \tag{2.8.1}$$

para um único vector $\mathbf{a} \in \mathbb{R}^n$ (que apenas depende de ϕ).

- Demonstração...

  A função definida por (2.8.1) é linear, como se deduz das propriedades do produto interno, que vimos no capítulo 1.

  Recìprocamente, dada uma forma linear $\phi : \mathbb{R}^n \to \mathbb{R}$, consideremos o vector $\mathbf{a} = (a_1, a_2, ..., a_n)$, cujas componentes na base canónica $\{\mathbf{e}_i\}$ de $\mathbb{R}^n$, são dadas por:

  $$a_i = \phi(\mathbf{e}_i) \qquad i = 1, 2, ..., n$$

  É fácil ver que (2.8.1) se verifica. De facto:

  $$\phi(\mathbf{x}) = \phi(\sum_{i=1}^{n} x^i \mathbf{e}_i) = \sum_{i=1}^{n} x^i \phi(\mathbf{e}_i) = \sum_{i=1}^{n} x^i a_i = \mathbf{x} \cdot \mathbf{a}$$

  Finalmente $\mathbf{a}$ é único: se $\mathbf{b}$ satisfaz também (2.8.1), então $\mathbf{x} \cdot (\mathbf{a} - \mathbf{b}) = 0, \quad \forall \mathbf{x} \in \mathbb{R}^n$, o que implica que $\mathbf{a} = \mathbf{b}$,

♣.

Uma **forma afim** em $\mathbb{R}^n$ é uma aplicação $T : \mathbb{R}^n \to \mathbb{R}$, do tipo:

$$T(\mathbf{x}) = \phi(\mathbf{x}) + b \tag{2.8.2}$$

onde ϕ é uma forma linear em $\mathbb{R}^n$ e $b \in \mathbb{R}$. Portanto, pela proposição anterior, T é afim se e só se é do tipo:

$$T(\mathbf{x}) = \mathbf{x} \cdot \mathbf{a} + b \tag{2.8.3}$$

2.8.2 Aplicações lineares simétricas. Formas quadráticas

Vamos começar por rever alguns conceitos de álgebra linear que serão úteis neste capítulo e em outros posteriores.

Consideremos um espaço vectorial real $\mathcal{V}$, e um endomorfismo linear $\mathbf{T} : \mathcal{V} \to \mathcal{V}$.

♣ **Definição 2.19** ... *Um escalar* $\lambda \in \mathbb{R}$*, diz-se um* **valor próprio** *de* $\mathbf{T}$*, se existe um vector* *não nulo* $\mathbf{x} \in \mathcal{V} - \{\mathbf{O}\}$*, tal que:*

$$\mathbf{T}(\mathbf{x}) = \lambda\,\mathbf{x} \tag{2.8.4}$$

Neste caso, o vector não nulo $\mathbf{x} \in \mathcal{V} - \{\mathbf{O}\}$*, que satisfaz (2.8.4), diz-se um* **vector próprio** *pertencente ao valor próprio* λ*.*

Nota...

Quando $\mathcal{V}$ é um espaço vectorial complexo, e T um endomorfismo linear (complexo) de $\mathcal{V}$, podemos dar exactamente a mesma definição, com $\mathbb{C}$ em vez de $\mathbb{R}$.

Quando $\mathcal{V} = \mathbb{R}^n$, munido do produto interno usual (1.1.4), podemos expressar o valor próprio λ, na forma:

$$\lambda = \frac{\mathbf{T}(\mathbf{x}) \cdot \mathbf{x}}{\|\mathbf{x}\|^2} \tag{2.8.5}$$

onde $\mathbf{x}$ é um vector próprio pertencente ao valor próprio λ. De facto:

$$\mathbf{T}(\mathbf{x}) = \lambda\,\mathbf{x} \quad \Rightarrow \quad \mathbf{T}(\mathbf{x}) \cdot \mathbf{x} = (\lambda\mathbf{x}) \cdot \mathbf{x} = \lambda\,\|\mathbf{x}\|^2$$

o que implica (2.8.5), já que $\mathbf{x} \neq \mathbf{O}$. Este resultado continua válido, quando $\mathcal{V}$ é um qualquer espaço vectorial real, munido de um produto interno.

É imediato verificar que o conjunto constituído pelo vector nulo $\mathbf{O}$, e por todos os vectores próprios, pertencentes ao mesmo valor próprio λ:

$$E(\lambda) \equiv \{\mathbf{x} \in \mathcal{V} : \mathbf{T}(\mathbf{x}) = \lambda\,\mathbf{x}\} \tag{2.8.6}$$

é um subespaço vectorial de $\mathcal{V}$, de dimensão $\dim E(\lambda) \geq 1$. Este subespaço chama-se o **espaço próprio** correspondente ao valor próprio λ.

♣ **Proposição 2.16** ... *Vectores próprios pertencentes a valores próprios distintos, são linearmente independentes, isto é: se* $\mathbf{T}(\mathbf{x}) = \lambda\,\mathbf{x}$ *(*$\mathbf{x} \neq \mathbf{O}$*),* $\mathbf{T}(\mathbf{y}) = \mu\,\mathbf{y}$ *(*$\mathbf{y} \neq \mathbf{O}$*), e* $\lambda \neq \mu$*, então* $\mathbf{x}$ *e* $\mathbf{y}$ *são linearmente independentes.*

- Demonstração...

  Com efeito, suponhamos que:

$$a\mathbf{x} + b\mathbf{y} = \mathbf{O} \qquad a, b \in \mathbb{R} \tag{2.8.7}$$

Então:

$$\mathbf{T}(a\mathbf{x}+b\mathbf{y}) = a\mathbf{T}(\mathbf{x})+b\mathbf{T}(\mathbf{y}) = a\lambda\mathbf{x}+b\mu\mathbf{y} = \mathbf{T}(\mathbf{O}) = \mathbf{O} \tag{2.8.8}$$

Multiplicando (2.8.7) por λ e subtraindo (2.8.8), obtemos:

$$b(\lambda-\mu)\mathbf{y} = \mathbf{O}$$

o que implica que $b=0$, uma vez que $\lambda \neq \mu$. Anàlogamente se obtem que $a=0$,

♣.

♣ **Corolário 2.2** ... *Se* $\dim \mathcal{V} = n$, *todo o endomorfismo linear* $\mathbf{T}: \mathcal{V} \to \mathcal{V}$ *tem quando muito* n *valores próprios distintos.*

Se $\mathbf{T}$ *tem exactamente* n *valores próprios distintos, então os correspondentes vectores próprios formam uma base de* $\mathcal{V}$, *e a matriz de* $\mathbf{T}$ *nessa base, é uma matriz diagonal cujas entradas da diagonal principal, são esses valores próprios.*

Um caso especialmente importante acontece quando $\mathcal{V} = \mathbb{R}^n$, munido do produto interno usual (ou mais geralmente, quando $\mathcal{V}$ é um qualquer espaço vectorial real, munido de um produto interno), e $\mathbf{T}$ é um endomorfismo simétrico:

♣ **Definição 2.20** ... *Um endomorfismo linear* $\mathbf{S}: \mathbb{R}^n \to \mathbb{R}^n$, *diz-se* **simétrico** *se* $\mathbf{S}$ *satisfaz a condição:*

$$\mathbf{S}(\mathbf{x})\cdot\mathbf{y} = \mathbf{x}\cdot\mathbf{S}(\mathbf{y}) \qquad \forall \mathbf{x},\mathbf{y} \in \mathbb{R}^n \tag{2.8.9}$$

♣ **Proposição 2.17** ... *Seja* $\mathbf{S}: \mathbb{R}^n \to \mathbb{R}^n$, *um endomorfismo simétrico em* $\mathbb{R}^n$ *(ou mais geralmente, num qualquer espaço vectorial real* $\mathcal{V}$, *munido de um produto interno). Suponhamos que* $\mathbf{x}$ *e* $\mathbf{y}$ *são vectores próprios, pertencentes respectivamente aos valores próprios distintos* λ *e* μ, *de* $\mathbf{S}$.

Então $\mathbf{x}$ *e* $\mathbf{y}$ *são ortogonais:* $\mathbf{x}\cdot\mathbf{y} = 0$.

- Demonstração...

Temos sucessivamente que:

$$\lambda(\mathbf{x}\cdot\mathbf{y}) = (\lambda\mathbf{x})\cdot\mathbf{y} = \mathbf{S}(\mathbf{x})\cdot\mathbf{y} = \mathbf{x}\cdot\mathbf{S}(\mathbf{y}) = \mathbf{x}\cdot(\mu\mathbf{y}) = \mu(\mathbf{x}\cdot\mathbf{y})$$

o que implica que $(\lambda-\mu)(\mathbf{x}\cdot\mathbf{y}) = 0$, e portanto $\mathbf{x}\cdot\mathbf{y} = 0$, já que $\lambda \neq \mu$,

♣.

Notemos que um endomorfismo linear real pode não ter valores próprios reais (por exemplo, uma rotação em $\mathbb{R}^2$). No entanto, é possível provar que um endomorfismo simétrico admite sempre um valor próprio (real). Além disso todos os seus valores próprios são reais (atendendo a (2.8.5)).

O facto de maior interesse sobre os endomorfismos simétricos, é que eles podem ser diagonalizados. Mais precisamente, é válida a seguinte proposição:

♣ **Proposição 2.18** *(***Diagonalização de endomorfismos simétricos***)*

... Seja $\mathbf{S} : \mathbb{R}^n \to \mathbb{R}^n$*, um endomorfismo simétrico em* $\mathbb{R}^n$ *(ou mais geralmente, num qualquer espaço vectorial real* $\mathcal{V}$*, munido de um produto interno, com* $\dim \mathcal{V} = n$*).*

Então existe uma base ortonormal $\{\mathbf{u}_1, \mathbf{u}_2, ..., \mathbf{u}_n\}$*, para* $\mathbb{R}^n$ *(ou* $\mathcal{V}$*), constituída por vectores próprios de* $\mathbf{S}$*.*

A matriz de $\mathbf{S}$ *nessa base é portanto a matriz diagonal* $\mathbf{diag}(\lambda_1, \lambda_2, ..., \lambda_n)$*, onde* λ_k *é o valor próprio correspondente ao vector próprio* $\mathbf{u}_k$*, para* $(k = 1, ..., n)$*.*

- Demonstração...

  A demonstração faz-se por indução sobre a dimensão n. Se $n = 1$, o resultado é trivial. Suponhamos que ele é válido, para todo o espaço vectorial real, munido de um produto interno, com dim $\leq n-1$.

  Como se referiu acima, $\mathbf{S}$ admite sempre um valor próprio (real) λ_1. Seja $\mathbf{u}_1 \neq \mathbf{O}$ um vector próprio pertencente ao valor próprio λ_1: $\mathbf{S}(\mathbf{u}_1) = \lambda_1 \mathbf{u}_1$. Podemos supôr que $\|\mathbf{u}_1\| = 1$. Seja $\mathcal{S}$ o subespaço ortogonal a $\mathbf{u}_1$, de tal forma que:
  $$\mathcal{V} = \mathbb{R}\,\mathbf{u}_1 \oplus \mathcal{S} \tag{2.8.10}$$

  Então $\mathbf{S}$ deixa $\mathcal{S}$ invariante: $\mathbf{S}(\mathcal{S}) \subseteq \mathcal{S}$ (porquê?). Além disso, $\mathcal{S}$ é um espaço vectorial com um produto interno, de dimensão $n-1$, e $\mathbf{S}\,|_{\mathcal{S}}$ é simétrico. Resta aplicar a hipótese de indução para concluir a prova,

♣.

Suponhamos novamente que $\mathcal{V}$ é um espaço vectorial real, com $\dim \mathcal{V} = n$, e vejamos um processo prático para o cálculo dos valores próprios de um endomorfismo linear $\mathbf{T} : \mathcal{V} \to \mathcal{V}$.

Para isso, comecemos por notar que a condição (2.8.4), é equivalente à condição de que:
$$(\mathbf{T} - \lambda \mathrm{Id})\,\mathbf{x} = \mathbf{O}$$
para algum vector não nulo $\mathbf{x}$, o que significa que o endomorfismo linear $\mathbf{T} - \lambda\mathrm{Id}$ é singular (tem um núcleo não trivial).

Como sabemos, se $[\mathbf{T}]$ é uma representação matricial de $\mathbf{T}$, numa dada base de $\mathcal{V}$, e se representamos por Id a matriz identidade $n \times n$, então o endomorfismo linear $\mathbf{T} - \lambda\mathrm{Id}$ é singular, se e só se:
$$\det\left([\mathbf{T}] - \lambda \mathrm{Id}\right) = 0 \tag{2.8.11}$$

Concluindo, se $\lambda \in \mathbb{R}$ é um valor próprio de $\mathbf{T}$, então λ satisfaz a equação (2.8.11). Recìprocamente, qualquer solução real $\lambda \in \mathbb{R}$ da referida equação (2.8.11), é um valor próprio de $\mathbf{T}$.

♣ **Definição 2.21** *... O polinómio em λ:*

$$p(\lambda) = det([\mathbf{T}] - \lambda\, Id) \tag{2.8.12}$$

diz-se o **polinómio característico** *do endomorfismo linear* $\mathbf{T}$*, e a equação (2.8.11):* $p(\lambda) = 0$*, diz-se a* **equação característica** *de* $\mathbf{T}$.

Nota...

Notemos que o polinómio característico do endomorfismo linear $\mathbf{T}$, não depende da representação matricial de $\mathbf{T}$. De facto, qualquer outra representação matricial de $\mathbf{T}$, é do tipo $C^{-1}[\mathbf{T}]C$, onde C é uma matriz inversível, e tem-se que:

$$\begin{aligned} \det(C^{-1}[\mathbf{T}]C - \lambda\,\mathrm{Id}) &= \det(C^{-1}[\mathbf{T}]C - \lambda C^{-1}C) = \det(C^{-1}([\mathbf{T} - \lambda\,\mathrm{Id})C) \\ &= \det([\mathbf{T}] - \lambda\,\mathrm{Id}) = p(\lambda) \end{aligned}$$

Portanto para encontrar os valores próprios (reais) de $\mathbf{T}$, devemos calcular as soluções reais da equação característica de $\mathbf{T}$:

$$p(\lambda) = \det([\mathbf{T}] - \lambda\mathrm{Id}) = 0 \tag{2.8.13}$$

Para encontrar os vectores próprios pertencentes a um dado valor próprio λ, devemos encontrar os vectores não nulos $\mathbf{x}$, vistos como um "vector coluna", do tipo:

$$[\mathbf{x}] = \begin{bmatrix} x^1 \\ x^2 \\ \vdots \\ x^n \end{bmatrix} \tag{2.8.14}$$

que satisfazem a equação matricial:

$$([\mathbf{T}] - \lambda\,\mathrm{Id})\,[\mathbf{x}] = \mathbf{O} \tag{2.8.15}$$

que não é mais do que um sistema de n equações lineares para as componentes $x^1, ..., x^n$ de $\mathbf{x}$.

Representemos por:

$$[\mathbf{E}] \equiv [\mathbf{e}_1 \;\; \mathbf{e}_2 \;\; ... \;\; \mathbf{e}_n] \tag{2.8.16}$$

a matriz (de entradas vectoriais) $(1 \times n)$, cujas entradas são os vectores da base canónica $[\mathbf{E}]$, de $\mathbb{R}^n$.

Suponhamos agora que:

$$[\mathbf{U}] \equiv [\mathbf{u}_1 \;\; \mathbf{u}_2 \;\; ... \;\; \mathbf{u}_n] \tag{2.8.17}$$

é a base ortonormada $[\mathbf{U}]$, para $\mathbb{R}^n$, constituída por vectores próprios de $\mathbf{T}$, e relativamente à qual a matriz de $\mathbf{T}$ é a matriz diagonal:

$$\mathbf{diag}(\lambda_1, \lambda_2, ..., \lambda_n)$$

onde λ_k é o valor próprio correspondente ao vector próprio $\mathbf{u}_k$, para $(k = 1, ..., n)$.

As equações:

$$\mathbf{u}_k = \sum_{i=1}^{n} p_k^i \mathbf{e}_i \qquad k = 1, ..., n \tag{2.8.18}$$

que exprimem cada vector da base $[\mathbf{U}]$, como combinação linear dos vectores da base $[\mathbf{E}]$, podem simplesmente ser escritas na forma matricial:

$$[\mathbf{U}] = [\mathbf{E}]\, P \tag{2.8.19}$$

onde P é a transposta da matriz $[p_{ik}]$, definida por (2.8.18). P é a chamada **matriz de passagem** da base $[\mathbf{E}]$ para a base $[\mathbf{U}]$, e portanto, a sua k-coluna é constituída pelas componentes de $\mathbf{u}_k$, relativamente à base $[\mathbf{E}]$.

Notemos que no caso presente, a matriz P é ortogonal, isto é:

$$P^{tr} = P^{-1} \tag{2.8.20}$$

já que é a matriz de passagem de uma base ortonormada $[\mathbf{E}]$, para uma outra base $[\mathbf{U}]$, igualmente ortonormada.

Dado um vector $\mathbf{x}$, podemos escrever:

$$\mathbf{x} = \sum_{i=1}^{n} x^i\, \mathbf{e}_i \quad \text{e também} \quad \mathbf{x} = \sum_{k=1}^{n} y^k\, \mathbf{u}_k$$

de tal forma que os correspondentes vectores coluna, das componentes de $\mathbf{x}$, nas bases $[\mathbf{E}]$ e $[\mathbf{U}]$, são dados respectivamente por:

$$[\mathbf{x}]_{\mathbf{E}} \equiv \begin{bmatrix} x^1 \\ x^2 \\ \vdots \\ x^n \end{bmatrix}$$

e:

$$[\mathbf{x}]_{\mathbf{U}} \equiv \begin{bmatrix} y^1 \\ y^2 \\ \vdots \\ y^n \end{bmatrix}$$

Atendendo agora a (2.8.18), temos que:

$$\mathbf{x} = \sum_{i=1}^{n} x^i\, \mathbf{e}_i = \sum_{k=1}^{n} y^k\, \mathbf{u}_k = \sum_{i=1}^{n} \sum_{k=1}^{n} y^k p_k^i \mathbf{e}_i$$

donde se deduz que:

$$x^i = \sum_{k=1}^{n} p_k^i y^k$$

ou em forma matricial:

$$[\mathbf{x}]_{\mathbf{E}} = P\, [\mathbf{x}]_{\mathbf{U}} \tag{2.8.21}$$

Portanto:

$$[\mathbf{x}]_{\mathbf{U}} = P^{-1}\,[\mathbf{x}]_{\mathbf{E}} \tag{2.8.22}$$

o que permite deduzir as coordenadas de $\mathbf{x}$ na base $[\mathbf{U}]$, conhecendo as coordenadas na base $[\mathbf{E}]$.

Designemos agora por $[\mathbf{T}]_{\mathbf{E}}$, a matriz de um endomorfismo linear $\mathbf{T}$, relativamente à base canónica $[\mathbf{E}]$. Recordemos como se define esta matriz: exprimimos cada $\mathbf{T}(\mathbf{e}_k)$, como combinação linear dos vectores da base $[\mathbf{E}]$:

$$\mathbf{T}(\mathbf{e}_k) = \sum_{i=1}^{n} t_k^i \mathbf{e}_i \qquad k = 1, ..., n$$

e a matriz $[\mathbf{T}]_{\mathbf{E}}$, é por definição a <u>transposta da matriz $[t_k^i]$</u>. Desta forma, a equação $\mathbf{y} = \mathbf{T}(\mathbf{x})$ escreve-se na forma matricial:

$$[\mathbf{y}]_{\mathbf{E}} = [\mathbf{T}]_{\mathbf{E}}\,[\mathbf{x}]_{\mathbf{E}} \tag{2.8.23}$$

Notemos que esta matriz é simétrica (=à sua transposta), quando $\mathbf{T} = \mathbf{S}$ é endomorfismo simétrico (de facto, $s_{ik} = \mathbf{S}(\mathbf{e}_k) \cdot \mathbf{e}_i = \mathbf{e}_k \cdot \mathbf{S}(\mathbf{e}_i) = s_{ki}$).

Consideremos agora a matriz de $\mathbf{T}$ na base $[\mathbf{U}]$: $[\mathbf{T}]_{\mathbf{U}}$, e vejamos como é que esta matriz se relaciona com $[\mathbf{T}]_{\mathbf{E}}$.

Atendendo a (2.8.23), podemos escrever:

$$[\mathbf{y}]_{\mathbf{U}} = [\mathbf{T}]_{\mathbf{U}}\,[\mathbf{x}]_{\mathbf{U}}$$

Mas por (2.8.21) e (2.8.22), vem que:

$$[\mathbf{T}]_{\mathbf{E}}\,[\mathbf{x}]_{\mathbf{E}} = [\mathbf{y}]_{\mathbf{E}} = P\,[\mathbf{y}]_{\mathbf{U}} = P\,[\mathbf{T}]_{\mathbf{U}}\,[\mathbf{x}]_{\mathbf{U}} = P\,[\mathbf{T}]_{\mathbf{U}}\,P^{-1}\,[\mathbf{x}]_{\mathbf{E}}$$

isto é, $[\mathbf{T}]_{\mathbf{E}} = P\,[\mathbf{T}]_{\mathbf{U}}\,P^{-1}$, ou de forma equivalente:

$$[\mathbf{T}]_{\mathbf{U}} = P^{-1}\,[\mathbf{T}]_{\mathbf{E}}\,P \tag{2.8.24}$$

<u>**Exemplo**</u> ...

Seja $\mathbf{S}$ o endomorfismo simétrico em $\mathbb{R}^3$, cuja matriz na base canónica de $\mathbb{R}^3$ é (a matriz simétrica):

$$[\mathbf{S}]_{\mathbf{E}} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 2 \\ 0 & 2 & 1 \end{bmatrix}$$

A equação característica é:

$$p(\lambda) = \det\left([\mathbf{S}] - \lambda \mathrm{Id}\right) = \begin{bmatrix} 1-\lambda & 0 & 0 \\ 0 & 1-\lambda & 2 \\ 0 & 2 & 1-\lambda \end{bmatrix} = 0$$

isto é:

$$(1-\lambda)[(1-\lambda)^2 - 4] = 0$$

Os valores próprios de T, são portanto $\lambda = 1, -1, 3$.

Calculemos uma base ortonormada de vectores próprios. Para isso substituímos sucessivamente λ por $1, -1$ e 3, na equação matricial seguinte:

$$\begin{bmatrix} 1-\lambda & 0 & 0 \\ 0 & 1-\lambda & 2 \\ 0 & 2 & 1-\lambda \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

Resolvendo os correspondentes sistemas de equações, e tendo o cuidado de normalizar os vectores próprios para que eles tenham norma 1, obtemos a base seguinte:

$$\begin{aligned} \mathbf{u}_1 &= \mathbf{e}_1 = (1,0,0) && \text{pertencente ao valor próprio } \lambda = 1 \\ \mathbf{u}_2 &= \frac{1}{\sqrt{2}}(0,1,-1) && \text{pertencente ao valor próprio } \lambda = -1 \\ \mathbf{u}_3 &= \frac{1}{\sqrt{2}}(0,1,-1) && \text{pertencente ao valor próprio } \lambda = 3 \end{aligned}$$

A matriz de passagem da base canónica $[\mathbf{E}] = [\mathbf{i}\ \ \mathbf{j}\ \ \mathbf{k}]$, para a base $[\mathbf{U}] = [\mathbf{u}_1\ \ \mathbf{u}_2\ \ \mathbf{u}_3]$, é portanto a matriz ortogonal ($P^{-1} = P^{tr}$):

$$P = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ 0 & -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}$$

Podemos verificar directamente que:

$$P^{tr}[T]_{\mathbf{E}}P = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 3 \end{bmatrix}$$

Regressemos finalmente ao tópico desta secção, nomeadamente o de formas quadráticas em $\mathbb{R}^n$.

♣ <u>**Definição**</u> **2.22** ... *Seja* $\mathbf{S} : \mathbb{R}^n \to \mathbb{R}^n$ *um endomorfismo simétrico de* $\mathbb{R}^n$*. Define-se então a* **forma quadrática** *associada a* $\mathbf{S}$*, como sendo a função* $q : \mathbb{R}^n \to \mathbb{R}$*, definida por:*

$$q(\mathbf{x}) = \mathbf{x} \cdot \mathbf{S}(\mathbf{x}) \tag{2.8.25}$$

Se $\mathbf{x}$ é interpretado como o vector coluna $[\mathbf{x}] = [\mathbf{x}]_{\mathbf{E}}$ (ver (2.8.14)) (onde $(x^1, ..., x^n)$ representam as componentes de $\mathbf{x}$, na base canónica de $\mathbb{R}^n$), e se $[\mathbf{S}] = [\mathbf{S}]_{\mathbf{E}}$ é a representação matricial de $\mathbf{S}$, na base canónica de $\mathbb{R}^n$, podemos escrever (2.8.25) na forma:

$$q(\mathbf{x}) = [\mathbf{x}]^{tr}\,[\mathbf{S}]\,[\mathbf{x}] \tag{2.8.26}$$

onde $[\mathbf{x}]^{tr}$ designa a transposta de $[\mathbf{x}]$.

Consideremos agora a base ortonormada $[\mathbf{U}]$, de $\mathbb{R}^n$, constituída por vectores próprios de $\mathbf{S}$, e relativamente à qual a matriz de $\mathbf{S}$ é a matriz diagonal:

$$[\mathbf{S}]_{\mathbf{U}} = \mathbf{diag}(\lambda_1, \lambda_2, ..., \lambda_n)$$

onde λ_k é o valor próprio correspondente ao vector próprio $\mathbf{u}_k$, para $(k = 1, ..., n)$.

A forma quadrática associada a $\mathbf{S}$, que foi escrita na forma (ver (2.8.26)):

$$q(\mathbf{x}) = [\mathbf{x}]_{\mathbf{E}}^{tr}\,[\mathbf{S}]_{\mathbf{E}}\,[\mathbf{x}]_{\mathbf{E}}$$

pode ser agora reescrita na base $[\mathbf{U}]$:

$$\begin{aligned} q(\mathbf{x}) &= [\mathbf{x}]_{\mathbf{E}}^{tr}\,[\mathbf{S}]_{\mathbf{E}}\,[\mathbf{x}]_{\mathbf{E}} \\ &= (P\,[\mathbf{x}]_{\mathbf{U}})^{tr}\,P\,[\mathbf{S}]_{\mathbf{U}}\,P^{-1}(P\,[\mathbf{x}]_{\mathbf{U}}) \\ &= [\mathbf{x}]_{\mathbf{U}}^{tr}\,[\mathbf{S}]_{\mathbf{U}}\,[\mathbf{x}]_{\mathbf{U}} \\ &= [\mathbf{x}]_{\mathbf{U}}^{tr}\,\mathbf{diag}(\lambda_1, \lambda_2, ..., \lambda_n)\,[\mathbf{x}]_{\mathbf{U}} \end{aligned} \tag{2.8.27}$$

atendendo a que $P\,P^{tr} = \mathrm{Id}$. Pondo:

$$[\mathbf{x}]_{\mathbf{U}} \equiv \begin{bmatrix} y^1 \\ y^2 \\ \vdots \\ y^n \end{bmatrix}$$

podemos finalmente escrever q na forma:

$$q(\mathbf{x}) = q(y^1, ..., y^n) = \lambda_1 (y^1)^2 + \lambda_2 (y^2)^2 + ... + \lambda_n (y^n)^2 \tag{2.8.28}$$

que se diz a **forma diagonal ou canónica** da forma quadrática q.

Exemplo ...

Continuando o exemplo anterior, consideremos a forma quadrática associada ao endomorfismo simétrico aí referido:

$$\begin{aligned} q(x, y, z) &= [x \;\; y \;\; z] \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 2 \\ 0 & 2 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} \\ &= x^2 + y^2 + z^2 + 4yz \end{aligned}$$

Se designamos por (u, v, w) as coordenadas de um vector $\mathbf{x}$, na base $\mathbf{U}$, então, se as coordenadas desse mesmo vector, na base $\mathbf{E}$, são (x, y, z), vem que:

$$[\mathbf{x}]_{\mathbf{E}} = \begin{bmatrix} x \\ y \\ z \end{bmatrix} = P \begin{bmatrix} u \\ v \\ w \end{bmatrix}$$

isto é:

$$\begin{aligned} x &= u \\ y &= \frac{1}{\sqrt{2}}v + \frac{1}{\sqrt{2}}w \\ z &= -\frac{1}{\sqrt{2}}v + \frac{1}{\sqrt{2}}w \end{aligned}$$

e nas novas coordenadas (u, v, w), q escreve-se na forma:

$$q(u, v, w) = u^2 - v^2 + 3w^2$$

como aliás pode ser verificado directamente.

♣ <u>**Definição**</u> **2.23** ... *Uma forma quadrática em* $\mathbb{R}^n$, $q(\mathbf{x}) = \mathbf{S}\mathbf{x} \cdot \mathbf{x}$, *diz-se:*

(i)... **definida positiva**, *se* $q(\mathbf{x}) > 0$, $\forall \mathbf{x} \neq \mathbf{O}$.

(ii)... **definida negativa**, *se* $q(\mathbf{x}) < 0$, $\forall \mathbf{x} \neq \mathbf{O}$.

(iii)... **indefinida**, *se* q *toma valores positivos e negativos.*

A proposição seguinte é consequência imediata da possibilidade de reduzir uma forma quadrática à forma diagonal.

♣ <u>**Proposição**</u> **2.19** ... *Uma forma quadrática em* $\mathbb{R}^n$, $q(\mathbf{x}) = \mathbf{S}\mathbf{x} \cdot \mathbf{x}$, *é:*

(i)... **definida positiva**, *se todos os valores próprios de* $\mathbf{S}$ *são estritamente positivos.*

(ii)... **definida negativa**, *se todos os valores próprios de* $\mathbf{S}$ *são estritamente negativos.*

(iii)... **indefinida**, *se os valores próprios de* $\mathbf{S}$ *são alguns positivos e alguns negativos (eventualmente nulos).*

2.9 Exercícios

<u>**Exercício**</u> **1** ... Relativamente às funções que se seguem, defina e faça um esboço do respectivo domínio:

(i). $f(x, y) = \frac{y}{x}$, (ii). $f(x, y) = \frac{x+y}{x-y}$, (iii). $f(x, y) = \frac{x+y}{x^2+y^2-1}$, (iv). $f(x, y, z) = \frac{2x+y-z}{x^2+y^2+z^2-1}$, (v). $f(x, y, z) = \frac{z}{x^2-4y^2-1}$

<u>**Exercício**</u> **2** ... Defina e faça um esboço do gráfico das funções seguintes:

(i). $f(x, y) = x - y + 2$, (ii). $f(x, y) = 3x$, (iii). $f(x, y) = x^2 + y^2$, (iv). $f(x, y) = -y^2$

<u>**Exercício**</u> **3** ... Descreva e esboce alguns conjuntos de nível associados às seguintes funções:

(i). $f(x, y) = x^2 + 2y^2$, (ii). $f(x, y) = y - x^2$, (iii). $f(x, y) = x - y + 2$, (iv). $f(x, y, z) = x^2 + 3y^2 + z^2$

<u>**Exercício**</u> **4** ... Defina e faça um esboço da superfície definida por cada uma das seguintes equações:

(i). $z = x^2 + 2$, (ii). $z^2 + x^2 = 4$, (iii). $z = |y|$, (iv). $x^2 + y^2 = z^2$

Exercício 5 ... Relativamente às formas quadráticas seguintes, determine: (i). a correspondente aplicação linear simétrica **S**, (ii). os valores próprios de **S**, (iii). uma base ortonormal de vectores próprios, (iv). a correspondente forma canónica (diagonal):

(1)... $4x^2 + 4xy + y^2$, (2)... xy, (3). $x^2 + 2xy - y^2$, (4). $34x^2 - 24xy + 41y^2$, (5). $x^2 + xy + yz + xz$, (6). $2x^2 + 4xz + y^2 - z^2$, (7). $3x^2 + 4xy + 8xz + 4yz + 3z^2$

Classifique as formas quadráticas anteriores, e descreva os correspondentes conjuntos de nível.

Exercício 6 ... Considere a função $f(x,y) = \frac{x^2y^2}{x^2y^2+(x-y)^2}$.

(i)... Determine o domínio de f.

(ii)... Mostrar que:

$$\lim_{x\to 0}\,(\lim_{y\to 0} f(x,y)) = 0 = \lim_{y\to 0}\,(\lim_{x\to 0} f(x,y))$$

mas que não existe o limite de f em $(0,0)$.

Exercício 7 ... Mostre, usando a definição, que:

$$(i).\ \lim_{(x,y)\to(0,0)} \frac{x^2y}{x^2+y^2} = 0, \qquad (ii).\ \lim_{(x,y)\to(1,2)} (2x+3y) = 8$$

Exercício 8 ... Determine caso existam, os seguintes limites:

$$(i).\ \lim_{(x,y)\to(0,0)} \frac{xy}{x^2+y^2+2} \qquad (ii).\ \lim_{(x,y)\to(0,0)} \frac{e^{xy}}{x+1}$$

$$(iii).\ \lim_{(x,y)\to(0,0)} \frac{\sin(x^2-y^2)}{x-y} \qquad (iv).\ \lim_{(x,y)\to(0,0)} \frac{xy^2}{x^2+y^4}$$

$$(v).\ \lim_{(x,y)\to(0,0)} \frac{x-y}{x+3y} \qquad (vi).\ \lim_{(x,y)\to(0,0)} \frac{x^2+2xy+y^2}{x^2+y^2}$$

Exercício 9 ... Relativamente a cada uma das funções f que se seguem, determine o conjunto dos pontos nos quais f é contínua .

$$(i).f(x,y) = x^4 - y^3 + 2xy^2 \qquad (ii).f(x,y) = \log(x^2) + y^2$$

$$(iii).f(x,y) = \arctan(\tfrac{y}{x}) \qquad (iv).f(x,y) = \frac{e^{x+y}}{x^2-y^2}$$

Exercício 10 ... Considere o seguinte campo escalar:

$$f(x,y) = \begin{cases} 0 & \text{se } y \leq 0 \text{ ou } y \geq x^2 \\ 1 & \text{se } 0 < y < x^2 \end{cases}$$

Mostre que $f(x,y) \to 0$, sempre que $(x,y) \to 0$ ao longo de qualquer recta que passe pela origem. Determine uma curva que passe pela origem, e ao longo da qual f seja constante e igual a 1 (com excepção da origem). f será contínuo na origem?

Exercício 11 ... Considere o seguinte campo escalar:

$$f(x,y) = \begin{cases} x^3 \sin(\frac{1}{x}) + y^2 & \text{se } x \neq 0 \\ y^2 & \text{se } x = 0 \end{cases}$$

Mostre que f é contínuo em $\mathbb{R}^2$.

Exercício 12 ... Mostre que o campo escalar:

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

é descontínuo em $(0,0)$.

Exercício 13 ... Seja $f : \mathbb{R}^2 \to \mathbb{R}$ uma função contínua em (a,b) e tal que $f(a,b) > 0$. Mostre que existe uma bola centrada em (a,b), na qual $f > 0$.

Exercício 14 ... Diga se é possível definir uma extensão contínua a todo o $\mathbb{R}^3$, das funções que se seguem:

(i). $f(x,y,z) = \frac{xy+yz}{x^2+y^2+z^2}$, $(x,y,z) \neq (0,0,0)$

(ii). $f(x,y) = \frac{xy \sin x}{x^2+y^2}$, $(x,y) \neq (0,0)$

Exercício 15 ... Calcule, usando a definição, a derivada direccional da função $f(x,y) = x - 2x^2y + y^2$, $(x,y) \in \mathbb{R}^2$, no ponto $(1,1)$ e na direcção do vector $(-1,3)$.

<u>Exercício</u> 16 ... Considere a função:

$$f(x,y) = \begin{cases} x+y & \text{se } x=0 \text{ ou } y=0 \\ 1 & \text{se } x\neq 0 \text{ e } y \neq 0 \end{cases}$$

Verifique se existem, e em caso afirmativo calcule:

(i). a derivada de f em $(0,0)$, segundo a direcção $(1,0)$.

(ii). a derivada de f em $(0,0)$, segundo a direcção $(0,1)$.

(iii). a derivada de f em $(0,0)$, segundo a direcção (a,b), com $a \neq 0$ e $b \neq 0$.

<u>Exercício</u> 17 ... Calcule as derivadas parciais de primeira ordem, de cada uma das seguintes funções:

$$(i).f(x,y) = e^{xy} \qquad (ii).f(x,y) = (x^2+y^2)\log(x^2+y^2)$$

$$(iii).f(x,y) = \cos(y\,e^{xy})\sin x \qquad (iv).f(x,y,z) = x(\cos y)(\cos z)$$

(v). $f(x,y) = \log(\sqrt{1+xy})$ nos pontos $(1,2)$ e $(0,0)$

(vi). $f(x,y,z) = x^{y^2} + \tan(xz) + 3y^2 - 2x$

<u>Exercício</u> 18 ... Calcule os gradientes das funções que se seguem:

(i). $f(x,y,z) = \sqrt{x^2+y^2+z^2}$, (ii). $f(x,y,z) = xy + yz + xz$

(iii). $f(x,y) = x\,e^{x^2+y^2}$, (iv). $f(x,y,z) = \log\sin\sqrt{x^2+y^2+z^2}$

<u>Exercício</u> 19 ... Calcule as derivadas direccionais das funções que se seguem, nos pontos e segundo as direcções indicadas:

(i). $f(x,y) = x + 2x^2 - 3xy \qquad \mathbf{p} = (1,1) \qquad \mathbf{v} = (\frac{3}{5}, \frac{4}{5})$

(ii). $f(x,y) = \log\sqrt{x^2+y^2} \qquad \mathbf{p} = (1,0) \qquad \mathbf{v} = (\frac{2\sqrt{5}}{5}, \frac{\sqrt{5}}{5})$

(iii). $f(x,y) = xyz \qquad \mathbf{p} = (1,1,1) \qquad \mathbf{v} = (\frac{1}{\sqrt{2}}, 0, \frac{1}{\sqrt{2}})$

(i). $f(x,y,z) = e^x + yz \qquad \mathbf{p} = (1,1,1) \qquad \mathbf{v} = (\frac{1}{\sqrt{3}}, -\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}})$

<u>Exercício</u> 20 ... Considere a função $f(x,y,z) = \sin(xy)\,e^{-z^2}$. Em que direcção, a partir do ponto $(1,\pi,0)$, nos devemos deslocar de modo a que f cresça mais ràpidamente?

Exercício 21 ... Suponha que a temperatura T, em cada ponto do espaço, é dada pela função $T(x, y, z) = e^{-x} + e^{-2y} + e^{3z}$.

Se estivermos no ponto $(1, 1, 1)$, em que direcção nos devemos deslocar de modo a que T decresça o mais ràpidamente possível?

Exercício 22 ... Uma chapa de metal aquecida está situada num plano xy, de tal forma que a temperatura T, em cada ponto, é inversamente proporcional à distância desse ponto à origem.

Se a temperatura em $\mathbf{p} = (3, 4)$ é de 100^o, determine a derivada de T em $\mathbf{p}$, na direcção do vector $(1, 1)$.

Em que direcção T cresce mais ràpidamente, a partir do ponto $\mathbf{p}$?

Em que direcção a variação de T é nula, a partir do ponto $\mathbf{p}$?

Exercício 23 ... Verifique se a função:

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

é contínua em $(-1, 1)$. É diferenciável nesse ponto?

Exercício 24 ... Considere o seguinte campo escalar:

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

(i). Mostre que f é descontínua em $(0, 0)$.

(ii). Defina as funções $\frac{\partial f}{\partial x}$ e $\frac{\partial f}{\partial y}$.

(iii). Calcule $df_{(1,1)}(-1, -1)$.

(iv)... Justifique que f é diferenciável em $(1, 2)$ e defina a respectiva diferencial $df_{(1,2)}$, nesse ponto.

Exercício 25 ... Calcule os valores das constantes a, b, c, de modo a que a derivada direccional da função:

$$f(x, y, z) = a\,xy^2 + b\,yz + c\,z^2x^3$$

no ponto $(1, 2, -1)$, tenha o valor 64, segundo a direcção do vector $(0, 0, 1)$.

Exercício 26 ... Considere o seguinte campo escalar:

$$f(x, y) = \begin{cases} x^3 \sin(\frac{1}{x}) + y^2 & \text{se } x \neq 0 \\ y^2 & \text{se } x = 0 \end{cases}$$

(i). Calcule as derivadas parciais de f em $(0, 0)$.

(ii). Prove que f é diferenciável em todos os pontos em que $x \neq 0$. Defina a diferencial de f, no ponto $(1, 3)$.

(iii). Verifique se f é diferenciável em $(0, 0)$.

(iv)... Mostre que $\frac{\partial f}{\partial x}$ é contínua em $(0, 0)$.

Exercício 27 ... Considere as funções seguintes:

(i). $f(x,y,z) = (x^2+y^2+z^2)^2$ (ii). $f(x,y,z) = \sin x \cdot \cos y \cdot z$

(iii). $f(x,y) = \frac{x+y}{x-y} \quad \forall (x,y) : x \neq y$

(a)... Relativamente a cada uma das referidas funções, diga em que pontos existem as respectivas derivadas parciais e calcule-as nesses pontos.

(b)... Em que pontos as referidas funções são diferenciáveis? Defina a diferencial nesses pontos.

Exercício 28 ... Considere o seguinte campo escalar:

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

(i). Calcule as derivadas parciais de f, nos pontos em que existem.

(ii). f é diferenciável em $(0,0)$? E em $(1,-1)$?

Exercício 29 ... Considere o seguinte campo escalar:

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

(i). f é contínuo em $(-2,0)$.

(ii)... Calcule $\frac{\partial f}{\partial x}(-2,0)$ e $\frac{\partial f}{\partial y}(-2,0)$. f é diferenciável em $(-2,0)$?

Exercício 30 ... A função:

$$f(x,y,z) = \begin{cases} \frac{x^3-y^2}{x^2+y^2+z^2} & \text{se } (x,y,z) \neq (0,0,0) \\ 0 & \text{se } (x,y,z) = (0,0,0) \end{cases}$$

é diferenciável no ponto $(0,0,0)$?

Calcule, caso exista, a derivada de f em $(0,0,0)$, na direcção do vector $(1,0,1)$.

Exercício 31 ... Determine o hiperplano tangente ao gráfico de cada uma das funções seguintes, nos pontos indicados:

(i). $f(x,y) = \frac{x^2+y^2}{xy}$ no ponto $(1,2,f(1,2))$.

(ii). $f(x,y) = e^x\, y$ no ponto $(-1,1,f(-1,1))$.

(iii). $f(x,y,z) = 2x^2 + z - zy^3$ no ponto $(1,0,1,f(1,0,1))$.

Exercício 32 ... Considere a curva $\alpha(t) = (e^t \sin t, e^{2t} \cos t, t^3)$, $t \in \mathbb{R}$, e a função $f(x, y, z) = xz + 1 + y$. Calcule a derivada de $f \circ \alpha$ em $t = 2\pi$.

Exercício 33 ... Considere a curva $\alpha(t) = (t, R\cos t, R\sin t)$, $t \in \mathbb{R}$, e a função $f(x, y, z) = \frac{x}{\sqrt{y^2+z^2}}$. Calcule a derivada de $(f \circ \alpha)'$.

Exercício 34 ... Relativamente a cada uma das funções que se seguem, calcule $F'(t)$, onde $F = f \circ \alpha$:

(i). $f(x, y) = x^2 + y^2$ *e* $\alpha(t) = (t, t^2)$

(ii). $f(x, y) = e^{xy} \cos(xy^2)$ *e* $\alpha(t) = (\cos t, \sin t)$

(iii). $f(x, y) = \log\left[\frac{1+e^{x^2}}{1+e^{y^2}}\right]$ *e* $\alpha(t) = (e^t, e^{-t})$

Exercício 35 ... Em cada das alíneas seguintes, calcule a derivada direccional de f, nos pontos e segundo as direcções indicadas:

(i). $f(x, y, z) = 3x - 5y + 2z$ no ponto $(2, 2, 1)$, segundo a direcção da normal à esfera de equação $x^2 + y^2 + z^2 = 9$.

(ii). $f(x, y, z) = x^2 - y^2$ num ponto genérico (a, b, c) da superfície $x^2 + y^2 + z^2 = 4$, segundo a direcção da normal a essa superfície.

(iii). $f(x, y, z) = x^2 + y^2 - z^2$ num ponto $(3, 4, 5)$, segundo a direcção da tangente à curva de intersecção das superfícies $2x^2 + 2y^2 - z^2 = 25$ e $x^2 + y^2 = z^2$.

Exercício 36 ... Determine o hiperplano tangente à superfície de equação $xy + yz + zx = 11$, no ponto $(1, 2, 3)$.

Exercício 37 ... Determine o hiperplano tangente à superfície de equação $z = \frac{x^2}{4} + \frac{y^2}{9}$, que é normal à recta de equação $x = z$, $y = 0$.

Exercício 38 ... Determine as equações da recta normal e do plano tangente à superfície de equação $x^2y + 2xz = 4$, no ponto $(2, -2, 3)$.

Exercício 39 ... Determine as equações da recta normal e do plano tangente à superfície de equação $x^2y^2 + xz - 2y^3 = 10$, no ponto $(2, 1, 4)$.

Exercício 40 ... Dada a superfície de equação $x^2 + y^2 - z^2 - x = 0$, determine os pontos em que o plano tangente é paralelo ao plano de equação $y = 0$.

Exercício 41 ... Relativamente a cada uma das funções que se seguem, determine a respectiva fórmula de Taylor de ordem 2, em cada um dos pontos indicados:

(i). $f(x, y) = x^3 + y^2 + xy^2$ no ponto $(1, 2)$.

(ii). $f(x, y) = \log(x + y), \quad x > 0\, e\, y > 0$ no ponto $(1, 1)$.

(iii). $f(x, y, z) = e^{a(x+y+z)}, \quad a \neq 0$ no ponto $(0, 0, 0)$.

Exercício 42 ... Relativamente a cada uma das funções que se seguem, determine e classifique os respectivos pontos críticos:

(i). $f(x, y) = x^2 + (y - 1)^2$ (ii). $f(x, y) = (x - y + 1)^2$

(iii). $f(x, y) = 2x^2 - xy - 3y^2 - 3x + 7y$ (iv). $f(x, y) = x^2y^3(6 - x - y)$

(v). $f(x, y) = x^3 + y^3 - 3xy$ (vi). $f(x, y, z) = x^2 + y^2 - z^2$

(vii). $f(x, y, z) = xy + yz + zx$ (viii). $f(x, y, z) = x^2 + y^2 + z^2 + xy$

Exercício 43 ... Calcule o valor máximo e o valor mínimo da função $f(x, y, z) = x + y + z$, quando restrita ao elipsóide $x^2 + 2y^2 + 3z^2 = 1$.

Exercício 44 ... Determine as distâncias máxima e mínima da origem à elipse de equação $5x^2 + 6xy + 5y^2 = 8$.

Exercício 45 ... Estudar os extremos da função $f(x, y) = x^2 + y^2$, quando as variáveis x e y, verificam a condição $x^2 + y = 1$.

Exercício 46 ... Considere os paralelipípedos rectangulares de volume constante V. Mostre que o cubo é o que tem menor área.

<u>Exercício</u> 47 ... Determine a altura e o raio da base do cilindro recto, com maior área, que pode ser inscrito numa esfera de raio a.

<u>Exercício</u> 48 ... Determine as dimensões de uma lata cilindrica, com uma tampa e com um volume de 1 litro, de tal forma a que se gaste um mínimo de material na sua construção.

<u>Exercício</u> 49 ... Determine os pontos onde a função $f(x, y) = x^2 + y^2 - xy + x + y$ atinge o máximo e o mínimo absolutos, no compacto:

$$K = \{(x, y) \in \mathbb{R}^2 : x \leq 0,\, y \leq 0 \, e \, x + y \geq -3\}$$

<u>Exercício</u> 50 ... Seja $\mathcal{U} \subseteq \mathbb{R}^n$ um aberto, e $f : \mathcal{U} \to \mathbb{R}$, uma função diferenciável. Seja $h : \mathbb{R} \to \mathbb{R}$ uma função crescente (ou decrescente), e $G = h \circ f$.

(i). Demontre que os pontos onde existem extremos, são os mesmos para f e G.

(ii). Aplique o resultado anterior, para determinar os extremos locais da função G, definida em $\mathcal{U} = \{(x, y) \in \mathbb{R}^2 : x > 0\}$, pela fórmula:

$$G(x, y) = e^{[(\log x)^2 + y^2]}$$

<u>Exercício</u> 51 ... Seja $f(x, y, z) = x^2 + y^2 + bxy + az$, com $a, b \in \mathbb{R}$.

(i). Obtenha uma relação entre a e b, que seja uma condição necessária para que o ponto $(1, 1, 1)$ seja um extremo relativo de f, quando restrita à esfera $x^2 + y^2 + z^2 = 3$.

(ii). Suponha que se verifica a condição anterior. Estude para que valores de a e b, o ponto $(1, 1, 1)$ é: (a). um máximo relativo, (b). um mínimo relativo, (c). não é extremo relativo.

<u>Exercício</u> 52 ... Calcule os extremos absolutos da função $f(x, y) = 4x^2 + y^2 - 4x - 3y$, sobre o compacto $K = \{(x, y) \in \mathbb{R}^2 : y \geq 0 \, e \, 4x^2 + y^2 \leq 4\}$.

Capítulo 3

Campos Vectoriais. Variedades em $\mathbf{R}^n$

3.1 Campos Vectoriais. Campos de Vectores. Exemplos

Por **campo vectorial**, designamos uma qualquer função vectorial de variável vectorial:

$$\mathbf{F} : \mathcal{D} \subseteq \mathbb{R}^n \to \mathbb{R}^m$$

onde $m, n \geq 1$.

Para uma tal função, definimos as respectivas funções componentes, como sendo os campos escalares $F^1, ..., F^m$, dados por:

$$\mathbf{F}(\mathbf{x}) = \sum_{i=1}^{m} F^i(\mathbf{x})\,\mathbf{e}_i \qquad \forall \mathbf{x} \in \mathcal{D}$$

onde $\{\mathbf{e}_i\}_{i=1,\cdots,m}$, designa a base canónica de $\mathbb{R}^m$.

