ÍNDICE

INTRODUÇÃO AO CÁLCULO DIFERENCIAL PARA FUNÇÕES REAIS DE MAIS DE UMA VARIÁVEL REAL

II.1

INTRODUÇÃO AO CÁLCULO DIFERENCIAL PARA FUNÇÕES REAIS DE MAIS DE UMA VARIÁVEL REAL

II.1

INTRODUÇÃO AO CÁLCULO DIFERENCIAL PARA FUNÇÕES REAIS DE MAIS DE UMA VARIÁVEL REAL

1. Função real de *n* variáveis

Consideremos, por exemplo, a expressão

$$z = \sqrt{1 - x^2 - y^2}\,,$$

sendo x, y, z variáveis reais (isto é, variáveis que só podem tomar como valores números reais).

Atribuindo às variáveis x, y um par de valores tais que

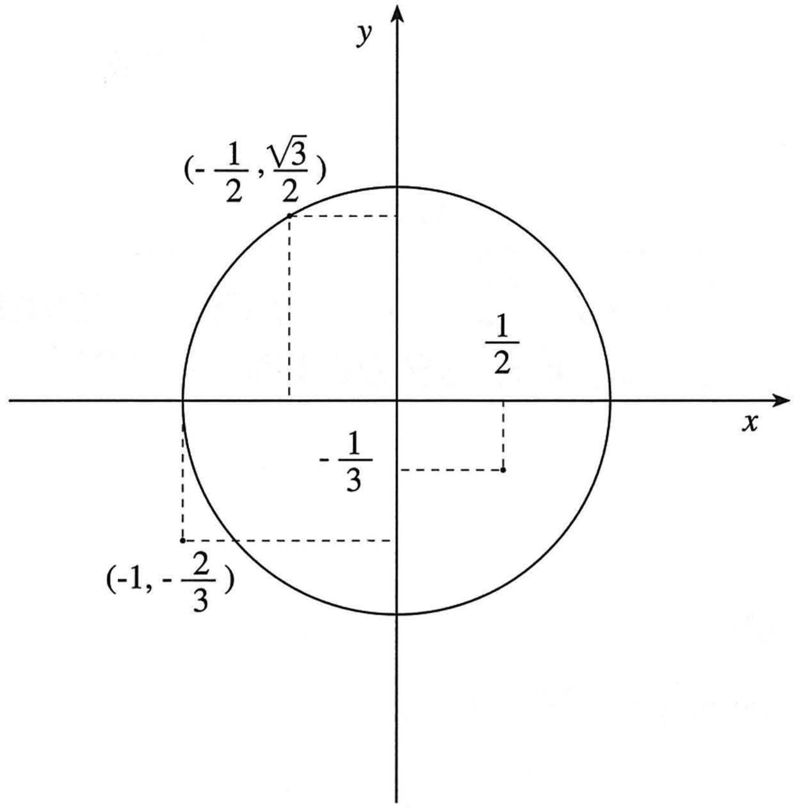
$$1 - x^2 - y^2 \geq 0, \quad \text{ou seja,} \quad x^2 + y^2 \leq 1,$$

(por exemplo o par $(1/2, -1/3)$), resultará para z um determinado valor real; porém, se o par de valores atribuídos a x, y não verifica a condição $x^2 + y^2 \leq 1$, é claro que a expressão do segundo membro deixa de ter sentido no campo real e não haverá portanto nenhum valor real de z que corresponda a tal sistema de valores de x, y.

Pois bem, exprime-se este facto dizendo que a variável z é uma função (*unívoca*) *das variáveis* x, y, *definida para os sistemas de valores de* x, y *tais que* $x^2 + y^2 \leq 1$.

Notemos ainda que, fixado no plano um referencial cartesiano ortonormal, os sistemas (x, y) que satisfazem à condição $x^2 + y^2 \leq 1$

representam os pontos cuja distância à origem é ≤ 1, ou sejam, os pontos do círculo com centro na origem e de raio 1. Diremos, então, que este círculo *é o campo de existência* da função considerada.



Outro exemplo. Seja a expressão:

$$z = y + \log(3x - x^2 - 2),$$

em que se admite excepcionalmente para z o valor $-\infty$; é fácil ver que o segundo membro só tem sentido no campo real quando