É usual escrever então que:

$$\mathbf{F} = (F^1, \cdots, F^m)$$

Como veremos, o estudo dos campos vectoriais é essencialmente uma generalização óbvia do estudo feito sobre campos escalares.

Quando $m = n$, $\mathbf{F}$ diz-se um **campo de vectores**:

♣ **Definição 3.1** ... *Um* **campo de vectores** *definido num subconjunto* $\mathcal{U} \subseteq \mathbb{R}^n$*, é uma aplicação* $\mathbf{F}$*, que associa a cada ponto* $\mathbf{x} \in \mathcal{U}$*, um vector* $\mathbf{F}(\mathbf{x}) \in \mathbb{R}^n$*:*

$$\mathbf{x} \in \mathcal{U} \subseteq \mathbb{R}^n \mapsto \mathbf{F}(\mathbf{x}) \in \mathbb{R}^n \tag{3.1.1}$$

Gràficamente, representamos o vector $\mathbf{F}(\mathbf{x})$ aplicado no ponto $\mathbf{x}$, em cada ponto $\mathbf{x} \in \mathcal{U}$ (ver a figura 4.13).

Exemplos ...

Figure 3.1: Campo de vectores $\mathbf{x} \mapsto \mathbf{F}(\mathbf{x}) \in \mathbf{R}^n$

.

(i)... Um **campo de vectores uniforme (ou constante)**, é do tipo:

$$\mathbf{F}(\mathbf{x}) \equiv \mathbf{F}_o \in \mathbb{R}^n \ \ (=\text{constante}) \qquad \forall \mathbf{x} \in \mathcal{U} \tag{3.1.2}$$

(ii)... Um **campo linear** em $\mathbb{R}^n$, é um campo de vectores da forma:

$$\mathbf{F}(\mathbf{x}) = \mathbf{A}\,\mathbf{x} \qquad\qquad \mathbf{x} \in \mathbb{R}^n \tag{3.1.3}$$

onde $\mathbf{A} : \mathbb{R}^n \to \mathbb{R}^n$ é uma transformação linear.

(iii)... Um **campo central** (de centro $\mathbf{O}$), é um campo de vectores da forma (ver a figura 4.15):

$$\mathbf{F}(\mathbf{x}) \equiv f(\|\mathbf{x}\|)\,\mathbf{x} \tag{3.1.4}$$

onde f é uma função real definida e de classe C^∞ no intervalo $\{r \in \mathbb{R} : r > 0\}$. Em geral, $\mathbf{F}$ está apenas definido em $\mathbb{R}^n - \{\mathbf{O}\}$. $\mathbf{F}$ diz-se **atractor** ou **repulsor**, conforme $f(r) < 0$ ou $f(r) > 0$, $\forall r > 0$, respectivamente.

Figure 3.2: Campo de forças central

.

(iv)... A força de atracção que a terra exerce sobre uma massa m, pode ser descrita por um campo de vectores em $\mathbb{R}^3$, o campo gravitacional, que, de acordo com a lei de Newton, é igual a:

$$\mathbf{F}(\mathbf{r}) \ = \ -G\frac{Mm}{r^2}\,\mathbf{e}_r \tag{3.1.5}$$

$$= \quad -G\frac{Mm}{r^3}\,\mathbf{r} \tag{3.1.6}$$

onde $\mathbf{r}(x,y,z) = (x,y,z)$ e $r = \|\mathbf{r}\|$.

$\mathbf{F}$ é um campo de gradientes $\mathbf{F} = -\nabla V$ onde V é dado por $V = -\frac{mMG}{r}$. Em coordenadas cartesianas $\mathbf{F}$ é dado por:

$$\mathbf{F}(x,y,z) = \Big(\frac{-mMG}{r^3}\,x, \frac{-mMG}{r^3}\,y, \frac{-mMG}{r^3}\,z\Big)$$

(v)... De acordo com a lei de Coulomb, a força exercida sobre uma carga e, situada no ponto $\mathbf{r}$, força essa exercida por uma carga Q situada na origem, é dada por:

$$\mathbf{F} = \frac{\epsilon Qe}{r^3}\mathbf{r} = -\nabla V \tag{3.1.7}$$

onde $V = \frac{\epsilon Qe}{r}$. Para $Qe > 0$, $\mathbf{F}$ é repulsor, enquanto que para $Qe < 0$, $\mathbf{F}$ é atractor.

♣ **Definição 3.2** ... *Seja* $\mathbf{F} : \mathcal{U} \subseteq \mathbb{R}^n \to \mathbb{R}^n$ *um campo de vectores.*

Uma **linha de fluxo** *(***linha de corrente** *ou* **curva integral***, de* $\mathbf{F}$ *é uma curva derivável:*

$$\alpha : \mathbf{I} \subseteq \mathbb{R} \to \mathcal{U}$$

tal que (ver a figura 3.3):

$$\dot{\alpha}(t) = \mathbf{F}(\alpha(t)) \tag{3.1.8}$$

Figure 3.3: Linha de Fluxo

.

Analiticamente, o problema que consiste em calcular a linha de fluxo α, que no instante inicial $t = 0$, passa num ponto $\mathbf{x}_o$, envolve a resolução da **equação diferencial**:

$$\dot{\alpha}(t) = \mathbf{F}(\alpha(t)) \tag{3.1.9}$$

com a condição inicial:

$$\alpha(0) = x_o$$

A integração de equações diferenciais será tratado com detalhe num capítulo posterior.

Em $\mathbb{R}^3$, usando as coordenadas cartesianas usuais, e escrevendo:

$$\mathbf{F} = (F^1, F^2, F^3) \qquad e \qquad \alpha(t) = (x(t), y(t), z(t))$$

a equação diferencial (3.1.9), escreve-se na forma do seguinte sistema de equações diferenciais:

$$\begin{cases} x'(t) = F^1(x(t), y(t), z(t)) \\ y'(t) = F^2(x(t), y(t), z(t)) \\ z'(t) = F^3(x(t), y(t), z(t)) \end{cases}$$

e as condições iniciais na forma:

$$(x(0), y(0), z(0)) = (x_o, y_o, z_o) \tag{3.1.10}$$

Exemplos ...

(i)... Seja $\mathbf{F}$ o seguinte campo linear em $\mathbb{R}^3$:

$$\mathbf{F}(\mathbf{x}) = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -\theta \\ 0 & \theta & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ -\theta\, z \\ \theta\, y \end{bmatrix} \tag{3.1.11}$$

Neste caso, o sistema de equações diferenciais tem a forma:

$$\begin{cases} x'(t) = 0 \\ y'(t) = -\theta\, z(t) \\ z'(t) = \theta\, y(t) \end{cases}$$

A linha de fluxo α, que satisfaz a condição inicial $\alpha(0) = \mathbf{x}_o$, é dada por:

$$\alpha(t) = (x_o, y_o \cos\theta t - z_o \sin\theta t, y_o \sin\theta t + z_o \cos\theta t) \tag{3.1.12}$$

ou em forma matricial:

$$\alpha(t) = \begin{bmatrix} x(t) \\ y(t) \\ z(t) \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\theta t & -\sin\theta t \\ 0 & \sin\theta t & \cos\theta t \end{bmatrix} \begin{bmatrix} x_o \\ y_o \\ z_o \end{bmatrix} \tag{3.1.13}$$

Geomètricamente, o ponto $\mathbf{x}_o = (x_o, y_o, z_o)$ é submetido a uma rotação de ângulo θt, em torno do eixo dos xx (no plano yz).

(i)... Seja $\mathbf{F}$ o seguinte campo linear em $\mathbb{R}^3$:

$$\mathbf{F}(\mathbf{x}) = \mathbf{A}\mathbf{x} = \begin{bmatrix} d_1 & 0 & 0 \\ 0 & d_2 & 0 \\ 0 & 0 & d_3 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} d_1\, x \\ d_2\, y \\ d_3\, z \end{bmatrix} \tag{3.1.14}$$

Neste caso, o sistema de equações diferenciais tem a forma:

$$\begin{cases} x'(t) = d_1\, x(t) \\ y'(t) = d_2\, y(t) \\ z'(t) = d_3\, z(t) \end{cases}$$

A linha de fluxo α, que satisfaz a condição inicial $\alpha(0) = \mathbf{x}_o$, é dada por:

$$\alpha(t) = (x_o\, e^{d_1\, t}, y_o\, e^{d_2\, t}, z_o\, e^{d_3\, t}) \tag{3.1.15}$$

ou em forma matricial:

$$\alpha(t) = \begin{bmatrix} x(t) \\ y(t) \\ z(t) \end{bmatrix} = \begin{bmatrix} e^{d_1\, t} & 0 & 0 \\ 0 & e^{d_2\, t} & 0 \\ 0 & 0 & e^{d_3\, t} \end{bmatrix} \begin{bmatrix} x_o \\ y_o \\ z_o \end{bmatrix} \tag{3.1.16}$$

Geomètricamente, a i-componente do vector $\mathbf{x}_o = (x_o, y_o, z_o)$ é expandida (ou contraída) de um factor igual a $e^{d_i\, t}$ $(i = 1, 2, 3)$.

O volume de um paralelipípedo de lados paralelos aos eixos coordenados e de comprimentos respectivamente iguais a l_1, l_2, l_3, é igual a $V(0) = l_1 l_2 l_3$. No instante t, após a expansão (ou contracção) esse volume é igual a:

$$V(t) = l_1 e^{d_1 t} l_2 e^{d_2 t} l_3 e^{d_3 t} = V(0)\, e^{d_1 t} e^{d_2 t} e^{d_3 t}$$

A taxa de variação desse volume é portanto igual a:

$$\boxed{\frac{V'(0)}{V(0)} = d_1 + d_2 + d_3 = \textbf{Traço A}} \tag{3.1.17}$$

Quando **Traço A** $= 0$ o volume do referido paralelipípedo permanece portanto constante.

Estes exemplos serão úteis posteriormente, quando definirmos a divergência e o rotacional de um campo de vectores.

3.2 Limites. Continuidade

♣ **Definição 3.3** ... *Seja* $\mathbf{F} : \mathcal{U} \subseteq \mathbb{R}^n \to \mathbb{R}^m$, *um campo vectorial,* $\mathbf{p} \in \mathbb{R}^n$ *um ponto limite de* $\mathcal{U}$, *e* $\mathbf{l} \in \mathbb{R}^m$ *um vector em* $\mathbb{R}^m$.

Diz-se que o **limite de F é l, quando x converge para p**, *e nota-se por* $\lim_{\mathbf{x}\to\mathbf{p}} \mathbf{F}(\mathbf{x}) = \mathbf{l}$, *se se verifica a seguinte condição:*

$$\forall \epsilon > 0 \quad \exists \delta > 0 : \forall \mathbf{x} \in \mathcal{U},\ 0 < \|\mathbf{x} - \mathbf{p}\| < \delta \Rightarrow \|\mathbf{F}(\mathbf{x}) - \mathbf{l}\| < \epsilon \tag{3.2.1}$$

ou mais sucintamente se:

$$\lim_{\|\mathbf{x}-\mathbf{p}\|\to 0} \|\mathbf{F}(\mathbf{x}) - \mathbf{l}\| = 0 \tag{3.2.2}$$

♣ **Definição 3.4** ... *Seja* $\mathbf{F} : \mathcal{U} \subseteq \mathbb{R}^n \to \mathbb{R}^m$, *um campo vectorial, e* $\mathbf{p} \in \mathcal{U}$ *um ponto limite de* $\mathcal{U}$.

Diz-se que **F** *é* **contínua em p**, *se se verifica a seguinte condição:*

$$\lim_{\mathbf{x}\to\mathbf{p}} \mathbf{F}(\mathbf{x}) = \mathbf{F}(\mathbf{p}) \tag{3.2.3}$$

F *diz-se contínua em* $\mathcal{U}$, *se todo o ponto de* $\mathcal{U}$ *é ponto limite de* $\mathcal{U}$, *e se* **F** *é contínua em todo o ponto de* $\mathcal{U}$.

♣ <u>**Proposição**</u> **3.1** ... *Seja* $\mathbf{F} = (F^1, \cdots, F^m) : \mathcal{U} \subseteq \mathbb{R}^n \to \mathbb{R}^m$*, um campo vectorial, e* $\mathbf{p} \in \mathcal{U}$ *um ponto limite de* $\mathcal{U}$*.*

Então $\mathbf{F} = (F^1, \cdots, F^m)$ *é contínuo em* $\mathbf{p}$*, se e só se cada função componente* F^i *(*$i = 1, \cdots, m$*), é contínua em* $\mathbf{p}$*.*

3.3 Diferenciabilidade

3.3.1 Derivadas direccionais vectoriais e derivadas parciais vectoriais

♣ <u>**Definição**</u> **3.5** ... *Seja* $\mathbf{F} : \mathcal{U} \subseteq \mathbb{R}^n \to \mathbb{R}^m$ *um campo vectorial definido num subconjunto* $\mathcal{U}$ *de* $\mathbb{R}^n$*,* $\mathbf{p}$ *um ponto interior de* $\mathcal{U}$*, e* $\mathbf{v} \neq \mathbf{O}$ *um vector de* $\mathbb{R}^n$*.*

Define-se a **derivada direccional vectorial de F, em p, na direcção de v**, *notada por* $D_{\mathbf{v}}\mathbf{F}(\mathbf{p})$*, através de:*

$$D_{\mathbf{v}}\mathbf{F}(\mathbf{p}) = \lim_{t \to 0} \frac{\mathbf{F}(\mathbf{p} + t\mathbf{v}) - \mathbf{F}(\mathbf{p})}{t} \quad \in \mathbb{R}^m \tag{3.3.1}$$

De especial interesse é o caso em que $\mathbf{v} = \mathbf{e}_i$, onde $\mathbf{e}_i, i = 1, ..., n$, é a base canónica de $\mathbb{R}^n$. Neste caso, $D_{\mathbf{e}_i}\mathbf{F}(\mathbf{p})$ diz-se a i**-derivada parcial vectorial de F em p**, e nota-se por $\partial_i\mathbf{F}(\mathbf{p})$, ou por $\frac{\partial \mathbf{F}}{\partial x^i}(\mathbf{p})$:

$$\frac{\partial \mathbf{F}}{\partial x^i}(\mathbf{p}) = \partial_i\mathbf{F}(\mathbf{p}) = D_{\mathbf{e}_i}\mathbf{F}(\mathbf{p}) \in \mathbb{R}^m$$

<u>Note que agora, quer $D_{\mathbf{v}}\mathbf{F}(\mathbf{p})$, quer cada $\frac{\partial \mathbf{F}}{\partial x^i}(\mathbf{p})$ são vectores em $\mathbb{R}^m$.</u>

É fácil ver que, se $\mathbf{F} = (F^1, \cdots, F^n)$, então cada derivada parcial vectorial $\frac{\partial \mathbf{F}}{\partial x^j}(\mathbf{p})$ ($j = 1, \cdots, n$), existe sse existem as derivadas parciais $\frac{\partial F^i}{\partial x^j}(\mathbf{p})$ ($i = 1, \cdots, m$), e nesse caso, tem-se que:

$$\begin{aligned} \frac{\partial \mathbf{F}}{\partial x^j}(\mathbf{p}) &= \sum_{i=1}^m \frac{\partial F^i}{\partial x^j}(\mathbf{p})\,\mathbf{e}_i \\ &= \Big(\frac{\partial F^1}{\partial x^j}(\mathbf{p}), \cdots, \frac{\partial F^m}{\partial x^j}(\mathbf{p})\Big) \end{aligned} \tag{3.3.2}$$

para cada $j = 1, \cdots, n$.

Como antes, a noção de derivada direccional vectorial é manifestamente insuficiente, e por isso, somos conduzidos à noção de diferencial, que a seguir trataremos.

3.3.2 Diferencial. Matriz Jacobiana

♣ **Definição 3.6** ... *Seja* $\mathbf{F}: \mathcal{U} \to \mathbb{R}^m$ *um campo vectorial, definido num subconjunto aberto* $\mathcal{U} \subseteq \mathbb{R}^n$. *Fixemos um qualquer ponto* $\mathbf{p} \in \mathcal{U}$.

Diz-se que $\mathbf{F}$ *é* **diferenciável (ou derivável) em p**, *se existe uma aplicação linear:*

$$d\mathbf{F}_\mathbf{p}: \mathbb{R}^n \to \mathbb{R}^m$$

(que depende de $\mathbf{F}$ *e de* $\mathbf{p}$*), tal que:*

$$\boxed{\lim_{\|\mathbf{h}\|\to 0} \frac{\|\mathbf{F}(\mathbf{p}+\mathbf{h})-\mathbf{F}(\mathbf{p})-d\mathbf{F}_\mathbf{p}(\mathbf{h})\|}{\|\mathbf{h}\|} = 0} \qquad (3.3.3)$$

Esta aplicação linear $d\mathbf{F}_\mathbf{p}: \mathbb{R}^n \to \mathbb{R}^m$ *(que é única), diz-se a* **diferencial** *de* $\mathbf{F}$ *em* $\mathbf{p}$.

$\mathbf{F}$ *diz-se* **diferenciável** *em* $\mathcal{U}$, *se o é em todo o ponto* $\mathbf{p} \in \mathcal{U}$, *e neste caso diz-se que* $S = \mathbf{gr}\,\mathbf{F}$ *é uma* **variedade diferenciável** *de dimensão* n, *em* $\mathbb{R}^{n+m}$, *de equação* $\mathbf{z} = \mathbf{F}(\mathbf{x})$, $\mathbf{x} \in \mathcal{U}$.

A fórmula (3.3.3), pode ainda ser escrita na forma:

$$\boxed{\mathbf{F}(\mathbf{p}+\mathbf{h}) = \mathbf{F}(\mathbf{p}) + d\mathbf{F}_\mathbf{p}(\mathbf{h}) + \mathbf{o}(\|\mathbf{h}\|)} \qquad (3.3.4)$$

onde $\lim_{\|\mathbf{h}\|\to 0} \frac{\mathbf{o}(\|\mathbf{h}\|)}{\|h\|} = \mathbf{O}$. Esta última fórmula diz-se a **fórmula de Taylor de primeira ordem** para $\mathbf{F}$, em $\mathbf{p}$.

Exemplo ...

Se $\mathbf{F} = \mathbf{L}: \mathbb{R}^n \to \mathbb{R}^m$ é uma aplicação linear, então $\mathbf{L}$ é diferenciável em todo o ponto $\mathbf{p} \in \mathbb{R}^n$ e:

$$d\mathbf{L}_\mathbf{p} = \mathbf{L} \qquad \forall \mathbf{p} \in \mathbb{R}^n \qquad (3.3.5)$$

Com efeito, o numerador em (3.3.3) é neste caso igual a (pondo $d\mathbf{F}_\mathbf{p} = d\mathbf{L}_\mathbf{p} = \mathbf{L}$):

$$\|\mathbf{F}(\mathbf{p}+\mathbf{h}) - \mathbf{F}(\mathbf{p}) - d\mathbf{F}_\mathbf{p}(\mathbf{h})\| = \|\mathbf{L}(\mathbf{p}+\mathbf{h}) - \mathbf{L}(\mathbf{p}) - \mathbf{L}(\mathbf{h})\| = 0$$

uma vez que estamos a supôr que $\mathbf{L}$ é linear (e portanto, $\mathbf{L}(\mathbf{p}+\mathbf{h}) = \mathbf{L}(\mathbf{p}) + \mathbf{L}(\mathbf{h})$).

Vejamos agora algumas propriedades da diferencial. Aplicando (3.3.4), fàcilmente se deduz a seguinte proposição:

♣ **Proposição 3.2** ... *Se* $\mathbf{F}$ *é diferenciável em* $\mathbf{p}$, *então* $\mathbf{F}$ *é contínua em* $\mathbf{p}$.

♣ **Proposição 3.3** ... *Seja* $\mathbf{F}: \mathcal{U} \to \mathbb{R}^m$ *um campo vectorial, definido num subconjunto aberto* $\mathcal{U} \subseteq \mathbb{R}^n$. *Fixemos um qualquer ponto* $\mathbf{p} \in \mathcal{U}$.

Então $\mathbf{F} = (F^1, \cdots, F^m)$, *é diferenciável em* $\mathbf{p}$ *se e só se cada função componente* F^i, *o é, e nesse caso, tem-se que:*

$$\begin{aligned} d\mathbf{F}_\mathbf{p}(\mathbf{h}) &= \sum_{i=1}^m d(F^i)_\mathbf{p}(\mathbf{h})\,\mathbf{e}_i \\ &= \sum_{i=1}^m (\nabla F^i(\mathbf{p}) \cdot \mathbf{h})\,\mathbf{e}_i \qquad \forall \mathbf{h} \in \mathbb{R}^n \end{aligned} \qquad (3.3.6)$$

♣ **Proposição 3.4** ... *Suponhamos que* $\mathbf{F}$ *é diferenciável em* $\mathbf{p}$*, com diferencial* $d\mathbf{F}_{\mathbf{p}} \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$.

Então a derivada direccional vectorial $D_{\mathbf{v}}\mathbf{F}(\mathbf{p})$ *existe, para todo o vector* $\mathbf{v} \in \mathbb{R}^n$*, e tem-se que:*

$$D_{\mathbf{v}}\mathbf{F}(\mathbf{p}) = d\mathbf{F}_{\mathbf{p}}(\mathbf{v}) \in \mathbb{R}^m \qquad \forall \mathbf{v} \in \mathbb{R}^n \tag{3.3.7}$$

- Demonstração...

  Se $\mathbf{v} = \mathbf{O}$, então $D_{\mathbf{O}}\mathbf{F}(\mathbf{p}) = 0 = df_{\mathbf{p}}(\mathbf{O})$. Podemos por isso supôr que $\mathbf{v} \neq \mathbf{O}$. Façamos $\mathbf{h} = t\mathbf{v}$ na fórmula de Taylor (3.3.4), com $t \neq 0$, para obter:

  $$\mathbf{F}(\mathbf{p} + t\mathbf{v}) - \mathbf{F}(\mathbf{p}) = d\mathbf{F}_{\mathbf{p}}(t\mathbf{v}) + \mathbf{o}(|t|\|\mathbf{v}\|) = t\, d\mathbf{F}_{\mathbf{p}}(\mathbf{v}) + \mathbf{o}(|t|\|\mathbf{v}\|)$$

  Dividamos por t, e façamos $t \to 0$, para obter (3.3.7),

♣.

Suponhamos agora que $\mathbf{v} = \sum_{j=1}^n v^j \mathbf{e}_j = (v^1, ..., v^n) \in \mathbb{R}^n$, e que $\mathbf{F} = (F^1, \cdots, F^m)$. Utilizando (3.3.6) e a linearidade de cada $(dF^i)_{\mathbf{p}}$, obtemos então que:

$$\begin{aligned} d\mathbf{F}_{\mathbf{p}}(\mathbf{v}) &= \sum_{i=1}^m d(F^i)_{\mathbf{p}}(\mathbf{v})\, \mathbf{e}_i \\ &= \sum_{i=1}^m \Big(d(F^i)_{\mathbf{p}}(\sum_{j=1}^n v^j\, \mathbf{e}_j) \Big)\, \mathbf{e}_i \\ &= \sum_{i=1}^m \Big(\sum_{j=1}^n v^j\, d(F^i)_{\mathbf{p}}(\mathbf{e}_j) \Big)\, \mathbf{e}_i \\ &= \sum_{i=1}^m \Big(\sum_{j=1}^n v^j\, \frac{\partial F^i}{\partial x^j}(\mathbf{p}) \Big)\, \mathbf{e}_i \end{aligned} \tag{3.3.8}$$

Em particular, se $\mathbf{v} = \mathbf{e}_j \in \mathbb{R}^n$, obtemos:

$$d\mathbf{F}\mathbf{p}(\mathbf{e}_j) = \sum_{i=1}^m \frac{\partial F^i}{\partial x^j}(\mathbf{p})\, \mathbf{e}_i \tag{3.3.9}$$

o que significa que a matriz da aplicação linear $d\mathbf{F}_p \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$, relativamente às bases canónicas de $\mathbb{R}^n$ e $\mathbb{R}^m$, é a matriz $(m \times n)$:

$$\mathbf{JacF}(\mathbf{p}) = [\frac{\partial F^i}{\partial x^j}(\mathbf{p})]_{i,j} = \begin{bmatrix} \frac{\partial F^1}{\partial x^1}(\mathbf{p}) & \frac{\partial F^1}{\partial x^2}(\mathbf{p}) & \cdots & \frac{\partial F^1}{\partial x^n}(\mathbf{p}) \\ \frac{\partial F^2}{\partial x^1}(\mathbf{p}) & \frac{\partial F^2}{\partial x^2}(\mathbf{p}) & \cdots & \frac{\partial F^2}{\partial x^n}(\mathbf{p}) \\ \vdots & \vdots & \cdots & \vdots \\ \vdots & \vdots & \cdots & \vdots \\ \frac{\partial F^m}{\partial x^1}(\mathbf{p}) & \frac{\partial F^m}{\partial x^2}(\mathbf{p}) & \cdots & \frac{\partial F^m}{\partial x^n}(\mathbf{p}) \end{bmatrix} \tag{3.3.10}$$

Esta matriz diz-se a **matriz Jacobiana de F em p**. Note que as colunas desta matriz são as componentes das derivadas parciais vectoriais $\frac{\partial \mathbf{F}}{\partial x^j}(\mathbf{p})$ $(j = 1, \cdots, n)$, na base canónica de $\mathbb{R}^m$.

♣ **Definição 3.7** ... *Um campo vectorial* $\mathbf{F} : \mathcal{U} \subset \mathbb{R}^n \to \mathbb{R}^m$, *definido no aberto* $\mathcal{U} \subseteq \mathbb{R}^n$, *diz-se de classe* C^k *em* $\mathcal{U}$ $(k = 0, 1, 2, ...)$, *se todas as derivadas parciais até à ordem* k *(inclusivé), das funções componentes de* $\mathbf{F}$, *existem e são contínuas em* $\mathcal{U}$.

♣ **Proposição 3.5** ... *Se* $\mathbf{F} : \mathcal{U} \subset \mathbb{R}^n \to \mathbb{R}^m$, *é de classe* C^1 *em* $\mathcal{U}$, *então* $\mathbf{F}$ *é diferenciável em todo o ponto de* $\mathcal{U}$.

3.4 Funções complexas de variável complexa. Funções holomorfas. Equações de Cauchy-Riemann

Nesta secção identificamos $\mathbb{R}^2$ com o plano complexo $\mathbb{C}$, através do isomorfismo (real) que associa a cada vector $(x, y) \in \mathbb{R}^2$, o número complexo $x + i\,y \in \mathbb{C}$:

$$(x, y) \in \mathbb{R}^2 \longleftrightarrow x + i\,y \in \mathbb{C}$$

O conjugado do número complexo $z = x + i\,y \in \mathbb{C}$, é o número complexo $\bar{z} = x - i\,y \in \mathbb{C}$.

O módulo do número complexo $z = x + i\,y \in \mathbb{C}$, é o número real positivo $|z| = \sqrt{x^2 + y^2} \geq 0$.

Recorde que $z\bar{z} = |z|^2$.

Qualquer função definida num subconjunto $\mathcal{U} \subseteq \mathbb{R}^2$, e com valores em $\mathbb{R}^2$:

$$f : (x, y) \in \mathcal{U} \subseteq \mathbb{R}^2 \mapsto f(x, y) = (u(x, y), v(x, y)) \in \mathbb{R}^2$$

pode ser escrita em notação complexa na forma:

$$f : z = x + i\,y \in \mathcal{U} \subseteq \mathbb{R}^2 \cong \mathbb{C} \mapsto f(z) = u(x, y) + i\,v(x, y) \in \mathbb{C}$$

e vice-versa, como é evidente.

Exemplos ...

(i)... $f(z) = \bar{z} \qquad z \in \mathbb{C}$

(ii)... $f(z) = \frac{1}{z} = \frac{x}{x^2+y^2} - i\,\frac{y}{x^2+y^2} \qquad z \in \mathbb{C}^* = \mathbb{C} - \{0\}$

(iii)... $f(z) = z^2 = (x + i\,y)^2 = (x^2 - y^2) + i\,(2xy) \qquad z = x + i\,y \in \mathbb{C}$

(iv)... $f(z) = e^z \equiv e^x\,\cos y + i\,e^x\,\sin y \qquad z = x + i\,y \in \mathbb{C}$

(v)... $f(z) = \sin z \equiv \frac{e^{i\,z} - e^{-i\,z}}{2i} \qquad z \in \mathbb{C}$

(vi)... $f(z) = \cos z \equiv \frac{e^{i\,z} + e^{-i\,z}}{2} \qquad z \in \mathbb{C}$

(vii)... $f(z) = \sinh z \equiv \frac{e^z - e^{-z}}{2} \qquad z \in \mathbb{C}$

(viii)... $f(z) = \cosh z \equiv \frac{e^z + e^{-z}}{2} \qquad z \in \mathbb{C}$

♣ **Definição 3.8** ... *Seja* $f : \mathcal{U} \to \mathbb{C}$ *uma função complexa definida num aberto* $\mathcal{U} \subseteq \mathbb{C} \cong \mathbb{R}^2$*.*
f *diz-se* **holomorfa** *(ou* **$\mathbb{C}$-diferenciável***) num ponto* $\mathbf{z} \in \mathcal{U}$*, se existir o limite:*

$$\lim_{h\to 0} \frac{f(z+h) - f(z)}{h} \tag{3.4.1}$$

onde $h \in \mathbb{C}^*$*. Nesse caso, o limite nota-se por* $f'(z)$ *e diz-se a* **derivada de** f **em** z*.* f *diz-se* **holomorfa em** $\mathcal{U}$ *, se o fôr em todo o ponto* $z \in \mathcal{U}$*.*

O limite (3.4.1) deve entender-se de acordo com a definição (3.2.1), e pode ser escrito na forma:

$$\lim_{h\to 0} \frac{|f(z+h) - f(z) - f'(z)\,h|}{|h|} = 0 \tag{3.4.2}$$

Comparando (3.4.2) com (3.3.3), e atendendo a que a aplicação:

$$h \in \mathbb{C} \to f'(z)\,h \tag{3.4.3}$$

é uma aplicação $\mathbb{R}$-linear, podemos concluir que se f é holomorfa num ponto $\mathbf{z} \in \mathcal{U}$, então f é $\mathbb{R}$-diferenciável em z, e a respectiva diferencial é a aplicação $\mathbb{R}$-linear definida por (3.4.3).

Pondo $f'(z) = a + i\,b$ e $h = c + i\,d$ em (3.4.3), podemos escrever:

$$\begin{aligned} f'(z)\,h &= (a+i\,b)(c+i\,d) \\ &= (ac-bd) + i\,(ad+bc) \\ &\cong (ac-bd, ad+bc) \in \mathbb{R}^2 \end{aligned}$$

donde se conclui, após um cálculo simples, que a matriz da aplicação $\mathbb{R}$-linear (3.4.3), é a matriz:

$$\begin{bmatrix} a & -b \\ b & a \end{bmatrix} \tag{3.4.4}$$

Como esta é também a matriz jacobiana de f em $z = x + i\,y$, matriz essa que é da forma:

$$\begin{bmatrix} \frac{\partial u}{\partial x} & \frac{\partial u}{\partial y} \\ \frac{\partial v}{\partial x} & \frac{\partial v}{\partial y} \end{bmatrix}$$

onde escrevemos f na forma $f(x,y) = (u(x,y), v(x,y))$, com $z = x + i\,y = (x,y)$, concluímos que se f é holomorfa num ponto $\mathbf{z} \in \mathcal{U}$, então devem ser válidas as seguintes **equações de Cauchy-Riemann**, no ponto z:

$$\begin{cases} \frac{\partial u}{\partial x}(z) &= \frac{\partial v}{\partial y}(z) \\ \frac{\partial u}{\partial y}(z) &= -\frac{\partial v}{\partial x}(z) \end{cases} \tag{3.4.5}$$

Concluindo, é agora fácil demonstrar o seguinte resultado essencial:

♣ **Proposição 3.6** ...
Seja $f : \mathcal{U} \to \mathbb{C}$ *uma função complexa definida num aberto* $\mathcal{U} \subseteq \mathbb{C} \cong \mathbb{R}^2$*.*

Então f *é holomorfa num ponto* $\mathbf{z} \in \mathcal{U}$*, se e só se* f *é* $\mathbb{R}$*-diferenciável em* z*, e se em* z*, se verificam as equações de Cauchy-Riemann (3.4.5).*

Exemplos ...

(i)... A função $f(z) = \bar{z}$ não é holomorfa em qualquer ponto $z \in \mathbb{C}$.

Com efeito, na forma real, f é a aplicação:

$$z = x + i\,y = (x, y) \to (x, -y)$$

que é $\mathbb{R}$-diferenciável $\forall (x, y) \in \mathbb{R}^2$, com matriz jacobiana constante e igual a:

$$\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$

que **não é** da forma (3.4.4).

(ii)... Se $f(z) = z^n$, então $f'(z) = nz^{n-1}$, $\forall n \in \mathbf{N}, n \geq 1$.

(iii)... Se $f(z) = e^z$, então $f'(z) = e^z$, $\forall z \in \mathbb{C}$.

(iv)... Se $f(z) = \frac{1}{z}$, então $f'(z) = \frac{-1}{z^2}$, $\forall z \in \mathbb{C}^*$.

3.5 Regra da Cadeia II. Derivação Implícita

3.5.1 Regra da Cadeia II

♣ **Proposição 3.7** (**Regra da Cadeia II**) ... *Seja* $\mathbf{G} : \mathcal{U} \subseteq \mathbb{R}^n \to \mathbb{R}^m$*, um campo vectorial, definido num aberto* $\mathcal{U} \subseteq \mathbb{R}^n$*, e* $\mathbf{F} : \mathcal{V} \subseteq \mathbb{R}^m \to \mathbb{R}^k$*, um outro campo vectorial, definido num aberto* $\mathcal{V} \subseteq \mathbb{R}^m$*, tal que* $\mathbf{G}(\mathcal{U}) \subset \mathcal{V}$.

Se $\mathbf{G}$ *é diferenciável em* $\mathbf{p} \in \mathcal{U}$*, e se* $\mathbf{F}$ *é diferenciável em* $\mathbf{G}(\mathbf{p}) \in \mathcal{V}$*, então* $\mathbf{F} \circ \mathbf{G} : \mathcal{U} \to \mathbb{R}^k$ *é diferenciável em* $\mathbf{p}$*, e:*

$$d(\mathbf{F} \circ \mathbf{G})_{\mathbf{p}} = d\mathbf{F}_{\mathbf{G}(\mathbf{p})} \circ d\mathbf{G}_{\mathbf{p}} \quad \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^k) \tag{3.5.1}$$

Neste caso, a matriz Jacobiana de $\mathbf{F} \circ \mathbf{G}$*, em* $\mathbf{p}$*, é igual ao produto das matrizes Jacobianas:*

$$\mathbf{Jac}(\mathbf{F} \circ \mathbf{G})(\mathbf{p}) = \mathbf{Jac}\mathbf{F}(\mathbf{G}(\mathbf{p})).\mathbf{Jac}\mathbf{G}(\mathbf{p}) \tag{3.5.2}$$

- Demonstração...

  Pretende-se provar que:

$$(\mathbf{F} \circ \mathbf{G})(\mathbf{p} + \mathbf{h}) - (\mathbf{F} \circ \mathbf{G})(\mathbf{p}) = d\mathbf{F}_{\mathbf{G}(\mathbf{p})}(d\mathbf{G}_{\mathbf{p}}(\mathbf{h})) + \mathbf{o}(\|\mathbf{h}\|) \tag{3.5.3}$$

  Mas como $\mathbf{G}$ e $\mathbf{F}$ são diferenciáveis em $\mathbf{p}$ e $\mathbf{G}(\mathbf{p})$, respectivamente, tem-se que:

$$\mathbf{G}(\mathbf{p} + \mathbf{h}) = \mathbf{G}(\mathbf{p}) + d\mathbf{G}_{\mathbf{p}}(\mathbf{h}) + \mathbf{o}(\|\mathbf{h}\|) \tag{3.5.4}$$

  e:

$$\mathbf{F}(\mathbf{G}(\mathbf{p}) + \mathbf{v}) = \mathbf{F}(\mathbf{G}(\mathbf{p})) + d\mathbf{F}_{\mathbf{G}(\mathbf{p})}(\mathbf{v}) + \mathbf{o}(\|\mathbf{v}\|) \tag{3.5.5}$$

Portanto, fazendo $\mathbf{v} = \mathbf{G}(\mathbf{p}+\mathbf{h}) - \mathbf{G}(\mathbf{p})$ em (3.5.5), e usando (3.5.4), obtemos que:

$$\begin{aligned}
\mathbf{F}(\mathbf{G}(\mathbf{p}+\mathbf{h})) &= \mathbf{F}(\mathbf{G}(\mathbf{p})) + d\mathbf{F}_{\mathbf{G}(\mathbf{p})}(\mathbf{G}(\mathbf{p}+\mathbf{h}) - \mathbf{G}(\mathbf{p})) + \mathbf{o}(\|\mathbf{G}(\mathbf{p}+\mathbf{h}) - \mathbf{G}(\mathbf{p})\|) \\
&= \mathbf{F}(\mathbf{G}(\mathbf{p})) + d\mathbf{F}_{\mathbf{G}(\mathbf{p})}(d\mathbf{G}_{\mathbf{p}}(\mathbf{h}) + \mathbf{o}(\|\mathbf{h}\|)) + \mathbf{o}(\|\mathbf{G}(\mathbf{p}+\mathbf{h}) - \mathbf{G}(\mathbf{p})\|) \\
&\qquad \text{por } (3.5.4) \\
&= \mathbf{F}(\mathbf{G}(\mathbf{p})) + d\mathbf{F}_{\mathbf{G}(\mathbf{p})}(d\mathbf{G}_{\mathbf{p}}(\mathbf{h})) + d\mathbf{F}_{\mathbf{G}(\mathbf{p})}(\mathbf{o}(\|\mathbf{h}\|)) \\
&\qquad + \mathbf{o}(\|\mathbf{G}(\mathbf{p}+\mathbf{h}) - \mathbf{G}(\mathbf{p})\|) \\
&= \mathbf{F}(\mathbf{G}(\mathbf{p})) + d\mathbf{F}_{\mathbf{G}(\mathbf{p})}(d\mathbf{G}_{\mathbf{p}}(\mathbf{h})) + \mathbf{o}(\|\mathbf{h}\|)
\end{aligned}$$

onde, na última igualdade se utilizou o facto de $\mathbf{G}$ ser contínua em $\mathbf{p}$ (por ser diferenciável em $\mathbf{p}$),

♣.

Exemplo ...

Sejam $\mathbf{G} : \mathbb{R}^2 \to \mathbb{R}^3$, e $\mathbf{F} : \mathbb{R}^3 \to \mathbb{R}^2$, dadas respectivamente por:

$$\mathbf{G}(u,v) = (u^2, uv, e^v) \qquad e \qquad \mathbf{F}(x,y,z) = (x^2 + y - z, x - yz)$$

Então $\mathbf{G}(1,0) = (1,0,1)$, e:

$$\mathbf{Jac}\mathbf{G}(1,0) = \begin{bmatrix} 2u & 0 \\ v & u \\ 0 & e^v \end{bmatrix}_{(u=1,v=0)} = \begin{bmatrix} 2 & 0 \\ 0 & 1 \\ 0 & 1 \end{bmatrix}$$

enquanto que:

$$\mathbf{Jac}\mathbf{F}(1,0,1) = \begin{bmatrix} 2x & 1 & -1 \\ 1 & -z & -y \end{bmatrix}_{(x=1,y=0,z=1)} = \begin{bmatrix} 2 & 1 & -1 \\ 0 & -1 & 0 \end{bmatrix}$$

Portanto:

$$\begin{aligned}
\mathbf{Jac}(\mathbf{F}\circ\mathbf{G})(1,0) = \mathbf{Jac}\mathbf{F}(1,0,1).\mathbf{Jac}\mathbf{G}(1,0) &= \begin{bmatrix} 2 & 1 & -1 \\ 0 & -1 & 0 \end{bmatrix} . \begin{bmatrix} 2 & 0 \\ 0 & 1 \\ 0 & 1 \end{bmatrix} \\
&= \begin{bmatrix} 4 & 0 \\ 0 & -1 \end{bmatrix}
\end{aligned}$$

Exemplo ...

Se $\mathbf{G} : \mathbb{R}^2 \to \mathbb{R}^2$, é dada por:

$$\mathbf{G}(r,\theta) = (r\cos\theta, r\sin\theta)$$

e $f : \mathbb{R}^2 \to \mathbb{R}$ é uma função diferenciável, então:

$$\mathbf{Jac}\mathbf{G}(r,\theta) = \begin{bmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{bmatrix}$$

enquanto que:

$$\mathbf{Jac}f(x,y) = \begin{bmatrix} \frac{\partial f}{\partial x} & \frac{\partial f}{\partial y} \end{bmatrix}$$

Portanto:

$$\begin{aligned}\mathbf{Jac}(f\circ\mathbf{G})(r,\theta) &= \mathbf{Jac}f(\mathbf{G}(r,\theta)).\mathbf{Jac}\mathbf{G}(r,\theta)\\ &= \begin{bmatrix}\frac{\partial f}{\partial x} & \frac{\partial f}{\partial y}\end{bmatrix}_{(x=r\cos\theta,y=r\sin\theta)}.\begin{bmatrix}\cos\theta & -r\sin\theta\\ \sin\theta & r\cos\theta\end{bmatrix}\end{aligned}$$

donde se deduz que:

$$\frac{\partial f}{\partial r} \equiv \frac{\partial(f\circ\mathbf{G})}{\partial r} = \cos\theta\frac{\partial f}{\partial x}+\sin\theta\frac{\partial f}{\partial y} \tag{3.5.6}$$

$$\frac{\partial f}{\partial \theta} \equiv \frac{\partial(f\circ\mathbf{G})}{\partial \theta} = -r\sin\theta\frac{\partial f}{\partial x}+r\cos\theta\frac{\partial f}{\partial y} \tag{3.5.7}$$

onde as derivadas parciais $\frac{\partial f}{\partial x}$ e $\frac{\partial f}{\partial y}$ devem ser calculadas no ponto $(r\cos\theta, r\sin\theta)$.

3.5.2 Derivação Implícita

Uma aplicação importante da regra da cadeia (3.5.2), é na chamada **derivação implícita**, como passamos a explicar.

Consideremos uma função diferenciável:

$$\mathbf{F}:\mathbb{R}^{n+m}\cong\mathbb{R}^n\times\mathbb{R}^m\to\mathbb{R}^m$$

Um elemento de $\mathbb{R}^{n+m}\cong\mathbb{R}^n\times\mathbb{R}^m$, pode ser expresso na forma:

$$(x^1,\cdots,x^n,y^1,\cdots,y^m)$$

ou como um par de vectores $(\mathbf{x},\mathbf{y})$, onde $\mathbf{x}=(x^1,\cdots,x^n)\in\mathbb{R}^n$ e $\mathbf{y}=(y^1,\cdots,y^m)\in\mathbb{R}^m$, e portanto a função $\mathbf{F}$, pode ser considerada como uma função de duas variáveis vectoriais $\mathbf{x}$ e $\mathbf{y}$.

Diz-se que uma função $\mathbf{f}:\mathcal{D}\subseteq\mathbb{R}^n\to\mathbb{R}^m$ está **definida implicitamente** pela equação:

$$\mathbf{F}(\mathbf{x},\mathbf{y})=\mathbf{O} \tag{3.5.8}$$

se:

$$\mathbf{F}(\mathbf{x},\mathbf{f}(\mathbf{x}))=\mathbf{O} \qquad \forall\mathbf{x}\in\mathcal{D} \tag{3.5.9}$$

Geomètricamente, a condição (3.5.9), significa que o conjunto de nível $\mathbf{O}$, da função $\mathbf{F}$, contem o gráfico da função $\mathbf{f}$ (ver a figura (3.4)):

$$\mathbf{gr}\,\mathbf{f}\subset\mathbf{F}^{-1}(\mathbf{O})$$

Exemplo ...

As funções $f(x)=\sqrt{1-x^2}$, $-1\le x\le 1$ e $g(x)=-\sqrt{1-x^2}$, $-1\le x\le 1$, estão ambas definidas implicitamente pela equação:

$$F(x,y)=x^2+y^2-1=0 \qquad (x,y)\in\mathbb{R}^2$$

Figure 3.4: Função definida implicitamente

O teorema da função implícita, que será estudado em breve, dá condições para a existência de uma função diferenciável $\mathbf{f}$, definida implicitamente pela equação $\mathbf{F}(\mathbf{x}, \mathbf{f}(\mathbf{x})) = \mathbf{O}$.

Supondo que existe uma função diferenciável $\mathbf{f}$, definida implicitamente pela equação $\mathbf{F}(\mathbf{x}, \mathbf{f}(\mathbf{x})) = \mathbf{O}$, vamos discutir para já, o problema de calcular a diferencial de $\mathbf{f}$.

Comecemos pela situação simples em que $F : \mathbb{R}^2 \to \mathbb{R}$, e $f : \mathcal{D} \subseteq \mathbb{R} \to \mathbb{R}$, verificam portanto:

$$F(x, f(x)) = 0 \qquad \forall x \in \mathcal{D} \tag{3.5.10}$$

Consideremos a seguinte composição de aplicações:

$$x \in \mathcal{D} \xrightarrow{G} (x, f(x)) \in \mathbb{R}^2 \xrightarrow{F} 0 \in \mathbb{R}$$

e apliquemos a regra da cadeia para calcular a derivada $(F \circ G)'(x)$, que identificamos com $\mathbf{Jac}(F \circ G)(x)$ (e que é 0, atendendo a (3.5.10)):

$$\begin{aligned} 0 = (F \circ G)'(x) &= \mathbf{Jac}F(x, f(x)).\mathbf{Jac}G(x) \\ &= \left[\begin{array}{cc} \frac{\partial F}{\partial x} & \frac{\partial F}{\partial y} \end{array} \right]_{(x,y=f(x))} . \left[\begin{array}{c} 1 \\ f'(x) \end{array} \right] \\ &= \frac{\partial F}{\partial x}(x, f(x)) + f'(x)\, \frac{\partial F}{\partial y}(x, f(x)) \end{aligned} \tag{3.5.11}$$

Portanto, se $\frac{\partial F}{\partial y}(x, f(x)) \neq 0$, deduzimos que:

$$f'(x) = -\frac{\frac{\partial F}{\partial x}(x, f(x))}{\frac{\partial F}{\partial y}(x, f(x))} \tag{3.5.12}$$

Podemos utilizar essencialmente o mesmo argumento, para deduzir a fórmula para $d\mathbf{f}_{\mathbf{x}}$, na situação geral.

Assim, suponhamos que existe uma função diferenciável $\mathbf{f} : \mathcal{U} \subseteq \mathbb{R}^n \to \mathbb{R}^m$, definida implicitamente pela equação:

$$\mathbf{F}(\mathbf{x}, \mathbf{f}(\mathbf{x})) = \mathbf{O} \tag{3.5.13}$$

onde $\mathbf{F} : \mathbb{R}^{n+m} \to \mathbb{R}^m$ é diferenciável, e consideremos a seguinte composição de aplicações:

$$\mathbf{x} \in \mathcal{D} \subseteq \mathbb{R}^n \xrightarrow{\mathbf{G}} (\mathbf{x}, \mathbf{y} = \mathbf{f}(\mathbf{x})) \in \mathbb{R}^{n+m} \xrightarrow{\mathbf{F}} \mathbf{O} \in \mathbb{R}^m \tag{3.5.14}$$

Apliquemos a regra da cadeia para calcular a matriz Jacobiana $\mathbf{Jac}(\mathbf{F} \circ \mathbf{G})(\mathbf{x})$, que é identicamente nula, atendendo a (3.5.13)):

$$\begin{aligned} \mathbf{O} = \mathbf{Jac}(\mathbf{F} \circ \mathbf{G})(\mathbf{x}) &= \mathbf{JacF}(\mathbf{G}(\mathbf{x})).\mathbf{JacG}(\mathbf{x}) \\ &= \begin{bmatrix} \mathbf{Jac_x F}(\mathbf{x}, \mathbf{f}(\mathbf{x})) & \mathbf{Jac_y F}(\mathbf{x}, \mathbf{f}(\mathbf{x})) \end{bmatrix} . \begin{bmatrix} \mathrm{Id}_n \\ \mathbf{Jacf}(\mathbf{x}) \end{bmatrix} \\ &= \mathbf{Jac_x F}(\mathbf{x}, \mathbf{f}(\mathbf{x})) + \mathbf{Jac_y F}(\mathbf{x}, \mathbf{f}(\mathbf{x})).\, \mathbf{Jacf}(\mathbf{x}) \end{aligned} \tag{3.5.15}$$

onde pusemos $\mathbf{F} = (F^1, \cdots, F^m)$, $\mathbf{f} = (f^1, \cdots, f^m)$ e:

$$\begin{bmatrix} \mathrm{Id}_n \\ \mathbf{Jacf}(\mathbf{x}) \end{bmatrix} = \begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \dots & \vdots \\ \vdots & \vdots & \dots & \vdots \\ 0 & 0 & \dots & 1 \\ \frac{\partial f^1}{\partial x^1} & \frac{\partial f^1}{\partial x^2} & \dots & \frac{\partial f^1}{\partial x^n} \\ \frac{\partial f^2}{\partial x^1} & \frac{\partial f^2}{\partial x^2} & \dots & \frac{\partial f^2}{\partial x^n} \\ \vdots & \vdots & \dots & \vdots \\ \vdots & \vdots & \dots & \vdots \\ \frac{\partial f^m}{\partial x^1} & \frac{\partial f^m}{\partial x^2} & \dots & \frac{\partial f^m}{\partial x^n} \end{bmatrix}_{\mathbf{x}}$$

$$\mathbf{Jac_x F}(\mathbf{x}, \mathbf{f}(\mathbf{x})) = \begin{bmatrix} \frac{\partial F^1}{\partial x^1} & \frac{\partial F^1}{\partial x^2} & \dots & \frac{\partial F^1}{\partial x^n} \\ \frac{\partial F^2}{\partial x^1} & \frac{\partial F^2}{\partial x^2} & \dots & \frac{\partial F^2}{\partial x^n} \\ \vdots & \vdots & \dots & \vdots \\ \vdots & \vdots & \dots & \vdots \\ \frac{\partial F^m}{\partial x^1} & \frac{\partial F^m}{\partial x^2} & \dots & \frac{\partial F^m}{\partial x^n} \end{bmatrix}_{(\mathbf{x}, \mathbf{f}(\mathbf{x}))} \tag{3.5.16}$$

e ainda:

$$\mathbf{Jac_y F}(\mathbf{x}, \mathbf{f}(\mathbf{x})) = \begin{bmatrix} \frac{\partial F^1}{\partial y^1} & \frac{\partial F^1}{\partial y^2} & \dots & \frac{\partial F^1}{\partial y^m} \\ \frac{\partial F^2}{\partial y^1} & \frac{\partial F^2}{\partial y^2} & \dots & \frac{\partial F^2}{\partial y^m} \\ \vdots & \vdots & \dots & \vdots \\ \vdots & \vdots & \dots & \vdots \\ \frac{\partial F^m}{\partial y^1} & \frac{\partial F^m}{\partial y^2} & \dots & \frac{\partial F^m}{\partial y^m} \end{bmatrix}_{(\mathbf{x}, \mathbf{f}(\mathbf{x}))} \tag{3.5.17}$$

Quando esta última matriz quadrada ($m \times m$), tem determinante não nulo (isto é, é inversível), podemos escrever a partir de (3.5.15), que:

$$\mathbf{Jacf}(\mathbf{x}) = -\left(\mathbf{Jac_y F}(\mathbf{x}, \mathbf{f}(\mathbf{x}))\right)^{-1}.\, \mathbf{Jac_x F}(\mathbf{x}, \mathbf{f}(\mathbf{x})) \tag{3.5.18}$$

Exemplo ...

Seja $F : \mathbb{R}^3 \to \mathbb{R}$ diferenciável, e suponhamos que $f : \mathcal{U} \subseteq \mathbb{R}^2 \to \mathbb{R}$ está definida implicitamente por $F(x, y, z) = 0$, isto é:

$$F(x, y, f(x, y)) = 0 \qquad \forall (x, y) \in \mathcal{U}$$

Por outras palavras, é possível explicitar (localmente), na equação $F(x, y, z) = 0$, a variável z, como função de x e y: $z = f(x, y)$.

Neste caso, o diagrama de aplicações (3.5.14), tem o aspecto seguinte:

$$(x, y) \in \mathcal{U} \subseteq \mathbb{R}^2 \xrightarrow{\mathbf{G}} (x, y, z = f(x, y)) \in \mathbb{R}^3 \xrightarrow{F} 0 \in \mathbb{R} \tag{3.5.19}$$

e portanto:

$$\begin{aligned} O = \mathbf{Jac}(F \circ \mathbf{G})(x, y) &= \mathbf{Jac}F(x, y, f(x, y)).\mathbf{JacG}(x, y) \\ &= \begin{bmatrix} \frac{\partial F}{\partial x} & \frac{\partial F}{\partial y} & \frac{\partial F}{\partial z} \end{bmatrix}_{(x,y,z=f(x,y))} . \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ \frac{\partial f}{\partial x} & \frac{\partial f}{\partial y} \end{bmatrix} \end{aligned}$$

donde se deduz que:

$$\begin{aligned} \frac{\partial F}{\partial x} + \frac{\partial f}{\partial x}\frac{\partial F}{\partial z} &= 0 \\ \frac{\partial F}{\partial y} + \frac{\partial f}{\partial y}\frac{\partial F}{\partial z} &= 0 \end{aligned}$$

onde as derivadas parciais de F, devem ser calculadas no ponto $(x, y, z = f(x, y))$. Portanto:

$$\frac{\partial f}{\partial x} = -\frac{\frac{\partial F}{\partial x}(x, y, z = f(x, y))}{\frac{\partial F}{\partial z}(x, y, z = f(x, y))} \tag{3.5.20}$$

$$\frac{\partial F}{\partial y} = -\frac{\frac{\partial F}{\partial y}(x, y, z = f(x, y))}{\frac{\partial F}{\partial z}(x, y, z = f(x, y))} \tag{3.5.21}$$

nos pontos em que $\frac{\partial F}{\partial z}(x, y, z = f(x, y)) \neq 0$.

Exemplo ...

Suponhamos que $z = f(x), y = g(x)$, $x \in \mathbf{I} \subset \mathbb{R}$ estão definidas implicitamente pelas duas equações:

$$F^1(x, y, z) = 0 \qquad e \qquad F^2(x, y, z) = 0$$

onde $(x, y, z) \in \mathbb{R}^3$, e calculemos $f'(x)$ e $g'(x)$.

Para isso, consideremos o diagrama:

$$\begin{aligned} &x \in \mathbf{I} \subseteq \mathbb{R} \xrightarrow{\mathbf{G}} (x, y = f(x), z = g(\mathbf{x})) \in \mathbb{R}^3 \\ &\qquad \xrightarrow{\mathbf{F}} (F^1(x, y = f(x), z = g(x)), F^2(x, y = f(x), z = g(x)) = (O, 0) \in \mathbb{R}^2 \end{aligned}$$

e apliquemos a regra da cadeia, para obter:

$$\begin{aligned}
\begin{bmatrix} 0 & 0 \end{bmatrix} &= \mathbf{Jac}(\mathbf{F}\circ\mathbf{G})(x) = \mathbf{JacF}(x, y=f(x), z=g(x)).\mathbf{JacG}(x) \\
&= \begin{bmatrix} \frac{\partial F^1}{\partial x} & \frac{\partial F^1}{\partial y} & \frac{\partial F^1}{\partial z} \\ \frac{\partial F^2}{\partial x} & \frac{\partial F^2}{\partial y} & \frac{\partial F^2}{\partial z} \end{bmatrix}_{(x,y=f(x),z=g(x))} . \begin{bmatrix} 1 \\ f'(x) \\ g'(x) \end{bmatrix} \\
&= \begin{bmatrix} \frac{\partial F^1}{\partial x} + f'(x)\,\frac{\partial F^1}{\partial y} + g'(x)\,\frac{\partial F^1}{\partial z} \\ \frac{\partial F^2}{\partial x} + f'(x)\,\frac{\partial F^2}{\partial y} + g'(x)\,\frac{\partial F^2}{\partial z} \end{bmatrix}
\end{aligned}$$

onde as derivadas parciais de F^1, F^2 devem ser calculadas em $(x, y = f(x), z = g(x))$. Daqui se deduz que:

$$\begin{aligned}
\frac{\partial F^1}{\partial x} + f'(x)\,\frac{\partial F^1}{\partial y} + g'(x)\,\frac{\partial F^1}{\partial z} &= 0 \\
\frac{\partial F^2}{\partial x} + f'(x)\,\frac{\partial F^2}{\partial y} + g'(x)\,\frac{\partial F^2}{\partial z} &= 0
\end{aligned}$$

que é um sistema de duas equações lineares. Em todos os pontos em que o determinante desse sistema é não nulo, ele admite uma solução única, dada por:

$$f'(x) = \frac{\frac{\partial(F^1,F^2)}{\partial(x,z)}}{\frac{\partial(F^1,F^2)}{\partial(z,y)}} \equiv \frac{\det \begin{bmatrix} \frac{\partial F^1}{\partial x} & \frac{\partial F^1}{\partial z} \\ \frac{\partial F^2}{\partial x} & \frac{\partial F^2}{\partial z} \end{bmatrix}}{\det \begin{bmatrix} \frac{\partial F^1}{\partial z} & \frac{\partial F^1}{\partial y} \\ \frac{\partial F^2}{\partial z} & \frac{\partial F^2}{\partial y} \end{bmatrix}} \tag{3.5.22}$$

e:

$$g'(x) = \frac{\frac{\partial(F^1,F^2)}{\partial(x,y)}}{\frac{\partial(F^1,F^2)}{\partial(y,z)}} \equiv \frac{\det \begin{bmatrix} \frac{\partial F^1}{\partial x} & \frac{\partial F^1}{\partial y} \\ \frac{\partial F^2}{\partial x} & \frac{\partial F^2}{\partial y} \end{bmatrix}}{\det \begin{bmatrix} \frac{\partial F^1}{\partial y} & \frac{\partial F^1}{\partial z} \\ \frac{\partial F^2}{\partial y} & \frac{\partial F^2}{\partial z} \end{bmatrix}} \tag{3.5.23}$$

Exemplo ...

Suponhamos que $z = f(x, y)$ está definida implicitamente pela equação:

$$H(x + z, yz) = z$$

e calculemos $\frac{\partial f}{\partial x}$ e $\frac{\partial f}{\partial z}$.

Para isso, consideremos o diagrama:

$$(x, y) \in \mathbb{R}^2 \xrightarrow{\mathbf{G}} (u = x + f(x, y), v = y\,f(x, y), w = f(x, y)) \in \mathbb{R}^3 \xrightarrow{\mathbf{F}} H(u, v) - w = 0$$

a última igualdade $(= 0)$, válida quando $u = x + f(x, y), v = y\,f(x, y), w = f(x, y)$. Apliquemos a regra da cadeia, para obter:

$$[0\ \ 0] = \begin{bmatrix} \frac{\partial H}{\partial u} & \frac{\partial H}{\partial v} & -1 \end{bmatrix} . \begin{bmatrix} 1 + \frac{\partial f}{\partial x} & \frac{\partial f}{\partial y} \\ y\,\frac{\partial f}{\partial x} & f(x, y) + y\,\frac{\partial f}{\partial y} \\ \frac{\partial f}{\partial x} & \frac{\partial f}{\partial y} \end{bmatrix}$$

onde as derivadas parciais de H devem ser calculadas em $(u = x + f(x,y), v = y\, f(x,y))$. Após efectuar os cálculos, deduzimos que:

$$\begin{aligned}
\frac{\partial f}{\partial x} &= \frac{-f(x,y)\,\frac{\partial H}{\partial v}}{\frac{\partial H}{\partial u} + y\,\frac{\partial H}{\partial v} - 1} \\
\frac{\partial f}{\partial y} &= \frac{-\frac{\partial H}{\partial u}}{\frac{\partial H}{\partial u} + y\,\frac{\partial H}{\partial v} - 1}
\end{aligned}$$

onde as derivadas parciais de H devem ser calculadas em $(u = x + f(x,y), v = y\, f(x,y))$.

3.6 Teoremas da Função Inversa e da Função Implícita

Comecemos por recordar o que acontece para funções reais de variável real.

Assim, suponhamos que $f : \mathcal{U} \subseteq \mathbb{R} \to \mathbb{R}$ é uma função de classe C^1, no aberto $\mathcal{U} \subseteq \mathbb{R}$, e seja $p \in \mathcal{U}$ um ponto onde $f'(p) \neq 0$. Se por exemplo $f'(p) > 0$, então $f'(x) > 0$, $\forall x$ em algum intervalo aberto $\mathbf{I} \subseteq \mathcal{U}$, que contem p. Portanto f é estritamente crescente em $\mathbf{I}$, e existe uma inversa local g, definida em algum intervalo aberto $\mathbf{J}$, que contem $f(p)$, isto é $g : \mathbf{J} \to \mathbf{I}$ é uma funcao tal que (ver a figura (3.5)):

$$g(f(x)) = x \qquad \forall x \in \mathbf{I}$$

e:

$$f(g(y)) = y \qquad \forall y \in \mathbf{J}$$

Figure 3.5: Função inversa

.

Além disso, g é de classe C^1 em $\mathbf{J}$ e:

$$g'(y) = \frac{1}{f'(g(y))} \qquad \forall y \in \mathbf{J} \tag{3.6.1}$$

Uma situação análoga ocorre para funções de várias variáveis, embora a demonstração seja bastante mais elaborada.

Mais precisamente é válido o seguinte teorema:

♣ **Proposição 3.8** *(Teorema da função inversa)*

... Suponhamos que $\mathbf{F} : \mathcal{U} \subseteq \mathbb{R}^n \to \mathbb{R}^n$ *é uma função de classe* C^1*, no aberto* $\mathcal{U} \subseteq \mathbb{R}^n$*, e seja* $\mathbf{p} \in \mathcal{U}$ *um ponto onde:*

$$det\,\mathbf{JacF}(\mathbf{p}) \neq 0 \tag{3.6.2}$$

de tal forma que a diferencial $d\mathbf{F_p} : \mathbb{R}^n \to \mathbb{R}^n$ *é uma aplicação linear inversível (ou um isomorfismo linear).*

Então $\mathbf{F}$ *é* **localmente inversível***, isto é, existe um aberto* $\mathcal{V} \subseteq \mathcal{U}$*, que contem* $\mathbf{p}$*, um aberto* $\mathcal{W}$ *que contem* $\mathbf{F}(\mathbf{p})$*, e uma aplicação* $\mathbf{G} : \mathcal{W} \to \mathcal{V}$*, de classe* C^1*, tal que:*

$$\begin{aligned} \mathbf{G}(\mathbf{F}(\mathbf{x})) &= \mathbf{x} \qquad \forall \mathbf{x} \in \mathcal{V} \\ \mathbf{F}(\mathbf{G}(\mathbf{y})) &= \mathbf{y} \qquad \forall \mathbf{y} \in \mathcal{W} \end{aligned}$$

Além disso:

$$d\mathbf{G_y} = [d\mathbf{F}(\mathbf{G}(\mathbf{y}))]^{-1} \qquad \forall \mathbf{y} \in \mathcal{W} \tag{3.6.3}$$

ou, de forma equivalente:

$$\mathbf{JacG}(\mathbf{y}) = [\mathbf{JacF}(\mathbf{G}(\mathbf{y}))]^{-1} \qquad \forall \mathbf{y} \in \mathcal{W} \tag{3.6.4}$$

Exemplo ...

A função $\mathbf{F} : \mathbb{R}^2 \to \mathbb{R}^2$, dada por:

$$\mathbf{F}(x, y) = (e^x \cos y, e^x \sin y)$$

é de classe C^1 em $\mathbb{R}^2$, e:

$$\mathbf{JacF}(x, y) = \begin{bmatrix} e^x \cos y & -e^x \sin y \\ e^x \sin y & e^x \cos y \end{bmatrix}$$

o que implica que:

$$\det \mathbf{JacF}(x, y) = e^x \neq 0 \qquad \forall (x, y) \in \mathbb{R}^2$$

e portanto $\mathbf{F}$ é localmente inversível em todo o ponto de $\mathbb{R}^2$. No entanto $\mathbf{F}$ não é globalmente inversível, já que não é sequer injectiva.

Passemos agora a discutir o Teorema da função implícita, que, como já referimos, fornece condições para a existência de uma função $\mathbf{f}$, definida implicitamente por uma equação do tipo $\mathbf{F}(\mathbf{x}, \mathbf{y}) = \mathbf{O}$.

Mais precisamente, consideremos de novo uma função diferenciável:

$$\mathbf{F} : \mathbb{R}^{n+m} \cong \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^m$$

considerada como antes, como uma função de duas variáveis vectoriais $\mathbf{x} = (x^1, \cdots, x^n) \in \mathbb{R}^n$ e $\mathbf{y} = (y^1, \cdots, y^m) \in \mathbb{R}^m$.

Recordemos que uma função $\mathbf{f} : \mathcal{D} \subseteq \mathbb{R}^n \to \mathbb{R}^m$ está **definida implicitamente** pela equação:

$$\mathbf{F}(\mathbf{x}, \mathbf{y}) = \mathbf{O} \tag{3.6.5}$$

se:

$$\mathbf{F}(\mathbf{x}, \mathbf{f}(\mathbf{x})) = \mathbf{O} \qquad \forall \mathbf{x} \in \mathcal{D} \tag{3.6.6}$$

o que geomètricamente significa que o conjunto de nível $\mathbf{O}$, da função $\mathbf{F}$, contem o gráfico da função $\mathbf{f}$ (ver a figura (3.4)):

$$\mathbf{gr}\,\mathbf{f} \subset \mathbf{F}^{-1}(\mathbf{O})$$

O Teorema da função implícita fornece portanto condições para que esse conjunto de nível $\mathbf{F}^{-1}(\mathbf{O})$, seja descrito localmente, como o gráfico da função $\mathbf{f}$.

♣ Proposição 3.9 *(Teorema da Função Implícita)*

... Seja:

$$\mathbf{F} : \mathcal{O} \subseteq \mathbb{R}^{n+m} \cong \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^m$$

uma função de classe C^1, definida num aberto $\mathcal{O} \subseteq \mathbb{R}^{n+m} \cong \mathbb{R}^n \times \mathbb{R}^m$, que contem um ponto $(\mathbf{x}_o, \mathbf{y}_o)$ onde:

$$\mathbf{F}(\mathbf{x}_o, \mathbf{y}_o) = \mathbf{O}$$

Suponhamos que se verifica a condição:

$$det\,\mathbf{Jac_y F}(\mathbf{x}_o, \mathbf{y}_o) = det \begin{bmatrix} \frac{\partial F^1}{\partial y^1} & \frac{\partial F^1}{\partial y^2} & \cdots & \frac{\partial F^1}{\partial y^m} \\ \frac{\partial F^2}{\partial y^1} & \frac{\partial F^2}{\partial y^2} & \cdots & \frac{\partial F^2}{\partial y^m} \\ \vdots & \vdots & \cdots & \vdots \\ \vdots & \vdots & \cdots & \vdots \\ \frac{\partial F^m}{\partial y^1} & \frac{\partial F^m}{\partial y^2} & \cdots & \frac{\partial F^m}{\partial y^m} \end{bmatrix}_{(\mathbf{x}_o, \mathbf{y}_o)} \neq 0 \tag{3.6.7}$$

Então existe um aberto $\mathcal{U} \subset \mathbb{R}^n$ que contem $\mathbf{x}_o$, um aberto $\mathcal{V} \subset \mathbb{R}^m$, que contem $\mathbf{y}_o$, e uma função:

$$\mathbf{f} : \mathcal{U} \subset \mathbb{R}^n \to \mathcal{V} \subset \mathbb{R}^m$$

de classe C^1, que verifica a condição:

$$\mathbf{F}(\mathbf{x}, \mathbf{f}(\mathbf{x})) = \mathbf{O} \qquad \forall \mathbf{x} \in \mathcal{U}$$

- Demonstração...

  Definamos uma nova função $\mathbf{G} : \mathcal{O} \subseteq \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^n \times \mathbb{R}^m$, através de:

$$\mathbf{G}(\mathbf{x}, \mathbf{y}) = (\mathbf{x}, \mathbf{F}(\mathbf{x}, \mathbf{y}))$$

  Temos então que:

$$\mathbf{JacG}(\mathbf{x}_o, \mathbf{y}_o) = \begin{bmatrix} \mathrm{Id}_n & \mathbf{O} \\ \mathbf{Jac_x F}(\mathbf{x}_o, \mathbf{y}_o) & \mathbf{Jac_y F}(\mathbf{x}_o, \mathbf{y}_o) \end{bmatrix} \tag{3.6.8}$$

  e portanto $\det \mathbf{JacG}(\mathbf{x}_o, \mathbf{y}_o) = \det \mathbf{Jac_y F}(\mathbf{x}_o, \mathbf{y}_o) \neq 0$.

  Podemos então aplicar o teorema da função inversa, para concluir que existe um aberto $\mathcal{W}$, em $\mathbb{R}^n \times \mathbb{R}^m$, que contem $\mathbf{G}(\mathbf{x}_o, \mathbf{y}_o) = (\mathbf{x}_o, \mathbf{O})$, um aberto em $\mathbb{R}^n \times \mathbb{R}^m$, que contem $(\mathbf{x}_o, \mathbf{y}_o)$, e que podemos supôr ser da forma $\mathcal{U} \times \mathcal{V}$, tal que $\mathbf{G} : \mathcal{U} \times \mathcal{V} \to \mathcal{W}$ admite uma inversa $\mathbf{H} : \mathcal{W} \to \mathcal{U} \times \mathcal{V}$, de classe C^1.

Claramente que $\mathbf{H}$ é da forma $\mathbf{H}(\mathbf{x}, \mathbf{y}) = (\mathbf{x}, \mathbf{k}(\mathbf{x}, \mathbf{y}))$, para uma certa função $\mathbf{k}$, de classe C^1.

Designemos por $\pi : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^m$, a projecção no segundo factor,i.e., $\pi(\mathbf{x}, \mathbf{y}) = \mathbf{y}$. Temos então que, atendendo a que $\pi \circ \mathbf{G} = \mathbf{F}$:

$$\mathbf{F}(\mathbf{x}, \mathbf{k}(x, y)) = \mathbf{F} \circ \mathbf{H}(\mathbf{x}, \mathbf{y}) = (\pi \circ \mathbf{G}) \circ \mathbf{H}(\mathbf{x}, \mathbf{y}) = \pi \circ (\mathbf{G}) \circ \mathbf{H})(\mathbf{x}, \mathbf{y}) = \pi(\mathbf{x}, \mathbf{y}) = \mathbf{y} \tag{3.6.9}$$

Portanto $\mathbf{F}(\mathbf{x}, \mathbf{k}(x, \mathbf{O})) = \mathbf{O}$, e podemos tomar finalmente $\mathbf{f}(\mathbf{x}) = \mathbf{k}(\mathbf{x}, \mathbf{O})$,

♣.

Figure 3.6: Teor. da função implícita, quando $F : \mathbf{R}^2 \to \mathbf{R}$. Nos pontos em que $\nabla F = (\frac{\partial F}{\partial x}, \frac{\partial F}{\partial y})$ é horizontal, i.e., nos quais $\frac{\partial F}{\partial y} = 0$, não é possível exprimir localmente $F^{-1}(0)$, como o gráfico de uma função f

.

Exemplo ...

A equação:

$$x + y + z - \sin(xyz) = 0$$

admite a solução $(0, 0, 0)$. Podemos aplicar o teorema da função implícita par justificar o facto de que, perto da origem $\mathbf{O} = (0, 0, 0) \in \mathbb{R}^3$, podemos exprimir z como uma função de x e y. Noutras palavras, perto da origem $\mathbf{O} = (0, 0, 0) \in \mathbb{R}^3$, o conjunto de nível:

$$\{(x, y, z) \in \mathbb{R}^3 : \, x + y + z - \sin(xyz) = 0\}$$

pode ser descrito como o gráfico de uma função de classe C^1, $z = f(x, y)$.

Com efeito, definindo $F : \mathbb{R}^3 \to \mathbb{R}$, através de :

$$F(x, y, z) = x + y + z - \sin(xyz)$$

vemos que F é de classe C^1, que $F(0, 0, 0) = 0$ e ainda que:

$$\frac{\partial F}{\partial z}(0, 0, 0) = 1 \neq 0$$

Portanto existe um aberto $\mathcal{U} \subset \mathbb{R}^2$, que contem $(0,0)$, um aberto $\mathcal{V} \subset \mathbb{R}$, que contem 0, e uma função de classe C^1, $f : \mathcal{U} \subset \mathbb{R}^2 \to \mathcal{V}$, tal que:

$$F(x, y, f(x, y)) = x + y + f(x, y) - \sin(xyf(x, y)) = 0$$

$\forall (x, y) \in \mathcal{U}$.

Exemplo ...

Consideremos as equações:

$$z^3 x + w^2 y^3 + 2xy = 0 \qquad e \qquad xyzw - 1 = 0$$

e demonstremos que perto do ponto $(-1, -1, 1, 1)$, podemos exprimir localmente, z e w como funções de x e y.

Para isso definamos a função:

$$\mathbf{F}(x, y, z, w) = (z^3 x + w^2 y^3 + 2xy, xyzw - 1)$$

Temos então que $\mathbf{F}$ é de classe C^1, que:

$$\mathbf{F}((-1, -1), (1, 1)) = (0, 0)$$

e ainda que:

$$\begin{aligned}
\det \mathbf{Jac_y F}((-1, -1), (1, 1)) &= \det \begin{bmatrix} \frac{\partial F^1}{\partial z} & \frac{\partial F^1}{\partial w} \\ \frac{\partial F^2}{\partial z} & \frac{\partial F^2}{\partial w} \end{bmatrix}_{((-1,-1),(1,1))} \\
&= \det \begin{bmatrix} 3z^2 x & 2wy^3 \\ xyw & xyz \end{bmatrix}_{((-1,-1),(1,1))} \\
&= \det \begin{bmatrix} -3 & -2 \\ 1 & 1 \end{bmatrix} = -1 \neq 0
\end{aligned}$$

Portanto existe um aberto $\mathcal{U} \subset \mathbb{R}^2$, que contem $(-1, -1)$, um aberto $\mathcal{V} \subset \mathbb{R}^2$, que contem $(1, 1)$, e uma função de classe C^1, $\mathbf{f} = (f, g) : \mathcal{U} \subset \mathbb{R}^2 \to \mathcal{V}$, tal que:

$$F(x, y, f(x, y), g(x, y)) = (0, 0) \qquad \forall (x, y) \in \mathcal{U}$$

.

3.7 Variedades parametrizadas. Parametrizações

Nesta secção vamos supôr por simplicidade, que todas as aplicações referidas são de classe C^∞.

Vamos agora definir variedades parametrizadas em $\mathbb{R}^n$. Intuitivamente, uma variedade parametrizada de dimensão k, em $\mathbb{R}^n$, é um subconjunto M de $\mathbb{R}^n$, em que cada ponto $\mathbf{p} \in M$ necessita de k números para que a sua posição seja univocamente determinada.

♣ **Definição 3.9** ... *Um subconjunto $M \subset \mathbb{R}^n$ diz-se uma* **variedade parametrizada** *de dimensão k, em $\mathbb{R}^n$, se e só se para cada ponto $\mathbf{p} \in M$, existe um aberto $\mathcal{U} \subset \mathbb{R}^n$, que contem $\mathbf{p}$, um aberto $\mathcal{W} \subset \mathbb{R}^k$, e uma função $\Phi : \mathcal{W} \to \mathbb{R}^n$, tal que (ver a figura(3.7)):*

(i)... Φ é injectiva

(i)... $\Phi(\mathcal{W}) = M \cap \mathcal{U}$

M diz-se uma **variedade parametrizada regular** *se além disso, se verifica a condição seguinte:*

(iii)... $d\Phi_{\mathbf{u}}$ tem característica k, $\forall \mathbf{u} \in \mathcal{W} \subset \mathbb{R}^k$, isto é, $d\Phi_{\mathbf{u}}$ é injectiva, $\forall \mathbf{u} \in \mathcal{W} \subset \mathbb{R}^k$.

Uma tal função Φ diz-se uma **parametrização local (regular)** *da variedade M, em torno de* $\mathbf{p}$.

As coordenadas $(u^1, \cdots, u^k)$ de cada ponto $\mathbf{u} = \Phi^{-1}(\mathbf{q}) \in \mathcal{W} \subset \mathbb{R}^k$, onde $\mathbf{q} \in M \cap \mathcal{U}$, dizem-se as **coordenadas locais (intrínsicas)** *de $\mathbf{q}$, associadas à parametrização local Φ (ver a figura(3.7)).*

Figure 3.7: Coordenadas locais (intrínsicas)

.

Geomètricamente, uma variedade (parametrizada regular) de dimensão k, em $\mathbb{R}^n$, é um subconjunto de $\mathbb{R}^n$, que é localmente como um aberto de $\mathbb{R}^k$, deformado de "maneira regular".

Como casos particulares extremos da definição anterior, temos: (i)... um conjunto discreto de pontos em $\mathbb{R}^n$, que é uma variedade em $\mathbb{R}^n$, de dimensão 0, e (ii)... um aberto de $\mathbb{R}^n$, que é uma variedade em $\mathbb{R}^n$, de dimensão n. Em particular, é usual utilizar as seguintes coordenadas locais:

(i)... **Coordenadas polares em $\mathbb{R}^2$**

Neste caso podemos por exemplo tomar $\mathcal{W} \subset \mathbb{R}^2_{(r,\theta)}$, $\mathcal{V} \subset \mathbb{R}^2_{(x,y)}$, definidos por:

$$\mathcal{W} = \{(r,\theta) : \, r > 0 \quad e \quad 0 < \theta < 2\pi\}$$
$$\mathcal{V} = \mathbb{R}^2 - \{(x,0) : \, x \geq 0\}$$

e a parametrização $\Phi : \mathcal{W} \to \mathcal{V}$, definida por:

$$\Phi(r,\theta) = (r \, \cos\theta, r \, \sin\theta) \tag{3.7.1}$$

(r, θ) dizem-se as coordenadas polares de $(x, y) = \Phi(r, \theta)$.

(ii)... **Coordenadas esféricas em $\mathbb{R}^3$**

Neste caso podemos por exemplo tomar $\mathcal{W} \subset \mathbb{R}^3_{(r,\theta,\varphi)}$, $\mathcal{V} \subset \mathbb{R}^3_{(x,y,z)}$, definidos por:

$$\mathcal{W} = \{(r, \theta, \varphi) : \ r > 0, \quad 0 < \theta < \pi \ \ e \ \ 0 < \varphi < 2\pi\}$$
$$\mathcal{V} = \mathbb{R}^3 - \{(x, 0, z) : \ x \geq 0\}$$

e a parametrização $\Phi : \mathcal{W} \to \mathcal{V}$, definida por:

$$\Phi(r, \theta, \varphi) = (r \ \sin\theta \ \cos\varphi, r \ \sin\theta \ \sin\varphi, r \ \cos\theta) \tag{3.7.2}$$

(r, θ) dizem-se as coordenadas esféricas de $(x, y, z) = \Phi(r, \theta, \varphi)$ (ver a figura (3.8.13)).

Figure 3.8: Coordenadas esféricas

.

(iii)... **Coordenadas cilindricas em $\mathbb{R}^3$**

Neste caso podemos por exemplo tomar $\mathcal{W} \subset \mathbb{R}^3_{(r,\theta,z)}$, $\mathcal{V} \subset \mathbb{R}^3_{(x,y,z)}$, definidos por:

$$\mathcal{W} = \{(r, \theta, z) : \ r > 0, \quad 0 < \theta < 2\pi\}$$
$$\mathcal{V} = \mathbb{R}^3 - \{(x, 0, z) : \ x \geq 0\}$$

e a parametrização $\Phi : \mathcal{W} \to \mathcal{V}$, definida por:

$$\Phi(r, \theta, z) = (r \ \cos\theta, r \ \sin\theta, z) \tag{3.7.3}$$

(r, θ) dizem-se as coordenadas cilindricas de $(x, y, z) = \Phi(r, \theta, z)$ (ver a figura (3.9)).

Exemplos mais interessantes podem ser obtidas aplicando a proposição seguinte, cuja demonstração omitimos:

♣ **Proposição 3.10** ... *Seja* $\mathbf{F} : \mathcal{O} \subseteq \mathbb{R}^n \to \mathbb{R}^m$*, uma aplicação de classe* C^∞*, definida num aberto* $\mathcal{O} \subseteq \mathbb{R}^n$*, com* $n \geq m$*.*

Suponhamos que $\mathbf{c} \in \mathbb{R}^m$ *é* **valor regular** *de* $\mathbf{F}$*, isto é, que:*

$$d\mathbf{F}_{\mathbf{x}} \quad \textit{é sobrejectiva} \ \ \forall \mathbf{x} \in M \equiv \mathbf{F}^{-1}(\mathbf{c}) \tag{3.7.4}$$

Então $M \equiv \mathbf{F}^{-1}(\mathbf{c})$ *é uma variedade (parametrizada regular) em* $\mathbb{R}^n$*, de dimensão* $k = n - m$*.*

Figure 3.9: Coordenadas cilindricas

.

Notemos que se $\mathbf{F} = (F^1, \cdots, F^m)$, então $\mathbf{c}$ é valor regular de $\mathbf{F}$ se e só se, em cada ponto $\mathbf{x} \in M \equiv \mathbf{F}^{-1}(\mathbf{c})$, os vectores gradiente $\nabla F^1(\mathbf{x})$,...,$\nabla F^m(\mathbf{x})$ são linearmente independentes.

Em particular, quando $m = 1$, esta condição significa que $\nabla F(\mathbf{x}) \neq 0$, $\forall \mathbf{x} \in M \equiv \mathbf{F}^{-1}(\mathbf{c})$, donde se conclui que uma hipersuperfície regular em $\mathbb{R}^n$, tal como foi definida no capítulo 3, é uma variedade de dimensão $n-1$, de acordo com a definição anterior.

Se $\Phi_1 : \mathcal{W}_1 \subset \mathbb{R}^k \to \mathbb{R}^n$ e $\Phi_2 : \mathcal{W}_2 \subset \mathbb{R}^k \to \mathbb{R}^n$, são duas parametrizações locais em torno de um ponto $\mathbf{p} \in M$, então é possível mostrar que as chamadas **aplicações de mudança de coordenadas** (ver a figura (3.10)):

$$\Phi_2^{-1} \circ \Phi_1 : \Phi_1^{-1}(\Phi_2(\mathcal{W}_2)) \to \mathbb{R}^k \tag{3.7.5}$$

e:

$$\Phi_1^{-1} \circ \Phi_2 : \Phi_2^{-1}(\Phi_1(\mathcal{W}_1)) \to \mathbb{R}^k \tag{3.7.6}$$

são de classe C^∞, com matrizes jacobianas não singulares.

Figure 3.10: Aplicações de mudança de coordenadas

.

Na secção seguinte veremos mais exemplos de variedades em $\mathbb{R}^n$.

3.8 Espaço Tangente. Método dos Multiplicadores de Lagrange II

♣ Definição 3.10 ... *Seja M uma variedade (parametrizada regular) de dimensão k, em $\mathbb{R}^n$. O* **espaço tangente a** M **num ponto p** $\in M$*, é o subespaço vectorial de $\mathbb{R}^n$, notado por $T_{\mathbf{p}}M$, e que pode ser descrito das seguintes formas equivalentes:*

(**A**). *Consideramos uma parametrização local:*

$$\Phi : \mathcal{W} \subset \mathbb{R}^k \to \mathbb{R}^n$$

de M em torno de $\mathbf{p}$. Se $\Phi(\mathbf{u}) = \mathbf{p}$, pômos então:

$$T_{\mathbf{p}}M \equiv d\Phi_u(\mathbb{R}^k) \tag{3.8.1}$$

(**B**). *Consideramos todos as curvas de classe C^∞, $\alpha : \mathbf{I} \subset \mathbb{R} \to \mathbb{R}^n$, tais que:*

$$\alpha(t) \in M, \quad \forall t \in \mathbf{I} \quad e \quad \alpha(0) = \mathbf{p}$$

Pômos então:

$$T_{\mathbf{p}}M \equiv \{\mathbf{V_p} = \dot{\alpha}(0) : \alpha \ \text{ nas condições indicadas}\} \tag{3.8.2}$$

Vejamos a equivalência das duas definições anteriores. A definição (**A**), apresenta $T_{\mathbf{p}}M$ como um subespaço vectorial de dimensão k, em $\mathbb{R}^n$, (uma vez que $d\Phi_{\mathbf{p}}$ é injectiva), mas tem o inconveniente de depender da parametrização Φ escolhida em (3.8.1). Não está claro que se tomarmos uma outra parametrização Ψ, com $\Psi(\mathbf{r}) = \mathbf{p}$, se tem:

$$d\Phi_{\mathbf{u}}(\mathbb{R}^k) = d\Psi_{\mathbf{r}}(\mathbb{R}^k) \tag{3.8.3}$$

Por outro lado, a definição (**B**), embora não dependa da parametrização, não torna claro que $T_{\mathbf{p}}M$ seja de facto um subespaço vectorial de dimensão k, em $\mathbb{R}^n$.

Ambos os inconvenientes ficam resolvidos, provando que (**A**) e (**B**), conduzem ao mesmo conjunto.

Com efeito, seja $\mathbf{V_p} \in d\Phi_{\mathbf{u}}(\mathbb{R}^k) \subset \mathbb{R}^n$. Temos então que $\mathbf{V_p} = d\Phi_{\mathbf{u}}(\mathbf{v})$ para algum vector $\mathbf{v} \in \mathbb{R}^k$, e é evidente que $\mathbf{V_p} = \dot{\alpha}(0)$, onde α é a curva:

$$\alpha(t) = \Phi(\mathbf{u} + t\,\mathbf{v}) \qquad t \in \mathbf{I}$$

que satisfaz as condições referidas em **B**.

Recìprocamente, seja $\alpha : \mathbf{I} \to \mathbb{R}^n$ uma curva de classe C^∞, tal que $\alpha(t) \in M, \forall t \in \mathbf{I}$, $\alpha(0) = \mathbf{p} \in M$ e $\dot{\alpha}(0) = \mathbf{V_p}$. Podemos supôr que $\mathbf{I}$ é suficientemente pequeno, para que $\alpha(\mathbf{I}) \subset \Phi(\mathcal{W}) = \mathcal{U} \cap M$ (ver a definição de parametrização local). Temos então que a curva:

$$\beta = \Phi^{-1} \circ \alpha : \mathbf{I} \to \mathbb{R}^k$$

é de classe C^∞, e como $\Phi \circ \beta = \alpha$, a regra da cadeia dá que:

$$\dot{\alpha}(0) = d\Phi_{\beta(0)}(\dot{\beta}(0)) = d\Phi_{\mathbf{u}}(\dot{\beta}(0))$$

isto é, $\mathbf{V_p} \equiv \dot{\alpha}(0) = d\Phi_{\mathbf{u}}(\dot{\beta}(0)) \in d\Phi_{\mathbf{u}}(\mathbb{R}^k)$, como se pretendia provar.

Habitualmente visualiza-se o espaço tangente $T_{\mathbf{p}}M$, como sendo o subespaço afim paralelo a $T_{\mathbf{p}}M$, passando por $\mathbf{p}$, como na figura (3.11). Mas não esqueçamos que $T_{\mathbf{p}}M$, tal como o definimos, é um subespaço vectorial de $\mathbb{R}^n$ (passando sempre na origem).

Figure 3.11: Espaço tangente $T_{\mathbf{p}}M$

.

Dada uma parametrização local:

$$\Phi : \mathcal{W} \subset \mathbb{R}^k \to \mathbb{R}^n$$

com $\Phi(\mathbf{u}) = \mathbf{p} \in M$, recordemos que a matriz da aplicação linear $d\Phi_u \in \mathcal{L}(\mathbb{R}^k, \mathbb{R}^n)$, relativamente às bases canónicas de $\mathbb{R}^k$ e $\mathbb{R}^n$, é a matriz Jacobiana $(n \times k)$:

$$\mathbf{Jac}\Phi(\mathbf{u}) = [\frac{\partial \Phi_i}{\partial u^j}(\mathbf{u})]_{i,j} = \begin{bmatrix} \frac{\partial \Phi_1}{\partial u^1}(\mathbf{u}) & \frac{\partial \Phi_1}{\partial u^2}(\mathbf{u}) & \dots & \frac{\partial \Phi_1}{\partial u^k}(\mathbf{u}) \\ \frac{\partial \Phi_2}{\partial u^1}(\mathbf{u}) & \frac{\partial \Phi_2}{\partial u^2}(\mathbf{u}) & \dots & \frac{\partial \Phi_2}{\partial u^k}(\mathbf{u}) \\ \vdots & \vdots & \dots & \vdots \\ \vdots & \vdots & \dots & \vdots \\ \frac{\partial \Phi_n}{\partial u^1}(\mathbf{u}) & \frac{\partial \Phi_n}{\partial u^2}(\mathbf{u}) & \dots & \frac{\partial \Phi_n}{\partial u^k}(\mathbf{u}) \end{bmatrix} \tag{3.8.4}$$

cujas colunas são as componentes das derivadas parciais vectoriais $\frac{\partial \Phi}{\partial u^i}(\mathbf{p})$ $(i = 1, \cdots, k)$, na base canónica de $\mathbb{R}^n$.

Como $d\Phi_{\mathbf{u}}$ tem característica k, $\forall \mathbf{u} \in \mathcal{W}$, as colunas dessa matriz são vectores linearmente independentes, em $\mathbb{R}^n$.

Podemos por isso definir uma base para o espaço tangente $T_{\mathbf{p}}M$, da seguinte forma:

$$\Phi_{u^1}(\mathbf{p}) \equiv \frac{\partial \Phi}{\partial u^1}(\mathbf{u}), \cdots, \Phi_{u^k}(\mathbf{p}) \equiv \frac{\partial \Phi}{\partial u^k}(\mathbf{u}) \tag{3.8.5}$$

onde $\mathbb{R}^k$ está munido das coordenadas cartesianas $u^1, \cdots, u^k$, e $\Phi_{u^i}(\mathbf{p}) = \frac{\partial \Phi}{\partial u^i}(\mathbf{u})$ é a derivada parcial vectorial de Φ, em ordem a u^i, $(i = 1, \cdots, k)$.

As coordenadas de um vector $\mathbf{V_p} \in T_{\mathbf{p}}M$, na base (3.8.5), ditas **coordenadas intrínsicas de $\mathbf{V_p}$**, são determinadas da seguinte forma: como vimos, $\mathbf{V_p}$ é da forma $\mathbf{V_p} = \dot{\alpha}(0)$ para alguma curva de classe C^∞, da forma $\alpha(t) = \Phi(\beta(t))$, $t \in \mathbf{I}$, com $\beta(0) = \mathbf{u} = \Phi^{-1}(\mathbf{p})$ e onde:

$$\beta(t) = (u^1, \cdots, u^k(t))$$

é a chamada **expressão local nas coordenadas locais u^i, da curva** α (ver a figura (3.12)). Temos então que, pela regra da cadeia:

$$\begin{aligned}
\mathbf{V_p} &= \dot{\alpha}(0) = \frac{d}{dt}\Big|_{t=0} (\Phi \circ \beta) = \frac{d}{dt}\Big|_{t=0} \Phi(u^1(t), \cdots, u^k(t)) \\
&= u_1'(0)\, \frac{\partial \Phi}{\partial u^1}(\mathbf{u}) + \cdots + u_k'(0)\, \frac{\partial \Phi}{\partial u^k}(\mathbf{u}) \\
&= u_1'(0)\, \Phi_{u^1}(\mathbf{p}) + \cdots + u_k'(0)\, \Phi_{u^k}(\mathbf{p}) \qquad (3.8.6)
\end{aligned}$$

Portanto na base $\{\Phi_{u^i}(\mathbf{p})\}_{i=1,\cdots,k}$ para $T_{\mathbf{p}}M$, associada à parametrização local Φ, as coordenadas intrínsicas de um vector $\mathbf{V_p} \in T_{\mathbf{p}}M$, são $(u_1'(0), \cdots, u_k'(0))$, onde $(u^1(t), \cdots, u^k(t))$, é a expressão local nas coordenadas locais u^i, de uma curva α nas condições indicadas.

Figure 3.12: Coordenadas intrínsicas de $\mathbf{V_p} \in T_{\mathbf{p}}M$

.

<u>**Exemplo**</u> ...

(i)... A base para o espaço tangente $T_{\mathbf{p}}(\mathbb{R}^2) \cong \mathbb{R}^2$, num ponto $\mathbf{p} = \Phi(r, \theta) \in \mathbb{R}^2$, associada à parametrização local em coordenadas polares (ver (3.7.1)), é constituída pelos dois vectores seguintes:

$$\begin{aligned}
\Phi_r &\equiv \frac{\partial \Phi}{\partial r}(r, \theta) = (\cos\theta, \sin\theta) \\
\Phi_\theta &\equiv \frac{\partial \Phi}{\partial \theta}(r, \theta) = (-r\, \sin\theta, r\, \cos\theta) \qquad (3.8.7)
\end{aligned}$$

Por vezes utilizam-se também os vectores unitários (ver a figura (**??**):

$$\begin{aligned}
\mathbf{e}_r &\equiv \frac{\Phi_r}{\|\Phi_r\|} = (\cos\theta, \sin\theta) \\
\mathbf{e}_\theta &\equiv \frac{\Phi_\theta}{\|\Phi_\theta\|} = (-\sin\theta, \cos\theta) \qquad (3.8.8)
\end{aligned}$$

(ii)... A base para o espaço tangente $T_{\mathbf{p}}(\mathbb{R}^3) \cong \mathbb{R}^3$, num ponto $\mathbf{p} = \Phi(r, \theta, \varphi) \in \mathbb{R}^3$, associada à parametrização local em coordenadas esféricas (ver (3.7.2)), é constituída pelos três vectores seguintes:

$$\Phi_r \equiv \frac{\partial \Phi}{\partial r}(r, \theta, \varphi) = (\sin\theta\, \cos\varphi, \sin\theta\, \sin\varphi, \cos\theta)$$

$$\begin{aligned} \Phi_\theta &\equiv \frac{\partial \Phi}{\partial \theta}(r,\theta,\varphi) = (r\,\cos\theta\cos\varphi, r\,\cos\theta\,\sin\varphi, -r\,\cos\theta) \\ \Phi_\varphi &\equiv \frac{\partial \Phi}{\partial \varphi}(r,\theta,\varphi) = (-r\,\sin\theta\sin\varphi, r\,\sin\theta\,\cos\varphi, 0) \end{aligned} \tag{3.8.9}$$

Por vezes utilizam-se também os vectores unitários (ver a figura (3.8.13):

$$\begin{aligned} \mathbf{e}_r &\equiv \frac{\Phi_r}{\|\Phi_r\|} = (\sin\theta\,\cos\varphi, \sin\theta\,\sin\varphi, \cos\theta) \\ \mathbf{e}_\theta &\equiv \frac{\Phi_\theta}{\|\Phi_\theta\|} = (\cos\theta\cos\varphi, \cos\theta\,\sin\varphi, \cos\theta) \\ \mathbf{e}_\varphi &\equiv \frac{\Phi_\varphi}{\|\Phi_\varphi\|} = (-\sin\varphi, \cos\varphi, 0) \end{aligned} \tag{3.8.10}$$

(iii)... A base para o espaço tangente $T_{\mathbf{p}}(\mathbb{R}^3) \cong \mathbb{R}^3$, num ponto $\mathbf{p} = \Phi(r,\theta,z) \in \mathbb{R}^3$, associada à parametrização local em coordenadas cilindricas (ver (3.7.3)), é constituída pelos três vectores seguintes:

$$\begin{aligned} \Phi_r &\equiv \frac{\partial \Phi}{\partial r}(r,\theta,z) = (\cos\theta, \sin\theta, 0) \\ \Phi_\theta &\equiv \frac{\partial \Phi}{\partial \theta}(r,\theta,z) = (-r\,\sin\theta, r\,\cos\theta, 0) \\ \Phi_z &\equiv \frac{\partial \Phi}{\partial z}(r,\theta,z) = (0,0,1) \end{aligned} \tag{3.8.11}$$

Por vezes utilizam-se também os vectores unitários (ver a figura (3.9):

$$\begin{aligned} \mathbf{e}_r &\equiv \frac{\Phi_r}{\|\Phi_r\|} = (\cos\theta, \sin\theta, 0) \\ \mathbf{e}_\theta &\equiv \frac{\Phi_\theta}{\|\Phi_\theta\|} = (-\sin\theta, \cos\theta, 0) \\ \mathbf{e}_z &\equiv \frac{\Phi_z}{\|\Phi_z\|} = (0,0,1) \end{aligned} \tag{3.8.12}$$

<u>**Exemplo**</u> ...

Sabemos já que a esfera de raio 1, $S^2 \subset \mathbb{R}^3$:

$$S^2 = \{\mathbf{x} \in \mathbb{R}^3 : \|\mathbf{x}\|^2 = 1\}$$

é uma variedade de dimensão 2 (uma superfície) em $\mathbb{R}^3$.

Em muitos cálculos práticos, é útil utilizar a seguinte parametrização local em coordenadas "esféricas" (ou "geográficas") (ver a figura (3.13):

$$\Phi(\theta,\varphi) = (\sin\theta\,\cos\varphi, \sin\theta\,\sin\varphi, \cos\theta) \tag{3.8.13}$$

definidas por exemplo no aberto:

$$\mathcal{U} = \{(\theta,\varphi) : 0 < \theta < \pi \quad 0 < \varphi < 2\pi\} \subset \mathbb{R}^2_{(\theta,\varphi)}$$

de tal forma que Φ é uma bijecção sobre o aberto:

$$\Phi(\mathcal{U}) = S^2 - \{(x,y,z) \in \mathbb{R}^3 : y = 0 \;\; x \geq 0\} \subset S^2$$

Figure 3.13: Coordenadas esféricas. θ é a colatitude, e φ a longitude

.

A matriz Jacobiana de Φ em $(\theta, \varphi) \in \mathcal{U}$, é igual a:

$$\mathbf{Jac}\Phi(\theta, \varphi) = \begin{bmatrix} \cos\theta\,\cos\varphi & -\sin\theta\,\sin\varphi \\ \cos\theta\,\sin\varphi & -\sin\theta\,\cos\varphi \\ -\sin\theta & 0 \end{bmatrix} \tag{3.8.14}$$

e tem característica 2, em todo o ponto $(\theta, \varphi) \in \mathcal{U}$.

Com efeito, a característica de $\mathbf{Jac}\Phi(\theta, \varphi)$ é 2, sse pelo menos um dos seus menores de ordem 2 fôr $\neq 0$. Os menores de ordem 2, são:

$$\cos\theta\,\sin\theta, \qquad \sin^2\theta\,\cos\varphi \quad e \quad \sin^2\theta\,\sin\varphi$$

Se eles se anulassem simultâneamente, então viria que:

$$\cos^2\theta\,\sin^2\theta + \sin^4\theta\,\cos^2\varphi + \sin^4\theta\,\sin^2\varphi = \sin^2\theta = 0$$

o que é impossível em $\mathcal{U}$, onde $0 < \theta < \pi$.

A base para o espaço tangente $T_{\mathbf{p}}S^2$, num ponto $\mathbf{p} = \Phi(\theta, \varphi) \in S^2$, associada à parametrização local Φ, é constituída pelos dois vectores seguintes:

$$\begin{aligned} \Phi_\theta &\equiv \frac{\partial\Phi}{\partial\theta}(\theta, \varphi) = (\cos\theta\,\cos\varphi, \cos\theta\,\sin\varphi, -\sin\theta) \\ \Phi_\varphi &\equiv \frac{\partial\Phi}{\partial\varphi}(\theta, \varphi) = (-\cos\theta\,\sin\varphi, \sin\theta\,\cos\varphi, 0) \end{aligned} \tag{3.8.15}$$

Exemplo ...