$$x^2 - 3x + 2 \leq 0,$$

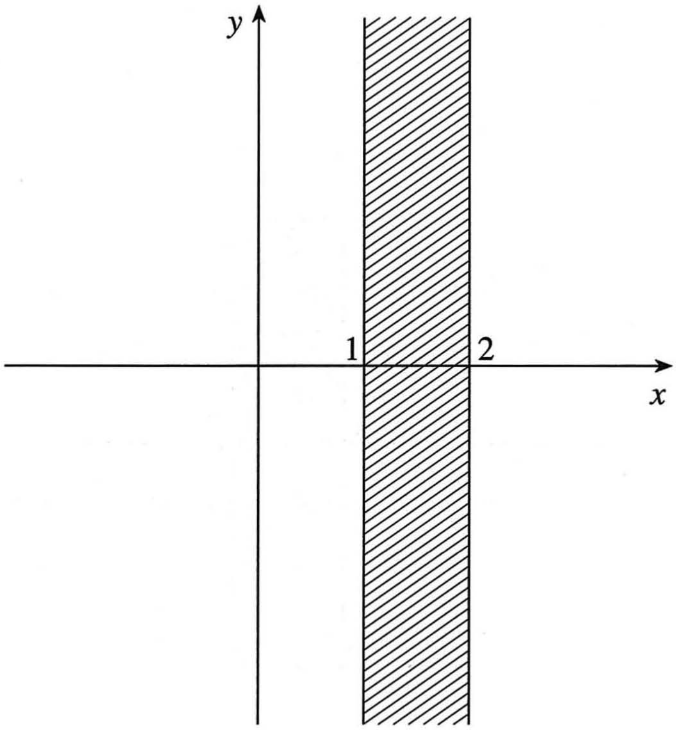
ou seja, quando:

$$1 \leq x \leq 2,$$

podendo y tomar qualquer valor real; por outro lado, os sistemas (x, y) que verificam esta condição representam no plano os pontos da faixa comprendida entre as rectas $x = 1$ e $x = 2$. Diremos, então, que z é uma função (unívoca) de x, y, definida na referida faixa.

Se em vez da anterior expressão se tiver esta outra

$$z = \log(3x - x^2 - 2) + \log(4 - y^2),$$



então tratar-se-á de uma função de x, y, definida para os pares de valores destas variáveis tais que

$$1 \leq x \leq 2\,, \quad -2 \leq y \leq 2\,,$$

os quais representam no plano o rectângulo compreendido entre as rectas $x=1$, $x=2$, $y=2$, $y=-2$; este rectângulo dir-se-á, pois, o campo de existência da função considerada.

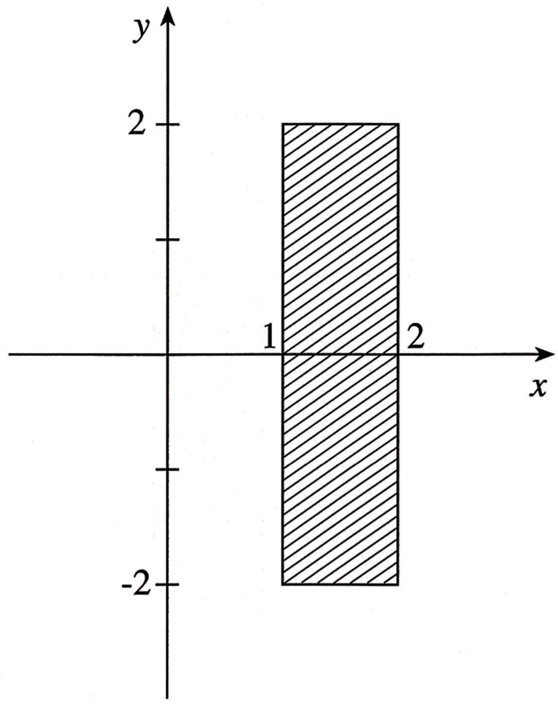
Analogamente, diremos que a expressão

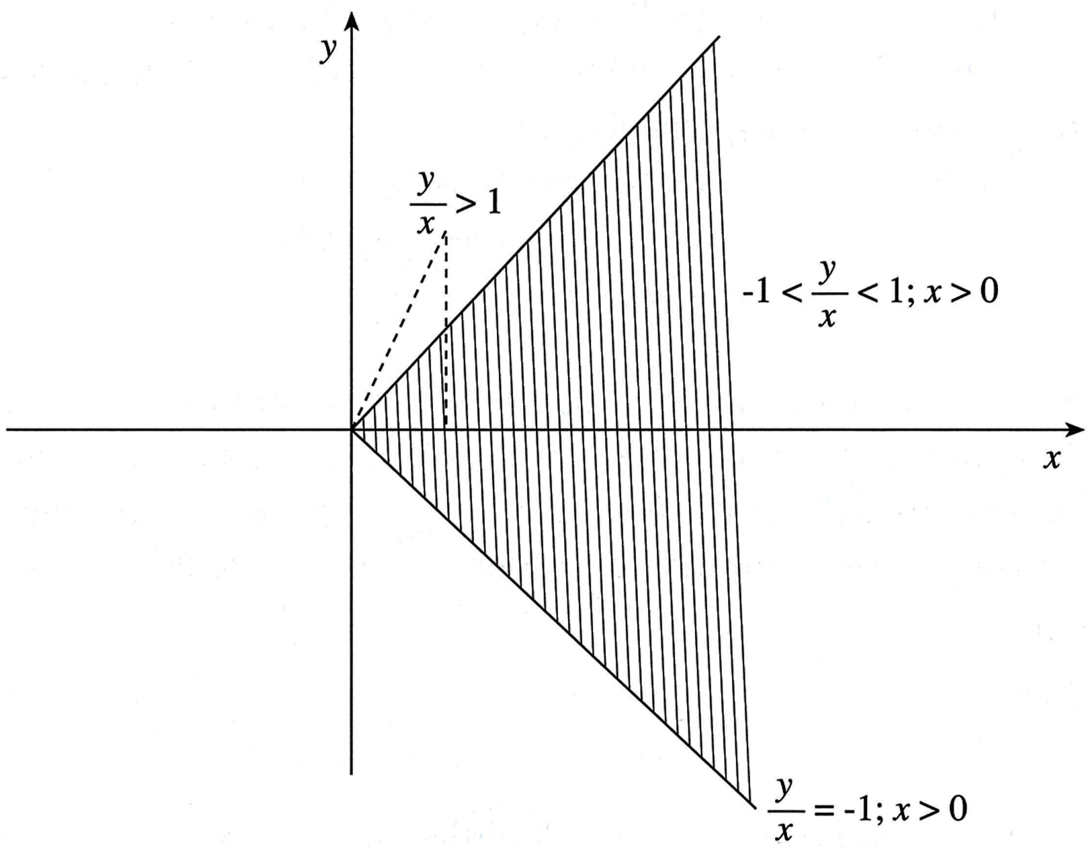
$$z = \log x + \sqrt{1 - \left(\frac{y}{x}\right)^2}$$