Considere um círculo no plano yz, com centro em $(0, a, 0)$ $(a > 0)$, e raio r, com $0 < r < a$. Este círculo é dado pelas equações:

$$(y - a)^2 + z^2 = r^2 \quad e \quad x = 0$$

e os pontos da superfície em $\mathbb{R}^3$, obtida rodando este círculo em torno do eixo dos zz, satisfazem a equação:

$$z^2 = r^2 - (\sqrt{x^2 + y^2} - a)^2$$

Se consideramos a função, definida em $\mathbb{R}^3$, através de:

$$f(x, y, z) = r^2 - (\sqrt{x^2 + y^2} - a)^2 - z^2$$

é fácil ver que 0 é valor regular de f, e que portanto $T^2 \equiv f^{-1}(0)$, é uma variedade de dimensão 2 em $\mathbb{R}^3$, chamada um "toro" bidimensional.

Uma parametrização local para T^2, é por exemplo (ver a figura (3.14)):

$$\Phi(u,v) = ((a + r\,\cos u)\,\cos v, (a + r\,\cos u)\,\sin v, r\,\sin u) \tag{3.8.16}$$

definida no aberto $\mathcal{U} \subset \mathbb{R}^2_{(u,v)}$:

$$\mathcal{U} = \{(u,v) \in \mathbb{R}^2 : 0 < u < 2\pi \quad 0 < v < 2\pi\}$$

Figure 3.14: Parametrização local de um toro

.

A base para o espaço tangente $T_{\mathbf{p}}T^2$, num ponto $\mathbf{p} = \Phi(u,v) \in T^2$, associada à parametrização local Φ, é constituída pelos dois vectores seguintes:

$$\begin{aligned} \Phi_u &\equiv \frac{\partial \Phi}{\partial u}(u,v) = (-r\sin u\,\cos v, -r\sin u\,\sin v, r\cos u) \\ \Phi_v &\equiv \frac{\partial \Phi}{\partial v}(u,v) = (-(a + r\cos u)\,\sin v, (a + r\cos u)\,\cos v, 0) \end{aligned} \tag{3.8.17}$$

<u>**Exemplo**</u> ...

Seja $\mathbf{f} : \mathcal{U} \subseteq \mathbb{R}^n \to \mathbb{R}^m$, uma função de classe C^∞, definida num aberto $\mathcal{U} \subseteq \mathbb{R}^n$, e considere o gráfico de $\mathbf{f}$:

$$M = \mathbf{gr}\,\mathbf{f} = \{(\mathbf{x},\mathbf{y}) \in \mathbb{R}^n \times \mathbb{R}^m : \mathbf{y} = \mathbf{f}(\mathbf{x}) \quad \mathbf{x} \in \mathcal{U}\}$$

É fácil verificar que $M = \mathbf{gr}\,\mathbf{f}$ é uma variedade (parametrizada regular) de dimensão n, em $\mathbb{R}^{n+m}$, e que M pode ser parametrizada (globalmente) por:

$$\Phi : \mathbf{x} \to (\mathbf{x}, \mathbf{f}(\mathbf{x})) \qquad \mathbf{x} \in \mathcal{U}$$

A base para o espaço tangente $T_{\mathbf{p}}(\mathbf{gr}\,\mathbf{f})$, num ponto $\mathbf{p} = (\mathbf{x}, \mathbf{f}(x))$, associada à parametrização Φ, é constituída pelos n vectores seguintes de $\mathbb{R}^{n+m}$:

$$\begin{aligned} \Phi_{x^1} &\equiv \frac{\partial \Phi}{\partial x^1}(\mathbf{x}, \mathbf{f}(x)) = (\mathbf{e}_1, \frac{\partial \mathbf{f}}{\partial x^1}(\mathbf{x})) \\ \Phi_{x^2} &\equiv \frac{\partial \Phi}{\partial x^2}(\mathbf{x}, \mathbf{f}(x)) = (\mathbf{e}_2, \frac{\partial \mathbf{f}}{\partial x^2}(\mathbf{x})) \\ &\cdots \\ \Phi_{x^n} &\equiv \frac{\partial \Phi}{\partial x^n}(\mathbf{x}, \mathbf{f}(x)) = (\mathbf{e}_n, \frac{\partial \mathbf{f}}{\partial x^n}(\mathbf{x})) \end{aligned} \tag{3.8.18}$$

onde $\{\mathbf{e}_j\}_{j=1,\cdots,n}$ é a base canónica de $\mathbb{R}^n$.

Exemplo ...

Uma **superfície de revolução** M, é obtida rodando uma curva plana regular C, em torno de uma linha nesse plano, que não intersecte a curva C. Tomemos o referido plano, como sendo o plano xz, e o eixo da rotação como sendo o eixo dos zz.

Suponhamos agora que:

$$x = f(v) \quad z = g(v) \quad \text{com} \;\; a < v < b \;\; e \;\; f(v) > 0$$

é uma parametrização regular para a curva C, e representemos por φ o ângulo da rotação em torno do eixo dos zz. Obtemos então uma parametrização local de M, através de (ver a figura (3.15)):

$$\Phi(\varphi, v) = (f(v)\,\cos\varphi, f(v)\,\sin\varphi, g(v))$$

definida no aberto $\mathcal{U} \subset \mathbb{R}^2_{(\varphi,v)}$:

$$\mathcal{U} = \{(\varphi, v) : 0 < \varphi < \pi \quad e \quad a < v < b\}$$

A base para o espaço tangente $T_{\mathbf{p}}M$, num ponto $\mathbf{p} = \Phi(\varphi, v) \in M$, associada à parametrização local Φ, é constituída pelos dois vectores seguintes:

$$\begin{aligned} \Phi_\varphi &\equiv \frac{\partial \Phi}{\partial \varphi}(\varphi, v) = (-f(v)\,\sin\varphi, f(v)\,\cos\varphi, 0) \\ \Phi_v &\equiv \frac{\partial \Phi}{\partial v}(\varphi, v) = (f'(v)\,\cos\varphi, f'(v)\,\sin\varphi, g'(v)) \end{aligned} \tag{3.8.19}$$

Figure 3.15: Superfície de revolução

.

Quando a variedade M é dada como imagem inversa de um valor regular, o espaço tangente pode ser calculado através da seguinte proposição:

♣ **Proposição 3.11** ... *Seja* $\mathbf{F} : \mathcal{O} \subseteq \mathbb{R}^n \to \mathbb{R}^m$*, uma aplicação de classe* C^∞*, definida num aberto* $\mathcal{O} \subseteq \mathbb{R}^n$*, com* $n \geq m$*.*

Suponhamos que $\mathbf{c} \in \mathbb{R}^m$ *é* **valor regular** *de* $\mathbf{F}$*, (isto é,* $d\mathbf{F}_{\mathbf{x}}$ *é sobrejectiva* $\forall \mathbf{x} \in M \equiv \mathbf{F}^{-1}(\mathbf{c})$*), de tal forma que* $M \equiv \mathbf{F}^{-1}(\mathbf{c})$ *é uma variedade (parametrizada regular) em* $\mathbb{R}^n$*, de dimensão* $k = n - m$ *(ver a figura(**??**)).*

Então $\forall \mathbf{p} \in M$*:*

$$T_{\mathbf{p}}M = \mathbf{Ker}\, d\mathbf{F}_{\mathbf{p}} \tag{3.8.20}$$

• Demonstração...

Seja $\mathbf{V_p} \in T_{\mathbf{p}}M$. Então $\mathbf{V_p} = \dot{\alpha}(0)$ para alguma curva $\alpha : \mathbf{I} \to \mathbb{R}^n$ tal que $\alpha(t) \in M$, $\forall t \in \mathbf{I}$ e $\alpha(0) = \mathbf{p}$.

Portanto, atendendo a que $\mathbf{F} \circ \alpha \equiv \mathbf{c}$ (constante), obtemos, aplicando a regra da cadeia, que:

$$d\mathbf{F}_{\alpha(0)}(\dot{\alpha}(0)) = d\mathbf{F_p}(\mathbf{V_p}) = \mathbf{O}$$

o que significa que $\mathbf{V_p} \in \mathbf{Ker}\, d\mathbf{F_p}$, $\forall \mathbf{V_p} \in T_{\mathbf{p}}M$. Finalmente, atendendo a que as dimensões de $T_{\mathbf{p}}M$ e $\mathbf{Ker}\, d\mathbf{F_p}$ são iguais (a $k = n - m$), obtemos (3.8.20),

♣.

Recordemos que se $\mathbf{F} = (F^1, \cdots, F^m)$, então $\mathbf{c}$ é valor regular de $\mathbf{F}$ se e só se, em cada ponto $\mathbf{p} \in M \equiv \mathbf{F}^{-1}(\mathbf{c})$, os vectores gradiente $\nabla F^1(\mathbf{p})$,...,$\nabla F^m(\mathbf{p})$ são linearmente independentes. A demonstração anterior mostra que estes vectores são ortogonais ao espaço tangente a M em $\mathbf{p} \in M$. Portanto, podemos ainda escrever que:

$$T_{\mathbf{p}}M = < \nabla F^1(\mathbf{p}), \cdots, \nabla F^m(\mathbf{p}) >^{\perp} \tag{3.8.21}$$

Esta observação permite generalizar o método dos multiplicadores de Lagrange, para a pesquisa de extremos condicionados, para a situação em que existem várias restrições.

Assim, temos a seguinte proposição, cuja demonstração é análoga à do teorema dos multiplicadores de Lagrange I:

♣ **Proposição 3.12** *(***Teorema dos multiplicadores de Lagrange II***)*

... Seja $\mathbf{G} : \mathcal{O} \subseteq \mathbb{R}^n \to \mathbb{R}^m$*, uma aplicação de classe* C^∞*, definida num aberto* $\mathcal{O} \subseteq \mathbb{R}^n$*, com* $n \geq m$*.*

Suponhamos que $\mathbf{c} \in \mathbb{R}^m$ *é* **valor regular** *de* $\mathbf{F}$*, (isto é, em cada ponto* $\mathbf{p} \in M \equiv \mathbf{F}^{-1}(\mathbf{c})$*, os vectores gradiente* $\nabla G_1(\mathbf{p}), \cdots, \nabla G_m(\mathbf{p})$ *são linearmente independentes), de tal forma que* $M \equiv \mathbf{F}^{-1}(\mathbf{c})$ *é uma variedade em* $\mathbb{R}^n$*, de dimensão* $k = n - m$*.*

Seja $f : \mathcal{O} \subseteq \mathbb{R}^n \to \mathbb{R}$ *um campo escalar diferenciável em* $\mathcal{O}$*.*

Se a restrição de f *à variedade* M*,* $f\,|_M$*, tem um máximo ou um mínimo local num ponto* $\mathbf{p}_o \in M$*, então existem um números reais* $\lambda_1, \cdots, \lambda_m$ *tais que:*

$$\nabla f(\mathbf{p}_o) = \lambda_1 \nabla G_1(\mathbf{p}_o) + \cdots + \lambda_m \nabla G_m(\mathbf{p}_o) \tag{3.8.22}$$

$\lambda_1, \cdots, \lambda_m$ *dizem-se* **multiplicadores de Lagrange***.*

Para calcular os pontos $\mathbf{p}_o = (x^1, \cdots, x^n) \in M \subset \mathbb{R}^n$, onde $f\,|_M$, tem um máximo ou um mínimo local, devemos resolver em ordem a $x^1, \cdots, x^n, \lambda_1, \cdots, \lambda_m$, o seguinte sistema de $n+m$ equações escalares:

$$\begin{aligned}
G_1(x^1, \cdots, x^n) &= 0 \\
G_2(x^1, \cdots, x^n) &= 0 \\
&\vdots \\
G_m(x^1, \cdots, x^n) &= 0 \\
\nabla f(x^1, \cdots, x^n) &= \sum_{i=1}^{m} \lambda_i \nabla G_i(x^1, \cdots, x^n)
\end{aligned} \tag{3.8.23}$$

<u>**Exemplo**</u> ...

Suponhamos que se pretende maximizar $f(x, y, z) = x$, no círculo C de intersecção do plano $z = 1$ com a esfera $x^2 + y^2 + z^2 = 4$.

Neste caso, definimos $\mathbf{G} : \mathbb{R}^3 \to \mathbb{R}^2$, através de $\mathbf{G} = (G_1, G_2)$ onde $G_1(x, y, z) = z - 1$ e $G_2(x, y, z) = x^2 + y^2 + z^2 - 4$, de tal forma que $C = \mathbf{G}^{-1}(\mathbf{O})$.

Como $\nabla f = (1, 0, 0)$, $\nabla G_1 = (0, 0, 1)$ e $\nabla G_2 = (2x, 2y, 2z)$, o sistema (3.8.23) consiste das equações seguintes:

$$z = 1 \qquad\qquad x^2 + y^2 + z^2 = 4$$

$$1 = 2\lambda_2\, x \quad\; 0 = 2\lambda_2\, y \quad\; 0 = \lambda_1 + 2\lambda_2\, z$$

Obtemos então duas soluções $(\pm\sqrt{3}, 0, 1)$, e o máximo é $\sqrt{3}$ e o mínimo $-\sqrt{3}$.

3.9 Métricas Riemannianas

Comecemos por recordar que um **produto interno** num espaço vectorial real $\mathcal{V}$, é uma aplicação:

$$\mathbf{g} : \mathcal{V} \times \mathcal{V} \to \mathbb{R} \tag{3.9.1}$$

que verifica as condições seguintes:

(i)... $\mathbf{g}$ é simétrica:

$$\mathbf{g}(\mathbf{V}, \mathbf{W}) = \mathbf{g}(\mathbf{W}, \mathbf{V}) \qquad \forall \mathbf{V}, \mathbf{W} \in \mathcal{V} \tag{3.9.2}$$

(ii)... $\mathbf{g}$ é bilinear:

$$\begin{aligned}
\mathbf{g}(\mathbf{U} + \mathbf{V}, \mathbf{W}) &= \mathbf{g}(\mathbf{U}, \mathbf{W}) + \mathbf{g}(\mathbf{V}, \mathbf{W}) & (3.9.3)\\
\mathbf{g}(\mathbf{U}, \mathbf{V} + \mathbf{W}) &= \mathbf{g}(\mathbf{U}, \mathbf{V}) + \mathbf{g}(\mathbf{V}, \mathbf{W}) & (3.9.4)\\
\mathbf{g}(\lambda\, \mathbf{U}, \mathbf{V}) &= \lambda\, \mathbf{g}(\mathbf{U}, \mathbf{V}) & (3.9.5)\\
\mathbf{g}(\mathbf{U}, \lambda\, \mathbf{V}) &= \lambda\, \mathbf{g}(\mathbf{U}, \mathbf{V}) & (3.9.6)
\end{aligned}$$

(iii)... $\mathbf{g}$ é não degenerada e definida positiva:

$$\mathbf{g}(\mathbf{U}, \mathbf{U}) \geq 0 \qquad e \qquad \mathbf{g}(\mathbf{U}, \mathbf{U}) = 0 \Leftrightarrow \mathbf{U} = \mathbf{O} \tag{3.9.7}$$

♣ **Definição 3.11** ... *Seja M uma variedade (parametrizada regular) de dimensão k, em $\mathbb{R}^n$. Uma* **métrica riemanniana** *em M, é uma aplicação $\mathbf{g}$, que a cada ponto $\mathbf{p} \in M$, associa um produto interno $\mathbf{g_p}$ no espaço tangente $T_\mathbf{p}M$, e que varia diferenciàvelmente com $\mathbf{p}$, no sentido seguinte:*

se $\Phi : \mathcal{W} \subset \mathbb{R}^k \to \mathbb{R}^n$ é uma parametrização local de M, em torno de $\mathbf{p} \in M$, consideremos a base de $T_\mathbf{p}M$, associada à parametrização Φ:

$$\Phi_{u^i}(\mathbf{q}) = \frac{\partial \Phi}{\partial u^i}(\mathbf{u}) \qquad i = 1, \cdots, k$$

onde $\Phi(\mathbf{u}) = \Phi(u^1, \cdots, u^k) = \mathbf{q} \in M$.

Definamos então as funções $g_{ij} : \mathcal{W} \to \mathbb{R}$, através de:

$$\boxed{g_{ij}(u^1, \cdots, u^k) \equiv \mathbf{g_q}(\Phi_{u^i}(\mathbf{q}), \Phi_{u^j}(\mathbf{q}))} \tag{3.9.8}$$

Exige-se então que estas funções sejam de classe C^∞.

As funções definidas por (3.9.8), dizem-se os **coeficientes da métrica g**, na parametrização Φ.

A **g** dá-se por vezes o nome de **tensor métrico** ou ainda **I forma fundamental**. É usual utilizar a notação seguinte (cujo significado analisaremos em breve), para a expressão local de **g**, nas coordenadas locais $(u^1, \cdots, u^k)$:

$$\boxed{ds^2 \equiv \mathbf{g}(u^1, \cdots, u^k) = \textstyle\sum_{i,j=1}^k g_{ij}(u^1, \cdots, u^k)\, du^i\, du^j} \tag{3.9.9}$$

onde as funções g_{ij} são dadas por (3.9.8).

Uma situação particularmente importante, é a seguinte. Seja M uma variedade (parametrizada regular) de dimensão k, em $\mathbb{R}^n$. Como sabemos, o espaço tangente $T_\mathbf{p}M$, em cada ponto $\mathbf{p} \in M$, é um subespaço vectorial de $\mathbb{R}^n$. Definamos então um produto interno em cada $T_\mathbf{p}M$, restringindo a $T_\mathbf{p}M$ o produto interno usual em $\mathbb{R}^n$, isto é:

$$\mathbf{g_p}(\mathbf{U_p}, \mathbf{V_p}) \equiv \mathbf{U}_p \cdot \mathbf{V_p} \qquad \forall \mathbf{U_p}, \mathbf{V_p} \in T_\mathbf{p}M \subseteq \mathbb{R}^n \tag{3.9.10}$$

Exemplo ...

(i)... A expressão local (3.9.9), para a métrica euclideana usual em $\mathbb{R}^2$, em coordenadas polares, tem o aspecto seguinte:

$$\boxed{ds^2 \equiv \mathbf{g}(r, \theta) = dr^2 + r^2\, d\theta^2} \tag{3.9.11}$$

(ii)... A expressão local (3.9.9), para a métrica euclideana usual em $\mathbb{R}^3$, em coordenadas esféricas, tem o aspecto seguinte:

$$\boxed{ds^2 \equiv \mathbf{g}(r, \theta, \varphi) = dr^2 + r^2\, d\theta^2 + r^2 \sin^2\theta\, d\varphi^2} \tag{3.9.12}$$

(ii)... A expressão local (3.9.9), para a métrica euclideana usual em $\mathbb{R}^3$, em coordenadas cilindricas, tem o aspecto seguinte:

$$\boxed{ds^2 \equiv \mathbf{g}(r, \theta, z) = dr^2 + r^2\, d\theta^2 + dz^2} \tag{3.9.13}$$

Note que nestes exemplos $g_{ij} = 0$ para $i \neq j$, isto é, os vectores das base para os espaços tangentes considerados, são ortogonais entre si. Neste caso diz-se que as coordenadas locais são ortogonais.

Quando M é uma superfície em $\mathbb{R}^3$, e:

$$\Phi : \mathcal{W} \subset \mathbb{R}^2_{(u,v)} \to \mathbb{R}^3$$

é uma parametrização local de M, os coeficientes da métrica definida por (3.9.10), são dados por:

$$\begin{aligned} E(u,v) &\equiv g_{11}(u,v) = \Phi_u(u,v) \cdot \Phi_u(u,v) \\ F(u,v) &\equiv g_{12}(u,v) = \Phi_u(u,v) \cdot \Phi_v(u,v) \\ G(u,v) &\equiv g_{22}(u,v) = \Phi_v(u,v) \cdot \Phi_v(u,v) \end{aligned} \tag{3.9.14}$$

e são funções diferenciáveis em $\mathcal{W}$, com $E > 0$, $G > 0$ e ainda $EG - F^2 > 0$. É usual escrever a expressão local da métrica $\mathbf{g}$, nas coordenadas locais (u,v), com a seguinte notação:

$$\mathbf{g}\,|_{\mathcal{W}} = ds^2 = E\,du^2 + 2\,F\,du\,dv + G\,dv^2 \tag{3.9.15}$$

Exemplo ...

Consideremos a esfera de raio 1, $S^2 \subset \mathbb{R}^3$, e a parametrização local em coordenadas "esféricas" (ou "geográficas"):

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

Como vimos antes, a base para o espaço tangente $T_{\mathbf{p}}S^2$, num ponto $\mathbf{p} = \Phi(\theta,\varphi) \in S^2$, associada à parametrização local Φ, é constituída pelos dois vectores seguintes:

$$\begin{aligned} \Phi_\theta &\equiv \frac{\partial\Phi}{\partial\theta}(\theta,\varphi) = (\cos\theta\,\cos\varphi, \cos\theta\,\sin\varphi, -\sin\theta) \\ \Phi_\varphi &\equiv \frac{\partial\Phi}{\partial\varphi}(\theta,\varphi) = (-\cos\theta\,\sin\varphi, \sin\theta\,\cos\varphi, 0) \end{aligned} \tag{3.9.16}$$

e portanto, os coeficientes da métrica usual em S^2, nestas coordenadas esféricas, são:

$$\begin{aligned} E(\theta,\varphi) &= \Phi_\theta \cdot \Phi_\theta = \cos^2\theta\,\cos^2\varphi + \cos^2\theta\,\sin^2\varphi + \sin^2\theta = 1 \\ F(\theta,\varphi) &= \Phi_\theta \cdot \Phi_\varphi = 0 \\ G(\theta,\varphi) &= \Phi_\varphi \cdot \Phi_\varphi = \sin^2\varphi \end{aligned} \tag{3.9.17}$$

e a expressão local da métrica $\mathbf{g}$, nas coordenadas locais (θ,φ), é:

$$\begin{aligned} \mathbf{g}\,|_{\mathcal{W}} = ds^2 &= E\,d\theta^2 + 2\,F\,d\theta\,d\varphi + G\,d\varphi^2 \\ &= d\theta^2 + \sin^2\theta\,d\varphi^2 \end{aligned} \tag{3.9.18}$$

Portanto, se $\mathbf{V_p}$ é um vector tangente à esfera, num ponto $\mathbf{p} = \Phi(\theta,\varphi)$, cujas coordenadas na base $\{\Phi_\theta, \Phi_\varphi\}$ de $T_{\mathbf{p}}(S^2)$, são:

$$\mathbf{V}_p = a\,\Phi_\theta + b\,\Phi_\varphi$$

então o quadrado do comprimento de $\mathbf{V_p}$ é igual a:

$$\begin{aligned} \|\mathbf{V_p}\|^2 &= E(\theta,\varphi)\,a^2 + 2F(\theta,\varphi)\,ab + G(\theta,\varphi)\,b^2 \\ &= a^2 + b^2\,\sin^2\theta \end{aligned} \tag{3.9.19}$$

Consideremos de novo uma variedade M, de dimensão k em $\mathbb{R}^n$, e uma curva $\alpha : [a, b] \to \mathbb{R}^n$ de classe C^1 por pedaços.

Suponhamos que $\alpha(t) \in M$, $\forall t \in [a, b]$, e que M está munida de uma métrica riemanniana $\mathbf{g}$.

Nestas condições define-se o **comprimento de** α, através de:

$$l(\alpha) \equiv \int_a^b \sqrt{\mathbf{g}_{\alpha(t)}(\dot{\alpha}(t), \dot{\alpha}(t))}\, dt \tag{3.9.20}$$

Suponhamos que $\Phi : \mathcal{W} \subset \mathbb{R}^k \to \mathbb{R}^n$ é uma parametrização local de M , tal que $\alpha([a, b]) \subset \Phi(\mathcal{W})$, e que:

$$\alpha(t) = \Phi(\beta(t)) = \Phi(u^1(t), \cdots, u^k(t)) \tag{3.9.21}$$

isto é, $\beta(t) = \Phi^{-1}(\alpha(t)) = (u^1(t), \cdots, u^k(t))$ é a expressão local da curva α, nas coordenadas locais $u^1, \cdots, u^k$.

Temos então que (pela regra da cadeia):

$$\begin{aligned} \dot{\alpha}(t) &= \sum_{i=1}^{k} \frac{du^i}{dt} \frac{\partial \Phi}{\partial u^i}(\alpha(t)) \\ &= \sum_{i=1}^{k} \frac{du^i}{dt}\, \Phi_{u^i}(\alpha(t)) \qquad \in T_{\alpha(t)}M \end{aligned} \tag{3.9.22}$$

e portanto:

$$\begin{aligned} \mathbf{g}_{\alpha(t)}(\dot{\alpha}(t), \dot{\alpha}(t)) &= \mathbf{g}_{\alpha(t)}\Big(\sum_{i=1}^{k} \frac{du^i}{dt}\, \Phi_{u^i}(\alpha(t)), \sum_{i=1}^{k} \frac{du^i}{dt}\, \Phi_{u^i}(\alpha(t)) \Big) \\ &= \sum_{i,j=1}^{k} \frac{du^i}{dt} \frac{du^j}{dt} \mathbf{g}_{\alpha(t)}\Big(\Phi_{u^i}(\alpha(t)), \Phi_{u^j}(\alpha(t)) \Big) \\ &= \sum_{i,j=1}^{k} g_{ij}(u^1(t), \cdots, u^k(t))\, \frac{du^i}{dt} \frac{du^j}{dt} \end{aligned} \tag{3.9.23}$$

Para uma curva α nas condições indicadas, define-se a função **comprimento de arco**, $s : [a, b] \to \mathbb{R}$, através de:

$$s(t) \equiv \int_a^t \sqrt{\mathbf{g}_{\alpha(t)}(\dot{\alpha}(t), \dot{\alpha}(t))}\, dt \tag{3.9.24}$$

Em coordenadas locais é dada por (atendendo a (3.9.23):

$$s(t) \equiv \int_a^t \sqrt{\sum_{i,j=1}^{k} g_{ij}(u^1(t), \cdots, u^k(t))\, \frac{du^i}{dt} \frac{du^j}{dt}}\, dt \tag{3.9.25}$$

Esta última expressão conduz à notação frequentemente utilizada para o tensor métrico $\mathbf{g}$, e que já referimos antes:

$$ds^2 = \sum_{i,j=1}^{k} g_{ij}(u^1, \cdots, u^k)\, du^i\, du^j$$

Exemplo ...

Consideremos a curva α na esfera S^2, cuja expressão local em coordenadas geográficas é:

$$\beta : t \to (\theta(t) = \frac{\pi}{2} - t, \varphi(t) = \log\cot(\frac{\pi}{4} - \frac{t}{2}))$$

com $t \in [0, \frac{\pi}{2}]$, e calculemos o seu comprimento (supondo S^2 munida da métrica riemanniana usual).

Atendendo a que a expressão local da métrica usual $\mathbf{g}$, nas coordenadas locais (θ, φ), é:

$$ds^2 = d\theta^2 + \sin^2\theta\, d\varphi^2$$

e como:

$$\frac{d\theta}{dt} = -1$$

e ainda:

$$\frac{d\varphi}{dt} = \frac{\operatorname{cosec}^2(\frac{\pi}{4} - \frac{t}{2})}{2\cot(\frac{\pi}{4} - \frac{t}{2})} = \frac{1}{\sin(\frac{\pi}{2} - t)}$$

vem que:

$$\begin{aligned}
l(\alpha) &\equiv \int_a^b \sqrt{\sum_{i,j=1}^k g_{ij}(u^1(t), \cdots, u^k(t))\, \frac{du^i}{dt}\, \frac{du^j}{dt}}\, dt \\
&= \int_0^{\frac{\pi}{2}} \sqrt{E(\theta(t), \varphi(t))(\frac{d\theta}{dt})^2 + 2F(\theta(t), \varphi(t))\frac{d\theta}{dt}\frac{d\varphi}{dt} + G(\theta(t), \varphi(t))(\frac{d\varphi}{dt})^2}\, dt \\
&= \int_0^{\frac{\pi}{2}} \sqrt{1 + \frac{\sin^2\theta(t)}{\sin^2(\frac{\pi}{2} - t)}}\, dt \\
&= \int_0^{\frac{\pi}{2}} \sqrt{2}\, dt \\
&= \frac{\pi}{\sqrt{2}}
\end{aligned}$$

Notemos ainda que a curva α, intersecta os paralelos $\theta \equiv c$ (constante), segundo um ângulo constante (α diz-se uma curva loxodrómica). De facto, designando esse ângulo por ϕ, temos que:

$$\begin{aligned}
\cos\phi &= \frac{\left(\frac{d\theta}{dt}\Phi_\theta + \frac{d\varphi}{dt}\Phi_\varphi\right) \cdot \Phi_\varphi}{\|\frac{d\theta}{dt}\Phi_\theta + \frac{d\varphi}{dt}\Phi_\varphi\|\|\Phi_\varphi\|} \\
&= \frac{\frac{d\theta}{dt}F + \frac{d\theta}{dt}G}{\|\dot{\alpha}\|\sqrt{G}} \\
&= \frac{\frac{1}{\sin(\frac{\pi}{2}-t)}\sin^2(\frac{\pi}{2} - t)}{\sqrt{2}\sin(\frac{\pi}{2} - t)} \\
&= \frac{1}{\sqrt{2}}
\end{aligned}$$

3.10 Divergência div $= \nabla\cdot$. Rotacional rot $= \nabla\times$. Laplaciano $\Delta = \nabla^2$.

3.10.1 Operador ∇

Seja $\mathcal{U}$ um aberto de $\mathbb{R}^n$, $C^\infty(\mathcal{U})$ o conjunto de todas as funções reais de classe C^∞, definidas em $\mathcal{U}$, e $\mathcal{X}^\infty(\mathcal{U})$ o conjunto de todos os campos de vectores de classe C^∞ definidos em $\mathcal{U}$.

Dada uma função $f \in C^\infty(\mathcal{U})$, definimos já o respectivo gradiente, como sendo o campo de vectores $\nabla f \in \mathcal{X}^\infty(\mathcal{U})$, que em coordenadas cartesianas se exprime na forma:

$$\nabla f(\mathbf{p}) = \Big(\frac{\partial f}{\partial x^1}(\mathbf{p}), \frac{\partial f}{\partial x^2}(\mathbf{p}), ..., \frac{\partial f}{\partial x^n}(\mathbf{p})\Big) \tag{3.10.1}$$

$$= \sum_{i=1}^n \frac{\partial f}{\partial x^i}(\mathbf{p})\, \mathbf{e}_i \tag{3.10.2}$$

Convem utilizar a seguinte notação, omitindo referência ao ponto $\mathbf{p}$, e colocando os vectores da base canónica de $\mathbb{R}^n$, à esquerda das derivadas parciais:

$$\nabla f = \sum_{i=1}^n \mathbf{e}_i \frac{\partial f}{\partial x^i} \tag{3.10.3}$$

de tal forma que ∇ aparece como um operador (diferencial):

$$\nabla : C^\infty(\mathcal{U}) \longrightarrow \mathcal{X}^\infty(\mathcal{U}) \tag{3.10.4}$$

que em coordenadas cartesianas se exprime na forma:

$$\nabla = \sum_{i=1}^n \mathbf{e}_i \frac{\partial}{\partial x^i} \tag{3.10.5}$$

As propriedades seguintes são de verificação imediata:

(i)... ∇ é um operador linear, i.e.:

$$\nabla(f+g) = \nabla f + \nabla g \tag{3.10.6}$$

$$\nabla(\lambda f) = \lambda(\nabla f) \qquad \forall \lambda \in \mathbb{R} \tag{3.10.7}$$

(ii)... ∇ satisfaz a **regra de Leibniz**:

$$\nabla(fg) = f\nabla g + g\nabla f \tag{3.10.8}$$

A proposição seguinte, refere outras propriedades do operador ∇, e demonstra-se utilizando a regra da cadeia.

♣ **Proposição 3.13** ... *(i). Seja $f \in \mathcal{X}^\infty(\mathcal{U})$ e $h \in \mathcal{X}^\infty(\mathbb{R})$. Então:*

$$\nabla(h \circ f)(\mathbf{x}) = h'(f(\mathbf{x}))\, \nabla f(\mathbf{x}) \tag{3.10.9}$$

(ii). Seja:

$$\Phi : \mathcal{V} \subseteq \mathbb{R}^n_{(u^1,\cdots,u^n)} \to \mathcal{U} \subseteq \mathbb{R}^m_{(x^1,\cdots,x^m)}$$

uma aplicação de classe C^∞, e $f \in C^\infty(\mathcal{U})$.

Então, pondo $\Phi = (\phi^1, \cdots, \phi^m)$, tem-se que:

$$\nabla(f \circ \Phi)(\mathbf{u}) = \sum_{k=1}^m \frac{\partial f}{\partial x^k}(\Phi(\mathbf{u}))\, \nabla\phi^k(\mathbf{u}) \tag{3.10.10}$$

- Demonstração...

(ii)... Tem-se sucessivamente que:

$$\begin{aligned}
\nabla(f \circ \Phi)(\mathbf{u}) &= \sum_{i=1}^n \mathbf{e}_i\, \frac{\partial(f\circ\Phi)}{\partial u^i}(\mathbf{u}) \\
&= \sum_{i=1}^n \mathbf{e}_i \sum_{k=1}^m \frac{\partial f}{\partial x^k}(\Phi(\mathbf{u}))\, \frac{\partial \phi^k}{\partial u^i}(\mathbf{u}) \\
&= \sum_{k=1}^m \frac{\partial f}{\partial x^k}(\Phi(\mathbf{u})) \sum_{i=1}^n \mathbf{e}_i\, \frac{\partial \phi^k}{\partial u^i}(\mathbf{u}) \\
&= \sum_{k=1}^m \frac{\partial f}{\partial x^k}(\Phi(\mathbf{u}))\, \nabla\phi^k(\mathbf{u})
\end{aligned}$$

♣.

3.10.2 Divergência div $= \nabla\cdot$

Seja $\mathcal{U}$ um aberto de $\mathbb{R}^n$, $C^\infty(\mathcal{U})$ o conjunto de todas as funções reais de classe C^∞, definidas em $\mathcal{U}$, e $\mathcal{X}^\infty(\mathcal{U})$ o conjunto de todos os campos de vectores de classe C^∞ definidos em $\mathcal{U}$.

Define-se o operador divergência:

$$\textbf{div} : \mathcal{X}^\infty(\mathcal{U}) \longrightarrow C^\infty(\mathcal{U}) \tag{3.10.11}$$

como sendo o operador que a cada campo de vectores $\mathbf{F} \in \mathcal{X}^\infty(\mathcal{U})$ asssocia a função que em coordenadas cartesianas é dada por:

$$\textbf{div}\, \mathbf{F}(\mathbf{x}) = \frac{\partial F^1}{\partial x^1}(\mathbf{x}) + \frac{\partial F^2}{\partial x^2}(\mathbf{x}) + \cdots + \frac{\partial F^n}{\partial x^n}(\mathbf{x}) \tag{3.10.12}$$

onde pusemos $\mathbf{F} = (F^1, \cdots, F^n)$.

É conveniente escrever o operador divergência formalmente na forma:

$$\textbf{div} = \nabla\cdot \tag{3.10.13}$$

onde $\cdot$ designa o produto interno usual em $\mathbb{R}^n$.

A proposição seguinte refere algumas propriedades da divergência de um campo de vectores, e é de fácil demonstração:

♣ **Proposição 3.14** ... *Sejam* $\mathbf{F}, \mathbf{G} \in \mathcal{X}^\infty(\mathcal{U})$, *e* $f \in C^\infty(\mathcal{U})$. *Então:*

(i)... $\nabla\cdot(\mathbf{F}+\mathbf{G}) = \nabla\cdot\mathbf{F} + \nabla\cdot\mathbf{G}$

(ii)... $\nabla\cdot(\lambda\,\mathbf{F}) = \lambda\,\nabla\cdot\mathbf{F} \qquad \lambda\in\mathbb{R}$

(iii)... $\nabla\cdot(f\,\mathbf{F}) = f(\nabla\cdot\mathbf{F}) + (\nabla f)\cdot\mathbf{F}$

3.10.3 Rotacional rot $= \nabla\times$, em $\mathbf{R}^3$

Seja $\mathcal{U}$ um aberto de $\mathbb{R}^3$, $C^\infty(\mathcal{U})$ o conjunto de todas as funções reais de classe C^∞, definidas em $\mathcal{U}$, e $\mathcal{X}^\infty(\mathcal{U})$ o conjunto de todos os campos de vectores de classe C^∞ definidos em $\mathcal{U}$.

Define-se o operador rotacional:

$$\textbf{rot} : \mathcal{X}^\infty(\mathcal{U}) \longrightarrow \mathcal{X}^\infty(\mathcal{U}) \tag{3.10.14}$$

como sendo o operador que a cada campo de vectores $\mathbf{F}\in\mathcal{X}^\infty(\mathcal{U})$ asssocia um outro campo de vectores, que em coordenadas cartesianas é obtido desenvolvendo segundo a primeira linha, o determinante formal:

$$\begin{aligned}\textbf{rot}\,\mathbf{F} &= \nabla\times\mathbf{F} = \det\begin{bmatrix}\mathbf{e}_1 & \mathbf{e}_2 & \mathbf{e}_3\\ \frac{\partial}{\partial x^1} & \frac{\partial}{\partial x^2} & \frac{\partial}{\partial x^3}\\ F^1 & F^2 & F^3\end{bmatrix}\\ &= (\frac{\partial F^3}{\partial x^2} - \frac{\partial F^2}{\partial x^3})\mathbf{e}_1 - (\frac{\partial F^3}{\partial x^1} - \frac{\partial F^1}{\partial x^3})\mathbf{e}_2 + (\frac{\partial F^2}{\partial x^1} - \frac{\partial F^1}{\partial x^2})\mathbf{e}_3 \end{aligned} \tag{3.10.15}$$

onde pusemos $\mathbf{F} = (F^1, F^2, F^3)$.

A proposição seguinte refere algumas propriedades do rotacional de um campo de vectores, e é de fácil demonstração:

♣ **Proposição 3.15** ... *Sejam* $\mathbf{F}, \mathbf{G} \in \mathcal{X}^\infty(\mathcal{U})$, *e* $f \in C^\infty(\mathcal{U})$, *onde* $\mathcal{U}$ *é um aberto de* $\mathbb{R}^3$. *Então:*

(i)... $\nabla\times(\mathbf{F}+\mathbf{G}) = \nabla\times\mathbf{F} + \nabla\times\mathbf{G}$

(ii)... $\nabla\times(\lambda\,\mathbf{F}) = \lambda\,\nabla\times\mathbf{F} \qquad \lambda\in\mathbb{R}$

(iii)... $\nabla\times(f\,\mathbf{F}) = f(\nabla\times\mathbf{F}) + (\nabla f)\times\mathbf{F}$

A proposição seguinte refere algumas propriedades quando se combinam os operadores atrás definidos:

♣ **Proposição 3.16** ... *Sejam* $\mathbf{F}, \mathbf{G} \in \mathcal{X}^\infty(\mathcal{U})$, *e* $f \in C^\infty(\mathcal{U})$, *onde* $\mathcal{U}$ *é um aberto de* $\mathbb{R}^3$. *Então:*

(i)... $\nabla \cdot (\mathbf{F} \times \mathbf{G}) = \mathbf{G} \cdot (\nabla \times \mathbf{F}) - \mathbf{F} \cdot (\nabla \times \mathbf{G})$

(ii)... $\nabla \times (\mathbf{F} \times \mathbf{G}) = \mathbf{F}\,(\nabla \cdot \mathbf{G}) + (\mathbf{G} \cdot \nabla)\,\mathbf{F} - \mathbf{G}\,(\nabla \cdot \mathbf{F}) - (\mathbf{F} \cdot \nabla)\,\mathbf{G}$

(iii)... $\nabla(\mathbf{F} \cdot \mathbf{G}) = (\mathbf{F} \cdot \nabla)\mathbf{G} + (\mathbf{G} \cdot \nabla)\mathbf{F} + \mathbf{F} \times (\nabla \times \mathbf{G}) + \mathbf{G} \times (\nabla \times \mathbf{F})$

(iv)... **O rotacional de um campo de gradientes é nulo**:

$$\nabla \times (\nabla f) = 0$$

(v)... **A divergência de um campo rotacional é nula**:

$$\nabla \cdot (\nabla \times \mathbf{F}) = 0$$

3.10.4 Interpretação geométrica da divergência e rotacional de um campo de vectores em $\mathbf{R}^3$

Consideremos um campo de vectores $\mathbf{F} \in \mathcal{X}^\infty(\mathcal{U})$, onde $\mathcal{U}$ é um aberto de $\mathbb{R}^3$.

Fixemos um ponto $\mathbf{p} \in \mathcal{U}$, e consideremos a diferencial de $\mathbf{F}$ em $\mathbf{p}$, e a respectiva matriz jacobiana:

$$\mathbf{J_p} = \mathbf{JacF}(\mathbf{x}) = \begin{bmatrix} \frac{\partial F^1}{\partial x^1}(\mathbf{p}) & \frac{\partial F^1}{\partial x^2}(\mathbf{p}) & \frac{\partial F^1}{\partial x^3}(\mathbf{p}) \\ \frac{\partial F^2}{\partial x^1}(\mathbf{p}) & \frac{\partial F^2}{\partial x^2}(\mathbf{p}) & \frac{\partial F^2}{\partial x^3}(\mathbf{p}) \\ \frac{\partial F^3}{\partial x^1}(\mathbf{p}) & \frac{\partial F^3}{\partial x^2}(\mathbf{p}) & \frac{\partial F^3}{\partial x^3}(\mathbf{p}) \end{bmatrix} \tag{3.10.16}$$

Definamos agora os endomorfismos de $\mathbb{R}^3$ cujas matrizes (na base canónica usual de $\mathbb{R}^3$) são respectivamente:

$$\mathbf{S_p} = \frac{1}{2}(\mathbf{J_p} + \mathbf{J_p}^{Tr}) \tag{3.10.17}$$

e:

$$\mathbf{A_p} = \frac{1}{2}(\mathbf{J_p} - \mathbf{J_p}^{Tr}) \tag{3.10.18}$$

É claro que $\mathbf{S_p}$ é simétrica, isto é, $\mathbf{S_p} = \mathbf{S_p}^{Tr}$, enquanto que $\mathbf{A_p}$ é antissimétrica, i.e., $\mathbf{A_p} = -\mathbf{A_p}^{Tr}$. Além disso:

$$\mathbf{JacF}(\mathbf{p}) = \mathbf{S_p} + \mathbf{A_p} \tag{3.10.19}$$

Um cálculo rápido mostra que:

$$\mathbf{A_p} = \frac{1}{2} \begin{bmatrix} 0 & \frac{\partial F^1}{\partial x^2}(\mathbf{p}) - \frac{\partial F^2}{\partial x^1}(\mathbf{p}) & \frac{\partial F^1}{\partial x^3}(\mathbf{p}) - \frac{\partial F^3}{\partial x^1}(\mathbf{p}) \\ -\frac{\partial F^1}{\partial x^2}(\mathbf{p}) + \frac{\partial F^2}{\partial x^1}(\mathbf{p}) & 0 & \frac{\partial F^2}{\partial x^3}(\mathbf{p}) - \frac{\partial F^3}{\partial x^2}(\mathbf{p}) \\ \frac{\partial F^3}{\partial x^1}(\mathbf{p}) - \frac{\partial F^1}{\partial x^3}(\mathbf{p}) & \frac{\partial F^3}{\partial x^2}(\mathbf{p}) - \frac{\partial F^2}{\partial x^3}(\mathbf{p}) & 0 \end{bmatrix} \tag{3.10.20}$$

Se $\mathbf{h} = (h_1, h_2, h_3) \in \mathbb{R}^3$, então um cálculo simples mostra que:

$$\mathbf{A_p}\begin{bmatrix} h_1 \\ h_2 \\ h_3 \end{bmatrix} = \frac{1}{2}\,\mathbf{rot}\,\mathbf{F}(\mathbf{p}) \times \mathbf{h} \tag{3.10.21}$$

que representa uma "*rotação infinitesimal*" em torno do eixo $\mathbf{rot}\,\mathbf{F}(\mathbf{p})$ (veja a secção 3.1).

Quanto ao endomorfismo simétrico definido pela matriz $\mathbf{S_p}$, também chamado o **tensor das deformações**, sabemos já que é possível escolher uma base ortonormada de $\mathbb{R}^3$, $\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$, tal que $\mathbf{S_p}(\mathbf{u}_i) = d_i(\mathbf{p})\,\mathbf{u}_i$ $(i = 1, 2, 3)$, e portanto na qual a matriz de $\mathbf{S_p}$ é diagonal:

$$\mathbf{S_p} = \begin{bmatrix} d_1(\mathbf{p}) & 0 & 0 \\ 0 & d_2(\mathbf{p}) & 0 \\ 0 & 0 & d_3(\mathbf{p}) \end{bmatrix} \tag{3.10.22}$$

Note que:

$$\begin{aligned} \textbf{Traço}\,\mathbf{JacF}(\mathbf{p}) &= \textbf{Traço}\,\mathbf{S_p} \\ &= d_1(\mathbf{p}) + d_2(\mathbf{p}) + d_3(\mathbf{p}) \\ &= \mathbf{div}\,\mathbf{F}(\mathbf{p}) \end{aligned} \tag{3.10.23}$$

Como se viu na secção 3.1, $\mathbf{div}\,\mathbf{F_p} = d_1(\mathbf{p}) + d_2(\mathbf{p}) + d_3(\mathbf{p}) = \frac{V'(0)}{V(0)}$, representa a taxa de variação do volume de um paralelipípedo de lados paralelos aos vectores $\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$.

Quando $\mathbf{div}\,\mathbf{F}(\mathbf{p}) = 0$ o volume do referido paralelipípedo permanece portanto constante, e o campo de vectores $\mathbf{F}$ diz-se **incompressível** (em $\mathbf{p}$).

3.10.5 Expressões em coordenadas curvilíneas ortogonais

Seja:

$$\Phi : \mathcal{U} \subseteq \mathbb{R}^3_{(u^1,u^2,u^3)} \longrightarrow \mathcal{V} \subseteq \mathbb{R}^3_{(x^1,x^2,x^3)} \tag{3.10.24}$$

um difeomorfismo de classe C^∞:

$$\mathbf{u} = (u^1, u^2, u^3) \mapsto \Phi(\mathbf{u}) = (\phi^1(u^1, u^2, u^3), \phi^2(u^1, u^2, u^3), \phi^3(u^1, u^2, u^3)) \tag{3.10.25}$$

com inversa $\Psi = \Phi^{-1}$:

$$\Psi : \mathcal{V} \subseteq \mathbb{R}^3_{(x^1,x^2,x^3)} \longrightarrow \mathcal{U} \subseteq \mathbb{R}^3_{(u^1,u^2,u^3)} \tag{3.10.26}$$

com:

$$\mathbf{x} = (x^1, x^2, x^3) \mapsto \Psi(\mathbf{x}) = (\psi^1(x^1, x^2, x^3), \psi^2(x^1, x^2, x^3), \psi^3(x^1, x^2, x^3)) \tag{3.10.27}$$

Como já sabemos, em cada ponto $\mathbf{x} = \Phi(\mathbf{u}) \in \mathcal{V}$, passam as três linhas coordenadas (ver a figura 3.16):

$$u^i \mapsto \Phi(u^1, u^2, u^3) \qquad i = 1, 2, 3$$

Figure 3.16: Linhas coordenadas

.

Vamos supôr nesta secção que essas linhas coordenadas se intersectam sempre ortogonalmente, i.e., que o sistema de coordenadas u^i é um **sistema de coordenadas ortogonal** (veja os exemplos das secções 3.7 e 3.9).

O tensor métrico usual, no sistema de coordenadas u^i, tem então a forma:

$$ds^2 = \Big(h_1(u^1, u^2, u^3)\Big)^2\,(du^1)^2 + \Big(h_2(u^1, u^2, u^3)\Big)^2\,(du^2)^2 + \Big(h_3(u^1, u^2, u^3)\Big)^2\,(du^3)^2 \qquad (3.10.28)$$

onde:

$$\begin{aligned}\Big(h_i(u^1, u^2, u^3)\Big)^2 &= (h_i(\mathbf{u}))^2 = \Phi_{u^i}(\mathbf{u}) \cdot \Phi_{u^i}(\mathbf{u}) \\ &= \frac{\partial \Phi}{\partial u^i}(\mathbf{u}) \cdot \frac{\partial \Phi}{\partial u^i}(\mathbf{u}) \end{aligned} \qquad (3.10.29)$$

para $i = 1, 2, 3$.

<u>**Exemplos**</u> ...

(i)... A expressão local (3.10.28), para a métrica euclideana usual em $\mathbb{R}^3$, em coordenadas esféricas $(u^1, u^2, u^3) = (r, \theta, \varphi)$, tem o aspecto seguinte:

$$ds^2 \equiv \mathbf{g}(r, \theta, \varphi) = dr^2 + r^2\, d\theta^2 + r^2\, \sin^2\theta\, d\varphi^2 \qquad (3.10.30)$$

Portanto aqui:

$$h_1(r, \theta, \varphi) \equiv 1 \qquad h_2(r, \theta, \varphi) = r \qquad h_3(r, \theta, \varphi) = r\, \sin\theta$$

(ii)... A expressão local (3.10.28), para a métrica euclideana usual em $\mathbb{R}^3$, em coordenadas cilindricas $(u^1, u^2, u^3) = (r, \theta, z)$, tem o aspecto seguinte:

$$ds^2 \equiv \mathbf{g}(r, \theta, z) = dr^2 + r^2\, d\theta^2 + dz^2 \qquad (3.10.31)$$

Portanto aqui:

$$h_1(r, \theta, z) \equiv 1 \qquad h_2(r, \theta, z) = r \qquad h_3(r, \theta, z) \equiv 1$$

Aplicando o teorema da função inversa, é possível deduzir que em cada ponto $\mathbf{x} = \Phi(\mathbf{u})$, se tem:

$$\begin{aligned} \nabla \psi^i(\mathbf{x}) &= \nabla \psi^i(\Phi(\mathbf{u})) \\ &= \frac{1}{h_i^2}\,\Phi_{u^i}(\mathbf{u}) \\ &= \frac{1}{h_i}\,\mathbf{e}_{u^i} \end{aligned} \tag{3.10.32}$$

para cada $i = 1, 2, 3$, e onde pusemos:

$$\mathbf{e}_{u^i} \equiv \frac{\Phi_{u^i}}{\|\Phi_{u^i}\|} = \frac{1}{h_i}\,\Phi_{u^i}(\mathbf{u})$$

Destas equações deduzimos ainda que:

$$\mathbf{e}_{u^1} \cdot (\mathbf{e}_{u^2} \times \mathbf{e}_{u^3}) = h_1 h_2 h_3\, \nabla\psi^1 \cdot (\nabla\psi^2 \times \nabla\psi^3) \tag{3.10.33}$$

e portanto:

$$\nabla\psi^1 \cdot (\nabla\psi^2 \times \nabla\psi^3) = \frac{1}{h_1 h_2 h_3} \tag{3.10.34}$$

Após estes preliminares estamos aptos a deduzir as expressões em coordenadas curvilíneas ortogonais, dos operadores considerados na secção anterior.

Assim, com as notações de (3.10.24), (3.10.25) e (3.10.27), suponhamos que:

$$f : \mathcal{V} \subseteq \mathbb{R}^3_{(x^1,x^2,x^3)} \longrightarrow \mathbb{R}$$

é um campo escalar, de tal forma que $f \circ \Phi$ é esse mesmo campo escalar mas agora expresso nas coordenadas u^i. Vejamos qual a expressão do respectivo gradiente.

Aplicando a regra da cadeia temos sucessivamente que:

$$\begin{aligned} \nabla f &= \sum_{i=1}^{3} \Big(\sum_{k=1}^{3} \frac{\partial (f\circ\Phi)}{\partial u^k} \frac{\partial \psi^k}{\partial x^i} \Big)\, \mathbf{e}_i \\ &= \sum_{k=1}^{3} \frac{\partial (f\circ\Phi)}{\partial u^k}\, \nabla\psi^k \end{aligned} \tag{3.10.35}$$

Finalmente, aplicando a equação (3.10.32), vem que:

$$\nabla f = \sum_{k=1}^{3} \frac{1}{h_k} \frac{\partial (f\circ\Phi)}{\partial u^k}\, \mathbf{e}_{u^k} \tag{3.10.36}$$

Como foi dito, $f \circ \Phi$ é o campo escalar f, expresso nas coordenadas u^i. No entanto é usual continuar a notá-lo ainda por f, simplificando as notações nas fórmulas anteriores. Assim por exemplo, a fórmula (3.10.36), escreve-se:

$$\nabla f = \sum_{k=1}^{3} \frac{1}{h_k} \frac{\partial f}{\partial u^k}\, \mathbf{e}_{u^k} \tag{3.10.37}$$

De aqui em diante utilizaremos estas simplificações de notação.

Exemplos ...

(i)... Em coordenadas esféricas em $\mathbb{R}^3$, tem-se que:

$$\nabla f = \frac{\partial f}{\partial r}\,\mathbf{e}_r + \frac{1}{r}\frac{\partial f}{\partial \theta}\,\mathbf{e}_\theta + \frac{1}{r\sin\theta}\frac{\partial f}{\partial \varphi}\,\mathbf{e}_\varphi \tag{3.10.38}$$

(ii)... Em coordenadas cilindricas em $\mathbb{R}^3$, tem-se que:

$$\nabla f = \frac{\partial f}{\partial r}\,\mathbf{e}_r + \frac{1}{r}\frac{\partial f}{\partial \theta}\,\mathbf{e}_\theta + \frac{\partial f}{\partial z}\,\mathbf{e}_z \tag{3.10.39}$$

Suponhamos agora, ainda com as notações de (3.10.24), (3.10.25) e (3.10.27), que:

$$\mathbf{F} : \mathcal{V} \subseteq \mathbb{R}^3_{(x^1,x^2,x^3)} \longrightarrow \mathbb{R}^3$$

é um campo de vectores em $\mathcal{V}$, e que:

$$\mathbf{F} = F_{u^1}\,\mathbf{e}_{u^1} + F_{u^2}\,\mathbf{e}_{u^2} + F_{u^3}\,\mathbf{e}_{u^3} \tag{3.10.40}$$

onde:

$$F_{u^i} = \mathbf{F}\cdot\mathbf{e}_{u^i} \qquad\qquad i = 1,2,3$$

Para deduzir a expressão da divergência $\mathbf{div}\,\mathbf{F} = \nabla\cdot\mathbf{F}$, nas coordenadas u^i, utilizemos as seguintes relações:

$$\begin{array}{rcl} \mathbf{e}_{u^1} &=& h_2h_3\nabla\psi^2\times\nabla\psi^3 \\ \mathbf{e}_{u^2} &=& h_3h_1\nabla\psi^3\times\nabla\psi^1 \\ \mathbf{e}_{u^3} &=& h_1h_2\nabla\psi^1\times\nabla\psi^2 \end{array}$$

substituindo-as em (3.10.40), e actuando a expressão resultante com o operador de divergência $\nabla\cdot$. Obtemos então:

$$\begin{array}{rcl} \nabla\cdot\mathbf{F} &=& \nabla\cdot(F_{u^1}\,h_2h_3\nabla\psi^2\times\nabla\psi^3) + \nabla\cdot(F_{u^2}\,h_3h_1\nabla\psi^3\times\nabla\psi^1) + \\ && \nabla\cdot(F_{u^3}\,h_1h_2\nabla\psi^1\times\nabla\psi^2) \end{array} \tag{3.10.41}$$

Mas, para o primeiro dos três termos do membro direito de (3.10.41), temos que:

$$\begin{array}{l} \nabla\cdot(F_{u^1}\,h_2h_3\nabla\psi^2\times\nabla\psi^3) = \nabla(F_{u^1}\,h_2h_3)\cdot(\nabla\psi^2\times\nabla\psi^3) + \\ \qquad +F_{u^1}\,h_2h_3\nabla\cdot(\psi^2\times\nabla\psi^3) \end{array} \tag{3.10.42}$$

Por outro lado:

$$\nabla(F_{u^1}\,h_2h_3) = \frac{\partial}{\partial u^1}(F_{u^1}\,h_2h_3)\,\nabla\psi^1 + \frac{\partial}{\partial u^2}(F_{u^1}\,h_2h_3)\,\nabla\psi^2 + \frac{\partial}{\partial u^3}(F_{u^1}\,h_2h_3)\,\nabla\psi^3 \tag{3.10.43}$$

e, como $\nabla\psi^1$, $\nabla\psi^2$ e $\nabla\psi^3$ são ortogonais entre si, obtemos:

$$\nabla(F_{u^1}\,h_2h_3)\cdot(\nabla\psi^2\times\nabla\psi^3) = \frac{\partial}{\partial u^1}(F_{u^1}\,h_2h_3)\,\nabla\psi^1\cdot(\nabla\psi^2\times\nabla\psi^3) \tag{3.10.44}$$

ou, por (3.10.34):

$$\nabla(F_{u^1}\, h_2h_3) \cdot (\nabla\psi^2 \times \nabla\psi^3) = \frac{1}{h_1h_2h_3}\frac{\partial}{\partial u^1}(F_{u^1}\, h_2h_3) \tag{3.10.45}$$

Agora:

$$\nabla \cdot (\nabla\psi^2 \times \nabla\psi^3) = \nabla\psi^3 \cdot (\nabla \cdot \nabla\psi^2) - \nabla\psi^2 \cdot (\nabla \times \nabla\psi^3) = 0 \tag{3.10.46}$$

atendendo a que o rotacional de um campo gradiente é nulo.

As equações (3.10.45) e (3.10.46), permitem agora escrever (3.10.42), na forma:

$$\nabla \cdot (F_{u^1}\, h_2h_3\nabla\psi^2 \times \nabla\psi^3) = \frac{1}{h_1h_2h_3}\frac{\partial}{\partial u^1}(F_{u^1}\, h_2h_3) \tag{3.10.47}$$

Esta relação, bem como duas outras semelhantes para os dois termos restantes do membro direito de (3.10.41), permitem escrever (3.10.41) na forma final:

$$\mathbf{div}\,\mathbf{F} = \nabla \cdot \mathbf{F} = \frac{1}{h_1h_2h_3}\Big[\frac{\partial}{\partial u^1}(F_{u^1}\, h_2h_3) + \frac{\partial}{\partial u^2}(F_{u^2}\, h_3h_1) + \frac{\partial}{\partial u^3}(F_{u^3}\, h_1h_2)\Big] \tag{3.10.48}$$

<u>**Exemplos**</u> ...

(i)... Em coordenadas esféricas, se $\mathbf{F} = F_r\,\mathbf{e}_r + F_\theta\,\mathbf{e}_\theta + F_\varphi\,\mathbf{e}_\varphi$, então:

$$\nabla \cdot \mathbf{F} = \frac{1}{r^2}\frac{\partial}{\partial r}(F_r\, r^2) + \frac{1}{r\,\sin\theta}\frac{\partial}{\partial\theta}(F_\theta\,\sin\theta) + \frac{1}{r\,\sin\theta}\frac{\partial F_\varphi}{\partial\varphi} \tag{3.10.49}$$

(ii)... Em coordenadas cilindricas, se $\mathbf{F} = F_r\,\mathbf{e}_r + F_\theta\,\mathbf{e}_\theta + F_z\,\mathbf{e}_z$, então:

$$\nabla \cdot \mathbf{F} = \frac{1}{r}\Big[\frac{\partial}{\partial r}(r\,F_r) + \frac{\partial F_\theta}{\partial\theta} + \frac{\partial}{\partial z}(r\,F_z)\Big] \tag{3.10.50}$$

Um cálculo análogo, permite deduzir a seguinte fórmula para o rotacional $\mathbf{rot}\,\mathbf{F} = \nabla \times \mathbf{F}$, de um campo de vectores:

$$\begin{aligned}\mathbf{rot}\,\mathbf{F} &= \nabla \times \mathbf{F} = \frac{1}{h_2h_3}\Big[\frac{\partial}{\partial u^2}(F_{u^3}\, h_3) - \frac{\partial}{\partial u^3}(F_{u^2}\, h_2)\Big]\mathbf{e}_{u^1}\\ &+ \frac{1}{h_3h_1}\Big[\frac{\partial}{\partial u^3}(F_{u^1}\, h_1) - \frac{\partial}{\partial u^1}(F_{u^3}\, h_3)\Big]\mathbf{e}_{u^2}\\ &+ \frac{1}{h_1h_2}\Big[\frac{\partial}{\partial u^1}(F_{u^2}\, h_2) - \frac{\partial}{\partial u^2}(F_{u^1}\, h_1)\Big]\mathbf{e}_{u^3}\end{aligned} \tag{3.10.51}$$

<u>**Exemplos**</u> ...

(i)... Em coordenadas cilindricas, se $\mathbf{F} = F_r\,\mathbf{e}_r + F_\theta\,\mathbf{e}_\theta + F_z\,\mathbf{e}_z$, então:

$$\nabla \times \mathbf{F} = \frac{1}{r}\det\begin{bmatrix}\mathbf{e}_r & r\,\mathbf{e}_\theta & \mathbf{e}_z\\ \frac{\partial}{\partial r} & \frac{\partial}{\partial\theta} & \frac{\partial}{\partial z}\\ F_r & r\,F_\theta & F_z\end{bmatrix} \tag{3.10.52}$$

(ii)... Em coordenadas esféricas, se $\mathbf{F} = F_r\,\mathbf{e}_r + F_\theta\,\mathbf{e}_\theta + F_\varphi\,\mathbf{e}_\varphi$, então:

$$\begin{aligned}
\nabla\times\mathbf{F} &= \Big[\frac{1}{r\,\sin\theta}\frac{\partial}{\partial\theta}(\sin\theta\,F_\varphi) - \frac{1}{r\,\sin\theta}\frac{\partial F_\theta}{\partial\varphi}\Big]\mathbf{e}_r \\
&+ \Big[\frac{1}{r\,\sin\theta}\frac{\partial F_r}{\partial\varphi} - \frac{1}{r}\frac{\partial}{\partial r}(r\,F_\varphi)\Big]\mathbf{e}_\theta \\
&+ \Big[\frac{1}{r}\frac{\partial}{\partial r}(r\,F_\theta) - \frac{1}{r}\frac{\partial F_r}{\partial\theta}\Big]\mathbf{e}_\varphi
\end{aligned} \tag{3.10.53}$$

3.10.6 O Laplaciano

O **Laplaciano** de uma função $f \in C^\infty(\mathcal{U})$, onde $\mathcal{U}$ é um aberto de $\mathbb{R}^n$, define-se através de:

$$\begin{aligned}
\Delta f &= \mathbf{div\,grad}\, f \\
&= \nabla\cdot\nabla f \\
&= \nabla^2 f
\end{aligned} \tag{3.10.54}$$

As expressões de Δf em vários sistemas de coordenadas são as seguintes:

(i)... em coordenadas cartesianas:

$$\Delta f = \frac{\partial^2 f}{\partial(x^1)^2} + \cdots + \frac{\partial^2 f}{\partial(x^n)^2} \tag{3.10.55}$$

(i)... em coordenadas cilindricas:

$$\Delta f = \frac{\partial^2 f}{\partial r^2} + \frac{1}{r^2}\frac{\partial^2 f}{\partial\theta^2} + \frac{\partial^2 f}{\partial z^2} + \frac{1}{r}\frac{\partial f}{\partial r} \tag{3.10.56}$$

(i)... em coordenadas esféricas:

$$\Delta f = \frac{\partial^2 f}{\partial r^2} + \frac{1}{r^2}\frac{\partial^2 f}{\partial\varphi^2} + \frac{1}{r^2\,\sin^2\varphi}\frac{\partial^2 f}{\partial\theta^2} + \frac{2}{r}\frac{\partial f}{\partial r} + \frac{\cot\varphi}{r^2}\frac{\partial f}{\partial\varphi} \tag{3.10.57}$$

3.11 Exercícios

Exercício 53 ... Relativamente às funções que se seguem, determine a respectiva matriz jacobiana, e defina a diferencial de f, nos pontos indicados:

(i). $\mathbf{F}(x,y,z) = (2x-y, xz, y^2-z^2)$ no ponto $(0,1,1)$.

(ii). $\mathbf{F}(x,y) = (x-y, x+y, x^2+y^2)$ no ponto $(1,-1)$.

(iii). $\mathbf{F}(x,y,z) = (x+e^z+y, yx^2)$ no ponto $(0,0,0)$.

Exercício 54 ... Considere as funções seguintes:

$$\mathbf{F}(x,y,z) = (\sin(xy),\ x^2-y^2) \quad e \quad \mathbf{G}(x,y) = (y,x,y)$$

Verifique a igualdade $\mathbf{Jac}(\mathbf{G}\circ\mathbf{F})(\mathbf{x}) = \mathbf{JacG}(\mathbf{F}(\mathbf{x}))\cdot\mathbf{JacF}(\mathbf{x})$.

Exercício 55 ... Considere as seguintes funções:

$$\mathbf{F}(x,y) = (xy, xy^2, x^2y) \quad e \quad \mathbf{G}(x,y,z) = (x-y, 2, 0, x)$$

(i). Defina $\mathbf{G}\circ\mathbf{F}$.

(ii). Mostre que $\mathbf{G}\circ\mathbf{F}$ é diferenciável em $\mathbb{R}^2$, e verifique a regra da cadeia neste exemplo particular.

Exercício 56 ... Seja $\alpha(t) = (\sin t, e^t\cos t, t^3)$ e $\mathbf{F}(x,y,z) = (xz+1, yz-1)$. Calcule a diferencial de $\mathbf{F}\circ\alpha$, usando a regra da cadeia.

Exercício 57 ... Mostre que se $f:\mathbb{R}\to\mathbb{R}$ é uma função diferenciável, e $h(x,y) = xy+f(x^2+y^2)$, então $\forall(x,y)\in\mathbb{R}^2$, tem-se:

$$y\frac{\partial h}{\partial x}(x,y) - \frac{\partial h}{\partial y}(x,y) = x^2-y^2$$

Exercício 58 ... Seja $\alpha:\mathbf{I}\to\mathbb{R}^2$ uma curva diferenciável , tal que $\alpha(t)$ pertence à hipérbole de equação $xy=1$. Considere as funções$f(x,y) = x^2+y^2$ e $g(x,y) = (x+y)^2$. Mostre que $f\circ\alpha$ e $g\circ\alpha$ têm a mesma derivada em qualquer ponto de $\mathbf{I}$.

Exercício 59 ... Calcule $\frac{\partial f}{\partial u}$ e $\frac{\partial f}{\partial v}$, em cada um dos casos seguintes, sem calcular a composição das funções referidas:

(i). $f(x,y) = x\log y + x^2y$ onde $x=2v$ e $y=uv$

(ii). $f(x,y) = e^y + x$ onde $x=u+v$ e $y=v-\log u$

(iii). $f(x,y,z) = xy^2 + y\log z$ onde $x=u+v,$ $y=u\,e^{-v}$ e $z=2v$

<u>Exercício</u> 60 ... (i). Suponha que $f = (u, v) : \mathcal{U} \subseteq \mathbb{C} \to \mathbb{C}$ é holomorfa em $\mathcal{U} \subseteq \mathbb{C}$, e que u, v são de classe C^2 em $\mathcal{U}$. Mostre que u e v são harmónicas em $\mathcal{U}$, isto é:

$$\Delta u = \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0$$

(e análogo para v).

(ii). Mostre que $u(x, y) = e^{-x}(x \sin y - y \cos y)$ é harmónica, e calcule uma função v tal que $f(z) = u(z) + iv(z)$ seja holomorfa em $\mathbb{C}$.

<u>Exercício</u> 61 ... Relativamente às funções que se seguem, calcule as derivadas parciais de ordem 1:

(i). $f(x, y) = (x \cos y, x \sin y, y)$

(ii). $f(x, y, z) = (x \cos z - y \sin z, x \sin z + y \cos z)$

(iii). $f(x, y) = (x + y, x - y, x^2 + y^3)$, no ponto $(1, 0)$.

<u>Exercício</u> 62 ... Relativamente às funções f que se seguem, calcule a derivada direccional de f nos pontos e segundo as direcções indicadas:

(i). $f(x, y, z) = (xy - \cos z^2, e^{xy} - x \cos(yz^2))$, no ponto $(1, 1, 0)$, na direcção do vector $(1, 1, 2)$.

(ii). $f(x, y) = (xy^3 - 2y, y^2, y^5 - 4x^2)$, no ponto $(1, -1)$, na direcção do vector $(1, 2)$.

<u>Exercício</u> 63 ... Mostre que a função $f(x, y) = (x \cos y, x \sin y)$ satisfaz as condições do teorema da função inversa em todo o ponto $(x, y) \in \mathbb{R}^2$, com $x > 0$, embora não seja injectiva nesse conjunto. Defina uma restrição de f que seja injectiva.

<u>Exercício</u> 64 ... Considere a função $f(x, y, z) = (e^x, \sin(x + y), e^z)$.

(i). Mostre que f é localmente inversível em $(0, 0, 0)$

(ii). Prove que existem pontos de $\mathbb{R}^3$ onde não se verificam as hipóteses do teorema da função inversa.

<u>Exercício</u> 65 ... Considere as funções $f(x, y) = (x^2 y, y - x$ e $g(x, y) = (x, y^3 x)$. Determine $\mathbf{Jac}(g \circ f)(1, 0)$. Prove que f é localmente inversível em $(1, 0)$ e que $g \circ f$ não é localmente inversível em $(1, 0)$.

Exercício 66 ... Considere a função $f(x, y, z) = (x^2 - y^2, 2xy)$.

(i). Demonstre que para todo o ponto $(x, y) \in \mathbb{R}^2 - \{(0, 0)\}$, a restrição de f a algum conjunto aberto que contenha (x, y), tem inverso.

(ii). Demonstre que se não restringirmos o seu domínio, então f não tem inverso.

(iii). Se f^{-1} é a inversa (local) de f numa vizinhança de $(1, 2)$, defina a diferencial de f no ponto $f(1, 2) = (-3, 4)$.

Exercício 67 ... Determine $df^{-1}[f(1, 1)]$, sendo f^{-1} a inversa de $f(x, y) = (x^2+2xy+y^2, x^2+y)$, numa vizinhança de $(1, 1)$.

Exercício 68 ... Considere a equação $F(x, y, z) = x^2 + y^2 + z^2 - xy = 0$. Diga em que pontos $(x_0, y_0, z_0) \in \mathbb{R}^3$ o teorema da função implícita garante que esta equação define implicitamente z como função de x, y: $z = g(x, y)$, e calcule numa vizinhança de (x_0, y_0, z_0), $\frac{\partial g}{\partial x}(x_0, y_0)$ e $\frac{\partial g}{\partial y}(x_0, y_0)$.

Exercício 69 ... Considere a equação $F(x, y, z) = x \cos z + yx = 0$. Diga em que pontos $(x_0, y_0, z_0) \in \mathbb{R}^3$ o teorema da função implícita garante que esta equação define implicitamente z como função de x, y: $z = g(x, y)$, numa vizinhança de (x_0, y_0, z_0), e calcule $\frac{\partial g}{\partial x}(x_0, y_0)$ e $\frac{\partial g}{\partial y}(x_0, y_0)$.

Exercício 70 ... Mostre que a equação $x \cos y - y \cos x = c$ (constante), define implicitamente y como função de x: $y = f(x)$, numa vizinhança de $(0, \frac{\pi}{2})$, e calcule $f'(0)$.

Exercício 71 ... Considere a equação $F(x, y, z) = 0$, para as funções F a seguir definidas. Diga se nos pontos $(x_0, y_0, z_0) \in \mathbb{R}^3$ o teorema da função implícita garante que essa equação define implicitamente z como função de x, y: $z = g(x, y)$, numa vizinhança de (x_0, y_0, z_0), e calcule $\frac{\partial g}{\partial x}(x_0, y_0)$, $\frac{\partial g}{\partial y}(x_0, y_0)$ e $\frac{\partial^2 g}{\partial x \partial y}(x_0, y_0)$.

(i). $F(x, y, z) = F(x, y, z) = x^2 + y^2 + z^2 - 3xy$ e $(x_0, y_0, z_0) = (1, 1, 1)$.

(ii). $F(x, y, z) = F(x, y, z) = (x + y + z) \cos z - 1$ e $(x_0, y_0, z_0) = (\pi, -1, \pi)$.

Exercício 72 ... Mostre que a equação $x^2 + xy + x \sin z + yz + z = 0$, define implicitamente z como função de x, y: $z = f(x, y)$, numa vizinhança de $(0, 0, 0)$.

Mostre que $\frac{\partial f}{\partial x}(0, 0) = 0 = \frac{\partial f}{\partial y}(0, 0)$, e averigue se $(0, 0)$ é máximo ou mínimo local de f.

Exercício 73 ... Mostre que existe uma vizinhança $\mathcal{U} \subseteq \mathbb{R}$ de $0 \in \mathbb{R}$, e uma única função $g : \mathcal{U} \to \mathbb{R}$ tal que $g(0) = 1$ e $x^2\, e^{g(x)} + [g(x)]^2\, e^x = 1$. Calcule $g''(0)$.

Exercício 74 ... Mostre que a equação $e^{x+y} - x\, e^{xz} = 0$ define implicitamente z como função de x, y: $z = f(x, y)$, numa vizinhança de $(1, -1, 0)$, e calcule $\frac{\partial^2 f}{\partial x \partial y}(1, -1)$

Exercício 75 ... Mostre que a equação $x^2 + e^x y + \sin y = 0$, $(x, y) \in \mathbb{R}^2$ define implicitamente y como função de x: $y = f(x)$, numa vizinhança de $(0, 0)$, e averigue se 0 é mínimo local de f.

Exercício 76 ...

(i). Prove que o sistema:

$$\begin{cases} xz^3 + y^2 w^3 = 1 \\ 2xy^3 + w^2 z = 0 \end{cases}$$

define implicitamente x e y como funções de z, w, numa vizinhança de $(z_0, w_0, x_0, y_0) = (0, 1, 0, 1)$.

(ii). Sejam $x = h(z, w)$ e $y = g(z, w)$ as funções referidas em (i). demonstre que a função $\mathbf{F}(z, w) = (h(z, w), g(z, w))$ admite inversa em torno do ponto $(0, 1)$.

Exercício 77 ... Mostre que o sistema:

$$\begin{cases} z - 2t + 2w + t^2 = -2 \\ z^2 - t^2 + w = 0 \end{cases}$$

é localmente resolúvel em ordem às variáveis z e t, numa vizinhança de $(t_0, z_0, w_0) = (1, -1, 0)$.

Calcule $z'(0)$ e $t'(0)$.

Exercício 78 ... Será possível resolver o sistema:

$$\begin{cases} xy^2 + xzu + yv^2 = 3 \\ u^2 yz + 2uv - u^2 v^2 = 2 \end{cases}$$

de tal forma a determinar $u = u(x, y, z)$ e $v = v(x, y, z)$ numa vizinhança de $(x, y, z) = (1, 1, 1)$, $(u, v) = (1, 1)$?

Exercício 79 ... Relativamente às superfícies $\mathbf{S}$ que se seguem, defina parametrizações locais:

(i). $\mathbf{S} = \{(x, y, z) \in \mathbb{R}^3 : z^2 - x^2 - y^2 = 1 \;\; e \;\; z > 0\}$

(ii). A parte do parabolóide elíptico $y = 6 - 3x^2 - 2z^2$ situada à direita do plano xz.

(iii). A parte da esfera $x^2 + y^2 + z^2 = 4$ situada acima do cone de equação $z = \sqrt{x^2 + y^2}$.

(iv). A parte do cilindro $x^2 + z^2 = 1$ situada entre os planos $y = -1$ e $y = 3$, e tal que $x > 0$.

Exercício 80 ... Nas alíneas que se seguem, determine uma equação cartesiana do espaço tangente a cada uma das superfícies e nos pontos indicados:

(i). $x = u + v$, $y = 3u^2$, $z = u - v$ em $\mathbf{p} = (2, 3, 0)$.

(ii). $x = u^2$, $y = u - v^2$, $z = v^2$, com $v \geq 0$, em $\mathbf{p} = (1, 0, 1)$.

(iii). $\Phi(u, v) = (uv, ue^v, ve^u)$ em $\mathbf{p} = (0, 0, 0)$.

(iv). $\Phi(u, v) = (u + v, u\cos v, v\sin u)$ em $\mathbf{p} = (1, 1, 0)$.

Exercício 81 ... Nas alíneas que se seguem, elimine os parâmetros u e v de modo a encontrar uma equação cartesiana da superfície que contem a imagem de cada uma das parametrizações indicadas:

(i). $\Phi(u, v) = (1 + u - v, 2 - 3u + v, 2u - v)$

(ii). $\Phi(u, v) = (a\cos u\cos v, a\sin u\cos v, a\sin v)$, com $u \in]0, 2\pi[$, $v \in] - \pi/2, \pi/2[$, $a > 0$.

(iii). $\Phi(u, v) = (v\sin\alpha\cos u, v\sin\alpha\sin u, v\cos\alpha)$, com $u \in]0, 2\pi[$, $v \in]0, h[$, $a > 0$.

(iv). $\Phi(u, v) = (a\sin u\cos v, b\sin u\sin v, c\cos u)$, com $a, b, c > 0$.

(v). $\Phi(u, v) = (u, a\sin v, a\cos v)$, com $u \in \mathbb{R}$, $v \in]0, 2\pi[$, $a > 0$.

Exercício 82 ... Determine o valor máximo e o valor mínimo da função $f(x, y, z) = x + y + z$, quando as variáveis estão ligadas pelas condições:

(i). $\frac{x^2}{4} + \frac{y^2}{5} + \frac{z^2}{25} = 1, \quad e \quad x + y = z$

(ii). $x^2 + y^2 = 2 \quad e \quad x + z = 1$

(iii). $x^2 - y^2 = 1 \quad e \quad 2x + z = 1$

Exercício 83 ... Relativamente às parametrizações locais que se seguem, calcule a expressão local do tensor métrico usual (**I** forma fundamental):

(i). $\Phi(u, v) = (u + v, u - v, uv)$.

(ii. $\Phi(\theta, t) = (f(t)\cos\theta, f(t)\sin\theta, g(t))$, sabendo que se trata de uma superfície de revolução.

(iii). $\Phi(u, v) = (u\cos v, u\sin v, \log\cos v + u)$, com $|v| < \pi/2$.

Exercício 84 ... Considere a seguinte parametrização local de um cone:

$$\Phi(\theta, \phi) = (\phi\cos\theta, \phi\sin\theta, \phi)$$

Calcule o comprimento da curva α, cuja expressão local nessa parametrização é:

$$\beta(t) = (\theta(t), \phi(t)) = (t, e^{t\cot(\gamma/\sqrt{2})})$$

onde $t \in [0, \pi]$ e γ é uma constante.

Exercício 85 ... Considere a seguinte parametrização local de uma esfera:

$$\Phi(\theta, \phi) = (\sin\phi\cos\theta, \sin\phi\sin\theta, \cos\phi)$$

Calcule o comprimento da curva α, cuja expressão local nessa parametrização é:

$$\beta(t) = (\theta(t), \phi(t)) = (\int_{\frac{\pi}{4}}^{t} \frac{1}{\sin R} dR, t)$$

onde $\pi/_4 \leq t \leq \pi/_2$ e γ é uma constante.

Capítulo 4

Funções vectoriais de variável real. Curvas em $\mathbf{R}^n$

4.1 Limites e continuidade

Comecemos o estudo das funções vectoriais de variável real, isto é de funções do tipo:

$$\alpha : t \in \mathcal{D} \subseteq \mathbb{R} \mapsto \alpha(t) \in \mathbb{R}^n \tag{4.1.1}$$

onde $\mathcal{D} \subseteq \mathbb{R}$ é um intervalo de $\mathbb{R}$.

Como para cada $t \in \mathcal{D}$, $\alpha(t)$ é um vector de $\mathbb{R}^n$, podemos escrever:

$$\begin{aligned} \alpha(t) &= \sum_{i=1}^{n} x^i(t)\mathbf{e}_i \\ &= (x^1(t), x^2(t), ..., x^n(t)) \end{aligned} \tag{4.1.2}$$

exprimindo assim a função α em coordenadas cartesianas. As funções x^i, $i = 1, ..., n$, definidas através de (4.1.2), dizem-se as **funções coordenadas (cartesianas)** , da função α.

Consideremos agora uma função vectorial de variável real $\alpha : \mathbf{I} \to \mathbb{R}^n$, definida num intervalo aberto $\mathbf{I} \subseteq \mathbb{R}$. Fixemos um qualquer ponto $t_o \in \mathbf{I}$ e seja $h \neq 0$ um número real suficientemente pequeno para que $t_o + h \in \mathbf{I}$. Se $\mathbf{v} \neq \mathbf{O}$ é um vector não nulo em $\mathbb{R}^n$, então:

$$\|\alpha(t_o + h) - (\alpha(t_o) + h\,\mathbf{v})\| \tag{4.1.3}$$

representa o "desvio" entre o valor exacto $\alpha(t_o + h)$ e o "valor aproximado" $\alpha(t_o) + h\,\mathbf{v}$, avaliado na recta que passa em $\alpha(t_o)$, e é paralela ao vector $\mathbf{v}$ (ver a figura 4.1).

Suponhamos que existe um vector $\mathbf{v} \in \mathbb{R}^n$, para o qual esse "desvio" converge mais ràpidamente para 0, do que h. Como sabemos, neste caso diz-se que α é diferenciável em t_o, e $\mathbf{v}$ diz-se a derivada de α em t_o. Mais formalmente:

♣ **Definição 4.1** ... *Seja* $\alpha : \mathbf{I} \to \mathbb{R}^n$ *uma função vectorial de variável real, definida num intervalo aberto* $\mathbf{I} \subseteq \mathbb{R}$*. Fixemos um qualquer ponto* $t_o \in \mathbf{I}$.

Figure 4.1: $\alpha(t_o + h) - (\alpha(t_o) + h\,\mathbf{v})$

.

Diz-se que α *é* **diferenciável (ou derivável) em** t_o*, se existe uma função afim:*

$$\mathbf{T}(h) = \alpha(t_o) + h\,\mathbf{v} \qquad h \in \mathbb{R} \tag{4.1.4}$$

tal que:

$$\lim_{h\to 0} \frac{\|\alpha(t_o + h) - \mathbf{T}(h)\|}{|h|} = \lim_{h\to 0} \frac{\|\alpha(t_o + h) - \alpha(t_o) - h\,\mathbf{v}\|}{|h|} = 0 \tag{4.1.5}$$

Neste caso, a função afim $\mathbf{T}$ *diz-se a* **aproximação afim óptima** *de* α *em* t_o*, e o vector* $\mathbf{v}$ *diz-se a* **derivada** *de* α *em* t_o*, e nota-se por* $\alpha'(t_o)$*.*

A parte linear de $\mathbf{T}$*, isto é, a função linear:*

$$h \mapsto h\,\mathbf{v} = h\,\alpha'(t_o) \tag{4.1.6}$$

diz-se a **diferencial** *de* α *em* t_o*, e nota-se por* $d\alpha_{t_o}$*.*

α *diz-se derivável em* $\mathbf{I}$ *se o é em todo o ponto* $t \in \mathbf{I}$*.*

É fácil ver que se existe uma função afim $\mathbf{T}$, que satisfaz (4.1.5), então ela é única, e portanto o vector $\mathbf{v} = \alpha'(t_o)$ está univocamente determinado. Por outro lado, atendendo a (**??**), vemos que o limite (4.1.5) se verifica sse:

$$\alpha'(t_o) = \lim_{h\to 0} \frac{\alpha(t_o + h) - \alpha(h)}{h} \tag{4.1.7}$$

o que permite deduzir a seguinte:

♣ **Proposição 4.1** ... *Seja* $\alpha : \mathbf{I} \to \mathbb{R}^n$ *uma função vectorial de variável real, definida num intervalo aberto* $\mathbf{I} \subseteq \mathbb{R}$*, e seja:*

$$t \mapsto (x^1(t), x^2(t), ..., x^n(t)) \tag{4.1.8}$$

a expressão de α *em coordenadas cartesianas.*

Então a derivada $\alpha'(t_o)$ *existe sse cada função* x^i *é derivável em* $t_o \in \mathbf{I}$ *e, neste caso:*

$$\alpha'(t_o) = ((x^1)'(t_o), ..., (x^n)'(t_0)) \tag{4.1.9}$$

Esta proposição reduz portanto o cálculo de derivadas de funções vectoriais de variável real, ao cálculo de derivadas de funções reais de variável real. Utilizando este facto, podemos fàcilmente deduzir as seguintes regras de derivação:

♣ **Proposição 4.2** ... *Sejam* $\alpha, \beta : \mathbf{I} \to \mathbb{R}^n$ *funções vectoriais de variável real, definidas e deriváveis num intervalo aberto* $\mathbf{I} \subseteq \mathbb{R}$, *e* $h : \mathbf{I} \to \mathbb{R}$ *uma função real também derivável em* $\mathbf{I}$. *Então as funções* $\alpha + \beta$, $h\,\alpha$, $\alpha \cdot \beta$ *e* $\alpha \times \beta$ *são também deriváveis em* $\mathbf{I}$, *sendo válidas as seguintes igualdades:*

$$(i) \quad ... \quad (\alpha + \beta)' = \alpha' + \beta' \tag{4.1.10}$$

$$(ii) \quad ... \quad (h\,\alpha)' = h'\,\alpha + h\,\alpha' \tag{4.1.11}$$

$$(iii) \quad ... \quad (\alpha \cdot \beta)' = \alpha' \cdot \beta + \alpha \cdot \beta' \tag{4.1.12}$$

$$(iv) \quad ... \quad (\alpha \times \beta)' = \alpha' \times \beta + \alpha \times \beta', \qquad \text{quando } n = 3 \tag{4.1.13}$$

Notemos que a definição (4.1.5), não exclui o caso em que $\mathbf{v} = \mathbf{O}$. Neste caso $\alpha'(t_o) = \mathbf{O}$, isto é, $\|\alpha(t_o + h) - \alpha(t_o)\|$ converge mais ràpidamente para 0 do que h.

♣ **Definição 4.2** ... *Seja* $\alpha : \mathbf{I} \to \mathbb{R}^n$ *uma função vectorial de variável real, definida num intervalo aberto* $\mathbf{I} \subseteq \mathbb{R}$.

Um ponto $t_o \in \mathbf{I}$ *diz-se* **regular** *se* $\alpha'(t_o)$ *existe e* $\alpha'(t_o) \neq \mathbf{O}$, *e* **singular** *se* $\alpha'(t_o)$ *existe e* $\alpha'(t_o) = \mathbf{O}$.

Se $\alpha : \mathbf{I} \to \mathbb{R}^n$ é uma função vectorial de variável real, definida e derivável num intervalo aberto $\mathbf{I} \subseteq \mathbb{R}$, então a função derivada:

$$\alpha' : t \in \mathbf{I} \mapsto \alpha'(t) \in \mathbb{R}^n \tag{4.1.14}$$

é uma função vectorial de variável real, bem definida em $\mathbf{I}$, e para a qual podemos calcular a respectiva derivada. Se esta existir, diz-se a segunda derivada de α, e nota-se por α''. Procedendo desta forma, definimos sucessivamente as derivadas de ordem k da função α, notadas por $\alpha^{(k)}$, $k = 0, 1, 2, ...$ (por convenção põe-se $\alpha^{(0)} = \alpha$, $\alpha^{(1)} = \alpha', \alpha^{(2)} = \alpha''$, etc....).

♣ **Definição 4.3** ... $\alpha : \mathbf{I} \to \mathbb{R}^n$ *diz-se de classe* C^m *em* $\mathbf{I}$, *se todas as derivadas* $\alpha^{(k)}$, $k = 0, 1, 2, ..., m$ *existem e são contínuas em* $\mathbf{I}$.

Por convenção, $\alpha : \mathbf{I} \to \mathbb{R}^n$ diz-se de classe C^0 em $\mathbf{I}$, se α é contínua em $\mathbf{I}$.

Quando $\mathbf{I} \subseteq \mathbb{R}$ é um intervalo fechado, ou semi-aberto, uma função $\alpha : \mathbf{I} \to \mathbb{R}^n$ diz-se de classe C^m em $\mathbf{I}$, se α é igual à restrição de uma função de classe C^m, definida num intervalo aberto que contem $\mathbf{I}$.

4.2 Curvas parametrizadas. Curvas regulares

Passemos agora ao estudo das curvas em $\mathbb{R}^n$. Seja $\mathbf{I} \subseteq \mathbb{R}$ um intervalo de $\mathbb{R}$.

♣ **Definição 4.4** ... *Uma função vectorial de variável real* $\alpha : \mathbf{I} \to \mathbb{R}^n$ *diz-se uma* **curva parametrizada**, *em* $\mathbb{R}^n$, *se* α *é de classe* C^1 *em* $\mathbf{I}$.

A variável $t \in \mathbf{I}$ *diz-se o* **parâmetro** *da curva parametrizada* α, *e a imagem* $\alpha(\mathbf{I}) \subset \mathbb{R}^n$ *diz-se o* **traço** *da curva parametrizada* α.

Uma curva parametrizada $\alpha : \mathbf{I} \to \mathbb{R}^n$ *diz-se* **regular**, *se todo o ponto* $t \in \mathbf{I}$ *é regular (i.e., se* $\frac{d\alpha}{dt} \neq \mathbf{O}$, $\forall t \in \mathbf{I}$*).*

É usual utilizar as seguintes notações e terminologia cinemática. $\alpha(t)$ descreve o movimento de um ponto (ou "partícula"), t é o tempo, e:

$$\begin{aligned}
\dot{\alpha}(t) = \mathbf{v}(t) &\equiv \frac{d\alpha}{dt} && \textbf{vector velocidade}\\
v(t) &\equiv \|\dot{\alpha}(t)\| && \textbf{velocidade}\\
\ddot{\alpha}(t) = \mathbf{a}(t) &\equiv \frac{d^2\alpha}{dt^2} && \textbf{vector aceleração}\\
a(t) &\equiv \|\ddot{\alpha}(t)\| && \textbf{aceleração}
\end{aligned} \tag{4.2.1}$$

do movimento desrito por α.

♣ **Definição 4.5** ... *Uma função* $\varphi : u \in \mathbf{J} \mapsto \varphi(u) \in \mathbb{R}$, *definida num intervalo* $\mathbf{J} \subseteq \mathbb{R}$, *diz-se uma* **mudança admissível de parâmetro** *se* φ *é de classe* C^1 *em* $\mathbf{J}$, *e se* $\frac{d\varphi}{du} \neq 0$, $\forall u \in \mathbf{J}$.

Como $\frac{d\varphi}{du}$ é contínua e $\frac{d\varphi}{du} \neq 0$, concluímos que, ou $\frac{d\varphi}{du} > 0$, $\forall u$ (i.e, φ é estritamente crescente em $\mathbf{J}$), ou $\frac{d\varphi}{du} < 0$ (i.e., é estritamente decrescente em $\mathbf{J}$). Em qualquer dos casos, φ é injectiva em $\mathbf{J}$, $\mathbf{I} \equiv \varphi(\mathbf{J})$ é um intervalo de $\mathbb{R}$ e, além disso, a função inversa $\psi \equiv \varphi^{-1} : \mathbf{I} \to \mathbf{J}$ é também uma mudança admissível de parâmetro (i.e., ψ é de classe C^1 em $\mathbf{I}$, e $\frac{d\psi}{dt} \neq 0$, $\forall t \in \mathbf{I}$).

Consideremos uma curva parametrizada $\alpha : t \in \mathbf{I} \mapsto \alpha(t) \in \mathbb{R}^n$, uma mudança admissível de parâmetro $\varphi : u \in \mathbf{J} \mapsto \varphi(u) = t \in \mathbf{I}$, e a seguinte composição de aplicações:

$$\mathbf{J} \xrightarrow{\varphi} \mathbf{I} \xrightarrow{\alpha} \mathbb{R}^n \tag{4.2.2}$$

Obtemos assim uma nova curva parametrizada $\alpha \circ \varphi : \mathbf{J} \to \mathbb{R}^n$ (regular, se α o é), com um novo parâmetro $u \in \mathbf{J}$, cujo traço é igual ao de α (ver a figura 4.2). Além disso, a orientação de $\alpha \circ \varphi$ será a mesma de α se $\frac{d\varphi}{du} > 0$, $\forall u$, e será a oposta se $\frac{d\varphi}{du} < 0$, $\forall u$.

O vector velocidade da nova curva parametrizada $\alpha \circ \varphi : \mathbf{J} \to \mathbb{R}^n$, é dado por:

$$\frac{d(\alpha \circ \varphi)}{du} = \frac{d\alpha}{dt}(\varphi(u))\,\frac{d\varphi}{du} \quad \forall u \in \mathbf{J} \tag{4.2.3}$$

como se deduz utilizando (4.1.9), e a regra da cadeia para funções reais de variável real.

Figure 4.2: Mudança admissível de parâmetro

.

É usual escrever a fórmula anterior (pondo $\beta \equiv \alpha \circ \varphi$ e $t = t(u) = \varphi(u)$), na forma (com abuso de notação):

$$\frac{d\beta}{du} = \frac{d\alpha}{dt}\frac{dt}{du} \tag{4.2.4}$$

Para terminar esta secção, vamos definir o conceito de curva em $\mathbb{R}^n$.

♣ Definição 4.6 ... *Duas curvas parametrizadas $\alpha : \mathbf{I} \to \mathbb{R}^n$ e $\beta : \mathbf{J} \to \mathbb{R}^n$, dizem-se* **equivalentes***, se existir uma mudança admissível de parâmetro $\varphi : u \in \mathbf{J} \mapsto \varphi(u) \in \mathbf{I}$, tal que $\beta = \alpha \circ \varphi$.*

É fácil ver que esta é uma relação de equivalência, no conjunto das curvas parametrizadas.

♣ Definição 4.7 ... *Uma* **curva de classe** C^1*, em $\mathbb{R}^n$, é uma classe de equivalência de curvas parametrizadas (de classe C^1).*

Uma curva de classe C^1, em $\mathbb{R}^n$, diz-se **regular** *se algum dos seus representantes é regular.*

Não é difícil adaptar a discussão anterior para definir **curvas de classe** C^m, em $\mathbb{R}^n$. Uma curva (parametrizada) diz-se de classe C^∞ se é de classe C^m, $\forall m$.

Se na definição de equivalência de curvas parametrizadas, permitirmos apenas mudanças admissíveis de parâmetro φ, com $\varphi' > 0$, obtemos anàlogamente uma relação de equivalência, cujas classes de equivalência se dizem **curvas orientadas**, em $\mathbb{R}^n$.

4.3 Comprimento de arco. Parametrização por arco

Consideremos uma curva parametrizada regular, de classe C^1, $\alpha : [a, b] \to \mathbb{R}^n$. Para calcular o respectivo comprimento, que notamos por $s(\alpha)$, é natural considerar aproximações poligonais de α. Para isso, consideremos uma partição $\mathcal{P}$ do intervalo $[a, b]$:

$$\mathcal{P} = \{a = t_o, t_1, ..., t_{k-1}, t_k = b\} \tag{4.3.1}$$

Figure 4.3: Aproximação poligonal de α

.

$\mathcal{P}$ define uma aproximação poligonal de α, nomeadamente a linha poligonal que une $\alpha(a)$ a $\alpha(b)$, com vértices sucessivos $\alpha(t_o)$, $\alpha(t_1)$,...,$\alpha(t_k)$ (ver a figura 4.5.16).

O comprimento:

$$s(\alpha, \mathcal{P}) = \sum_{i=1}^{k} \|\alpha(t_i) - \alpha(t_{i-1})\| \tag{4.3.2}$$

desta poligonal, é então uma aproximação para o comprimento $s(\alpha)$, de α.

Isto motiva a seguinte definição:

♣ **Definição 4.8** ... *Define-se o* **comprimento** $s(\alpha)$, *da curva (parametrizada)* $\alpha : [a, b] \to \mathbb{R}^n$, *através de :*

$$s(\alpha) \equiv \lim_{|\mathcal{P}|\to 0} s(\alpha, \mathcal{P}) \tag{4.3.3}$$

desde que este limite exista, no sentido em que, dado um $\epsilon > 0$ *arbitrário, deverá existir um* $\delta > 0$, *tal que:*

$$\forall \mathcal{P} : \quad |\mathcal{P}| < \delta \Rightarrow |s(\alpha) - s(\alpha, \mathcal{P})| < \epsilon \tag{4.3.4}$$

onde $|\mathcal{P}| \equiv \max_{1\leq i\leq k} |t_i - t_{i-1}|$ *designa a "norma" da partição* $\mathcal{P}$.

Quando $s(\alpha)$ *existe, diz-se que a curva (parametrizada)* $\alpha : [a, b] \to \mathbb{R}^n$ *é* **rectificável**.

Pode acontecer que uma curva (parametrizada) $\alpha : [a, b] \to \mathbb{R}^n$, não seja **rectificável**, se exigirmos apenas que α seja contínua! No entanto é possível provar a seguinte proposição muito útil para cálculos:

♣ **Proposição 4.3** ... *Seja* $\alpha : [a, b] \to \mathbb{R}^n$, *uma curva (parametrizada) regular de classe* C^1. *Então* α *é rectificável, e o seu comprimento* $s(\alpha)$, *é dado pelo integral:*

$$s(\alpha) = \int_a^b \|\dot{\alpha}(t)\| \, dt \tag{4.3.5}$$

Se $\alpha(t) = (x^1(t), ..., x^n(t))$, $t \in [a, b]$, é a representação de α em coordenadas cartesianas, então (4.3.5), é dado por:

$$s(\alpha) = \int_a^b \sqrt{\Big(\frac{dx^1}{dt}\Big)^2 + \Big(\frac{dx^2}{dt}\Big)^2 + ... + \Big(\frac{dx^n}{dt}\Big)^2}\ dt \tag{4.3.6}$$

Por exemplo, se $f : [a, b] \to \mathbb{R}$ é uma função de classe C^1, então o gráfico de f, é o traço da curva regular $\alpha : [a, b] \to \mathbb{R}^2$, dada por $\alpha(t) = (t, f(t))$, e portanto deduzimos de (4.3.6), a seguinte fórmula:

$$s(gr\, f) = s(\alpha) = \int_a^b \sqrt{1 + \Big(\frac{df}{dt}\Big)^2}\ dt \tag{4.3.7}$$

para o comprimento do gráfico $gr\, f$, de f.

Demonstremos que o comprimento de α é uma propriedade da curva e não da forma como está parametrizada, como aliás seria de prever. Para isso, comecemos por recordar a "fórmula da mudança de variáveis", em integrais de funções reais de variável real:

♣ **Proposição 4.4** (**Fórmula da mudança de variáveis**)... *Sejam $f : [a, b] \to \mathbb{R}$ contínua, $\phi : [c, d] \to \mathbb{R}$ derivável, com derivada $\phi' = \frac{d\phi}{du}$ integrável, e tal que $\phi([c, d]) \subset [a, b]$. Então:*

$$\int_{\phi(c)}^{\phi(d)} f(t)\, dt = \int_c^d f(\phi(u))\, \phi'(u)\, du \tag{4.3.8}$$

Podemos agora demonstrar a seguinte proposição:

♣ **Proposição 4.5** ... *Sejam $\alpha : [a, b] \to \mathbb{R}^n$ e $\beta : [c, d] \to \mathbb{R}^n$, duas curvas parametrizadas equivalentes, e seja $\phi : u \in [c, d] \mapsto \phi(u) \in [a, b]$, uma mudança admissível de parâmetro, de tal forma que $\beta = \alpha \circ \varphi$.*

Então:

$$s(\beta) = s(\alpha \circ \phi) = s(\alpha) \tag{4.3.9}$$

- Demonstração...