define z como função de x, y, que tem por campo de existência o ângulo convexo compreendido entre as bissectrizes do 1.º e do 4.º quadrantes, *excluído o vértice*; com efeito, o segundo membro daquela expressão só tem sentido no campo real quando se tiver ao mesmo tempo

$$x \geq 0 \quad \text{e} \quad -1 \leq \frac{y}{x} \leq 1;$$

se for simultaneamente $x=0$, $y=0$ (vértice do ângulo), é claro que $\frac{y}{x}$ não tem significado.


y
2
-2
1
2
x


y
$\frac{y}{x} > 1$
$-1 < \frac{y}{x} < 1; x > 0$
x
$\frac{y}{x} = -1; x > 0$

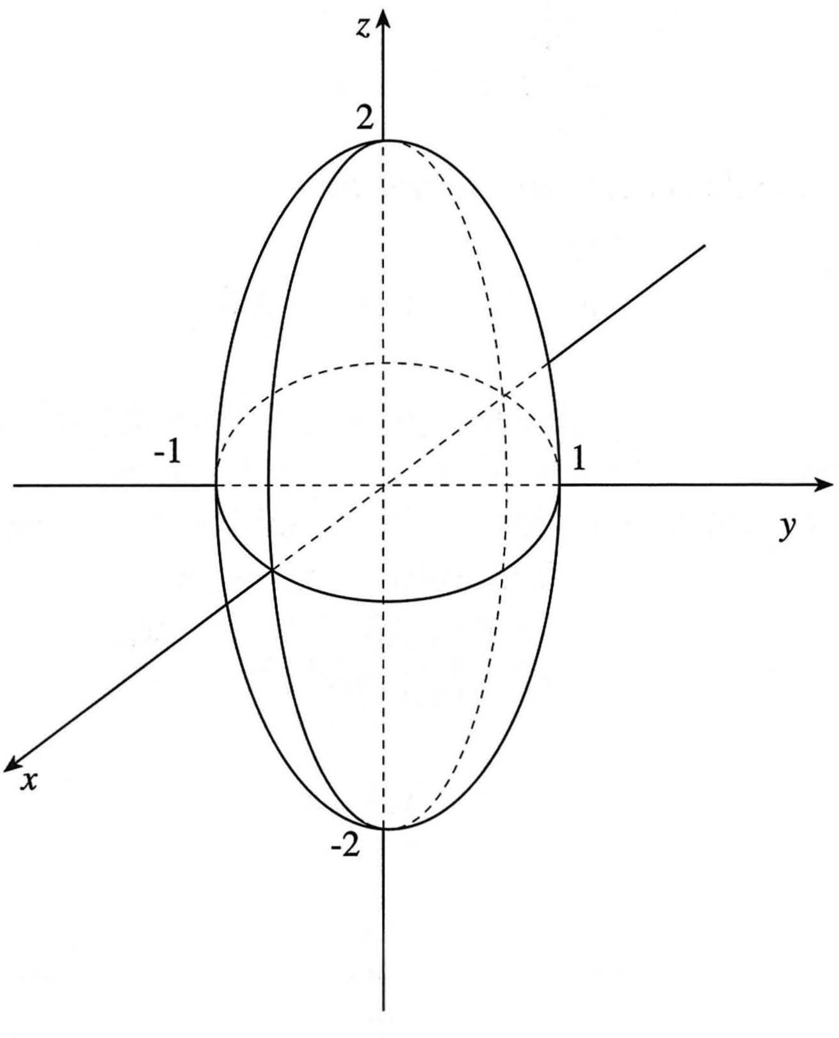
Finalmente, a expressão:

$$z = x - \frac{1}{x^2 + 3y^2}$$

definirá z como função (unívoca) de x, y, para *todos* os pares de valores reais de x, y; o campo de existência desta função será, pois, *todo* o *plano*.