  Suponhamos que $\phi' = \frac{d\phi}{du} > 0$, de tal forma que $\phi(c) = a$ e $\phi(d) = b$ (o caso $\phi' < 0$ trata-se de forma análoga). Então:

$$\begin{aligned} s(\beta) = s(\alpha \circ \phi) &= \int_c^d \|\frac{d\beta}{du}\|\, du \\ &= \int_c^d \|\frac{d\alpha}{dt}(\phi(u))\phi'(u)\|\, du \\ &= \int_c^d \|\frac{d\alpha}{dt}(\phi(u))\|\phi'(u)\, du \\ &= \int_{\phi(c)}^{\phi(d)} \|\frac{d\alpha}{dt}(t)\|\, dt \\ &= \int_a^b \|\dot{\alpha}(t)\|\, dt = s(\alpha) \end{aligned} \tag{4.3.10}$$

  pondo $t = \phi(u)$, e usando (4.3.8),

♣.

Consideremos novamente uma curva (parametrizada) regular de classe C^1, $\beta : \mathbf{I} \rightarrow \mathbb{R}^n$, definida num intervalo $\mathbf{I} \subseteq \mathbb{R}$, fixemos um ponto $t_o \in \mathbf{I}$, e consideremos a seguinte **função comprimento de arco**, definida em $\mathbf{I}$, através de:

$$s = s(t) \equiv \int_{t_o}^{t} \| \frac{d\beta}{du} \| \, du \tag{4.3.11}$$

onde, como é habitual (por abuso de notação), usamos o mesmo símbolo s para a variável dependente e para a função comprimento de arco.

Se $t \geq t_o$, então $s \geq 0$ é igual ao comprimento do arco da curva β, entre $\beta(t_o)$ e $\beta(t)$. Se $t < t_o$, então $s < 0$ será igual a -(comprimento do arco da curva β, entre $\beta(t_o)$ e $\beta(t)$).

Do teorema fundamental do cálculo, concluímos que a função comprimento de arco $s = s(t)$, é de classe C^1 em $\mathbf{I}$, com uma derivada:

$$\frac{ds}{dt} = \|\dot{\beta}(t)\| > 0 \tag{4.3.12}$$

estritamente positiva em $\mathbf{I}$. Portanto, $t \in \mathbf{I} \mapsto s = s(t)$ é uma mudança admissível de parâmetro (de classe C^m se β é de classe C^m). Notemos (com abuso de notação) por $t = t(s)$, a função inversa da função comprimento de arco $s = s(t)$, e por $\mathbf{J} = s(\mathbf{I})$.

Consideremos agora a curva $\alpha : s \in \mathbf{J} \mapsto \alpha(s) \in \mathbb{R}^n$, equivalente a β, definida por (ver a figura 4.4):

$$\alpha(s) \equiv \beta(t(s)) \qquad s \in \mathbf{J} = s(\mathbf{I}) \tag{4.3.13}$$

Figure 4.4: Reparametrização por arco. $\alpha(s) \equiv \beta(t(s))$

.

À curva α, definida anteriormente por (4.3.13), chama-se a **reparametrização por arco** da curva (regular) β.

Como $\frac{dt}{ds} = \frac{1}{\frac{ds}{dt}} = \frac{1}{\|\dot{\beta}(t)\|}$, deduzimos que:

$$\begin{aligned} \|\frac{d\alpha}{ds}\| &= \|\frac{d\beta}{dt}\frac{dt}{ds}\| \\ &= \frac{1}{\|\dot{\beta}(t)\|}\|\dot{\beta}(t)\| = 1 \end{aligned} \tag{4.3.14}$$

isto é, a reparametrização por arco, α, da curva (regular) β, é percorrida com velocidade constante e igual a 1.

Em geral, podemos adoptar a seguinte definição:

♣ **Definição 4.9** ... *Uma parametrização $\alpha : s \in \mathbf{J} \mapsto \alpha(s) \in \mathbb{R}^n$, de uma curva regular em $\mathbb{R}^n$, diz-se uma* **parametrização por arco***, ou uma* **parametrização natural** *dessa curva, se* $\|\frac{d\alpha}{ds}\| = 1, \quad \forall s \in \mathbf{J}$.

s diz-se então um **parâmetro natural** *ou um* **parâmetro de comprimento de arco***.*

Quando $\alpha : s \in \mathbf{J} \mapsto \alpha(s) \in \mathbb{R}^n$ é uma parametrização natural de uma curva regular em $\mathbb{R}^n$, então $|s_2 - s_1|$ é o comprimento do arco dessa curva, compreendido entre $\alpha(s_1)$ e $\alpha(s_2)$. Por outro lado, se $\bar{\alpha}(\bar{s})$ é uma outra parametrização natural dessa mesma curva, então $s = \pm\bar{s} + \text{constante}$.

Na secção seguinte as derivadas relativamente a um parâmetro natural s, serão representadas por:

$$\frac{d\alpha}{ds} = \alpha'(s), \qquad \frac{d^2\alpha}{ds^2} = \alpha''(s), ... \tag{4.3.15}$$

4.4 Curvatura e torção. Geometria local das curvas regulares em $\mathbb{R}^3$

Consideremos uma parametrização natural $\alpha : s \in \mathbf{J} \mapsto \alpha(s) \in \mathbb{R}^3$, de uma curva regular em $\mathbb{R}^3$, de classe C^m $(m \geq 3)$. Utilizemos as notações (4.3.15), para as sucessivas derivadas relativamente ao parâmetro natural s.

Como vimos antes, o vector velocidade $\alpha'(s)$ tem norma constante e igual a 1.

Este vector $\alpha'(s)$, diz-se o **vector unitário tangente** em s, à curva (orientada) representada por α, e nota-se por $\mathbf{t} = \mathbf{t}(s) \equiv \alpha'(s)$. Notemos que, por mudança de orientação, o vector tangente muda o seu sentido.

Como $\|\alpha'(s)\|^2 = \alpha'(s) \cdot \alpha'(s) \equiv 1 \ \ \forall s$, obtemos por derivação, que:

$$\alpha''(s) \cdot \alpha'(s) = \alpha''(s) \cdot \mathbf{t}(s) = 0 \quad \forall s \tag{4.4.1}$$

o que significa que o vector aceleração $\alpha''(s) = \mathbf{t}'(s)$, é sempre perpendicular ao vector tangente $\mathbf{t} = \alpha'$.

♣ **Definição 4.10** ... *Seja $\alpha : s \in \mathbf{J} \mapsto \alpha(s) \in \mathbb{R}^3$ uma parametrização natural de uma curva regular em $\mathbb{R}^3$, de classe C^m $(m \geq 3)$.*

Chama-se **curvatura** *de α em s, e nota-se por $k(s)$, ao número (≥ 0):*

$$k(s) \equiv \|\alpha''(s)\| = \|\mathbf{t}'(s)\| \tag{4.4.2}$$

Quando $k(s) \neq 0$, chama-se **raio de curvatura** *de α em s, ao número*

$$\rho(s) \equiv \frac{1}{k(s)} \tag{4.4.3}$$

Geomètricamente, a curvatura $k(s)$ fornece uma medida de quão ràpidamente a curva α, se afasta da sua linha tangente em s, numa vizinhança de s.

Assim por exemplo, se $\alpha(s) = \mathbf{p} + s\mathbf{v}$ (onde $\mathbf{p} \in \mathbb{R}^3$ e $\mathbf{v} \in \mathbb{R}^3$ são vectores fixos em $\mathbb{R}^3$, com $\|\mathbf{v}\| = 1$) é uma recta em $\mathbb{R}^3$, então $k \equiv 0$. Recìprocamente, se $k(s) = \|\alpha''(s)\| \equiv 0$, então por integração deduzimos que $\alpha(s) = \mathbf{p} + s\mathbf{v}$, e portanto α é uma linha recta.

Notemos que α'' e a curvatura permanecem invariantes, se mudarmos a orientação da curva α.

Nos pontos em que $k(s) \neq 0$, podemos definir um vector unitário $\mathbf{n}(s)$, na direcção do vector aceleração $\alpha''(s)$, através de:

$$\mathbf{n}(s) \equiv \frac{\alpha''(s)}{k(s)} \tag{4.4.4}$$

e que é, como já vimos, perpendicular ao vector tangente $\mathbf{t}(s) = \alpha'(s)$. O vector $\mathbf{n}(s)$ diz-se por isso, o **vector normal unitário**, em s, e o plano que passa em $\alpha(s)$ e é determinado por $\mathbf{t}(s)$ e $\mathbf{n}(s)$, diz-se o **plano osculador**, em s. Este plano consiste portanto dos pontos $\mathbf{x} \in \mathbb{R}^3$, tais que $\mathbf{x} - \alpha(s)$ é perpendicular a $\mathbf{t}(s) \times \mathbf{n}(s)$, isto é:

$$\begin{aligned} \mathbf{x} \in \mathbb{R}^3 : \qquad (\mathbf{x} - \alpha(s)) \cdot (\mathbf{t}(s) \times \mathbf{n}(s)) &= \\ = [\mathbf{x} - \alpha(s), \mathbf{t}(s), \mathbf{n}(s)] &= 0 \end{aligned} \tag{4.4.5}$$

ou, de forma equivalente, por:

$$\mathbf{x} \in \mathbb{R}^3 : \qquad [\mathbf{x} - \alpha(s), \alpha'(s), \alpha''(s)] = 0 \tag{4.4.6}$$

Nos pontos em que $k(s) = 0$ (que se dizem **pontos de inflexão**), o vector normal e o plano osculador não estão definidos.

Para prosseguir a análise local de α, vamos supôr que $k(s) \neq 0$, $\forall s$.

O vector unitário $\mathbf{b}(s) \equiv \mathbf{t}(s) \times \mathbf{n}(s)$, é perpendicular ao plano osculador, e chama-se o **vector binormal**, em s.

Calculemos $\mathbf{b}'(s)$. Para isso, observemos que, por um lado $\mathbf{b}'(s)$ é ortogonal a $\mathbf{b}(s)$ (uma vez que $\|\mathbf{b}(s)\|^2 = \mathbf{b}(s) \cdot \mathbf{b}(s) \equiv 1$), e por outro lado (atendendo a que $\mathbf{t}'(s) = \alpha''(s) = k(s)\mathbf{n}(s)$):

$$\begin{aligned} \mathbf{b}'(s) &= \mathbf{t}'(s) \times \mathbf{n}(s) + \mathbf{t}(s) \times \mathbf{n}'(s) \\ &= k(s)\mathbf{n}(s) \times \mathbf{n}(s) + \mathbf{t}(s) \times \mathbf{n}'(s) \\ &= \mathbf{t}(s) \times \mathbf{n}'(s) \end{aligned} \tag{4.4.7}$$

o que implica que $\mathbf{b}'(s)$ é perpendicular também ao vector tangente unitário $\mathbf{t}(s)$. Isto significa que $\mathbf{b}'(s)$ deve ser um múltiplo escalar de $\mathbf{n}(s)$, i.e., $\mathbf{b}'(s) = \tau(s)\,\mathbf{n}(s)$, para alguma função $\tau(s)$.

♣ **Definição 4.11** ... *Seja* $\alpha : s \in \mathbf{J} \mapsto \alpha(s) \in \mathbb{R}^3$ *uma parametrização natural de uma curva regular em* $\mathbb{R}^3$*, de classe* C^m $(m \geq 3)$*, tal que* $\alpha''(s) \neq \mathbf{O}$, $\forall s$.

Chama-se **torção** *de* α *em* s*, e nota-se por* $\tau(s)$*, ao número definido por:*

$$\mathbf{b}'(s) = \tau(s)\,\mathbf{n}(s) \tag{4.4.8}$$

Geomètricamente, $|\tau(s)| = \|\mathbf{b}'(s)\|$ fornece uma medida de quão ràpidamente a curva α, se afasta do seu plano osculador em s, numa vizinhança de s.

Por exemplo, se $\tau \equiv 0$ (e $k \neq 0$), então $\mathbf{b}(s) \equiv \mathbf{b}_o =$ constante, e portanto:

$$\frac{d}{ds}(\alpha(s)\cdot\mathbf{b}_o) = \alpha'(s)\cdot\mathbf{b}_o = \mathbf{t}(s)\cdot\mathbf{b}_o = 0$$

isto é, $\alpha(s)\cdot\mathbf{b}_o =$ constante, o que significa que $\alpha(s)$ está contida num plano perpendicular a $\mathbf{b}_o$, e portanto α é uma curva plana (contida no seu plano osculador). A recíproca é também válida.

Notemos que, por mudança de orientação, o vector binormal $\mathbf{b}$ muda de sinal, uma vez que $\mathbf{b} = \mathbf{t}\times\mathbf{n}$. Deduzimos por isso que $\mathbf{b}'$, e portanto a torção τ, permanecem invariantes sob mudança de orientação de α.

Vamos resumir o que fizemos até agora:

(i)... A cada valor do parâmetro natural s, associamos três vectores unitários, ortogonais entre si:

$$\begin{array}{rcll} \mathbf{t}(s) & = & \alpha'(s) & \text{vector unitário tangente} \\ \mathbf{n}(s) & = & \dfrac{\alpha''(s)}{k(s)} & \text{vector unitário normal} \\ \mathbf{b}(s) & = & \mathbf{t}(s)\times\mathbf{n}(s) & \text{binormal} \end{array}$$

O triedro $\{\mathbf{t}(s), \mathbf{n}(s), \mathbf{b}(s)\}$, diz-se o **triedro de Frenet** em s (ver a figura 4.5).

Figure 4.5: Triedro de Frenet

.

(ii)... Em seguida, exprimimos as derivadas $\mathbf{t}'(s)$ e $\mathbf{b}'(s)$, de $\mathbf{t}(s)$ e $\mathbf{b}(s)$, na base $\{\mathbf{t}(s), \mathbf{n}(s), \mathbf{b}(s)\}$:

$$\begin{aligned} \mathbf{t}'(s) &= k(s)\,\mathbf{n}(s) \\ \mathbf{b}'(s) &= \tau(s)\,\mathbf{n}(s) \end{aligned}$$

obtendo deste modo, certas entidades geométricas (a curvatura $k(s)$, e a torção $\tau(s)$), que dão informação sobre o comportamento de α, numa vizinhança de s.

(iii)... Calculemos finalmente a derivada $\mathbf{n}'(s)$, exprimindo-a na base $\{\mathbf{t}(s), \mathbf{n}(s), \mathbf{b}(s)\}$. Como $\mathbf{n} = \mathbf{b} \times \mathbf{t}$, tem-se que:

$$\begin{aligned} \mathbf{n}'(s) &= \mathbf{b}'(s) \times \mathbf{t}(s) + \mathbf{b}(s) \times \mathbf{t}'(s) \\ &= \tau(s)\,\mathbf{n}(s) \times \mathbf{t}(s) + \mathbf{b}(s) \times k(s)\,\mathbf{n}(s) \\ &= -\tau(s)\,\mathbf{b}(s) - k(s)\,\mathbf{t}(s) \end{aligned} \tag{4.4.9}$$

e obtemos de novo a curvatura e a torção.

As equações acima obtidas:

$$\left\{ \begin{array}{lclcc} \mathbf{t}'(s) & = & & k(s)\,\mathbf{n}(s) & \\ \mathbf{n}'(s) & = & -k(s)\,\mathbf{t}(s) & & -\tau(s)\,\mathbf{b}(s) \\ \mathbf{b}'(s) & = & & \tau(s)\,\mathbf{n}(s) & \end{array} \right. \tag{4.4.10}$$

ou em forma matricial:

$$\begin{bmatrix} \mathbf{t}' \\ \mathbf{n}' \\ \mathbf{b}' \end{bmatrix} = \begin{bmatrix} 0 & k & 0 \\ -k & 0 & -\tau \\ 0 & \tau & 0 \end{bmatrix} \begin{bmatrix} \mathbf{t} \\ \mathbf{n} \\ \mathbf{b} \end{bmatrix} \tag{4.4.11}$$

dizem-se as **equações de Frenet** da curva α.

Finalmente, o plano que passa em $\alpha(s)$ e é determinado pelo par $\{\mathbf{t}(s), \mathbf{b}(s)\}$, diz-se o **plano rectificante** em s, e o plano que passa em $\alpha(s)$ e é determinado pelo par $\{\mathbf{n}(s), \mathbf{b}(s)\}$, diz-se o **plano normal** em s (ver a figura 4.5).

A seguinte proposição, cuja demonstração omitimos, mostra que a curvatura e a torção descrevem completamente o comportamento local da curva, a menos de um movimento rígido em $\mathbb{R}^3$:

♣ **Proposição 4.6** (**Teorema fundamental da teoria local das curvas em** $\mathbb{R}^3$) ... *Dadas funções diferenciáveis $k(s) > 0$ e $\tau(s)$, $s \in \mathbf{I}$, existe uma curva parametrizada regular $\alpha : \mathbf{I} \to \mathbb{R}^3$, tal que s é o parâmetro comprimento de arco, $k(s)$ é a curvatura e $\tau(s)$ a torção de α.*

Além disso, qualquer outra curva $\bar{\alpha}$, que satisfaz as mesmas condições, difere de α por um movimento rígido em $\mathbb{R}^3$, isto é, existe uma transformação ortogonal $\mathcal{O} : \mathbb{R}^3 \to \mathbb{R}^3$ (com determinante positivo), e um vector $\mathbf{p} \in \mathbb{R}^3$, tais que $\bar{\alpha} = \mathcal{O} \circ \alpha + \mathbf{p}$.

Consideremos finalmente uma curva (parametrizada) regular de classe C^m, com $m \geq 3$, $\beta : t \in \mathbf{I} \mapsto \beta(t) \in \mathbb{R}^3$, não necessàriamente parametrizada por arco, e seja $\alpha : s \in \mathbf{J} \mapsto \alpha(s) \in \mathbb{R}^n$, a reparametrização por arco da curva β, definida por (ver a figura 4.4):

$$\alpha(s) \equiv \beta(t(s)) \qquad s \in \mathbf{J} = s(\mathbf{I}) \tag{4.4.12}$$

Podemos definir para β, as entidades geométricas que acabamos de descrever. Assim, por exemplo, o vector unitário normal $\mathbf{n}(t) = \mathbf{n}_\beta(t)$ à curva β, em t, define-se como sendo igual ao vector unitário normal $\mathbf{n}(s) = \mathbf{n}_\alpha(s)$ à curva α, em $s = s(t)$, a curvatura $k(t) = k_\beta(t)$, de β em t, é igual à curvatura $k(s) = k_\alpha(s)$, de α em $s = s(t)$, e anàlogamente a torção $\tau(t) = \tau_\beta(t)$, de β em t, é igual à torção $\tau(s) = \tau_\alpha(s)$, de α em $s = s(t)$:

$$\mathbf{n}(t) = \mathbf{n}_\beta(t) \equiv \mathbf{n}_\alpha(s(t)) = \mathbf{n}(s(t)) \tag{4.4.13}$$

$$k(t) = k_\beta(t) \equiv k_\alpha(s(t)) = k(s(t)) \tag{4.4.14}$$

$$\tau(t) = \tau_\beta(t) \equiv \tau_\alpha(s(t)) = \tau(s(t)) \tag{4.4.15}$$

Utilizando para as sucessivas derivadas em ordem a t, as notações (4.2.1):

$$\frac{d\beta}{dt} = \dot{\beta}, \qquad \frac{d^2\beta}{dt^2} = \ddot{\beta}, ... \tag{4.4.16}$$

é possível provar, após alguns cálculos, que a curvatura $k(t)$, de β em t, é dada por:

$$k(t) = \frac{\|\dot{\beta} \times \ddot{\beta}\|}{\|\dot{\beta}\|^3} \tag{4.4.17}$$

e que a torção $\tau(t)$, de β em t, é dada por:

$$\tau(t) = -\frac{[\dot{\beta}, \ddot{\beta}, \dddot{\beta}]}{\|\dot{\beta} \times \ddot{\beta}\|^2} \tag{4.4.18}$$

fórmulas que permitem o cálculo directo de k e τ.

Notemos finalmente que, uma vez que t não é necessàriamente um parâmetro natural, a velocidade $v(t) \equiv \|\dot{\beta}(t)\|$ não é necessàriamente constante. Portanto o vector aceleração $\mathbf{a}(t) \equiv \ddot{\beta}(t)$, não é necessàriamente perpendicular ao vector unitário tangente $\mathbf{t}(t) = \frac{\dot{\beta}(t)}{\|\dot{\beta}(t)\|} = \frac{1}{v(t)}\,\dot{\beta}(t)$.

Derivando esta última relação $\dot{\beta}(t) = v(t)\,\mathbf{t}(t)$, e utilizando as definições (4.4.13) e (4.4.14), tem-se que:

$$\begin{aligned}
\mathbf{a}(t) \equiv \ddot{\beta}(t) &= \dot{v}(t)\,\mathbf{t}(t) + v(t)\dot{\mathbf{t}}(t) \\
&= \dot{v}(t)\,\mathbf{t}(t) + v(t)\frac{\mathbf{t}'(s(t))}{\frac{dt}{ds}} \\
&= \dot{v}(t)\,\mathbf{t}(t) + v^2(t)k(s(t))\,\mathbf{n}(s(t)) \\
&= \dot{v}(t)\,\mathbf{t}(t) + v^2(t)k(t)\,\mathbf{n}(t)
\end{aligned} \tag{4.4.19}$$

obtendo assim uma decomposição do vector aceleração numa parte tangencial e outra normal. As componentes de $\mathbf{a}(t)$, segundo $\mathbf{t}(t)$ e $\mathbf{n}(t)$ respectivamente, dizem-se a **aceleração tangencial** e a **aceleração normal ou centrípeta**:

$$\mathbf{a}(t) \equiv \ddot{\beta}(t) = \underbrace{\dot{v}(t)}_{\text{acel. tangencial}}\,\mathbf{t}(t) + \underbrace{v^2(t)k(t)}_{\text{acel. centrípeta}}\,\mathbf{n}(t) \tag{4.4.20}$$

4.5 Curvas planas em coordenadas polares

4.5.1 Coordenadas polares

Consideremos duas cópias de $\mathbb{R}^2$, uma, notada por $\mathbb{R}^2_{(\theta,r)}$, munida das coordenadas designadas por (θ, r), e outra, notada por $\mathbb{R}^2_{(x,y)}$, munida das coordenadas designadas por (x, y). A aplicação:

$$(\theta, r) \stackrel{\Phi}{\longmapsto} (x = r\cos\theta, y = r\sin\theta) \tag{4.5.1}$$

quando restrita à faixa $\mathcal{U} =)0, 2\pi(\times)0, +\infty(\subset \mathbb{R}^2_{(\theta,r)}$, é uma aplicação bijectiva de $\mathcal{U}$, sobre $\mathcal{V} = \mathbb{R}^2_{(x,y)} - \{(x, y) : \ y = 0 \quad x \geq 0\}$ (ver a figura 4.6).

Figure 4.6: Coordenadas polares

.

Isto significa que qualquer ponto $\mathbf{p} = (x, y) \in \mathcal{V}$, fica unìvocamente determinado pelas suas **coordenadas polares** $(\theta, r) \in \mathcal{U}$, através da aplicação Φ, dada por (4.5.1). Geomètricamente, $r = \|\mathbf{p}\|$ e θ representa o ângulo orientado que o vector $\mathbf{p}$ faz com a parte positiva do eixo dos xx.

Por outro lado, se considerarmos as **curvas coordenadas**:

$$r \longmapsto (x = r\cos\theta, y = r\sin\theta) \quad \theta \text{ fixo} \tag{4.5.2}$$
$$\theta \longmapsto (x = r\cos\theta, y = r\sin\theta) \quad r \text{ fixo} \tag{4.5.3}$$

vemos que os respectivos vectores velocidade:

$$\mathbf{v}_r = (\cos\theta, \sin\theta) \tag{4.5.4}$$
$$\mathbf{v}_\theta = r(-\sin\theta, \cos\theta) \tag{4.5.5}$$

são perpendiculares (e em particular, linearmente independentes).

Representemos por:

$$\mathbf{e}_r \equiv \mathbf{v}_r = (\cos\theta, \sin\theta) = \cos\theta\mathbf{i} + \sin\theta\mathbf{j} \tag{4.5.6}$$
$$\mathbf{e}_\theta \equiv \frac{1}{r}\mathbf{v}_\theta = (-\sin\theta, \cos\theta) = -\sin\theta\mathbf{i} + \cos\theta\mathbf{j} \tag{4.5.7}$$

Figure 4.7: Os vectores $\mathbf{e}_r$ e $\mathbf{e}_\theta$

.

os dois vectores unitários tangentes (ortogonais entre si), às curvas coordenadas (4.5.2) e (4.5.3), respectivamente (ver a figura 4.7).

Dada uma curva parametrizada $\alpha : \mathbf{I} \subseteq \mathbb{R} \to \mathcal{U} \subset \mathbb{R}^2_{(\theta,r)}$:

$$\alpha : t \mapsto \alpha(t) = (\theta(t), r(t)) \tag{4.5.8}$$

a respectiva imagem sob Φ, define uma única curva parametrizada $\Phi \circ \alpha$ em $\mathcal{V} \subset \mathbb{R}^2_{(x,y)}$. α diz-se a **expressão de $\Phi \circ \alpha$, em coordenadas polares**. Recìprocamente, dada uma curva parametrizada $\beta : \mathbf{J} \subseteq \mathbb{R} \to \mathcal{V} \subset \mathbb{R}^2_{(x,y)}$, fica determinada uma única curva parametrizada $\Phi^{-1} \circ \beta$, em $\mathcal{U} \subset \mathbb{R}^2_{(\theta,r)}$, que se diz a **expressão de β, em coordenadas polares** (ver a figura 4.8).

Figure 4.8: Curva em coordenadas polares

.

Por exemplo, dada uma função $\theta \in]0, 2\pi[\mapsto r = r(\theta)$, o respectivo gráfico, é uma curva em $\mathcal{U} \subset \mathbb{R}^2_{(\theta,r)}$, que pode ser parametrizada por:

$$\theta \mapsto (\theta, r(\theta)) \tag{4.5.9}$$

e a respectiva imagem, sob Φ, em $\mathcal{V} \subset \mathbb{R}^2_{(x,y)}$, é uma curva que pode ser parametrizada por:

$$\theta \mapsto (r(\theta)\cos\theta, r(\theta)\sin\theta) \tag{4.5.10}$$

Figure 4.9: Equação polar de uma curva

.

É habitual chamar à equação $r = r(\theta)$, a **equação polar** da curva (4.5.10) (ver a figura 4.9).

Como exemplo concreto, recordemos as equações polares das cónicas.

4.5.2 Equações polares das cónicas

Uma das possíveis definições de cónica, é a seguinte:

$\clubsuit$ **Definição 4.12** ... *Seja D uma linha recta fixa no plano, F um ponto fixo, não pertencente a D, e $\epsilon > 0$ um número positivo. Representemos por $d(\mathbf{x}, D)$ a distância entre o ponto $\mathbf{x}$ e a recta D.*

Ao conjunto:

$$\{\mathbf{x} \in \mathbb{R}^2 : \quad \|\mathbf{x} - F\| = \epsilon\, d(\mathbf{x}, D)\} \tag{4.5.11}$$

chama-se **cónica** *com* **excentricidade** ϵ, **foco** F, *e* **directriz** D.

A cónica (4.5.11) diz-se uma:

$$\begin{array}{rll} \textbf{elipse} & \textit{se} & \epsilon < 1 \\ \textbf{parábola} & \textit{se} & \epsilon = 1 \\ \textbf{hipérbole} & \textit{se} & \epsilon > 1 \end{array}$$

Escolhamos um sistema de eixos cartesianos centrado no foco F (de tal forma que $F = \mathbf{O}$), e tal que a directriz D seja perpendicular ao eixo dos xx, e esteja situada à direita do foco $F = \mathbf{O}$, a uma distância d deste (isto é, D é a recta de equação $x = d$) (ver a figura 4.10).

Pondo $\mathbf{x} = (x, y)$ e fazendo os cálculos, atendendo a que $\|\mathbf{x} - F\| = \|\mathbf{x}\|$ e a que $d(\mathbf{x}, D) = |x - d|$, concluímos que a equação (4.5.11) pode ser escrita na forma:

$$\|\mathbf{x}\| = \epsilon\, |x - d| \tag{4.5.12}$$

Figure 4.10: Equação polar de uma cónica

.

Se designarmos por (θ, r) as coordenadas polares de $\mathbf{x} = (x, y)$, então como $\|\mathbf{x}\| = r$ e $x = r\cos\theta$, podemos escrever esta última equação na forma:

$$r = \epsilon\,|r\cos\theta - d| \tag{4.5.13}$$

Se $\mathbf{x}$ está situado à esquerda da directriz D, então $x = r\cos\theta < d$, e portanto $|r\cos\theta - d| = d - r\cos\theta$, e (4.5.13) fica na forma $r = \epsilon\,(d - r\cos\theta)$, ou, resolvendo em ordem a r:

$$r = r(\theta) = \frac{\epsilon\, d}{\epsilon\,\cos\theta + 1} \tag{4.5.14}$$

Se $\mathbf{x}$ está situado à direita da directriz D, então $x = r\cos\theta > d$, e obtemos de forma análoga à anterior:

$$r = r(\theta) = \frac{\epsilon\, d}{\epsilon\,\cos\theta - 1} \tag{4.5.15}$$

Como $r > 0$, esta última equação (4.5.17), implica que $\epsilon > 1$, isto é, existem pontos à direita de D apenas no caso da hipérbole. Resumindo a discussão anterior, podemos enunciar a seguinte proposição:

♣ **Proposição 4.7** ... *Seja $\mathcal{C}$ uma cónica com excentricidade ϵ, foco $F \equiv \mathbf{O}$, e directriz $D: x = d$, situada perpendicularmente ao eixo dos xx e a uma distância d, à direita de $F = \mathbf{O}$.*

Então se $0 < \epsilon < 1$, $\mathcal{C}$ é uma elipse, e se $\epsilon = 1$, uma parábola. Nestes dois casos, todo o ponto de $\mathcal{C}$ se situa à esquerda de D, e satisfaz a **equação polar***:*

$$r = r(\theta) = \frac{\epsilon\, d}{1 + \epsilon\,\cos\theta} \tag{4.5.16}$$

Se $\epsilon > 1$, $\mathcal{C}$ é uma hipérbole com um ramo em cada um dos lados de D. Pontos do ramo esquerdo satisfazem a equação polar (4.5.16), enquanto que os do ramo direito satisfazem a **equação polar***:*

$$r = r(\theta) = \frac{\epsilon\, d}{\epsilon\,\cos\theta\, - 1} \tag{4.5.17}$$

4.5.3 Velocidade e aceleração em coordenadas polares. Aceleração radial e transversa

Consideremos uma curva parametrizada $\alpha : \mathbf{I} \subseteq \mathbb{R} \to \mathcal{U} \subset \mathbb{R}^2_{(\theta,r)}$:

$$\alpha : t \mapsto \alpha(t) = (\theta(t), r(t)) \tag{4.5.18}$$

Como vimos antes, a imagem de α sob Φ, define uma única curva parametrizada $\beta \equiv \Phi \circ \alpha$ em $\mathcal{V} \subset \mathbb{R}^2_{(x,y)}$ (de tal forma que α é a expressão de $\beta = \Phi \circ \alpha$, em coordenadas polares). A curva β pode ser parametrizada por:

$$\beta : t \mapsto (r(t)\cos\theta(t), r(t)\sin\theta(t)) \tag{4.5.19}$$

Calculemos o vector velocidade da curva β:

$$\begin{aligned}
\dot{\beta}(t) = \frac{d\beta}{dt} &= (\dot{r}(t)\cos\theta(t) - r(t)\dot{\theta}(t)\sin\theta(t), \dot{r}(t)\sin\theta(t) + r(t)\dot{\theta}(t)\cos\theta(t)) \\
&= \dot{r}(t)(\cos\theta(t), \sin\theta(t)) + r(t)\dot{\theta}(t)(-\sin\theta(t), \cos\theta(t)) \\
&= \dot{r}(t)\,\mathbf{e}_r(t) + r(t)\dot{\theta}(t)\,\mathbf{e}_\theta(t)
\end{aligned} \tag{4.5.20}$$

onde utilizamos (4.5.6) e (4.5.7).

À expressão que acabamos de obter é usual chamar a expressão do vector velocidade em coordenadas polares:

$$\dot{\beta}(t) = \dot{r}(t)\,\mathbf{e}_r(t) + r(t)\dot{\theta}(t)\,\mathbf{e}_\theta(t) \tag{4.5.21}$$

À componente $r(t)$, segundo $\mathbf{e}_r$, chama-se **velocidade radial**, e à componente $r(t)\dot{\theta}(t)$, segundo $\mathbf{e}_\theta$, **velocidade transversa**.

Procedendo de forma análoga, podemos calcular a seguinte expressão do vector aceleração $\ddot{\beta}(t)$, em coordenadas polares:

$$\ddot{\beta}(t) = (\ddot{r}(t) - r(t)\dot{\theta}^2(t))\,\mathbf{e}_r(t) + (r(t)\ddot{\theta}(t) + 2\dot{r}(t)\dot{\theta}(t))\,\mathbf{e}_\theta(t) \tag{4.5.22}$$

As componentes deste vector aceleração $\ddot{\beta}(t)$, respectivamente segundo $\mathbf{e}_r$ e $\mathbf{e}_\theta$, dizem-se:

$$\textbf{aceleração radial} = \ddot{r}(t) - r(t)\dot{\theta}^2(t) \tag{4.5.23}$$

$$\begin{aligned}
\textbf{aceleração transversa} &= r(t)\ddot{\theta}(t) + 2\dot{r}(t)\dot{\theta}(t) \\
&= \frac{1}{r(t)}\frac{d}{dt}(r^2(t)\dot{\theta}(t))
\end{aligned} \tag{4.5.24}$$

Portanto:

$$\ddot{\beta}(t) = \underbrace{\ddot{r}(t) - r(t)\dot{\theta}^2(t)}_{\text{acel. radial}}\,\mathbf{e}_r + \underbrace{r(t)\ddot{\theta}(t) + 2\dot{r}(t)\dot{\theta}(t)}_{\text{acel. transversa}}\,\mathbf{e}_\theta \tag{4.5.25}$$

Na situação especial em que $t = \theta$, de tal forma que β pode ser descrita pela equação polar $r = r(\theta)$, as fórmulas anteriores para os vectores velocidade e aceleração, tomam o aspecto seguinte:

$$\dot{\beta}(\theta) = \dot{r}(\theta)\,\mathbf{e}_r(\theta) + r(\theta)\,\mathbf{e}_\theta(\theta) \tag{4.5.26}$$

$$\ddot{\beta}(\theta) = (\ddot{r}(\theta) - r(\theta))\,\mathbf{e}_r(\theta) + 2\dot{r}(\theta)\,\mathbf{e}_\theta(\theta) \tag{4.5.27}$$

4.5.4 Área de "conjuntos radiais"

Suponhamos que uma curva β, está descrita por uma equação polar $r = r(\theta) > 0$, $0 < \theta < 2\pi$ ($r(\theta)$ contínua), e calculemos a área do "conjunto radial" $\mathcal{R} \subset \mathbb{R}^2_{(x,y)}$, constituído pelos pontos $(x, y) \in \mathbb{R}^2_{(x,y)}$, cujas coordenadas polares (θ, r), verificam as desigualdades (ver a figura 4.11):

$$0 \le r \le r(\theta) \quad e \quad a \le \theta \le b \tag{4.5.28}$$

onde $0 \le a < b \le 2\pi$.

Figure 4.11: Conjunto radial $\mathcal{R}$

.

Para calcular a referida área, consideremos uma partição $\mathcal{P}$, do intervalo $[a, b]$, $\mathcal{P} = \{a = \theta_o < \theta_1 < \theta_2 < ... < \theta_n = b\}$, e utilizemos as notações seguintes:

$$\begin{aligned} m_j &\equiv \min\{r(\theta) : \theta_{j-1} \le \theta \le \theta_j\} \\ M_j &\equiv \max\{r(\theta) : \theta_{j-1} \le \theta \le \theta_j\} \\ \Delta\theta_j &\equiv \theta_j - \theta_{j-1} \\ A_j &\equiv \text{área do subconjunto radial } 0 \le r \le r(\theta),\ \theta_{j-1} \le \theta\theta_j \end{aligned}$$

Então é evidente que (ver a figura 4.12):

$$\frac{1}{2} m_j^2 \, \Delta\theta_j \le A_j \le \frac{1}{2} M_j^2 \, \Delta\theta_j \tag{4.5.29}$$

Quanto à área total $A(\mathcal{R}) \equiv \sum_{j=1}^n a_j$, de $\mathcal{R}$, tem-se que:

$$\sum_{j=1}^n \frac{1}{2} m_j^2 \, \Delta\theta_j \le A \le \sum_{j=1}^n \frac{1}{2} M_j^2 \, \Delta\theta_j \tag{4.5.30}$$

As somas que ocorrem na fórmula anterior (4.5.30), não são mais do que as somas de Riemann inferior e superior, para o integral de Riemann $\int_a^b \frac{1}{2} r^2(\theta) \, d\theta$. Como as desigualdades (4.5.30) são válidas para toda a partição $\mathcal{P}$, concluímos que:

$$A(\mathcal{R}) = \frac{1}{2} \int_a^b r^2(\theta) \, d\theta \tag{4.5.31}$$

Figure 4.12: Cálculo da área do conjunto radial $\mathcal{R}$

.

4.6 Mecânica orbital

4.6.1 Campos de forças. Lei de Newton

♣ **Definição 4.13** ... *Um* **campo de vectores** *definido num subconjunto* $\mathcal{U} \subseteq \mathbb{R}^n$*, é uma aplicação* $\mathbf{F}$*, que associa a cada ponto* $\mathbf{x} \in \mathcal{U}$*, um vector* $\mathbf{F}(\mathbf{x}) \in \mathbb{R}^n$*:*

$$\mathbf{x} \in \mathcal{U} \subseteq \mathbb{R}^n \mapsto \mathbf{F}(\mathbf{x}) \in \mathbb{R}^n \tag{4.6.1}$$

Gràficamente, representamos o vector $\mathbf{F}(\mathbf{x})$ aplicado no ponto $\mathbf{x}$, em cada ponto $\mathbf{x} \in \mathcal{U}$ (ver a figura 4.13).

Figure 4.13: Campo de vectores $\mathbf{x} \mapsto \mathbf{F}(\mathbf{x}) \in \mathbf{R}^n$

.

Nesta secção vamos supôr que é dado um campo de vectores $\mathbf{F}$, que representa um campo de forças, definido num certo subconjunto $\mathcal{U} \subseteq \mathbb{R}^3$. Isto é, cada vector $\mathbf{F}(\mathbf{x})$ representa a força exercida sobre uma partícula teste (de massa 1), situada no ponto $\mathbf{x} \in \mathcal{U}$.

De acordo com as leis de Newton, o movimento futuro (e passado) de uma partícula que num certo instante t_o, se situa num ponto $\mathbf{p}_o$ e tem velocidade $\mathbf{v}_o$, fica completamente determinado pelas forças que nela actuam. O movimento é então determinado pela lei: "o vector aceleração é directamente proporcional à força que actua na partícula."

Vamos nesta secção utilizar as notações (4.2.1), habituais em mecânica, representando a posição da referida partícula no instante t, pelo **vector de posição** $\mathbf{r}(t) \in \mathbb{R}^3$, e as sucessivas derivadas em ordem ao "tempo" t por:

$$\dot{\mathbf{r}}(t) = \mathbf{v}(t) = \frac{d\mathbf{r}}{dt} \quad \text{vector velocidade} \tag{4.6.2}$$

$$\ddot{\mathbf{r}}(t) = \mathbf{a}(t) = \frac{d^2\mathbf{r}}{dt^2} \quad \text{vector aceleração} \tag{4.6.3}$$

e assim sucessivamente.

Assim por exemplo, na ausência de quaisquer forças, teremos que $\mathbf{r}(t_o) = \mathbf{p}_o$, $\dot{\mathbf{r}}(t_o) = \mathbf{v}_o$ e $\ddot{\mathbf{r}}(t) = \mathbf{O}$, $\forall t \in \mathbb{R}$, e o movimento $\mathbf{r}(t)$ deve ser unìvocamente determinado por estas condições. Como veremos posteriormente, $\mathbf{r}$ é uma solução da **equação diferencial** $\ddot{\mathbf{r}} = \mathbf{O}$, que satisfaz as **condições iniciais** $\mathbf{r}(t_o) = \mathbf{p}_o$ e $\dot{\mathbf{r}}(t_o) = \mathbf{v}_o$. As leis de Newton exigem que esta solução seja única, o que de facto acontece, como veremos futuramente. A solução única é $\mathbf{r}(t) = \mathbf{p}_o + t\mathbf{v}_o$, $t \in \mathbb{R}$, o que significa que a partícula se move numa linha recta, com velocidade constante igual a $\|\mathbf{v}_o\|$.

Em geral, suponhamos que $\mathbf{F}$ é um campo de forças (que por simplicidade, supômos que não depende do tempo t), definido num certo subconjunto $\mathcal{U} \subseteq \mathbb{R}^3$. Uma partícula de massa m, situada num ponto $\mathbf{x} \in \mathcal{U}$, estará então sujeita à acção de uma força igual a $m\,\mathbf{F}(\mathbf{x})$, e de acordo com Newton, essa partícula "acelera" na direcção de $\mathbf{F}(\mathbf{x})$. A correspondente aceleração é dada pela lei de Newton:

♣ **Proposição 4.8** *(***Lei de Newton***)...*

$$m\mathbf{F} = m\mathbf{a} \tag{4.6.4}$$

isto é, "**Força exercida sobre uma partícula de massa** m=**massa** m. **aceleração**" (recordemos que cada vector $\mathbf{F}(\mathbf{x})$ representa a força exercida sobre uma partícula teste (de massa 1), situada no ponto $\mathbf{x} \in \mathcal{U}$).

Suponhamos que uma partícula de massa m (constante), está sujeita à acção de um campo de forças $\mathbf{F}$, e suponhamos que, no instante t_o, ela se situa na posição $\mathbf{r}(t_o) = \mathbf{r}_o$, com um vector velocidade $\dot{\mathbf{r}}(t_o) = \mathbf{v}_o$. Seja $\mathbf{r}(t)$ o vector posição que descreve o movimento da partícula, de acordo com a lei de Newton. Então, num qualquer instante t, a partícula estará na posição $\mathbf{r}(t)$ e por isso, estará sujeita a uma força igual a $m\,\mathbf{F}(\mathbf{r}(t))$. De acordo com (4.6.4), a **equação (diferencial) do movimento** é:

$$\ddot{\mathbf{r}}(t) = \mathbf{F}(\mathbf{r}(t)) \tag{4.6.5}$$

o que significa que $\mathbf{r}$ é uma solução da equação diferencial $\ddot{\mathbf{x}} = \mathbf{F}(\mathbf{x})$, que satisfaz as condições iniciais $\mathbf{r}(t_o) = \mathbf{r}_o$ e $\dot{\mathbf{r}}(t_o) = \mathbf{v}_o$. Como veremos num capítulo posterior, podemos garantir que, sob certas condições de "regularidade" de $\mathbf{F}$, essa solução é única, de acordo com a exigência de Newton.

Vamos de seguida analisar dois campos de forças particularmente importantes.

4.6.2 Movimento num campo uniforme (constante)

Por definição, um **campo de forças uniforme (ou constante)**, é um campo de vectores constante:

$$\mathbf{F}(\mathbf{x}) \equiv \mathbf{F}^o \in \mathbb{R}^3 \quad =\text{constante} \tag{4.6.6}$$

onde, de acordo com a nossa convenção, $\mathbf{F}(\mathbf{x}) \equiv \mathbf{F}^o$, representa a força (constante) exercida sobre uma partícula teste (<u>de massa 1</u>), situada no ponto $\mathbf{x} \in \mathbb{R}^3$).

Designando por $F^o = \|\mathbf{F}^o\| \geq 0$ a norma (intensidade) do campo $\mathbf{F}$, é usual supôr, por uma escolha conveniente dos eixos coordenados, que (ver a figura 4.14):

$$\mathbf{F}(\mathbf{x}) \equiv -F^o\,\mathbf{k} \tag{4.6.7}$$

Figure 4.14: Campo de forças uniforme

.

Se uma partícula de massa m (constante) está sujeita à acção de um campo de forças $\mathbf{F}$ uniforme, dado por (4.6.7), então de acordo com a lei de Newton (4.6.4), $-mF^o\,\mathbf{k} = m\,\mathbf{a}$, e portanto o respectivo vector aceleração $\mathbf{a} = \ddot{\mathbf{r}}$ é constante:

$$\mathbf{a}(t) = \ddot{\mathbf{r}}(t) \equiv -F^o\,\mathbf{k} \tag{4.6.8}$$

(independente de m), e o movimento diz-se uniformemente acelerado.

Suponhamos que, no instante inicial t_o, ela se situa na posição $\mathbf{r}(t_o) = \mathbf{r}_o$, com um vector velocidade $\dot{\mathbf{r}}(t_o) = \mathbf{v}_o$. Então o movimento é dado por:

$$\mathbf{r}(t) = -\frac{1}{2}F^o t^2\,\mathbf{k} + \mathbf{v}_o\,t + \mathbf{r}_o \tag{4.6.9}$$

Por exemplo, para um objecto de massa m (constante), que se move sob a influência exclusiva da força devida à gravidade (isto é, a única força a que está sujeito, é o seu próprio peso), o vector aceleração é igual a:

$$\mathbf{a}(t) = \ddot{\mathbf{r}}(t) \equiv -g\,\mathbf{k} \tag{4.6.10}$$

(independente de m), onde g é a aceleração da gravidade ($\approx 9.80\,m/seg^2$). O objecto diz-se então em "queda livre", e o movimento determinado pelas condições iniciais $\mathbf{r}(t_o) = \mathbf{r}_o$ e $\dot{\mathbf{r}}(t_o) = \mathbf{v}_o$, é dado por:

$$\mathbf{r}(t) = -\frac{1}{2}gt^2\,\mathbf{k} + \mathbf{v}_o\,t + \mathbf{r}_o \tag{4.6.11}$$

4.6.3 Movimento num campo central

Por definição, um **campo de forças central** (de centro $\mathbf{O}$), é um campo de vectores da forma (ver a figura 4.15):

$$\mathbf{F}(\mathbf{x}) \equiv f(\|\mathbf{x}\|)\,\mathbf{x} \tag{4.6.12}$$

onde f é uma função real definida e de classe C^∞ no intervalo $\{r \in \mathbb{R} : \; r > 0\}$. Em geral, $\mathbf{F}$ está apenas definido em $\mathbb{R}^3 - \{\mathbf{O}\}$. $\mathbf{F}$ diz-se **atractor** ou **repulsor**, conforme $f(r) < 0$ ou $f(r) > 0$, $\forall r > 0$, respectivamente.

Mais uma vez, de acordo com a nossa convenção anterior, $\mathbf{F}(\mathbf{x}) \equiv f(\|\mathbf{x}\|)\,\mathbf{x}$, representa a força exercida sobre uma partícula teste (de massa 1), situada no ponto $\mathbf{x} \in \mathbb{R}^3 - \{\mathbf{O}\}$.

Figure 4.15: Campo de forças central

.

Se uma partícula de massa m (constante) está sujeita à acção de um campo de forças central $\mathbf{F}$, dado por (4.6.12), então, num qualquer instante t, a partícula estará na posição $\mathbf{r}(t)$ e por isso, estará sujeita a uma força igual a $m\,\mathbf{F}(\mathbf{r}(t)) = mf(r(t))\,\mathbf{r}(t)$, onde $r(t) = \|\mathbf{r}(t)\|$. De acordo com (4.6.4), a **equação (diferencial) do movimento** é:

$$\ddot{\mathbf{r}}(t) = f(r(t))\,\mathbf{r}(t) \tag{4.6.13}$$

Vejamos algumas propriedades do movimento num tal campo de forças central.

♣ **Proposição 4.9** ... *O movimento $\mathbf{r}(t)$, de uma partícula sujeita à acção de um campo de forças central, da forma (4.6.12), é plano.*

- Demonstração...

  Calculemos a derivada $\frac{d}{dt}(\mathbf{r}(t) \times \dot{\mathbf{r}}(t))$.

  Atendendo a (4.6.12), temos que:

$$\begin{aligned} \frac{d}{dt}(\mathbf{r}(t) \times \dot{\mathbf{r}}(t)) &= \dot{\mathbf{r}}(t) \times \dot{\mathbf{r}}(t) + \mathbf{r}(t) \times \ddot{\mathbf{r}}(t) \\ &= r(t) \times \ddot{\mathbf{r}}(t) \\ &= \mathbf{r}(t) \times f(r(t))\,\mathbf{r}(t) \\ &= \mathbf{O} \end{aligned} \tag{4.6.14}$$

Isto significa que:

$$\mathbf{r}(t) \times \dot{\mathbf{r}}(t) \equiv \mathbf{u} \quad (= \text{constante}) \tag{4.6.15}$$

Se $\mathbf{u} \neq \mathbf{O}$, isto implica que $\mathbf{r}(t)$ e $\dot{\mathbf{r}}(t)$ permanecem sempre no plano ortogonal a $\mathbf{u}$, e portanto o movimento processa-se nesse plano. Se $\mathbf{u} = \mathbf{O}$, então $\dot{\mathbf{r}}(t) = h(t)\,\mathbf{r}(t)$ para alguma função real h, o que significa que o vector velocidade é sempre colinear com o vector de posição, e portanto o movimento processa-se nessa linha,

♣.

♣ <u>**Definição**</u> **4.14** ... *Dada uma partícula de massa m, com vector de posição $\mathbf{r}(t)$, define-se o* **momento angular (ou cinético) m**, *através de:*

$$\mathbf{m}(t) \equiv m\,(\mathbf{r}(t) \times \dot{\mathbf{r}}(t)) \tag{4.6.16}$$

A dedução anterior, nomeadamente a fórmula (4.6.14), prova que no movimento de uma partícula sujeita à acção de um campo de forças central, da forma (4.6.12), o momento angular $\mathbf{m}(t) \equiv m\,(\mathbf{r}(t) \times \dot{\mathbf{r}}(t))$, permanece constante ("**lei da conservação do momento angular**").

Uma vez que o movimento é plano e atendendo à simetria do campo central, é conveniente utilizar coordenadas polares em $\mathbb{R}^2$, para deduzir de uma forma mais simples, mais propriedades sobre esse movimento. Utilizemos pois os conceitos e notações da secção 2.6.