Nos exemplos anteriores trata-se apenas de funções de *duas variáveis*. Consideremos agora a expressão

$$u = \sqrt{1 - x^2 - y^2 - \frac{z^2}{4}}\,;$$



o segundo membro só tem sentido no campo real para os sistemas de valores de x, y, z, tais que

$$x^2 + y^2 + \frac{z^2}{4} \leq 1 .$$

Ora, os sistemas (x, y, z) nestas condições representam pontos situados sobre o elipsóide de equação

$$x^2 + y^2 + \frac{z^2}{4} = 1 ,$$

ou pontos interiores ao mesmo (quando se verifica a relação $<$). Trata-se, pois, duma função (unívoca) de x, y, z, que tem por campo de existência o referido conjunto de pontos.

Podem ainda apresentar-se, naturalmente, funções de 4 variáveis, 5 variáveis, etc. Seja, por exemplo, a expressão:

$$y = \log (x_1 - 3x_2 + x_3 + 2x_4) ;$$

é claro que, a cada sistema de valores de x_1, x_2, x_3, x_4, tais que:

$$x_1 - 3x_2 + x_3 + 2x_4 \geq 0 ,$$

corresponderá, por meio daquela expressão, um determinado valor de y (real ou infinito); ao passo que, se atribuirmos a x_1, x_2, x_3, x_4, um sistema de valores que não verifique tal condição, não resultará para y nenhum valor real ou infinito. Diremos, então, que a variável y é *uma função (unívoca) das variáveis* x_1, x_2, x_3, x_4, *definida no conjunto dos sistemas* (x_1, x_2, x_3, x_4) *tais que*

$$x_1 - 3x_2 + x_3 + 2x_4 \geq 0 .$$

Mas é claro que, neste caso, não podemos atribuir ao campo de existência da função nenhum significado geométrico intuitivo. No entanto, podemos dizer que o campo de existência da função considerada é *um conjunto de pontos do espaço* $\mathbf{R}^4$, pois que, no curso de Matemáticas Gerais, convencionámos chamar "pontos do espaço $\mathbf{R}^4$" *aos sistemas de quatro números reais* tais como

$$(0, -1, 3, 0), \quad (1/2, \sqrt{3}, -5, \sqrt{3}), \quad \text{etc., etc.}$$

Dum modo geral, convencionámos chamar *espaço* $\mathbf{R}^n$ ou espaço numérico a n dimensões reais, ao conjunto de todos os sistemas $(x_1, x_2, \ldots x_n)$ constituídos por n números reais; cada um destes sistemas passa, então, a chamar-se *ponto* ou *vector* do espaço $\mathbf{R}^n$. Mas é preciso nunca perder de vista que se trata aqui unicamente duma *convenção de linguagem*, tendente a simplificar e a tornar mais sugestivos os enunciados das proposições matemáticas.

Também já se viu que tal convenção foi sugerida pelo facto de ser possível, nos casos em que $n=1$, $n=2$ ou $n=3$, estabelecer uma correspondência biunívoca entre os elementos de $\mathbf{R}^n$ e os *pontos* da recta, do plano ou do espaço ordinário (conforme $n=1$, $n=2$ ou $n=3$); podendo ainda nesses casos estabelecer-se uma correspondência biunívoca entre os elementos de $\mathbf{R}^n$ e os *vectores* da recta, do plano ou do espaço ordinário. *Para* $n > 3$, *a situação muda radicalmente, do ponto de vista intuitivo, porquanto o espaço físico não tem mais de* 3 *dimensões.*

Contudo, por analogia, continua a dar-se aos elementos de $\mathbf{R}^n$ o nome de *pontos* ou *vectores*, mesmo no caso em que $n > 3$.

Esta convenção permite-nos, desde já, formular de maneira sugestiva o *Conceito Geral de Função Real de n variáveis Reais*:

Seja $\mathcal{C}$ um conjunto qualquer de pontos do espaço $\mathbf{R}^n$. Sejam, por outro lado, $x_1, x_2, \ldots, x_n$, variáveis reais e y uma variável que pode tomar, como valores, números reais, ou, eventualmente, um dos símbolos ∞, $+\infty$ ou $-\infty$. Se as variáveis $x_1, x_2, \ldots, x_n$ e y estão relacionadas entre si de modo tal que, a cada sistema de valores de $x_1, x_2, \ldots, x_n$, *pertencente a* $\mathcal{C}$ ficam a corresponder um ou mais valores de y, sem que o mesmo aconteça para os sistemas de valores de $x_1, x_2, \ldots, x_n$, *não pertencentes a* $\mathcal{C}$, então diremos que y é *uma função de* $x_1, x_2, \ldots, x_n$, *definida em* $\mathcal{C}$. Ao conjunto $\mathcal{C}$ dar-se-á, neste caso, o nome de *campo de existência* da função considerada; $x_1, x_2, \ldots, x_n$, dir-se-ão *variáveis independentes* (ou *argumentos*) e *y variável dependente* (ou *função*).

Se a cada sistema, $(x_1, x_2, \ldots, x_n) \in \mathcal{C}$, corresponder *um só* valor de *y*, a função dir-se-á *unívoca*. Normalmente, quando se falar de funções, ficará subentendido que se trata de funções unívocas, a não ser que se faça explícita advertência em contrário.

Genericamente, as funções de n variáveis $x_1, x_2, \ldots, x_n$, serão representadas por notações tais como:

$$f(x_1, x_2, \ldots, x_n), \quad \varphi(x_1, x_2, \ldots, x_n), \ldots$$

Diz-se que duas expressões

$$f(x_1, x_2, \ldots, x_n), \quad g(x_1, x_2, \ldots, x_n)$$

representam *funções idênticas* ou a *mesma função,* quando:

1) – São ambas definidas para os mesmos sistemas de valores de $x_1, x_2, \ldots, x_n$;

2) – Assumem valores iguais para cada sistema admissível de valores de $x_1, x_2, \ldots, x_n$.

Em tal hipótese escreve-se, para exprimir este facto:

$$f(x_1, x_2, \ldots, x_n) \equiv g(x_1, x_2, \ldots, x_n),$$

ou simplesmente $f = g$. Caso contrário, as duas funções dir-se-ão distintas e escrever-se-á

$$f(x_1, x_2, \ldots, x_n) \not\equiv g(x_1, x_2, \ldots, x_n),$$

ou simplesmente $f \neq g$.

Portanto, duas funções serão distintas quando se verificar uma *pelo menos* das seguintes hipóteses (sem ser necessário que se verifiquem as duas):

a) – Não têm o mesmo campo de existência;

b) – Existe *pelo menos um* sistema de valores de $x_1, x_2, \ldots, x_n$, para o qual as funções tomam valores diferentes.

Dadas duas funções $\varphi(x_1, x_2, \ldots, x_n)$, $\phi(x_1, x_2, \ldots, x_n)$, diz-se que a segunda é um *prolongamento* da primeira quando, representando por $\mathcal{C}$ e $\mathcal{C}^*$, respectivamente, os campos de existência de φ e de ϕ, se tem $\mathcal{C} \subset \mathcal{C}^*$ e as duas funções tomam o mesmo valor em cada ponto de $\mathcal{C}$; também se diz, neste caso, que φ é a *restrição de* ϕ *ao conjunto* $\mathcal{C}$.

2. Representação analítica de funções de mais de uma variável

O que se disse no curso de Matemáticas Gerais a respeito da representação analítica de funções de uma variável e da correlativa classificação das funções, aplica-se *mutatis mutandis* às funções de mais de uma variável. Chegou mesmo a fazer-se ali um esboço desse estudo para tais funções. Não há portanto necessidade de repetir o que então foi dito.

Convirá apenas salientar que, tal como acontece para as funções de uma só variável, pode uma função de n variáveis ser definida por meio de várias expressões analíticas dadas para diferentes partes do seu campo de existência (não se dirá, então, que está representada analiticamente). Tal é o caso, por exemplo, da função $\varphi(x, y)$ assim definida:

$$\varphi(x, y) = \begin{cases} x + y, & \text{se } x \neq y, \\ x^2, & \text{se } x = y. \end{cases}$$

Mas note-se, entretanto, que esta mesma função é susceptível da seguinte representação analítica:

$$\varphi(x, y) = x + y + \lim_{n \to \infty} \frac{x^2 - x - y}{[1 + (x - y)^2]^n}\,;$$

basta observar que o denominador da fracção tende para ∞ ou para 1, conforme for $x \neq y$ ou $x = y$.

NOTA. Do que se disse neste número e no final do precedente, ressalta a necessidade de fazer uma distinção entre “campo de existência duma função” e “campo de existência duma expressão analítica”.

Dada uma expressão analítica (de carácter designatório) contendo variáveis x, y, ..., o seu campo de existência será constituído pelos sistemas de valores das variáveis para os quais são *possíveis* todas as operações indicadas nessa expressão. Assim, por exemplo, o campo de existência da expressão $\sqrt{x-y}$, no domínio dos números reais, será o conjunto dos pares (x, y) tais que $x \geqslant y$, porque, só para esses pares de números, as operações indicadas são possíveis (com a exigência de os resultados serem números reais).

Quanto ao campo de existência duma função, *esse é fixado por um acto da nossa livre vontade*, ao definir a função juntamente com a lei que faz corresponder a cada ponto desse campo um determinado valor. Assim, por exemplo, nós podemos definir uma função $\varphi(x, y)$ do seguinte modo:

1) – *Campo de existência de* $\varphi(x, y)$:
Conjunto dos pares (x, y) tais que; $x > 0, y > 0, x \geqslant y$.
2) – *Lei de correspondência da função*:
$\varphi(x, y) = \sqrt{x-y}$ em todo o seu campo de existência.