Recordando a expressão (4.5.25) do vector aceleração em coordenadas polares, vemos que a equação (diferencial) do movimento (4.6.13), $\ddot{\mathbf{r}}(t) = f(r(t))\,\mathbf{r}(t)$, pode ser reescrita na forma:

$$\ddot{\mathbf{r}}(t) = \underbrace{\ddot{r}(t) - r(t)\dot{\theta}^2(t)}_{\text{acel. radial}}\,\mathbf{e}_r + \underbrace{r(t)\ddot{\theta}(t) + 2\dot{r}(t)\dot{\theta}(t)}_{\text{acel. transversa}}\,\mathbf{e}_\theta = f(r(t))r(t)\,\mathbf{e}_r \tag{4.6.17}$$

uma vez que $\mathbf{r}(t) = r(t)\,\mathbf{e}_r$.

Podemos portanto concluir que:

(i)... a aceleração transversa é nula:

$$r(t)\ddot{\theta}(t) + 2\dot{r}(t)\dot{\theta}(t) = \frac{1}{r(t)}\frac{d}{dt}(r^2(t)\dot{\theta}(t)) \equiv 0 \tag{4.6.18}$$

e ainda que:

(ii)... a equação (diferencial) do movimento se reduz à equação:

$$\ddot{r}(t) - r(t)\dot{\theta}^2(t) = f(r(t))r(t) \tag{4.6.19}$$

Recordando a expressão (4.5.21) do vector velocidade em coordenadas polares, $\dot{\mathbf{r}}(t) = \dot{r}(t)\,\mathbf{e}_r(t) + r(t)\dot{\theta}(t)\,\mathbf{e}_\theta(t)$, deduzimos que $\mathbf{r}(t) \times \dot{\mathbf{r}}(t) = r^2(t)\dot{\theta}(t)\,(\mathbf{e}_r \times \mathbf{e}_\theta)$. Portanto a equação (4.6.18), traduz novamente a conservação do momento angular.

Figure 4.16: Área $A(t)$ varrida pelo vector de posição $\mathbf{r}(t)$

.

Existe além disso, uma outra interpretação interessante, para a equação (4.6.18). Com efeito, representemos por $A(t)$ a área varrida pelo vector de posição $\mathbf{r}(t)$, desde um instante inicial $t = t_o$ até ao instante t (ver a figura 4.16).

Demonstremos que a "velocidade" $\frac{dA}{dt}$, com que varia esta área $A(t)$, é igual a:

$$\frac{dA}{dt} \equiv \dot{A}(t) = \frac{1}{2} r^2(t)\dot{\theta}(t) \tag{4.6.20}$$

De facto, podemos supôr que numa pequena vizinhança de t, a curva é representada por uma equação polar do tipo $r = r(t) = r(\theta(t))$. Assim, se t_o está nessa vizinhança, a área $A(t)$ do conjunto radial entre $\theta_o = \theta(t_o)$ e $\theta = \theta(t)$, é dada por (ver a secção 2.6.4, nomeadamente a fórmula (4.5.31)):

$$A(t) = \int_{\theta_o}^{\theta(t)} \frac{1}{2} r^2(\theta)\, d\theta$$

Derivando este integral, aplicando o teorema fundamental do cálculo e a regra da cadeia, obtemos exactamente (4.6.20), como se pretendia.

Notemos agora (4.6.18), é equivalente a

$$0 = \frac{d}{dt}(r^2(t)\dot{\theta}(t)) = 2\frac{d}{dt}\frac{dA}{dt}$$

o que permite enunciar a seguinte proposição:

♣ **Proposição 4.10** ... *No movimento de uma partícula sujeita à acção de um campo de forças central, o vector de posição* $\mathbf{r}(t)$ *varre área a uma velocidade constante e igual a:*

$$\frac{dA}{dt} \equiv \dot{A}(t) = \frac{1}{2} r^2(t)\dot{\theta}(t) \equiv c \quad (= \textit{constante}) \tag{4.6.21}$$

4.6.4 Leis de Kepler e Lei da atracção universal de Newton

No princípio do século XVII, J.Kepler, a partir de uma grande quantidade de observações astronómicas efectuadas por Tycho Brahe, formulou as seguintes três leis para o movimento dos planetas em torno do sol:

Primeira Lei de Kepler ... *"O vector de posição de cada planeta, relativamente ao sol, varre áreas iguais em tempos iguais"*.

Segunda Lei de Kepler ... *"Cada planeta move-se sobre uma elipse com o sol num dos focos"*.

Terceira Lei de Kepler ... *"O quadrado do período (i.e., o tempo necessário para completar uma órbita) é proporcional ao cubo do eixo maior da elipse, sendo a constante de proporcionalidade a mesmapara todo o planeta"*.

Cerca de 50 anos mais tarde, I.Newton utilizou a sua própria lei: "**Força exercida sobre uma partícula de massa** m=**massa** m", para deduzir, a partir das leis de Kepler, a natureza básica da força que mantém cada planeta na respectiva órbita. Vamos de seguida fazer esta dedução, utilizando os métodos de que já dispomos.

♣ **Proposição 4.11** ... *As duas primeiras Leis de Kepler implicam que a força exercida sobre cada planeta, é uma força central, cujo centro é o sol.*

Além disso, essa força é atractiva (dirigida em direcção ao sol), e a sua intensidade é proporcional a $\frac{1}{r^2}$, onde r é a distância do planeta ao sol.

- Demonstração...

A segunda lei de Kepler implica que o movimento é plano. Nesse plano escolhamos um sistema de eixos cartesianos centrado no sol, e façamos os cálculos em coordenadas polares, supondo que o movimento é descrito pela curva parametrizada (de classe C^∞):

$$t \mapsto (r(t), \theta(t)) \tag{4.6.22}$$

Seja $A(\theta)$ a área varrida pelo vector de posição $\mathbf{r}(t)$, quando este se desloca de t_o a t. Como vimos na secção anterior, a velocidade de variação desta área, é dada por:

$$\dot{A}(t) = \frac{dA}{dt} = \frac{1}{2}r^2(t)\dot{\theta}(t)$$

e a primeira lei de Kepler afirma que:

$$\frac{dA}{dt} = \frac{1}{2}r^2(t)\dot{\theta}(t) \equiv c \quad \text{(constante)}$$

.

Portanto:

$$\frac{d^2A}{dt^2} = r(t)\dot{r}(t)\dot{\theta}(t) + \frac{1}{2}r^2(t)\ddot{\theta}(t) \equiv 0 \tag{4.6.23}$$

Mas a aceleração é dada por (ver a secção 2.6.3, fórmula (4.5.25)):

$$\ddot{\mathbf{r}}(t) = \underbrace{\ddot{r}(t) - r(t)\dot{\theta}^2(t)}_{\text{acel. radial}}\,\mathbf{e}_r + \underbrace{r(t)\ddot{\theta}(t) + 2\dot{r}(t)\dot{\theta}(t)}_{\text{acel. transversa}}\,\mathbf{e}_\theta \tag{4.6.24}$$

Estas duas últimas equações implicam que a aceleração transversa se anula, e portanto:

$$\ddot{\mathbf{r}}(t) = (\ddot{r}(t) - r(t)\dot{\theta}^2(t))\,\mathbf{e}_r \tag{4.6.25}$$

o que, pela Lei de Newton, implica que a força é também radial.

Demonstremos agora a segunda parte da proposição. Para isso suponhamos que a órbita elíptica (de acordo com a segunda lei de Kepler), tem por equação polar (ver (4.5.16)):

$$r = r(\theta) = \frac{\epsilon\, d}{1 + \epsilon\, \cos\theta} \tag{4.6.26}$$

de tal forma que $r(t)$ e $\theta(t)$ satisfazem esta equação $\forall t$.

Pretende-se demonstrar que (atendendo a (4.6.25) e à lei de Newton):

$$\ddot{r}(t) - r(t)\dot{\theta}^2(t) = -\frac{C}{r^2(t)} \tag{4.6.27}$$

onde $C > 0$ é uma constante.

Para efectuar os cálculos de uma forma simples, é conveniente utilizar $u(t) = u(\theta(t)) \equiv \frac{1}{r(t)} = \frac{1}{r(\theta(t))}$.

Como vimos antes, a "velocidade de área" $\dot{A}(t) = \frac{1}{2}r^2(t)\dot{\theta}(t) \equiv c$ (constante), e portanto:

$$r^2(t)\dot{\theta}(t) \equiv 2c \Rightarrow \dot{r}(t) = -2c\frac{du}{d\theta}$$

Derivando novamente, obtemos:

$$\ddot{r}(t) = -4c^2u^2(t)\frac{d^2u}{d\theta^2}$$

e, substituindo na aceleração, obtemos:

$$\ddot{r}(t) - r(t)\dot{\theta}^2(t) = -4c^2u^2(t)\Big(\frac{d^2u}{d\theta^2} + u(t)\Big)$$

Por outro lado, usando $u(t) = \frac{1}{r(t)}$ em (4.6.26), e derivando, obtemos:

$$\frac{d^2u}{d\theta^2} + u(t) = \frac{1}{\epsilon d}$$

de tal forma que:

$$\begin{aligned} \ddot{r}(t) - r(t)\dot{\theta}^2(t) &= \frac{-4c^2}{\epsilon d}u^2(t) \\ &= -\frac{4c^2}{\epsilon d}\frac{1}{r^2(t)} \end{aligned} \tag{4.6.28}$$

e portanto $C = \frac{4c^2}{\epsilon d} > 0$, o que termina a demonstração,

♣.

♣ **Proposição 4.12** *(***Lei da atracção universal de Newton***)*

... As Leis de Kepler implicam que existe uma constante $k > 0$, a mesma para todo o planeta, tal que a aceleração radial é igual a $-\frac{k}{r^2}$. Portanto, a intensidade da força de atracção exercida sobre cada planeta, é igual a:

$$-\frac{k\,m}{r^2} \tag{4.6.29}$$

isto é, é proporcional à massa m desse planeta, e proporcional ao inverso do quadrado da sua distância ao sol.

- Demonstração...

O eixo maior $2a$, da elipse dada por (4.6.26), é igual à soma dos valores máximo e mínimo de r:

$$2a = r_{\max} + r_{\min} = \frac{\epsilon d}{1-\epsilon} + \frac{\epsilon d}{1+\epsilon} = \frac{2\epsilon d}{1-\epsilon^2}$$

Representemos o período de revolução por T (=tempo necessário para completar uma órbita), e por A, a área da elipse (4.6.26). Como já sabemos, a área é varrida a uma velocidade constante: $\dot{A}(t) = \frac{1}{2}r^2(t)\dot{\theta}(t) \equiv c$ (constante). Portanto a área total A, da elipse é igual a $A = cT$, já que é necessário um período T para varrer toda a elipse. Logo:

$$T = \frac{A}{c} \tag{4.6.30}$$

Por outro lado $A = \frac{\pi\epsilon^2 d^2}{(1-\epsilon^2)^{3/2}}$, e portanto:

$$T = \frac{\pi\epsilon^2 d^2}{c(1-\epsilon^2)^{3/2}} \tag{4.6.31}$$

Atendendo agora à fórmula para $2a$, obtemos que:

$$\frac{a^3}{T^2} = \frac{8\epsilon^3 d^3}{(1-\epsilon^2)^3}\frac{c^2(1-\epsilon^2)^3}{\pi^2\epsilon^4 d^4} = \frac{8c^2}{\pi^2\epsilon d}$$

e substituindo este resultado na fórmula (4.6.28), para a aceleração radial de um planeta a uma distância r do sol, concluímos que:

$$\begin{aligned}\ddot{r}(t) - r(t)\dot{\theta}^2(t) &= -\frac{4c^2}{\epsilon d}\frac{1}{r^2(t)} \\ &= -(\frac{\pi^2}{2}\frac{a^3}{T^2})\frac{1}{r^2(t)}\end{aligned} \tag{4.6.32}$$

Finalmente a terceira lei de Kepler afirma que $\frac{a^3}{T^2}$, é o mesmo para cada planeta, o que atendendo a (4.6.32), termina a demonstração,

♣.

É usual tomar a constante k, referida na proposição anterior, igual a $G\,M$, onde M é a massa do sol e G é a chamada constante de gravitação universal. Resumindo, a **Lei da atracção universal**

de Newton afirma que a força central de atracção exercida sobre um planeta de massa m, é dada por:

$$\begin{aligned} \mathbf{F}(\mathbf{r}) &= -G\frac{Mm}{r^2}\,\mathbf{e}_r & (4.6.33)\\ &= -G\frac{Mm}{r^3}\,\mathbf{r} & (4.6.34) \end{aligned}$$

Recìprocamente, é possível deduzir as leis de Kepler a partir da lei de atracção universal de Newton e da lei de Newton (ver Apostol, vol. 1, secção 14.20).

4.7 Integrais de linha

4.7.1 Integral de um campo escalar ao longo de uma curva

Antes de definirmos os chamados integrais de linha, convém generalizar um pouco o conceito de curva (regular) parametrizada de classe C^m.

♣ **Definição 4.15** ... *Uma função $\alpha : \mathbf{I} \subseteq \mathbb{R} \to \mathbb{R}^n$, diz-se uma* **curva** *(parametrizada)* **seccionalmente** *(de classe) C^m, se α é de classe C^m em $\mathbf{I}$, com a possível excepção de um número finito de pontos em $\mathbf{I}$.*

Todos os conceitos da secção 2.3, se generalizam de forma natural para curvas (parametrizadas) seccionalmente C^m. Assim por exemplo, permitimos mudanças admissíveis de parâmetro $\phi : \mathbf{J} \to \mathbf{I}$, seccionalmente C^m, i.e., ϕ é de classe C^m em $\mathbf{J}$, com a possível excepção de um número finito de pontos em $\mathbf{J}$, e com derivada ϕ' sempre > 0 (ou < 0), nos pontos em que existir. Com mudanças admissíveis de parâmetro deste tipo, podemos definir curvas seccionalmente C^m, através de uma relação de equivalência análoga à utilizada na secção 2.3. Por exemplo, uma curva regular seccionalmente C^1, será uma classe de equivalência de curvas parametrizadas regulares seccionalmente C^1.

Suponhamos agora que $f : \mathcal{U} \subseteq \mathbb{R}^n \to \mathbb{R}$, é uma função real de variável vectorial (isto é, f é um **campo escalar**), definida num subconjunto $\mathcal{U} \subseteq \mathbb{R}^n$, que contem o traço de uma curva parametrizada seccionalmente C^1, $\alpha : [a, b] \to \mathcal{U}$.

Imaginemos por exemplo, que $f(\alpha(t))$ representa a densidade (por unidade de comprimento) em $\alpha(t)$, do material de que é feito o "fio" α. Dada uma partição $\mathcal{P} = \{a = t_o, t_1, ..., t_{k-1}, t_k = b\}$ do intervalo $[a, b]$, é natural considerar a soma:

$$m(\alpha, f, \mathcal{P}) = \sum_{i=1}^{k} f(\alpha(t_i))\, \|\alpha(t_i) - \alpha(t_{i-1})\|$$

como uma aproximação para o valor da massa total do referido "fio".

Utilizando argumentos análogos aos que foram expostos na (ideia da) demonstração da fórmula (4.3.5), na seccao 2.4, podemos garantir que, sob certas condições (por exemplo, se $f \circ \alpha$ é contínua), o limite:

$$\lim_{|\mathcal{P}| \to 0} m(\alpha, f, \mathcal{P})$$

existe e é dado pelo integral:

$$\int_a^b f(\alpha(t)) \|\dot{\alpha}(t)\| \, dt$$

que por isso, pode ser tomado como o valor exacto da massa do referido "fio".

Isto motiva a seguinte definição:

♣ **Definição 4.16** ... *Seja* $f : \mathcal{U} \subseteq \mathbb{R}^n \to \mathbb{R}$, *um* **campo escalar**, *definido num subconjunto* $\mathcal{U} \subseteq \mathbb{R}^n$, *que contem o traço de uma curva parametrizada seccionalmente* C^1, $\alpha : [a, b] \to \mathcal{U}$. *Suponhamos que* $f \circ \alpha$ *é contínua.*

Define-se então o **integral de** f **ao longo de** α, *notado por* $\int_\alpha f \, ds$, *através de:*

$$\int_\alpha f \, ds \equiv \int_a^b f(\alpha(t)) \|\dot{\alpha}(t)\| \, dt \tag{4.7.1}$$

Quando $f \equiv 1$, recuperamos o conceito de comprimento de arco, discutido na seccao 2.4.

Podemos demonstrar, de forma completamente análoga à utilizada na proposição 13 (baseada na "fórmula da mudança de variáveis", em integrais de funções reais de variável real), que $\int_\alpha f \, ds$ não depende da forma como α está parametrizada, como aliás seria de prever.

Por exemplo, se α é uma curva plana, $\alpha(t) = (x(t), y(t)) \in \mathbb{R}^2$, e f é um campo escalar em $\mathbb{R}^2$, então:

$$\int_\alpha f \, ds = \int_a^b f(x(t), y(t)) \sqrt{\dot{x}(t) + \dot{y}(t)} \, dt \tag{4.7.2}$$

Quando $f(x, y) \geq 0$, este integral representa a área de uma faixa vertical cuja base é o traço de α, e a altura, num ponto (x, y) nesse traço, é igual a $f(x, y)$.

4.7.2 Integral de um campo de vectores ao longo de uma curva. Trabalho

Passemos agora à definição de um outro tipo de integral de linha, muito importante nas aplicações.

Para motivar essa definição, imaginemos que $\mathbf{F}$ é um campo de forças definido num certo subconjunto $\mathcal{U} \subseteq \mathbb{R}^n$, de tal forma que $\mathbf{F}(\mathbf{x})$ representa a força exercida sobre uma partícula teste de massa 1, situada em $\mathbf{x}$. Quando esse campo de forças $\mathbf{F}$, actua sobre uma partícula de massa m, deslocando-a ao longo de uma certa trajectória, realiza então uma certa quantidade de energia, a que chamamos **trabalho**.

Por exemplo, se uma partícula de massa m se move uma distância d, ao longo de uma linha recta, sob a acção de uma força de intensidade F (por unidade de massa), com a mesma direcção e sentido desse movimento, então o trabalho realizado é por definição, igual a $m \, F \, d$.

Ainda um outro exemplo: seja $\mathbf{F}$ um campo de forças uniforme (constante), definido num certo subconjunto $\mathcal{U} \subseteq \mathbb{R}^n$, e suponhamos que $\mathbf{F}$ actua sobre uma partícula de massa m, deslocando-a desde um ponto $\mathbf{p}$ até um ponto $\mathbf{q}$, ao longo do segmento de recta que une esses dois pontos. O trabalho realizado por $\mathbf{F}$ para efectuar esse deslocamento, deve apenas depender da componente de $\mathbf{F}$ na direcção do referido movimento. Portanto, se $\mathbf{q} - \mathbf{p} = d\,\mathbf{u}$, onde $\mathbf{u}$ é um vector unitário, esse trabalho é igual a:

$$m\,(\mathbf{F} \cdot \mathbf{u})\,d = m\,(\mathbf{F} \cdot (\mathbf{q} - \mathbf{p})) \tag{4.7.3}$$

já que $\mathbf{F} \cdot \mathbf{u}$, é exactamente a componente de $\mathbf{F}$ na direcção do referido movimento.

Retomemos a situação geral, em que $\mathbf{F}$ é um campo de forças definido num certo subconjunto $\mathcal{U} \subseteq \mathbb{R}^n$, que contem o traço de uma curva parametrizada seccionalmente C^1, $\alpha : [a, b] \to \mathcal{U}$, e suponhamos que se pretende calcular o trabalho realizado por $\mathbf{F}$, no deslocamento de uma partícula de massa $m = 1$, ao longo da trajectória representada por α. Representemos por $W(\alpha, \mathbf{F})$ esse trabalho.

Para calcular $W(\alpha, \mathbf{F})$, procedendo como antes, consideramos uma partição $\mathcal{P} = \{a = t_o, t_1, ..., t_{k-1}, t_k = b\}$ do intervalo $[a, b]$, e é natural afirmar que a soma:

$$W(\alpha, \mathbf{F}, \mathcal{P}) = \sum_{i=1}^{k} \mathbf{F}(\alpha(t_i)) \cdot \, (\alpha(t_i) - \alpha(t_{i-1}))$$

é uma aproximação para o valor desse trabalho, $W(\alpha, \mathbf{F})$.

Mais uma vez, utilizando argumentos análogos aos que foram expostos na (ideia da) demonstração da fórmula (4.3.5), na seccao 2.4, podemos garantir que, sob certas condições (por exemplo, se a função $t \mapsto \mathbf{F}(\alpha(t)) \cdot \dot{\alpha}(t)$ é contínua), o limite:

$$\lim_{|\mathcal{P}| \to 0} W(\alpha, \mathbf{F}, \mathcal{P})$$

existe e é dado pelo integral:

$$\int_a^b \mathbf{F}(\alpha(t)) \cdot \dot{\alpha}(t) dt$$

que por isso, pode ser tomado como o valor exacto referido trabalho.

Isto motiva a seguinte definição:

♣ **Definição 4.17** ... *Seja* $\mathbf{F} : \mathcal{U} \subseteq \mathbb{R}^n \to \mathbb{R}^n$, *um* **campo de vectores**, *definido num subconjunto* $\mathcal{U} \subseteq \mathbb{R}^n$, *que contem o traço de uma curva parametrizada seccionalmente* C^1, $\alpha : [a, b] \to \mathcal{U}$. *Suponhamos que a função* $t \mapsto \mathbf{F}(\alpha(t)) \cdot \dot{\alpha}(t)$ *é contínua em* $[a, b]$.

Define-se então o **integral de F ao longo de** α, *notado por* $\int_\alpha \mathbf{F} \cdot d\mathbf{s}$, *através de:*

$$\int_\alpha \mathbf{F} \cdot d\mathbf{s} \equiv \int_a^b \mathbf{F}(\alpha(t)) \cdot \dot{\alpha}(t)\, dt \tag{4.7.4}$$

Exemplo ...

Suponhamos novamente que uma partícula de massa m, sob a influência de um campo de forças $\mathbf{F}$, se desloca ao longo de uma curva $\alpha : [a, b] \to \mathbb{R}^3$. Seja $v(t) = \|\dot{\alpha}(t)\| = \|\mathbf{v}(t)\|$, a velocidade (escalar) no instante t.

Define-se então a **energia cinética** dessa partícula através de:

$$E(t) \equiv \frac{1}{2} m\, v^2(t) \tag{4.7.5}$$

Vamos provar que a variação da energia cinética da referida partícula, no intervalo de tempo $[a, b]$, é igual ao trabalho realizado por $\mathbf{F}$, ao deslocar a partícula ao longo de α, desde $\alpha(a)$ até $\alpha(b)$.

Como por definição, esse trabalho é igual a:

$$W(\alpha, \mathbf{F}) = \int_a^b m\, \mathbf{F}(\alpha(t)) \cdot \dot{\alpha}(t) dt$$

a afirmação traduz-se em que:

$$\begin{aligned} W(\alpha, \mathbf{F}) &= E(b) - E(a) \\ &= \frac{1}{2} m\, v^2(b) - \frac{1}{2} m\, v^2(a) \end{aligned} \tag{4.7.6}$$

De facto, de acordo com a lei de Newton:

$$m\mathbf{F}(\alpha(t)) = m\, \ddot{\alpha}(t) = m\, \dot{\mathbf{v}}(t)$$

e portanto:

$$\mathbf{F}(\alpha(t)) \cdot \dot{\alpha}(t) = \mathbf{F}(\alpha(t)) \cdot \mathbf{v}(t) = \dot{\mathbf{v}}(t) \cdot \mathbf{v}(t) = \frac{1}{2} \frac{d}{dt} (\mathbf{v}(t) \cdot \mathbf{v}(t)) = \frac{1}{2} \frac{d}{dt} v^2(t)$$

Substituindo isto no integral, vem que:

$$W(\alpha, \mathbf{F}) = m \int_a^b \mathbf{F}(\alpha(t)) \cdot \dot{\alpha}(t) dt = m \int_a^b \frac{1}{2} \frac{d}{dt} v^2(t) = \frac{1}{2} m\, v^2(b) - \frac{1}{2} m\, v^2(a)$$

como se pretendia.

Em componentes, se:

$$\mathbf{F}(\mathbf{x}) = (F^1(\mathbf{x}), F^2(\mathbf{x}), ..., F^n(\mathbf{x})) \qquad \mathbf{x} \in \mathcal{U} \subseteq \mathbb{R}^n$$

e se:

$$\alpha(t) = (x^1(t), x^2(t), ..., x^n(t)) \qquad t \in [a, b]$$

então o integral (4.7.4), escreve-se na forma:

$$\int_\alpha \mathbf{F} \cdot d\mathbf{s} = \int_a^b \left(F^1(\alpha(t)) \frac{dx^1}{dt} + F^2(\alpha(t)) \frac{dx^2}{dt} + ... + F^n(\alpha(t)) \frac{dx^n}{dt} \right) dt \tag{4.7.7}$$

ou ainda na forma simplificada:

$$\int_\alpha \mathbf{F} \cdot d\mathbf{s} = \int_\alpha F^1\, dx^1 + F^2\, dx^2 + ... + F^n\, dx^n \tag{4.7.8}$$

Quando α é regular ($\dot{\alpha}(t) \neq \mathbf{O} \ \ \forall t$), podemos escrever o integral (4.7.4) numa outra forma útil. Assim, se designamos por $\mathbf{t}(t) = \frac{\dot{\alpha}(t)}{\|\dot{\alpha}(t)\|}$, o vector unitário tangente à curva α, temos que:

$$\begin{aligned}
\int_\alpha \mathbf{F} \cdot d\mathbf{s} &= \int_a^b \mathbf{F}(\alpha(t)) \cdot \dot{\alpha}(t)\, dt \\
&= \int_a^b \Big(\mathbf{F}(\alpha(t)) \cdot \frac{\dot{\alpha}(t)}{\|\dot{\alpha}(t)\|}\Big) \|\dot{\alpha}(t)\|\, dt \\
&= \int_a^b (\mathbf{F}(\alpha(t)) \cdot \mathbf{t}(t))\, \|\dot{\alpha}(t)\|\, dt \\
&= \int_a^b (\mathbf{F}(\alpha(t)) \cdot \mathbf{t}(t))\, ds \\
&= \int_\alpha \mathbf{F} \cdot \mathbf{t}\, ds \qquad (4.7.9)
\end{aligned}$$

isto é, o integral do campo de vectores $\mathbf{F}$ ao longo de α, é igual ao integral da função $\mathbf{F} \cdot \mathbf{t}$ (="componente tangencial de $\mathbf{F}$, ao longo de α"), ao longo de α.

Como veremos num capítulo posterior, quando $\mathbf{V}$ é um campo de vectores que representa o **campo de velocidades** de um fluido, então o integral $\int_\alpha \mathbf{V} \cdot d\mathbf{s}$, de $\mathbf{V}$ ao longo da curva regular α, representa a chamada **circulação** do referido fluido ao longo de α. Com efeito, (4.7.9) diz que $\int_\alpha \mathbf{V} \cdot d\mathbf{s}$, é igual ao integral da componente tangencial da velocidade $\mathbf{V}$, relativamente ao comprimento de arco.

Vejamos agora algumas propriedades dos integrais de linha.

Em primeiro lugar, podemos demonstrar, de forma completamente análoga à utilizada na proposição 13 (baseada na "fórmula da mudança de variáveis", em integrais de funções reais de variável real), a seguinte proposição:

♣ **Proposição 4.13** ... *Seja $\mathbf{F} : \mathcal{U} \subseteq \mathbb{R}^n \to \mathbb{R}^n$, um campo de vectores, definido num subconjunto $\mathcal{U} \subseteq \mathbb{R}^n$, que contem o traço de uma curva parametrizada seccionalmente C^1, $\alpha : [a,b] \to \mathcal{U}$, e suponhamos que o integral (4.7.4) está definido.*

Seja $\phi : [c,d] \to [a,b]$, uma mudança admissível de parâmetro (ϕ é bijectiva, de classe C^1 e $\phi' \neq 0$). Então:

$$(i)\ldots \int_{\alpha\circ\phi} \mathbf{F} \cdot d\mathbf{s} = \int_\alpha \mathbf{F} \cdot d\mathbf{s} \qquad \text{se } \phi' > 0 \qquad (4.7.10)$$

$$(ii)\ldots \int_{\alpha\circ\phi} \mathbf{F} \cdot d\mathbf{s} = -\int_\alpha \mathbf{F} \cdot d\mathbf{s} \qquad \text{se } \phi' < 0 \qquad (4.7.11)$$

Consideremos agora uma curva (parametrizada) $\alpha : [a,b] \to \mathbb{R}^n$, e um ponto $c \in [a,b]$. Designando por α^1 e por α^2, respectivamente as curvas $\alpha^1 = \alpha\,|_{[a,c]}$ e $\alpha^2 = \alpha\,|_{[c,b]}$, diremos que α é a **justaposição** das curvas α^1 e α^2, e escrevemos $\alpha = \alpha^1 + \alpha^2$.

Mais geralmente, se $\mathcal{P} = \{t_o < t_1 < \ldots < t_k\}$ é uma partição de $[a,b]$, e se $\alpha^j = \alpha\,|_{[t_{j-1},t_j]}$ ($j = 1, 2, \ldots, k$), então diremos que α é a **justaposição** das curvas $\alpha^1, \alpha^2, \ldots, \alpha^k$, e escrevemos $\alpha = \alpha^1 + \alpha^2 + \ldots + \alpha^k$.

Nota...

Dadas duas curvas (parametrizadas) $\alpha : [a, b] \to \mathbb{R}^n$ e $\beta : [c, d] \to \mathbb{R}^n$, tais que $\alpha(b) = \beta(c)$, podemos definir a justaposição $\alpha + \beta$, da seguinte forma: reparametrizamos β no intervalo $[b, b + (d - c)]$ (isto é, consideramos uma mudança admissível de parâmetro $\phi : [b, b + (d - c)] \to [c, d]$, e a reparametrização $\hat{\beta} = \beta \circ \phi$), e definimos a justaposição $\alpha + \beta$, como sendo $\alpha + \hat{\beta}$. Esta definição generaliza-se naturalmente, para a justaposição de várias curvas.

Ainda uma outra definição: dada uma curva (parametrizada) $\alpha : [a, b] \to \mathbb{R}^n$, define-se o **inverso** de α, notado por $-\alpha$, como sendo a curva $-\alpha : [a, b] \to \mathbb{R}^n$, dada por:

$$-\alpha(t) = \alpha(a + b - t) \qquad t \in [a, b]$$

Posto isto, podemos enunciar a seguinte proposição, cuja demonstração é imediata a partir das definições:

♣ Proposição 4.14 ... *Seja $\mathbf{F} : \mathcal{U} \subseteq \mathbb{R}^n \to \mathbb{R}^n$, um campo de vectores, definido num subconjunto $\mathcal{U} \subseteq \mathbb{R}^n$, que contem o traço de uma curva parametrizada seccionalmente C^1, $\alpha : [a, b] \to \mathcal{U}$, e suponhamos que o integral (4.7.4) está definido.*

(i)... Se α é a **justaposição** *das curvas $\alpha^1, \alpha^2, ..., \alpha^k$: $\alpha = \alpha^1 + \alpha^2 + ... + \alpha^k$, então:*

$$\int_{\alpha^1+\alpha^2+...+\alpha^k} \mathbf{F} \cdot d\mathbf{s} = \sum_{j=1}^{k} \int_{\alpha^j} \mathbf{F} \cdot d\mathbf{s} \tag{4.7.12}$$

(ii)...

$$\int_{-\alpha} \mathbf{F} \cdot d\mathbf{s} = \int_{\alpha} \mathbf{F} \cdot d\mathbf{s} \tag{4.7.13}$$

4.7.3 Campos conservativos. Princípio da conservação da energia mecânica

Nesta secção todas as curvas mencionadas serão seccionalmente C^1. Uma curva (parametrizada) $\alpha : [a, b] \to \mathbb{R}^n$, diz-se fechada (ou um lacete) se $\alpha(a) = \alpha(b)$.

♣ Definição 4.18 ... *Um campo de vectores $\mathbf{F} : \mathcal{U} \subseteq \mathbb{R}^n \to \mathbb{R}^n$, definido num subconjunto $\mathcal{U} \subseteq \mathbb{R}^n$, diz-se* **conservativo** *se:*

$$W(\alpha, \mathbf{F}) \equiv \int_{\alpha} \mathbf{F} \cdot d\mathbf{s} = 0 \quad \forall \text{ curva fechada } \alpha, \text{ cujo traço está em } \mathcal{U} \tag{4.7.14}$$

Exemplo ...

Consideremos o campo de vectores:

$$\mathbf{F}(x,y) = \Big(\frac{-y}{x^2+y^2}, \frac{x}{x^2+y^2}\Big)$$

definido em $\mathcal{U} = \mathbb{R}^2 - \{\mathbf{O}\}$.

$\mathbf{F}$ não é conservativo. Com efeito, se calculamos o integral de $\mathbf{F}$, ao longo da circunferência de raio um:

$$\alpha(t) = (\cos t, \sin t) \quad 0 \leq t \leq 2\pi$$

obtemos:

$$\int_\alpha \mathbf{F} \cdot d\mathbf{s} = \int_0^{2\pi} \mathbf{F}(\alpha(t)) \cdot \dot{\alpha}(t)\, dt = \int_0^{2\pi} (\sin^2 t + \cos^2 t)\, dt = 2\pi \neq 0$$

.

Em geral, quando $\mathbf{F}$ é um campo de forças, então $\mathbf{F}$ é conservativo se e só se o trabalho realizado por $\mathbf{F}$, ao deslocar uma partícula ao longo de uma qualquer curva fechada α, é nulo: "$W(\alpha, \mathbf{F}) = 0$, $\forall$ curva fechada α". De forma equivalente, se e só se a variação de energia cinética da referida partícula, nesse deslocamento, é nula.

Quando $\mathbf{F}$ é conservativo, o trabalho realizado ao deslocar uma partícula, desde um ponto $\mathbf{p}_o$ até um ponto $\mathbf{p}$, não depende da curva α, ao longo da qual se faz esse deslocamento. De facto, sejam α, β duas curvas que unem p_o a $\mathbf{p}$. Então a curva $\alpha + (-\beta)$, obtida justapondo α ao inverso de β, é uma curva fechada. Portanto, pelas propriedades (4.7.12) e (4.7.13), vem que:

$$0 = W(\alpha + (-\beta), \mathbf{F}) = W(\alpha, \mathbf{F}) - W(\beta, \mathbf{F}) \Rightarrow W(\alpha, \mathbf{F}) = W(\beta, \mathbf{F}) \tag{4.7.15}$$

como se tinha afirmado.

♣ **Definição 4.19** ... *Seja $\mathbf{F} : \mathcal{U} \subseteq \mathbb{R}^n \to \mathbb{R}^n$, um campo de vectores, definido num subconjunto $\mathcal{U} \subseteq \mathbb{R}^n$.*

Uma função escalar $U : \mathcal{U} \subseteq \mathbb{R}^n \to \mathbb{R}$, diz-se um **potencial** *para o campo de vectores $\mathbf{F}$, se dados dois quaisquer pontos $\mathbf{p}_o, \mathbf{p} \in \mathcal{U}$, se tem que:*

$$W(\alpha, \mathbf{F}) = \int_\alpha \mathbf{F} \cdot d\mathbf{s} = U(\mathbf{p}) - U(\mathbf{p}_o) \tag{4.7.16}$$

para toda a curva α, que une $\mathbf{p}_o$ a $\mathbf{p}$ e cujo traço está contido em $\mathcal{U}$.

♣ **Proposição 4.15** ... *Seja $\mathcal{U} \subseteq \mathbb{R}^n$ um subconjunto* **conexo por arcos**, *em $\mathbb{R}^n$ (isto é, dados dois quaisquer pontos $\mathbf{p}_o, \mathbf{p} \in \mathcal{U}$, existe uma curva seccionalmente C^1, que une $\mathbf{p}_o$ a $\mathbf{p}$ e cujo traço está contido em $\mathcal{U}$.)*

(i)... Um campo de vectores $\mathbf{F} : \mathcal{U} \subseteq \mathbb{R}^n \to \mathbb{R}^n$, é conservativo se e só se existe um potencial $U : \mathcal{U} \subseteq \mathbb{R}^n \to \mathbb{R}$, para $\mathbf{F}$.

(ii)... Dois potenciais para um mesmo campo de vectores $\mathbf{F}$, diferem por uma constante.

• Demonstração...

(i)... É óbvio que se existe um potencial para $\mathbf{F}$, então $\mathbf{F}$ é conservativo.

Recìprocamente, suponhamos que $\mathbf{F}$ é conservativo. Fixemos um qualquer ponto $\mathbf{p}_o \in \mathcal{U}$. Seja $\mathbf{p} \in \mathcal{U}$, um ponto arbitrário em $\mathcal{U}$, e α uma curva qualquer (seccionalmente C^1), que une $\mathbf{p}_o$ a $\mathbf{p}$ e cujo traço está contido em $\mathcal{U}$. Como $\mathcal{U}$ é conexo por arcos, existe sempre uma curva nessas condições.

Definamos $U : \mathcal{U} \subseteq \mathbb{R}^n \to \mathbb{R}$, através de:

$$U(p) \equiv W(\alpha, \mathbf{F}) = \int_\alpha \mathbf{F} \cdot d\mathbf{s} \tag{4.7.17}$$

Pelas observações que fizemos acima, U está bem definido (não depende da curva α que une $\mathbf{p}_o$ a $\mathbf{p}$). Resta provar que U é um potencial para $\mathbf{F}$. Para isso, consideremos um outro ponto arbitrário $\mathbf{q} \in \mathcal{U}$, uma curva qualquer β, em $\mathcal{U}$, que une $\mathbf{p}_o$ a $\mathbf{q}$, e ainda uma curva qualquer γ, em $\mathcal{U}$, que une $\mathbf{p}$ a $\mathbf{q}$. Então $\alpha + \gamma + (-\beta)$ é uma curva fechada, e portanto:

$$W(\alpha + \gamma + (-\beta), \mathbf{F}) = 0 \quad \Rightarrow \quad W(\gamma, \mathbf{F}) = W(\beta, \mathbf{F}) - W(\alpha, \mathbf{F})$$

isto é, atendendo a (4.7.17):

$$W(\gamma, \mathbf{F}) = U(\mathbf{q}) - U(\mathbf{p})$$

o que significa que U é um potencial para $\mathbf{F}$.

(ii)... Se U' é um outro potencial para $\mathbf{F}$, e se α é uma curva qualquer, em $\mathcal{U}$, que une $\mathbf{p}_o$ a $\mathbf{p}$, então:

$$W(\alpha, \mathbf{F}) = U'(\mathbf{p}) - U'(\mathbf{p}_o)$$

Mas por definição (4.7.17), $U(\mathbf{p}) = W(\alpha, \mathbf{F})$, e portanto:

$$U(\mathbf{p}) = U'(\mathbf{p}) - U'(\mathbf{p}_o)$$

isto é, U' e U diferem pela constante $U'(\mathbf{p}_o)$,

♣.

Exemplo ...

Recordemos que a **Lei da atracção universal de Newton**, afirma que a força central de atracção, que uma partícula de massa M (o sol), situada na origem do sistema de coordenadas, exerce sobre uma outra de massa m (um planeta), é dada por:

$$\mathbf{F}(\mathbf{r}) = -G\frac{Mm}{r^3}\,\mathbf{r} = -\frac{k}{r^3}\mathbf{r} \tag{4.7.18}$$

pondo $k = G\,M$.

Vamos provar que este campo é conservativo com um potencial dado por:

$$U(\mathbf{r}) = \frac{k}{\|\mathbf{r}\|} = \frac{k}{r} \tag{4.7.19}$$

Para isso, consideremos dois quaisquer pontos $\mathbf{p}$ e $\mathbf{q}$, em $\mathbb{R}^3 - \{\mathbf{O}\}$, e uma curva qualquer $\alpha : [a, b] \to \mathbb{R}^3 - \{\mathbf{O}\}$, que os une, i.e., $\alpha(a) = \mathbf{p}$ e $\alpha(b) = \mathbf{q}$. Pretende-se provar que:

$$W(\alpha, \mathbf{F}) = U(\mathbf{q}) - U(\mathbf{p}) \tag{4.7.20}$$

isto é, que:

$$\begin{aligned}
\int_\alpha \mathbf{F} \cdot d\mathbf{s} &= \int_a^b -\frac{k}{\|\alpha(t)\|^3} \alpha(t) \cdot \dot{\alpha}(t)\, dt \\
&= \frac{k}{\|\alpha(b)\|} - \frac{k}{\|\alpha(a)\|} \\
&= \frac{k}{\|\mathbf{q}\|} - \frac{k}{\|\mathbf{p}\|}
\end{aligned} \tag{4.7.21}$$

Mas:

$$\frac{d}{dt} \frac{1}{\|\alpha(t)\|} = -\frac{\alpha(t)}{\|\alpha(t)\|^3} \cdot \dot{\alpha}(t)$$

como pode ser verificado directamente. Portanto:

$$\begin{aligned}
\int_\alpha \mathbf{F} \cdot d\mathbf{s} &= \int_a^b -\frac{k}{\|\alpha(t)\|^3} \alpha(t) \cdot \dot{\alpha}(t)\, dt \\
&= \int_a^b \frac{d}{dt} \frac{k}{\|\alpha(t)\|} \\
&= \frac{k}{\|\alpha(b)\|} - \frac{k}{\|\alpha(a)\|}
\end{aligned}$$

como se pretendia.

Exemplo ...

Consideremos novamente um campo de forças $\mathbf{F}$, definido num subconjunto conexo por arcos $\mathcal{U} \subseteq \mathbb{R}^n$, e suponhamos que $\mathbf{F}$ admite um potencial U. Então, por definição, o trabalho realizado ao deslocar uma partícula de um ponto $\mathbf{p}$ até um ponto $\mathbf{q}$, não depende da curva α que se utilizou nesse deslocamento. De facto, esse trabalho é dado exactamente pela diferença do potencial em $\mathbf{q}$ e $\mathbf{p}$:

$$W(\alpha, \mathbf{F}) = U(\mathbf{q}) - U(\mathbf{p}) \qquad \text{independente de } \alpha$$

Por outro lado, como vimos no exemplo da secção anterior, esse mesmo trabalho é igual à variação de energia cinética (ver a fórmula (4.7.6)):

$$\begin{aligned}
W(\alpha, \mathbf{F}) &= E(b) - E(a) \\
&= \frac{1}{2} m\, v^2(b) - \frac{1}{2} m\, v^2(a)
\end{aligned}$$

Pondo $\alpha : [a, b] \to \mathcal{U}$, com $\alpha(a) = \mathbf{p}$ e $\alpha(b) = \mathbf{q}$, e ainda $E(b) = E(\mathbf{q})$ e $E(a) = E(\mathbf{p})$, vemos portanto que:

$$U(\mathbf{q}) - U(\mathbf{p}) = W(\alpha, \mathbf{F}) = E(\mathbf{q}) - E(\mathbf{p})$$

isto é:

$$E(\mathbf{q}) - U(\mathbf{q}) = E(\mathbf{p}) - U(\mathbf{p}) \qquad \forall \mathbf{p}, \mathbf{q} \in \mathcal{U} \tag{4.7.22}$$

É usual chamar a $-U$, a **energia potencial** da partícula, e à soma $E + (-U)$, a respectiva **energia total**. Portanto, a fórmula (4.7.22), traduz o seguinte **Princípio da conservação da energia mecânica**: "A energia total $E + (-U)$ de uma partícula que se move sob a influência de um campo de forças conservativo, permanece sempre constante".

Capítulo 5

Fluxos. Campos de Vectores. Equações diferenciais

5.0.4 Fluxos e Campos de Vectores

Imaginemos um fluido em movimento numa certa região $\mathcal{U}$ de $\mathbb{R}^3$. Cada partícula do fluido descreve uma certa trajectória, isto é, para cada $\mathbf{x} \in \mathcal{U}$, existirá uma curva que representa a trajectória de $\mathbf{x}$.

Com o objectivo de estudar o movimento do fluido, englobando as trajectórias de todas as suas partículas, notemos por:

$$\mathbf{x} \equiv \Phi(t, \mathbf{x}_o)$$

a trajectória da partícula que no instante inicial $t = 0$, ocupava a posição $\mathbf{x}_o$ em $\mathcal{U}$.

Portanto, para cada $\mathbf{x}_o$ fixo, $\mathbf{x} = \Phi(t, \mathbf{x}_o) \in \mathcal{U}$ representa a posição "espacial", no instante t, da partícula que no instante inicial $t = 0$, ocupava a posição $\mathbf{x}_o$ em $\mathcal{U}$ (ver a figura 5.1).

A aplicação Φ diz-se a **aplicação de fluxo do fluido**, ou simplesmente o fluxo do fluido.

Figure 5.1: Trajectória de $\mathbf{x}_o$

.

Para cada $\mathbf{x}_o \in \mathcal{U}$ fixo, podemos então definir a curva parametrizada:

$$\alpha_{\mathbf{x}_o}(t) \equiv \Phi(t, \mathbf{x}_o) \qquad \mathbf{x}_o \text{ fixo} \tag{5.0.1}$$

que representa a trajectória da partícula que no instante inicial $t = 0$, ocupa a posição $\mathbf{x}_0$:$\alpha_{\mathbf{x}_o}(0) = \Phi(0, \mathbf{x}_o) = \mathbf{x}_o$.

Por outro lado, para cada t fixo, podemos definir uma aplicação $\phi_t : \mathcal{U} \to \mathcal{U}$, através de:

$$\phi_t(\mathbf{x}_o) \equiv \Phi(t, \mathbf{x}_o) \qquad t \text{ fixo} \tag{5.0.2}$$

tal que $\phi_0 = \mathrm{Id}$. É natural supôr ainda que cada ϕ_t é uma aplicação bijectiva (duas partículas não podem ocupar a mesma posição no mesmo instante de tempo).

É usual chamar a $\mathbf{x}_o$ **coordenada material**, e a $\mathbf{x} = \Phi(t, \mathbf{x}_o) = \phi_t(\mathbf{x}_o)$ **coordenada espacial**. As coordenadas espacial e material coincidem $t = 0$, e o facto de ϕ_t ser uma bijecção implica que podemos conhecer a posição inicial da partícula, conhecendo a sua posição num qualquer instante t.

Por definição o **campo de velocidades do fluido** é o campo de vectores $\mathbf{v}(t, \mathbf{x})$, dependente do tempo t, e que no ponto $\mathbf{x} = \Phi(t, \mathbf{x}_o)$ é igual ao vector velocidade, no instante t, da curva $\alpha_{\mathbf{x}_o}(t)$ (ver a figura 5.2):

$$\mathbf{v}(t, \Phi(t, \mathbf{x}_o)) \equiv \dot{\alpha}_{\mathbf{x}_o}(t) \tag{5.0.3}$$

Isto é, o campo de velocidades do fluido, no instante t e no ponto $\mathbf{x} = \Phi(t, \mathbf{x}_o)$, é igual à velocidade $\mathbf{v}(t, \mathbf{x})$ da partícula que está em $\mathbf{x}$, no instante t. Portanto, podemos escrever também que:

$$\mathbf{v}(t, \mathbf{x}) \equiv \frac{\partial \Phi(\mathbf{x}_o, t)}{\partial t} \Big|_{\mathbf{x}_o = \phi_t^{-1}(\mathbf{x})} \tag{5.0.4}$$

que se diz a expressão do campo de velocidades do fluido, em coordenadas espaciais.

Quando o campo de velocidades do fluido é independente de t, o fluxo (ou o fluido) diz-se **estacionário**.

Figure 5.2: Campo de velocidades do fluido

.

Exemplo ... 1

Consideremos um fluido em $\mathbb{R}^2$, cujo fluxo é dado por:

$$\Phi(t, x_o, y_o) = (x_o e^t, y_o e^{2t})$$

Neste caso:

$$\alpha_{(x_o,y_o)}(t) = (x_o e^t, y_o e^{2t})$$

e:

$$\dot{\alpha}_{(x_o,y_o)}(t) = (x_o e^t, 2y_o e^{2t})$$

Portanto:

$$\mathbf{v}(t, \Phi(t, x_o, y_o)) = \dot{\alpha}_{(x_o,y_o)}(t) = (x_o e^t, 2y_o e^{2t})$$

Pondo $(x, y) = \Phi(t, x_o, y_o) = (x_o e^t, y_o e^{2t})$, deduzimos que $x_o = xe^{-t}, y_o = ye^{-2t}$, e portanto a expressão do campo de velocidades do fluido, em coordenadas espaciais (ver (5.0.4)), é:

$$\begin{aligned} \mathbf{v}(t, x, y) &= (x_o e^t, 2y_o e^{2t}) \,|_{\{x_o = xe^{-t}, y_o = ye^{-2t}\}} \\ &= (xe^{-t}e^t, 2ye^{-2t}e^{2t}) \\ &= (x, 2y) \end{aligned}$$

o que significa que o fluxo é estacionário.

Exemplo ... 2

Consideremos um fluido em $\mathbb{R}^3$, cujo fluxo é dado por:

$$\Phi(t, x_o, y_o, z_o) = (x_o + t, y_o + t^2, z_o + t^3)$$

Neste caso:

$$\alpha_{(x_o,y_o,z_o)}(t) = (x_o + t, y_o + t^2, z_o + t^3)$$

e:

$$\dot{\alpha}_{(x_o,y_o)}(t) = (1, 2t, 3t^2)$$

e a expressão do campo de velocidades do fluido, em coordenadas espaciais (ver (5.0.4)), é:

$$\mathbf{v}(t, x, y, z) = (1, 2t, 3t^2)$$

independente da posição, mas dependente do tempo.