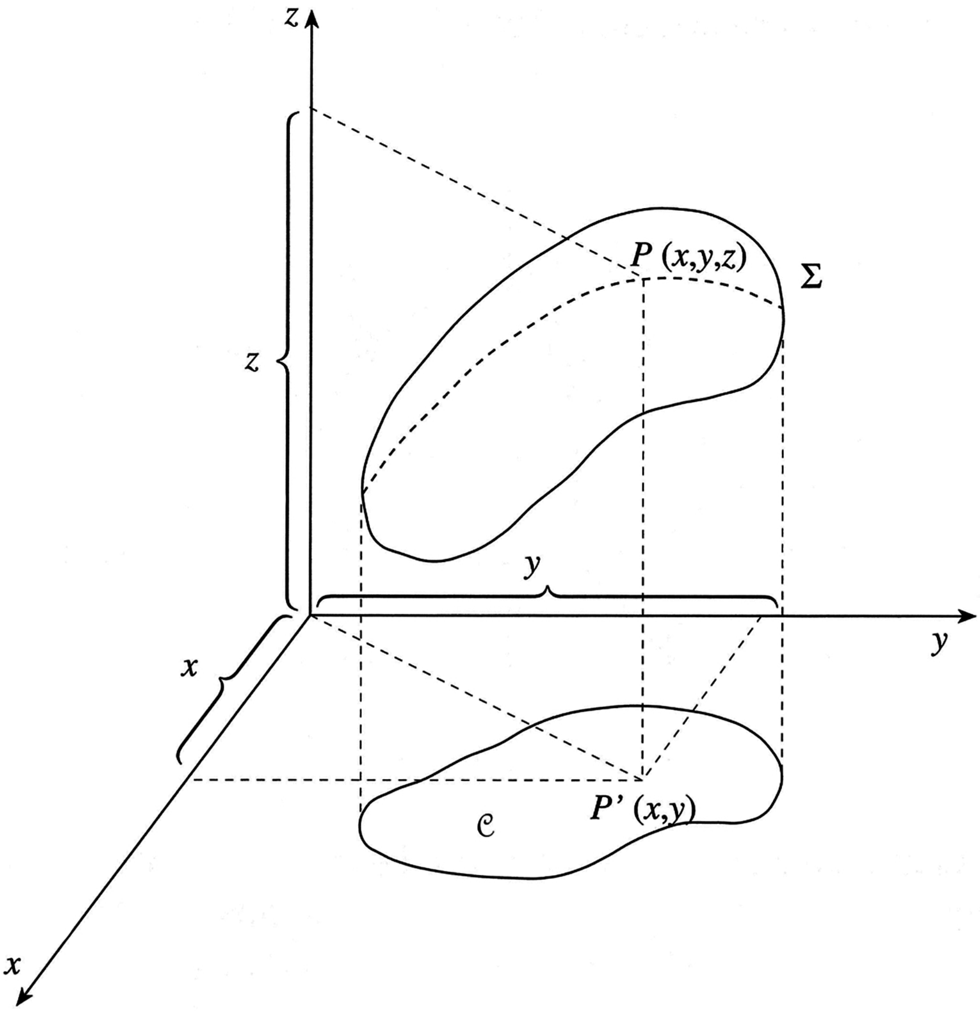
É claro que, neste exemplo, o campo de existência da função é mais restrito que o campo de existência da expressão analítica utilizada.

Em geral, quando se define uma função por meio duma expressão analítica, *sem nada acrescentar,* subentende-se que o campo de existência da função é o da sua expressão analítica.

3. Representação gráfica das funções de duas variáveis

Consideremos uma função real $z = \varphi(x, y)$ das duas variáveis reais x, y, definida num dado conjunto $\mathcal{C}$ de pontos do plano.

Então, a cada par $(x, y) \in \mathcal{C}$, corresponderá o valor de z dado por $\varphi(x, y)$; ora, como sabemos, fixado no espaço um referencial cartesiano (ortonormal) O x y z, o par (x, y) será representado por um ponto P' do plano O x y, enquanto o terno de valores (x, y, z), em que $z = \varphi(x, y)$, será representado por um ponto P do espaço tridimensional. Pois bem, ao conjunto Σ dos pontos $P(x, y, z)$ assim obtidos quando se atribuem a x, y, todos os possíveis pares de valores do conjunto $\mathcal{C}$, dá-se o nome de *gráfico* ou *imagem geométrica* da função $\varphi(x, y)$.



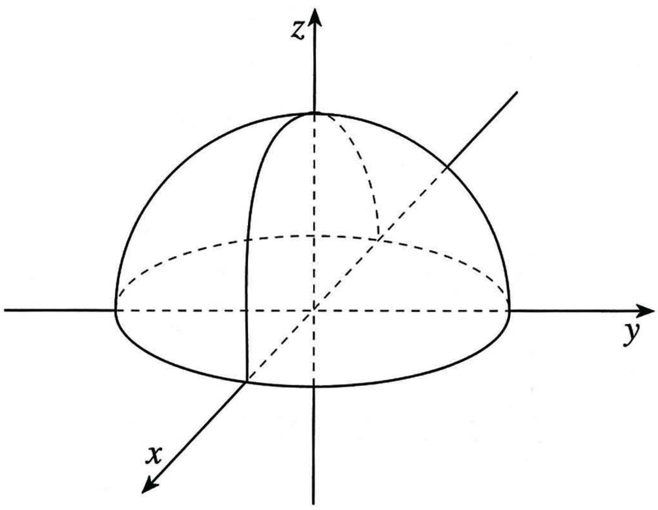
Nos casos que mais frequentemente se apresentam na prática, o gráfico de uma função de duas variáveis é aquilo que habitualmente se chama uma *superfície*. Mas nem sempre assim acontece.

Exemplos – 1) O gráfico duma função linear de x, y (função inteira do 1.º grau em x, y), isto é, duma função do tipo

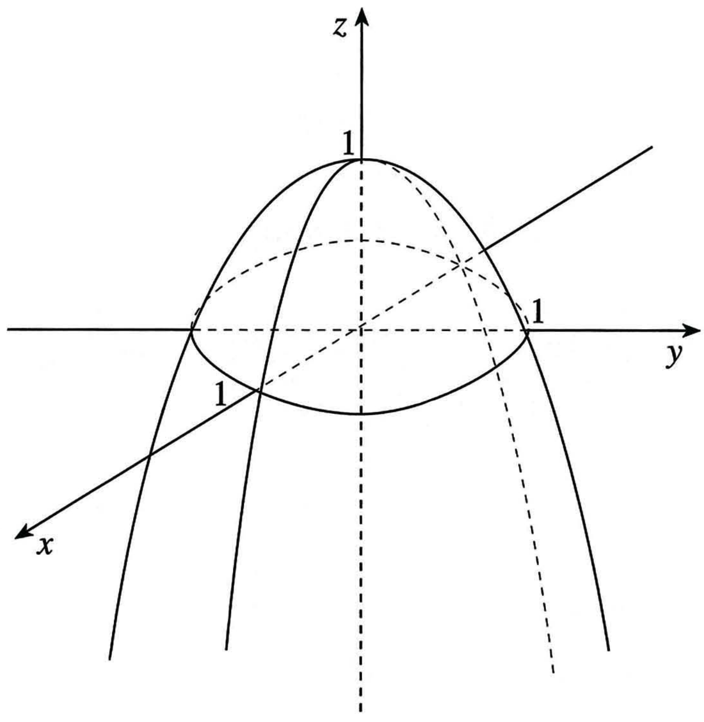
$$z = ax + by + c,$$

sendo a, b, c números reais quaisquer, será, como é fácil reconhecer, um plano não paralelo ao eixo dos zz. O campo de existência desta função é todo o plano dos $x\,y$.

2) O gráfico da função de x, y, $z = \sqrt{1 - x^2 - y^2}$ é a metade da superfície esférica de centro na origem e raio 1 situada acima do plano dos xy. O campo de existência da função é, como vimos atrás, o círculo de centro na origem e raio 1 situado no plano dos xy.



3) O gráfico da função de x, y, $z = 1 - x^2 - y^2$ é um parabolóide de revolução que tem por eixo de simetria o eixo dos zz.



Passando às funções de mais de duas variáveis, na mesma ordem de ideias, deixa de ser possível a representação gráfica em termos intuitivos. Assim, por exemplo, o gráfico duma função de 3 variáveis seria um conjunto de pontos do espaço $\mathbf{R}^4$ (isto é, de sistemas de 4 números reais); mas, como já dissemos por mais de uma vez, o espaço $\mathbf{R}^n$, para $n > 3$, não tem qualquer representação geométrica intuitiva.

4. Transformações pontuais entre espaços cartesianos

Consideremos uma função real y de n variáveis reais $x_1, x_2, \ldots, x_n$:

$$y = f(x_1, x_2, \ldots, x_n), \tag{1}$$

e seja $\mathcal{C}$ o seu campo de existência. Já sabemos que cada sistema de valores de $x_1, x_2, \ldots, x_n$, é, por definição, um ponto de $\mathbf{R}^n$. Ponhamos agora, para abreviar a escrita,

$$X = (x_1, x_2, \ldots, x_n);$$

é claro que sendo $x_1, x_2, \ldots, x_n$, variáveis reais, será X um *ponto* (*ou vector*) de $\mathbf{R}^n$, que varia com $x_1, x_2, \ldots, x_n$; diz-se, então, que X é *uma variável pontual* ou *uma variável vectorial* (sobre $\mathbf{R}^n$).

Por outro lado, a função (1) poderá escrever-se mais condensadamente sob a forma

$$y = f(X),$$

e também se poderá dizer que a variável *real* y é uma função da variável *pontual* X, definida no subconjunto $\mathcal{C}$ de $\mathbf{R}^n$.

Mais geralmente consideremos m funções reais $y_1, y_2, \ldots, y_m$, das n variáveis reais $x_1, x_2, \ldots, x_n$, todas com um mesmo campo de existência $\mathcal{C}$:

$$\begin{cases} y_1 = f_1(x_1, x_2, \ldots, x_n) \\ y_2 = f_2(x_1, x_2, \ldots, x_n) \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ y_m = f_m(x_1, x_2, \ldots, x_n) \end{cases} . \tag{2}$$

Pondo, para abreviar,

$$X = (x_1, x_2, \ldots, x_n)$$
$$Y = (y_1, y_2, \ldots, y_m),$$

segue-se que, por meio das funções (2), a cada valor de X pertencente a $\mathcal{C}$, fica a corresponder um determinado valor de $Y \in \mathbf{R}^m$ (estamos agora a supor que as variáveis $y_1, y_2, \ldots, y_m$, não assumem valores infinitos). Podemos portanto dizer que as funções (2) definem Y como *função pontual* (ou *vectorial*) da variável *pontual* (ou *vectorial*) X; e podemos escrever abreviadamente o sistema (2) sob a forma

$$Y = F(X).$$

É claro que o campo de existência da função pontual $F(X)$ será ainda $\mathcal{C}$. Chamaremos *contradomínio* de $F(X)$, naturalmente, aquele subconjunto de $\mathbf{R}^m$ constituído pelos valores que toma $F(X)$, quando X percorre $\mathcal{C}$; isto é, o conjunto de todos os pontos Y tais que

$$Y = F(X), \quad X \in \mathcal{C}.$$

Notemos ainda que a função F representa aquilo que se chama *uma transformação de* $\mathcal{C}$ *para* $\mathbf{R}^m$.