Exemplo ... 3

Consideremos um fluido em $\mathbb{R}^3$, cujo fluxo é dado por:

$$\Phi(t, x_o, y_o, z_o) = (x_o + t, y_o(1 + t), z_o e^t)$$

Neste caso:

$$\alpha_{(x_o,y_o,z_o)}(t) = (x_o + t, y_o(1 + t), z_o e^t)$$

e:

$$\dot{\alpha}_{(x_o,y_o)}(t) = (1, y_o, z_o e^t)$$

e a expressão do campo de velocidades do fluido, em coordenadas espaciais (ver (5.0.4)), é:

$$\mathbf{v}(t, x, y, z) = (1, \frac{y}{1+t}, z) \qquad t \neq 1$$

dependente da posição e do tempo.

5.0.5 Fluxos e equações diferenciais

Sob certas condições de regularidade impostas ao campo de velocidades de um fluido, é possível assegurar que esse campo de velocidades determina unìvocamente o fluxo Φ: a trajectória $\mathbf{x}(t) \equiv \alpha_{\mathbf{x}_o}(t) = \Phi(t, \mathbf{x}_o)$, da partícula que no instante inicial $t = 0$, ocupa a posição $\mathbf{x}_0$, deve ser a única solução da equação diferencial:

$$\dot{\mathbf{x}}(t) = \mathbf{v}(t, \mathbf{x}(t)) \tag{5.0.5}$$

que satisfaz a condição inicial:

$$\mathbf{x}(0) = \mathbf{x}_o \tag{5.0.6}$$

Portanto, deveremos ter:

$$\frac{\partial \Phi(t, \mathbf{x}_o)}{\partial t} = \mathbf{v}(t, \Phi(t, \mathbf{x}_o)) \tag{5.0.7}$$

<u>**Exemplo**</u> ... 4

A equação diferencial (5.0.5), correspondente ao fluxo do exemplo 1, acima, tem a forma seguinte:

$$\dot{\mathbf{x}}(t) = (\dot{x}(t), \dot{y}(t)) = (x(t), 2y(t))$$

ou, com uma notação mais simplificada:

$$\dot{\mathbf{x}} = (\dot{x}, \dot{y}) = (x, 2y)$$

Podemos ainda escrever esta equação diferencial, na forma matricial seguinte:

$$\dot{\mathbf{x}} = \begin{bmatrix} \dot{x} \\ \dot{y} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = A\,\mathbf{x}$$

A equação diferencial (5.0.5), correspondente ao fluxo do exemplo 3, acima, tem a forma seguinte:

$$\dot{\mathbf{x}}(t) = (\dot{x}(t), \dot{y}(t), \dot{z}(t)) = (1, \frac{y(t)}{1+t}, z(t))$$

ou, com notação mais simplificada:

$$\dot{\mathbf{x}} = (\dot{x}, \dot{y}, \dot{z}) = (1, \frac{y}{1+t}, z)$$

Podemos ainda escrever esta equação diferencial, na forma matricial seguinte:

$$\dot{\mathbf{x}} = \begin{bmatrix} \dot{x} \\ \dot{y} \\ \dot{z} \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & \frac{1}{1+t} & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} + \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = A(t)\,\mathbf{x} + \mathbf{e}_1$$

<u>**Exemplo**</u> ... 5

Consideremos um fluido em $\mathbb{R}^3$, cujo fluxo é dado por:

$$\Phi(t, x_o, y_o, z_o) = (x_o \cos t + y_o \sin t, y_o \cos t - x_o \sin t, z_o + t)$$

Neste caso:

$$\alpha_{(x_o,y_o,z_o)}(t) = (x_o \cos t + y_o \sin t, y_o \cos t - x_o \sin t, z_o + t)$$

e:

$$\dot{\alpha}_{(x_o,y_o)}(t) = (-x_o \sin t + y_o \cos t, -y_o \sin t - x_o \cos t, 1)$$

e a expressão do campo de velocidades do fluido, em coordenadas espaciais (ver (5.0.4)), é:

$$\mathbf{v}(t, x, y, z) = (y, -x, 1)$$

independente do tempo. Portanto o fluxo é estacionário.

A equação diferencial (5.0.5), correspondente a este fluxo tem a forma seguinte (com notação simplificada):

$$\dot{\mathbf{x}} = (\dot{x}, \dot{y}, \dot{z}) = \mathbf{v}(t, x, y, z) = (y, -x, 1)$$

Podemos ainda escrever esta equação diferencial, na forma matricial seguinte:

$$\dot{\mathbf{x}} = \begin{bmatrix} \dot{x} \\ \dot{y} \\ \dot{z} \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = A\,\mathbf{x} + \mathbf{e}_3$$

Exemplo ... 6

Consideremos um fluido em $\mathbb{R}^3$, cujo fluxo é dado por:

$$\Phi(t, x_o, y_o, z_o) = (x_o(1+t) + y_o(1 - e^t), y_o e^{-t}, z_o(1+t))$$

Neste caso:

$$\alpha_{(x_o,y_o,z_o)}(t) = (x_o(1+t) + y_o(1 - e^t), y_o e^{-t}, z_o(1+t))$$

e:

$$\dot{\alpha}_{(x_o,y_o)}(t) = x_o - y_o e^t, -y_o e^{-t}, z_o)$$

Uma vez que:

$$\phi_t^{-1}(x, y, z) = (\frac{x + y(e^{2t} - e^t)}{1+t}, ye^t, \frac{z}{1+t})$$

deduzimos que a expressão do campo de velocidades do fluido, em coordenadas espaciais (ver (5.0.4)), é:

$$\mathbf{v}(t, x, y, z) = (\frac{x - ye^t - (2+t)ye^{2t}}{1+t}, -y, \frac{z}{1+t})$$

(com $t \neq -1$) dependente do tempo. Portanto o fluxo não é estacionário.

A equação diferencial (5.0.5), correspondente a este fluxo tem a forma seguinte (com notação simplificada):

$$\dot{\mathbf{x}} = (\dot{x}, \dot{y}, \dot{z}) = (\frac{x - ye^t - (2+t)ye^{2t}}{1+t}, -y, \frac{z}{1+t})$$

(com $t \neq -1$).

Capítulo 6

Equações diferenciais ordinárias

6.1 Alguns modelos

Consideremos uma população de organismos de uma certa espécie, como por exemplo, átomos de uma substância radiactiva.

Designemos por $p(t)$ o número de membros dessa população no instante t. O quociente $\frac{\dot{p}(t)}{p(t)}$ representa então a **taxa de crescimento** no instante t.

Em geral a taxa de crescimento $\frac{\dot{p}(t)}{p(t)}$, será uma função do tempo t e também da população $p(t)$:

$$\frac{\dot{p}(t)}{p(t)} = F(t, p(t)) \tag{6.1.1}$$

Num sistema fechado, sem "migração", teremos que:

$$F(t, p(t)) = N(t, p(t)) - M(t, p(t)) \tag{6.1.2}$$

onde $N(t, p(t))$ é a **taxa de natalidade** e $M(t, p(t))$ a **taxa de mortalidade**.

Se N, M, ou mais geralmente F são funções conhecidas, a **evolução temporal** da população será descrita pela função $p = p(t)$, que deverá ser portanto uma **solução** da **equação diferencial** (6.1.1), isto é:

$$\dot{p}(t) = F(t, p(t))\, p(t) \tag{6.1.3}$$

ou mais sucintamente:

$$\dot{p} = F(t, p)\, p \tag{6.1.4}$$

O problema consiste agora em **integrar** essa equação, i.e., em determinar uma solução, uma função de classe C^1, $p : \mathbf{I} \to \mathbb{R}$, definida num intervalo $\mathbf{I} \subseteq \mathbb{R}$, que satisfaça a equação diferencial (6.1.3), e que num **instante inicial** $t_o \in \mathbf{I}$, verifique a condição:

$$p(t_o) = p_o \tag{6.1.5}$$

para algum valor dado p_o.

Um tal problema diz-se um **problema de Cauchy** ou um **problema de valor inicial** (PVI):

$$(PVI)_{(t_o,p_o)} \quad \begin{cases} \dot{p} = F(t,p)p \\ p(t_o) = p_o \end{cases}$$

Vejamos alguns exemplos concretos.

Exemplo ... (Taxa de crescimento constante)

Neste caso $F(t,p) \equiv k$, $\forall (t,p) \in \mathbb{R}^2$, onde $k \in \mathbb{R}$ é uma constante não nula $k \neq 0$, e o PVI toma a forma seguinte:

$$(PVI)_{(t_o,p_o)} \quad \begin{cases} \dot{p} = k\,p \\ p(t_o) = p_o \end{cases}$$

Se $p(t) \neq 0$, obtemos o seguinte:

$$\begin{aligned} \dot{p}(t) = k\,p(t) \quad \Leftrightarrow \quad \frac{\dot{p}(t)}{p(t)} = k \quad \Leftrightarrow \quad \frac{d}{dt} \log|p(t)| = k \\ \Leftrightarrow \quad \log|p(t)| = kt + a \Leftrightarrow |p(t)| = e^{kt+a} = C\,e^{kt} \end{aligned}$$

para alguma constante $a \in \mathbb{R}$ e $C = e^a$.

Da última equação deduzimos que $p(t)$ nunca muda de sinal, e como $p(t_o) = p_o$, podemos concluir que a solução do PVI, é dada por:

$$p(t) = p_o\,e^{k(t-t_o)} \qquad \forall t \in \mathbb{R} \tag{6.1.6}$$

Notemos que $\lim_{t\to\infty} p(t) = \infty$, quando $k > 0$ (crescimento ilimitado), e que $\lim_{t\to\infty} p(t) = 0$, quando $k < 0$ (extinção).

Exemplo ... (Equação com crescimento limitado)

O modelo do exemplo anterior não é muito realista. Por exemplo, em geral não se observa crescimento ilimitado. Vamos portanto supôr agora que existe uma população limite $\xi > 0$, que é tal que quando p excede o valor de ξ, a taxa de crescimento torna-se negativa, isto é, vamos supôr que:

$$F(t,p) \leq 0 \qquad \text{quando} \quad p \geq \xi \tag{6.1.7}$$

Uma situação particularmente simples ocorre quando F é uma função linear de p:

$$F(t,p) = \beta(\xi - p) \qquad \forall p \in \mathbb{R} \tag{6.1.8}$$

onde $\beta, \xi > 0$. Com estas hipóteses a equação de crescimento toma a forma seguinte:

$$\dot{p} = \alpha p - \beta p^2 = (\alpha - \beta p)\,p \tag{6.1.9}$$

com $\alpha = \beta\xi$, dita **equação de crescimento limitado**.

Para integrar a equação (6.1.9), usamos o chamado **método de separação de variáveis**, escrevendo (6.1.9) na forma:

$$\frac{\dot{p}}{(\alpha - \beta p)\,p} = 1 \tag{6.1.10}$$

onde excluímos os casos $p = 0$ e $p = \alpha/\beta = \xi$. Representando por G uma primitiva de $1/(\alpha - \beta p)\, p$, i.e, se:

$$G'(p) = \frac{1}{(\alpha - \beta p)\, p}$$

então (6.1.10) é equivalente a:

$$\frac{d}{dt}(G(p(t)) = 1$$

e portanto:

$$G(p(t)) = t + C \tag{6.1.11}$$

para alguma constante C. Usando fracções parciais, obtemos agora que:

$$G(p) = \int \frac{dp}{(\alpha - \beta p)\, p} = \frac{1}{\alpha} \int \frac{dp}{p} + \frac{\beta}{\alpha} \int \frac{dp}{\alpha - \beta p} = \log \mid \frac{p}{\alpha - \beta p} \mid^{1/\alpha}$$

Portanto (6.1.11) é equivalente a:

$$\mid \frac{p}{\alpha - \beta p} \mid = e^{\alpha(t+C)} = C_1\, e^{\alpha t} \tag{6.1.12}$$

onde $C_1 = e^{\alpha C}$. Como devemos ter $p(t_o) = p_o$, obtemos, se $p_o \neq 0$ e $p_o \neq \xi = \alpha/\beta$:

$$\mid \frac{p_o}{\alpha - \beta p_o} \mid = C_1 e^{\alpha t_o}$$

Isto implica que:

$$\mid \frac{p(t)}{p_o} \mid = \mid \frac{\alpha - \beta p(t)}{\alpha - \beta p_o} \mid e^{\alpha(t - t_o)}$$

De (6.1.12), podemos concluir que $p(t) \neq 0$ e $\alpha - \beta p(t) \neq 0$, $\forall t \in \mathbb{R}$, se essas condições se verificam para algum t_o. Em particular, $\frac{\alpha - \beta p(t)}{\alpha - \beta p_o} > 0$, donde se segue que:

$$\frac{p(t)}{p_o} = \frac{\alpha - \beta p(t)}{\alpha - \beta p_o} e^{\alpha(t - t_o)}$$

e finalmente:

$$p(t) = \frac{\alpha p_o}{\beta p_o + (\alpha - \beta p_o) e^{-\alpha(t - t_o)}} \quad \forall t \in \mathbb{R} \tag{6.1.13}$$

A dedução anterior mostra que, se $p_o \neq 0$ e $p_o \neq \xi = \alpha/\beta$, a função p, dada por (6.1.13), é a única solução do PVI:

$$(PVI)_{(t_o, p_o)} \qquad \begin{cases} \dot{p} = (\alpha - \beta p)\, p \\ p(t_o) = p_o \end{cases}$$

De (6.1.13), deduzimos ainda que:

$$p(t) \nearrow \xi \quad \text{quando} \;\; t \to \infty \qquad se \;\; 0 < p_o < \xi$$

e:

$$p(t) \searrow \xi \quad \text{quando} \;\; t \to \infty \qquad se \;\; p_o > \xi$$

Por outro lado, como $\ddot{p} = (\alpha - 2\beta p)(\alpha - \beta p)p$, vemos que $\ddot{p} > 0$, se $p \in [0, \xi/2] \cap [\xi, \infty]$, e $\ddot{p} < 0$, se $\xi/2 < p < \xi$. O gráfico de p, para valores diferentes de p_o, tem o aspecto ilustrado na figura 6.1.

Figure 6.1:

6.2 Métodos elementares de integração

Seja $\mathcal{U} \subseteq \mathbb{R}^n$ um aberto de $\mathbb{R}^n$, $\mathbf{I} \subseteq \mathbb{R}$ um intervalo aberto de $\mathbb{R}$, $(t_o, \mathbf{x}_o) \in \mathbf{I} \times \mathcal{U}$ um ponto dado em $\mathbf{I} \times \mathcal{U}$, e finalmente $\mathbf{F} : \mathbf{I} \times \mathcal{U} \to \mathbb{R}^n$ uma função contínua.

Consideremos o **problema de Cauchy**, ou **problema de valor inicial** $(PVI)_{(t_o,\mathbf{x}_o)}$, definido por:

$$(PVI)_{(t_o,\mathbf{x}_o)} \quad \begin{cases} \dot{\mathbf{x}} = \mathbf{F}(t, \mathbf{x}) \\ \mathbf{x}(t_o) = \mathbf{x}_o \end{cases}$$

Uma função $\alpha : \mathbf{J} \to \mathcal{U}$ diz-se uma **solução** do **problema de valor inicial** $(PVI)_{(t_o,\mathbf{x}_o)}$, se se verificam as condições seguintes:

(i)... $t_o \in \mathbf{J} \subseteq \mathbf{I}$.

(ii)... α é de classe C^1 em $\mathbf{J}$.

(iii)... $\dot{\alpha}(t) = \mathbf{F}(t, \alpha(t)) \qquad \forall t \in \mathbf{J}$.

(iv)... $\alpha(t_o) = \mathbf{x}_o$.

Se $\tilde{\alpha} : \tilde{\mathbf{J}} \to \mathcal{U}$ é uma outra solução do problema de valor inicial $(PVI)_{(t_o,\mathbf{x}_o)}$, com $\mathbf{J} \subset \tilde{\mathbf{J}}$ e $\tilde{\alpha} \mid_{\mathbf{J}} = \alpha$, diz-se que $\tilde{\alpha}$ é uma **extensão** ou um **prolongamento** de α. Uma solução diz-se **não prolongável** quando não admite qualquer prolongamento próprio. Finalmente, quando $\mathbf{J} = \mathbf{I}$, α diz-se uma **solução global**.

Geomètricamente, uma solução $\alpha : \mathbf{J} \to \mathcal{U}$ da equação diferencial $\dot{\mathbf{x}} = \mathbf{F}(t, \mathbf{x})$, é uma curva de classe C^1, em $\mathcal{U}$, cujo vector tangente no ponto $\alpha(t)$ é igual a $\mathbf{F}(t, \alpha(t))$.

O gráfico da solução α, isto é, o conjunto:

$$\mathbf{gr}\,\alpha = \{(t, \alpha(t)) : t \in \mathbf{J}\} \subset \mathbb{R} \times \mathbb{R}^n \tag{6.2.1}$$

é chamado uma **curva integral** da equação diferencial $\dot{\mathbf{x}} = \mathbf{F}(t, \mathbf{x})$.

Se $\mathbf{gr}\,\alpha$ é parametrizado por:

$$\phi : t \mapsto (t, \alpha(t)), \qquad t \in \mathbf{J}$$

Figure 6.2: Curva integral

.

então o vector tangente à curva $\mathbf{gr}\,\alpha$, no ponto $(t, \alpha(t))$, é igual ao vector $(1, \mathbf{F}(\phi(t)) = (1, \mathbf{F}(t, \alpha(t))$ (ver a figura 6.2).

Consideremos o caso unidimensional ($n = 1$). Neste caso é habitual considerar o chamado **campo de direcções** da equação diferencial $\dot{x} = f(t, x)$, obtido desenhando em cada ponto $(t, x) \in \mathbf{I} \times \mathcal{U} \subseteq \mathbb{R} \times \mathbb{R}$, uma pequena linha com declive igual a $f(t, x)$ (ver a figura 6.3).

Figure 6.3: Campo de direcções de $\dot{x} = f(t, x)$

.

Desta forma obtemos uma visão geral do aspecto das curvas integrais da equação diferencial.

6.2.1 Equações separáveis

Suponhamos que:

$$f(t, x) = g(t)\, h(x) \qquad (t, x) \in \mathbf{I} \times \mathcal{U} \subseteq \mathbb{R} \times \mathbb{R}$$

onde $g : \mathbf{I} \to \mathbb{R}$, $h : \mathcal{U} \subseteq \mathbb{R} \to \mathbb{R}$ são funções reais contínuas.

Neste caso a equação diferencial:

$$\dot{x} = g(t)\, h(x) \tag{6.2.2}$$

diz-se uma **equação diferencial de primeira ordem de variáveis separadas**.

Consideremos o $(PVI)_{(t_o,x_o)}$, definido por:

$$\begin{cases} \dot{x} = g(t)\, h(x) \\ x(t_o) = x_o \end{cases}$$

Podemos então provar o seguinte **teorema de existência e unicidade**:

♣ **Proposição 6.1** ...

(i). Se $h(x_o) = 0$, então $x(t) \equiv x_o, \quad \forall t \in \mathbf{I}$, é uma solução global da equação diferencial (6.2.2).

(ii). Se $h(x_o) \neq 0$, existe um intervalo aberto $\mathbf{J} \subseteq \mathbf{I}$, contendo t_o, e tal que a equação diferencial (6.2.2) tem uma única solução em $\mathbf{J}$.

• Demonstração...

(i). imediato.

(ii).

1. Consideremos o intervalo aberto maximal $\mathcal{D} \subseteq \mathcal{U} \subset \mathbb{R}$, que contem x_o e onde $h(\xi) \neq 0, \quad \forall \xi \in \mathcal{D}$. Para $x \in \mathcal{D}$ definimos:

$$H(x) = \int_{x_o}^{x} \frac{d\xi}{h(\xi)} \tag{6.2.3}$$

Tem-se que H é de classe C^1 em $\mathcal{D}$, $H' = 1/h$ e $H(x_o) = 0$. Pelo teorema da função inversa, H é um difeomorfismo de classe C^1, de $\mathcal{D}$ sobre um intervalo aberto $\mathcal{V} \subseteq \mathbb{R}$, que contem $0 = H(x_o)$.

2. Consideremos agora a (única) função:

$$G(t) = \int_{t_o}^{t} g(\tau)\, d\tau \qquad t \in \mathbf{I}$$

Tem-se que G é de classe C^1 em $\mathbf{I}$, $G' = g$ e $G(t_o) = 0$.

3. Finalmente, seja $\mathbf{J}$ o intervalo aberto maximal, contido em $G^{-1}(\mathcal{V})$, e que contem t_o. Em $\mathbf{J}$ definamos:

$$x(t) = H^{-1}(G(t)) \qquad \forall t \in \mathbf{J} \tag{6.2.4}$$

Temos então que $x(t)$ é de classe C^1 em $\mathbf{J}$, e que $x(t_o) = H^{-1}(G(t_o)) = x_o$. Por outro lado, derivando implicitamente $H(x(t)) = G(t)$, $t \in \mathbf{J}$, usando $H' = 1/h$, obtemos:

$$\dot{x}(t) = g(t)h(x(t)) \qquad \forall t \in \mathbf{J}$$

o que significa que $x = H^{-1} \circ G : \mathbf{J} \to \mathcal{U}$ resolve o PVI considerado.

Quanto à unicidade, se $\tilde{x} : \tilde{\mathbf{J}} \to \mathcal{U}$ é uma outra solução do mesmo PVI, para a qual $h(\tilde{x}(t)) \neq 0$, então $\tilde{x}(\tilde{\mathbf{J}}) \subseteq \mathcal{D}$, uma vez que $\tilde{x}(\tilde{\mathbf{J}})$ é um subconjunto conexo de $\mathcal{U}$, que contem x_o. Mas $H(\tilde{x}(t)) = G(t), \quad \forall t \in \tilde{J}$, e portanto:

$$\tilde{x}(t) = H^{-1}(G(t)) = x(t) \qquad \forall t \in \mathbf{J} \cap \tilde{\mathbf{J}}$$

Pela definição de $\mathbf{J}$, concluimos que $\tilde{\mathbf{J}} \subseteq \mathbf{J}$, donde $\mathbf{J} \cap \tilde{\mathbf{J}} = \tilde{\mathbf{J}}$, e portanto x será um prolongamento de $\tilde{x}$, o que prova a unicidade local,

♣.

Sucintamente, quando $h(x_o) \neq 0$, a solução do PVI, obtem-se dividindo ambos os membros de (6.2.2) por $h(x)$, integrando de t_o a $t \in \mathbf{J}$:

$$G(t) = \int_{t_o}^{t} g(\tau)\, d\tau = \int_{x_o}^{x} \frac{d\xi}{h(\xi)} = H(x) \tag{6.2.5}$$

e resolvendo em ordem a x esta última equação (6.2.5).

Vejamos alguns exemplos concretos:

Exemplo ...

Consideremos o $(PVI)_{(t_o,x_o)}$, definido por:

$$\begin{cases} \dot{x} = t\,x^3 \\ x(0) = 1 \end{cases}$$

Neste caso, $\mathbf{I} = \mathcal{U} = \mathbb{R}$, $g(t) = t$ e $h(x) = x^3$. Como $h(x_o) = h(1) = 1 \neq 0$, a proposição anterior garante portanto a existência de uma única solução, que pode ser obtida resolvendo em ordem a x:

$$G(t) = \int_0^t \tau d\tau = \int_1^x \frac{d\xi}{\xi^3} = H(x)$$

Obtemos:

$$H(x) = \int_1^x \frac{d\xi}{\xi^3} = -\frac{1}{2\xi^2} \Big|_1^x = \frac{1}{2} - \frac{1}{2x^2}$$

$\forall x \in]0,\infty[= \mathcal{D}$ (o intervalo maximal $\mathcal{D}$, em $\mathcal{U} = \mathbb{R}$, que contem $x_o = 1$, e onde $h(\xi) = \xi^3$ não se anula). H é um difeomorfismo deste intervalo sobre o intervalo $\mathcal{V} =]-\infty, 1/2[$, e $H(1) = 0$

Por outro lado:

$$G(t) = \int_0^t \tau d\tau = \frac{t^2}{2}$$

O intervalo aberto maximal $\mathbf{J}$, contido em $G^{-1}(\mathcal{V}) = G^{-1}(]-\infty, 1/2[)$, e que contem $t_o = 0$, é o intervalo $\mathbf{J} =]-1, 1[$. Em $\mathbf{J}$ definamos finalmente:

$$x(t) = H^{-1}(G(t)) = \sqrt{\frac{1}{1-t^2}} \qquad \forall t \in \mathbf{J} =]-1,1[\tag{6.2.6}$$

(tomamos a raiz positiva uma vez que $x(0) = 1 > 0$), que é a solução procurada.

Exemplo ...

Consideremos o $(PVI)_{(t_o,x_o)}$, definido por:

$$\begin{cases} \dot{x} = 1 + x^2 \\ x(t_o) = x_o \end{cases}$$

Neste caso, $\mathbf{I} = \mathcal{U} = \mathbb{R}$, $g \equiv 1$ e $h(x) = 1 + x^2 \neq 0, \quad \forall x \in \mathbb{R}$. A proposição anterior garante portanto a existência de uma única solução, que pode ser obtida resolvendo em ordem a x:

$$G(t) = \int_{t_o}^{t} d\tau = \int_{x_o}^{x} \frac{d\xi}{1+\xi^2} = H(x)$$

Obtemos:

$$x(t) = \tan(t - C) \qquad t \in]C - \frac{\pi}{2}, C - \frac{\pi}{2}[$$

onde $C = t_o - \arctan x_o$. Notemos que esta solução é não prolongável, embora não seja global (ver a figura 6.4).

Figure 6.4:

6.2.2 Equações diferenciais linear homogéneas de primeira ordem

Consideremos o $(PVI)_{(t_o,x_o)}$, definido por:

$$\begin{cases} \dot{x} = a(t)\,x \\ x(t_o) = x_o \end{cases}$$

onde $a : \mathbf{I} \subseteq \mathbb{R} \to \mathbb{R}$ é uma função de classe C^1, no intervalo aberto $\mathbf{I} \subseteq \mathbb{R}$. Neste caso $h(x) = x$ e $g(t) = a(t)$.

A equação $\dot{x} = a(t)\,x$ diz-se uma **equação diferencial linear homogénea de primeira ordem**.

Se $x_o \neq 0$, separando as variáveis obtemos:

$$G(t) = \int_{t_o}^{t} a(\tau)\,d\tau = \int_{x_o}^{x} \frac{d\xi}{\xi} = H(x)$$

e portanto:

$$\log \left| \frac{x}{x_o} \right| = \int_{t_o}^{t} a(\tau)\,d\tau$$

ou:

$$\left| \frac{x(t)}{x_o} \right| = e^{\int_{t_o}^{t} a(\tau)\,d\tau}$$

Daqui se conclui que $x(t) \neq 0 \ \ \forall t$ no intervalo de existência $\mathbf{J}$. Portanto $\mathbf{J} = \mathbf{I}$ e, uma vez que $x(0) = x_0$, $\mid \frac{x(t)}{x_o} \mid = \frac{x(t)}{x_o}$.

Portanto $\forall x_o \neq 0$, a função:

$$x(t) = x_o \, e^{\int_{t_o}^{t} a(\tau)\, d\tau} \tag{6.2.7}$$

é a única solução global do $(PVI)_{(t_o,x_o)}$ considerado.

Quando $x_o = 0$, a solução é a trivial $x(t) \equiv 0$, que é também única, uma vez que nenhuma das soluções não triviais (6.2.7) se anula.

Concluindo, (6.2.7) representa a única solução global do $(PVI)_{(t_o,x_o)}$, para cada $(t_o, x_o) \in \mathbf{I} \times \mathbb{R}$.

6.2.3 Equações exactas

Sejam $\mathbf{I}$ e $\mathcal{U}$ intervalos abertos em $\mathbb{R}$, e $f : \mathbf{I} \times \mathcal{U} \to \mathbb{R}$, uma função real contínua. Formalmente, a equação diferencial:

$$\frac{dy}{dx} = f(x, y) \tag{6.2.8}$$

pode ser escrita como uma "equação com diferenciais" na forma:

$$f(x, y)dx - dy = 0 \tag{6.2.9}$$

Para sermos mais precisos, consideremos uma "1-forma diferencial", de classe C^k $(k \geq 0)$, i.e, uma expressão da forma:

$$\alpha = A(x, y)\, dx + B(x, y)\, dy \tag{6.2.10}$$

onde A, B são duas funções de classe C^k num aberto $\mathcal{D} \subseteq \mathbb{R}^2$.

Em cada ponto $(x, y) \in \mathcal{D} \subseteq \mathbb{R}^2$, a equação:

$$\alpha = A(x, y)\, dx + B(x, y)\, dy = 0$$

define uma recta $r(x, y)$ no espaço tangente $T_{(x,y)}\mathbb{R}^2$, perpendicular ao vector $(A(x, y), B(x, y))$. Desta forma fica definido um campo de linhas:

$$\{r(x, y) : \quad (x, y) \in \mathcal{D}\}$$

em $\mathcal{D} \subseteq \mathbb{R}^2$ (ver a figura 6.5).

♣ **Definição 6.1** ...

Seja $\alpha = A(x, y)\, dx + B(x, y)\, dy$ uma 1-forma diferencial, de classe C^k $(k \geq 0)$, definida num aberto $\mathcal{D} \subseteq \mathbb{R}^2$, e $\{r(x, y) : \quad (x, y) \in \mathcal{D}\}$ o correspondente campo de linhas. Suponhamos ainda que α que nunca se anula em $\mathcal{D}$.

Uma curva regular Γ, de classe C^1, diz-se uma **solução** *da equação $\alpha = 0$, se existir uma parametrização de Γ, de classe C^1, $\phi : \mathbf{J} \to \mathcal{D}$, tal que $\dot{\phi}(t) \in r(\phi(t))$, $\forall t \in \mathbf{J}$, isto é:*

$$A(x(t), y(t))\, \dot{x}(t) + B(x(t), y(t))\, \dot{y}(t) = 0 \qquad \forall t \in \mathbf{J} \tag{6.2.11}$$

onde pusemos $\phi(t) = (x(t), y(t)) \quad t \in \mathbf{J} \subseteq \mathbb{R}$.

Figure 6.5: Campo de linhas definido por $A(x,y)\,dx + B(x,y)\,dy = 0$

♣ **Proposição 6.2** ... *Calcular uma curva integral para a equação diferencial:*

$$\frac{dy}{dx} = y' = f(x,y) \qquad (x,y) \in \mathbf{I} \times \mathcal{U} \subseteq \mathbb{R}^2 \tag{6.2.12}$$

é equivalente a calcular uma solução para a equação:

$$\alpha = f(x,y)dx - dy = 0 \qquad (x,y) \in \mathbf{I} \times \mathcal{U} \subseteq \mathbb{R}^2 \tag{6.2.13}$$

- Demonstração...

Seja $s : x \in \mathbf{J} \mapsto s(x) = y \in \mathcal{U}$ uma solução da equação diferencial (6.2.12).

Então $\phi(t) = (t, s(t)),\ \ t \in \mathbf{J}$ é uma curva parametrizada regular de classe C^1, em $\mathbf{I} \times \mathcal{U}$, para a qual se tem:

$$\dot{s}(t) - f(t, s(t)) = 0$$

Reciprocamente, seja $\phi : t \in \mathbf{J} \mapsto \phi(t) = (u(t), v(t)) \in \mathbf{I} \times \mathcal{U}$, uma curva parametrizada regular de classe C^1, em $\mathbf{I} \times \mathcal{U}$, tal que:

$$f(u(t), v(t))\dot{u}(t) - \dot{v}(t) = 0 \tag{6.2.14}$$

Se $\dot{u}(t) = 0$, para algum $t \in \mathbf{J}$, então também $\dot{v}(t) = 0$, o que contraria o facto de ϕ ser regular. Portanto $\dot{u}(t) \neq 0,\ \ \forall t \in \mathbf{J}$, e u é um difeomorfismo de $\mathbf{J}$ sobre um intervalo aberto $\mathcal{V} = u(\mathbf{J})$. Além disso, de (6.2.14), segue-se que:

$$\frac{\dot{v}(t)}{\dot{u}(t)} = f(u(t), v(t)) \qquad t \in \mathbf{J} \tag{6.2.15}$$

Finalmente, introduzindo uma nova variável $x = u(t)$, e pondo $s(x) = v[u^{-1}(x)]$, com $x \in \mathcal{V}$, obtemos pela regra da cadeia que:

$$\frac{du}{dx} = f(x, u(x))$$

♣.

Uma 1-forma diferencial $\alpha = A(x,y)\,dx + B(x,y)\,dy$, contínua num aberto $\mathcal{D} \subseteq \mathrm{I\!R}^2$, diz-se **exacta** se existe uma função $f : \mathcal{D} \to \mathrm{I\!R}$, de classe C^1, tal que $df = \alpha$, isto é:

$$\begin{aligned} df &= \frac{\partial f}{dx}\,dx + \frac{\partial f}{dy}\,dy \\ &= A(x,y)\,dx + B(x,y)\,dy \end{aligned} \tag{6.2.16}$$

ou ainda:

$$\frac{\partial f}{dx} = A(x,y) \qquad e \qquad \frac{\partial f}{dy} = B(x,y) \tag{6.2.17}$$

$\forall (x,y) \in \mathcal{D}$.

Uma 1-forma $\alpha = A(x,y)\,dx + B(x,y)\,dy$, de classe C^1 no aberto $\mathcal{D} \subseteq \mathrm{I\!R}^2$, diz-se **fechada** se:

$$A_y \equiv \frac{\partial A}{dy} = \frac{\partial B}{dx} \equiv B_x$$

♣ **Proposição 6.3** ... *Seja α uma 1-forma contínua num aberto $\mathcal{U} \subseteq \mathrm{I\!R}^2$, e $f : \mathcal{U} \to \mathrm{I\!R}$ uma função de classe C^1 tal que $df = \alpha$.*

Suponhamos que $c \in \mathrm{I\!R}$ é valor regular de f (isto é, $\nabla f(x,y) \neq \mathbf{O}$, $\forall (x,y) \in f^{-1}(c)$).

Então, $f^{-1}(c)$ é localmente uma reunião finita de curvas solução da equação $\alpha = 0$. Além disso, qualquer curva solução da equação $\alpha = 0$, está contida num conjunto de nível de f.

- Demonstração...

  Se $\phi : \mathbf{J} \subseteq \mathrm{I\!R} \to \mathcal{U}$ é uma parametrização regular de classe C^1, de uma curva Γ, contida em $f^{-1}(c)$, então $f \circ \phi \equiv c$, em $\mathbf{J}$, e portanto (uma vez que $df = \alpha$):

$$\begin{aligned} 0 &= \nabla f(\phi(t)) \cdot \dot{\phi}(t) = \\ &= \frac{\partial f}{dx}\,\dot{x}(t) + \frac{\partial f}{dy}\,\dot{y}(t) = \\ &= A(x(t),y(t))\,\dot{x}(t) + B(x(t),y(t))\,\dot{y}(t) = 0 \end{aligned} \tag{6.2.18}$$

  o que significa que Γ é uma curva solução de $\alpha = 0$.

  Reciprocamente, se $\phi : \mathbf{J} \subseteq \mathrm{I\!R} \to \mathcal{U}$ é uma parametrização regular de classe C^1, de uma curva solução Γ, de $\alpha = 0$, então (6.2.18) verifica-se, e portanto $f \circ \phi$ é constante em $\mathbf{J}$,

♣.

♣ **Definição 6.2** ... *Uma função $f : \mathcal{U} \subseteq \mathrm{I\!R}^2 \to \mathrm{I\!R}$, de classe C^1, diz-se um* **integral da equação** *$\alpha = 0$, se $df = \alpha$.*

Quando f é um integral de $\alpha = 0$, os segmentos regulares das curvas de nível de f, isto é, as curvas às quais se retira os pontos do conjunto:

$$K = \{(x, y) \in \mathcal{U} : \nabla f(x, y) = 0\}$$

representam todas as soluções regulares de $\alpha = 0$.

♣ **Proposição 6.4** ... *(***Lema de Poincaré***)*

Seja $\mathcal{R} \subseteq \mathbb{R}^2$ um rectângulo aberto e $\alpha = A(x, y)\, dx + B(x, y)\, dy$ uma 1-forma de classe C^1 em $\mathcal{R}$.

Então se α é fechada, α é exacta em $\mathcal{R}$.

- Demonstração...

  Suponhamos que $f : \mathcal{R} \to \mathbb{R}$ é uma função de classe C^1, que satisfaz $df = \alpha$. Temos então que:

$$f_x = \frac{\partial f}{\partial x} = A \quad e \quad f_y = \frac{\partial f}{\partial y} = B$$

  Integrando a primeira equação, de x_o a x, com y fixo, obtemos:

$$f(x, y) = \Big(\int_{x_o}^{x} A(t, y)\, dt\Big) + h(y) \tag{6.2.19}$$

  onde $h(y)$ é uma "constante de integração", de classe C^1 como função de y. Derivando (6.2.19) em ordem a y, obtemos:

$$B(x, y) = f_y(x, y) = \Big(\int_{x_o}^{x} A_y(t, y)\, dt\Big) + h'(y)$$

  e portanto:

$$h'(y) = B(x, y) - \int_{x_o}^{x} A_y(t, y)\, dt \tag{6.2.20}$$

  Uma vez que:

$$\frac{\partial}{\partial x}\Big[B(x, y) - \int_{x_o}^{x} A_y(t, y)\, dt\Big] = B_x(x, y) - A_y(x, y) = 0$$

  o membro direito de (6.2.20) é uma função apenas de y. Portanto h pode ser determinada por integração:

$$h(y) = \int_{y_o}^{y} \Big[B(x, s) - \int_{x_o}^{x} A_y(t, s)\, dt\Big]\, ds$$

  Finalmente obtemos:

$$f(x, y) = \int_{x_o}^{x} A(t, y)\, dt + \int_{y_o}^{y} \Big[B(x, s) - \int_{x_o}^{x} A_y(t, s)\, dt\Big]\, ds \tag{6.2.21}$$

  para um qualquer ponto arbitrário $(x_o, y_o) \in \mathcal{R}$.

  Concluindo, se $df = \alpha$, então f tem necessàriamente a forma (6.2.21). Reciprocamente, definindo $f : \mathcal{R} \to \mathbb{R}$, através de (6.2.21), para um qualquer ponto arbitrário $(x_o, y_o) \in \mathcal{R}$, é imediato verificar que $df = \alpha$,

♣.

Atendendo à proposição 3, o problema de integração da equação $\alpha = 0$, pode ser considerado resolvido se fôr possível determinar um integral f para a forma α. Neste caso, se se pretende uma solução que contenha um dado ponto (x_o, y_o), a **solução geral** é obtida na seguinte forma implícita:

$$f(x, y) = c$$

onde a "constante de integração" é determinada pela condição de que $f(x_o, y_o) = c$.

Exemplo ...

Calcular a solução geral da equação:

$$\alpha = (y \cos x + 2xe^y)\, dx + (\sin x + x^2 e^y + 2)\, dy = 0$$

Neste caso $\mathcal{R} = \mathbb{R}^2$ e como:

$$A_y = \cos x + 2xe^y = B_y$$

α é fechada, logo exacta, pela proposição anterior. Se f satisfaz $df = \alpha$, então:

$$f_x = y \cos x + 2xe^y \qquad e \qquad f_y = \sin x + x^2 e^y + 2 \tag{6.2.22}$$

Integrando a primeira equação em ordem a x, obtemos:

$$f(x, y) = y \sin x + x^2 e^y + h(y)$$

Derivando relativamente a y, e usando a segunda equação em (6.2.22), vem que:

$$\sin x + x^2 e^y + h'(y) = f_y = \sin x + x^2 e^y$$

donde se deduz que $h'(y) = 2$, isto é, $h(y) = 2y$. Portanto:

$$f(x, y) = y \sin x + x^2 e^y + 2y$$

e a solução geral de $\alpha = 0$ é:

$$f(x, y) = y \sin x + x^2 e^y + 2y = c \qquad c \in \mathbb{R}$$

6.2.4 Método dos factores de integração

♣ **Proposição 6.5** ... *Se α é uma 1-forma contínua no aberto $\mathcal{U} \subseteq \mathbb{R}^2$, e se $h : \mathcal{U} \to \mathbb{R}$ é uma função contínua que nunca se anula em $\mathcal{U}$, então as equações $\alpha = 0$ e $h\,\alpha = 0$ têm as mesmas curvas solução.*

- Demonstração...

  Cálculo directo,

♣.

♣ **Definição 6.3** ... *Seja α uma 1-forma contínua no aberto $\mathcal{U} \subseteq \mathbb{R}^2$, e $h : \mathcal{U} \to \mathbb{R}$ uma função contínua que nunca se anula em $\mathcal{U}$.*

Se $h\alpha$ é exacta, diz-se que h é um **factor de integração** *para α.*

Seja $\alpha = A(x,y)\,dx + B(x,y)\,dy$ uma 1-forma de classe C^1 no aberto $\mathcal{U} \subseteq \mathbb{R}^2$, e $h : \mathcal{U} \to \mathbb{R}$ um factor de integração para α. Então $h\alpha$ é exacta, logo fechada, e portanto:

$$(hA)_y = (hB)_x$$

isto é:

$$Ah_y - Bh_x + (A_y - B_x)h = 0 \tag{6.2.23}$$

Aplicando o Lema de Poincaré, quando $\mathcal{U}$ é um rectângulo em $\mathbb{R}^2$, podemos afirmar que:

"se $h : \mathcal{U} \to \mathbb{R}$ uma função de classe C^1, que nunca se anula em $\mathcal{U}$, e que satisfaz a equação (6.2.23), então h é um factor de integração para α."

Na prática, é muitas vezes vezes possível encontrar soluções para a equação (6.2.23), da forma $h = h(x)$, $h = h(y)$, $h = h(xy)$, etc...

Exemplo ...

Encontrar a solução geral da equação:

$$\frac{y^2}{2} + 2ye^x + (y + e^x)\frac{dy}{dx} = 0$$

O problema é equivalente a calcular a solução geral da equação:

$$\alpha = (\frac{y^2}{2} + 2ye^x)dx + (y + e^x)dy = 0$$

α não é exacta, já que:

$$A_y = y + 2e^x \neq e^x = B_x$$

Tentemos encontrar um factor de integração da forma $h = h(x)$. Neste caso a equação (6.2.23), reduz-se à equação diferencial ordinária:

$$-Bh' + (A_y - B_x)h = 0$$

onde $h' = dh/dx$, isto é:

$$h' = h$$

em que uma das soluções é:

$$h(x) = e^x$$

Portanto $h(x) = e^x$ é um factor de integração de α, o que significa que $e^x\alpha$ é exacta, isto é, existe uma função f, tal que $df = e^x\alpha$, i.e:

$$f_x = e^x(\frac{y^2}{2} + 2ye^x) \quad e \quad f_y = e^x(y + e^x)$$

Integrando a primeira equação em ordem a x, com y fixo, vem que:

$$f(x, y) = \frac{y^2}{2}e^x + ye^{2x} + g(y)$$

Derivando esta última, em ordem a y e usando a segunda equação, obtemos:

$$f_y = ye^x + e^{2x} + g'(y) = e^x(y + e^x)$$

donde se deduz que $g'(y) = 0$, i.e., $g(y) \equiv c$ (constante). Portanto:

$$f(x, y) = \frac{y^2}{2}e^x + ye^{2x} = C$$

e a solução geral pretendida é obtida resolvendo esta equação em ordem a y:

$$y(x) = -e^x \pm [e^{2x} + 2Ce{-}x]^{1/2}$$

6.2.5 Equações lineares de primeira ordem não homogéneas. Método da variação das constantes

Seja $\mathbf{I} \subseteq \mathbb{R}$ um interval aberto de $\mathbb{R}$, e $a, b : \mathbf{I} \to \mathbb{R}$ duas funções contínuas em $\mathbf{I}$.

Consideremos a seguinte **equação linear de primeira ordem não homogénea**:

$$\dot{x} = a(t)x + b(t) \tag{6.2.24}$$

A integração desta equação é equivalente à integração da equação:

$$\alpha = dx - [a(t)x + b(t)]dt = 0 \tag{6.2.25}$$

em $\mathcal{U} = \mathbb{R} \times \mathbf{I}$.

α não é fechada, e por isso vamos tentar encontrar um factor de integração da forma $h = h(t)$. Neste caso a equação (6.2.23), reduz-se à equação diferencial ordinária:

$$\dot{h} = -ah$$

cuja solução é:

$$h(t) = e^{-\int_{t_o}^{t} a(\tau)\, d\tau} \qquad t_o \in \mathbf{I}$$

Temos então que $h\alpha$ é exacta, e seguindo o método proposto anteriormente, determinamos a solução geral de (6.2.24), na forma $f(x,t) = c$, onde:

$$f(x,t) = h(t)x - \int_{t_0}^{t} h(\tau)b(\tau)d\tau \qquad t \in \mathbf{I}$$

Se $x_0 \in \mathbb{R}$ é dado, resolvendo a equação implícita:

$$f(x,t) = f(x_o, t_o) = h(t_o)x_o$$

em ordem a x, obtemos a função:

$$x(t) = \frac{h(t_o)}{h(t)}x_o + \int_{t_o}^{t} \frac{h(\tau)}{h(t)} b(\tau)d\tau$$

De (6.2.25), concluimos então que:

$$x(t) = e^{\int_{t_o}^{t} a(\tau)\,d\tau}\, x_o + \int_{t_o}^{t} e^{\int_{s}^{t} a(\tau)\,d\tau}\, b(s)ds \qquad t \in \mathbf{I} \tag{6.2.26}$$

é solução do $(PVI)_{(t_o,x_o)}$, definido por:

$$\begin{cases} \dot{x} = a(t)\,x + b(t) \\ x(t_o) = x_o \end{cases}$$

Se $y : \mathbf{I} \to \mathbb{R}$ é uma outra solução do mesmo PVI, então $u = x - y$ é solução da equação linear homogénea:

$$\begin{cases} \dot{u} = a(t)\,u \\ u(t_o) = 0 \end{cases}$$

e portanto $u \equiv 0$, isto é $x = y$.

Podemos assim enunciar o seguinte teorema:

♣ **Proposição 6.6** ... *Seja* $\mathbf{I} \subseteq \mathbb{R}$ *um interval aberto de* $\mathbb{R}$*, e* $a, b : \mathbf{I} \to \mathbb{R}$ *duas funções contínuas em* $\mathbf{I}$.

Então, $\forall (t_o, x_o) \in \mathbf{I} \times \mathbb{R}$*, o* $(PVI)_{(t_o,x_o)}$*, definido por:*

$$\begin{cases} \dot{x} = a(t)\,x + b(t) \\ x(t_o) = x_o \end{cases}$$

correspondente à **equação linear de primeira ordem não homogénea** $\dot{x} = a(t)x + b(t)$*, admite uma única solução global dada por:*

$$x(t) = U(t, t_o)x_o + \int_{t_o}^{t} U(t,s)b(s)ds \qquad t \in \mathbf{I} \tag{6.2.27}$$

onde U *é o chamado* **operador de evolução**, *definido por:*

$$U(t,s) = e^{\int_s^t a(\tau)d\tau} \qquad \forall s, t \in \mathbf{I} \tag{6.2.28}$$

O "truque" usual para resolver o $(PVI)_{(t_o,x_o)}$, definido por:

$$\begin{cases} \dot{x} = a(t)\,x + b(t) \\ x(t_o) = x_o \end{cases}$$

é o chamado **método da variação das constantes**. Neste método, começamos por determinar uma solução arbitrária da equação homogénea:

$$\dot{x} = a(t)x \tag{6.2.29}$$

como por exemplo:

$$u(t) = e^{\int_{t_o}^{t} a(\tau)d\tau} \tag{6.2.30}$$

Em seguida, tentamos encontrar uma solução da equação não homogénea:

$$\dot{x} = a(t)\,x + b(t) \tag{6.2.31}$$

que seja da forma:

$$x(t) = c(t)u(t) \tag{6.2.32}$$

para alguma função desconhecida c (a "constante variável"). Substituindo (6.2.32) em (6.2.31), obtemos:

$$\dot{c}u + \dot{u}c = acu + b$$

e portanto, usando (6.2.29), tem-se que:

$$\dot{c} = \frac{b}{u}$$

Integrando esta última equação e usando (6.2.30) bem como a condição inicial, obtemos finalmente a fórmula (6.2.27).

Exemplo ...

Calcular a solução do $(PVI)_{(t_o,x_o)}$, definido por:

$$\begin{cases} \dot{x} = -2tx - t \\ x(1) = 2 \end{cases}$$

Comecemos por determinar uma solução arbitrária da equação homogénea:

$$\dot{x} = -2tx$$

como por exemplo:

$$u(t) = e^{\int_1^t -2\tau d\tau} = e^{1-t^2}$$

Em seguida, tentamos encontrar uma solução da equação não homogénea, que seja da forma:

$$x(t) = c(t)u(t)$$

para alguma função desconhecida $c = c(t)$. Como antes, obtemos:

$$\dot{c} = \frac{b}{u} = \frac{-t}{e^{1-t^2}} = -te^{t^2-1}$$

Integrando esta última equação, vem que:

$$c(t) = -\frac{1}{2}e^{t^2-1} + C$$

Usando agora $x(t) = c(t)u(t)$ e a condição inicial, obtemos finalmente:

$$x(t) = -\frac{1}{2} + \frac{5}{2}e^{1-t^2}$$

Exemplo ...

Calcular a solução do $(PVI)_{(t_o,x_o)}$, definido por:

$$\begin{cases} \dot{x} = -x + \frac{1}{1+t^2} \\ x(2) = 3 \end{cases}$$

Comecemos por determinar uma solução arbitrária da equação homogénea:

$$\dot{x} = -x$$

como por exemplo:

$$u(t) = e^{\int_2^t -d\tau} = e^{2-t}$$

Em seguida, tentamos encontrar uma solução da equação não homogénea, que seja da forma:

$$x(t) = c(t)u(t)$$

para alguma função desconhecida $c = c(t)$. Como antes, obtemos:

$$\dot{c} = \frac{b}{u} = \frac{e^{t-2}}{1+t^2}$$

Integrando esta última equação, vem que:

$$c(t) = \int_2^t \frac{e^{\tau-2}}{1+\tau^2}d\tau + C$$

Usando agora $x(t) = c(t)u(t)$ e a condição inicial, obtemos finalmente:

$$x(t) = 3e^{2-t} + e^{2-t}\int_2^t \frac{e^{\tau-2}}{1+\tau^2}d\tau$$