Reciprocamente, toda a transformação F de $\mathcal{C}$ para $\mathbf{R}^m$ será representada por um sistema de m funções reais de n variáveis reais, todas definidas no conjunto $\mathcal{C}$.

No curso de Matemáticas Gerais estudou-se em detalhe o caso em que a transformação F se diz linear; neste caso a transformação é representável por um sistema de funções reais do tipo

$$\begin{cases} y_1 = a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n \\ y_2 = a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ y_m = a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n \end{cases}$$

as quais têm, pois, por campo de existência todo o espaço $\mathbf{R}^n$ (trata-se duma transformação de $\mathbf{R}^n$ para $\mathbf{R}^m$).

O aluno deverá ter o cuidado de recordar tudo o que se disse então a respeito de tais transformações e da respectiva representação matricial.

Exemplo duma transformação de $\mathbf{R}^2$ em $\mathbf{R}^2$ que não é linear:

$$\begin{cases} y_1 = x_1 \cos x_2 \\ y_2 = x_1 \operatorname{sen} x_2 \end{cases}.$$

Consideremos de novo a função vectorial

$$Y = F(X) \tag{3}$$

definida num subconjunto $\mathcal{C}$ de $\mathbf{R}^n$ e tomando os valores em $\mathbf{R}^m$ (isto é, sendo $Y \in \mathbf{R}^m$). Note-se que, em particular, pode ter-se $m = 1$ ou $n = 1$. No primeiro caso tem-se $\mathbf{R}^m = \mathbf{R}$ e, portanto, Y será uma variável real: a função (3) consiste, então, numa só função real das n variáveis reais $x_1, x_2, \ldots, x_n$:

$$y = f(x_1, x_2, \ldots, x_n).$$

No segundo caso tem-se $\mathbf{R}^n = \mathbf{R}$ e, portanto, X reduz-se a uma só variável real; trata-se, pois, dum sistema de m funções reais de uma variável real x:

$$\begin{cases} y_1 = f_1(x) \\ y_2 = f_2(x) \\ \dots\dots\dots\dots \\ y_m = f_m(x) \end{cases}.$$

De resto, estas duas possibilidades já se nos tinham apresentado a propósito das transformações lineares: correspondiam-lhes, então, respectivamente, as matrizes linhas e as matrizes colunas.

Pode mesmo acontecer que $m = n = 1$, e então recai-se no caso da função real duma só variável real.

5. Noções topológicas em $\mathbf{R}^n$

Para um estudo adequado das funções de n variáveis, no que se refere a campos de existência, limites, continuidade, etc., há que introduzir certas definições prévias.

DEFINIÇÃO 1. *Dado um vector* $X = (x_1, x_2, \ldots, x_n)$ *do espaço* $\mathbf{R}^n$, *chama-se módulo ou comprimento de X, e representa-se pela notação* $|X|$ *o número não negativo indicado pela fórmula*

$$|X| = \sqrt{x_1^2 + x_2^2 + \cdots + x_n^2}\,.$$

Como se vê imediatamente, esta definição é dada por analogia com o que se passa a respeito dos vectores do espaço ordinário.

DEFINIÇÃO 2. *Sejam* $X=(x_1, \ldots, x_n)$ *e* $Y = (y_1, \ldots, y_n)$ *dois pontos quaisquer de* $\mathbf{R}^n$. *Chama-se distância de X a Y o número não negativo*

$$|X - Y| = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2 + \cdots + (x_n - y_n)^2}\,.$$

DEFINIÇÃO 3. *Dados um ponto A de* $\mathbf{R}^n$ *e um número* $\rho > 0$, *chama-se esfera fechada (ou simplesmente esfera) de centro A e raio* ρ *ao conjunto de todos os pontos de* $\mathbf{R}^n$ *cuja distância a A é inferior ou igual a* ρ; *isto é, ao conjunto dos pontos X de* $\mathbf{R}^n$ *tais que:*

$$|X - A| \leqslant \rho\,.$$

Facilmente se reconhece que um tal conjunto vem a ser: um segmento com centro em A, quando $n = 1$; um círculo com centro em A, quando $n = 2$; uma esfera propriamente dita, com centro em A, quando $n = 3$.

Quando $n > 3$, também se usa o termo "hiperesfera" em vez de "esfera", para salientar que se trata duma esfera num *hiperespaço* (isto é, um espaço com mais de 3 dimensões).

DEFINIÇÃO 4. *Entende-se por esfera aberta de centro A e raio* ρ *o conjunto dos pontos X de* $\mathbf{R}^n$ *tais que*

$$|X - A| < \rho\,.$$

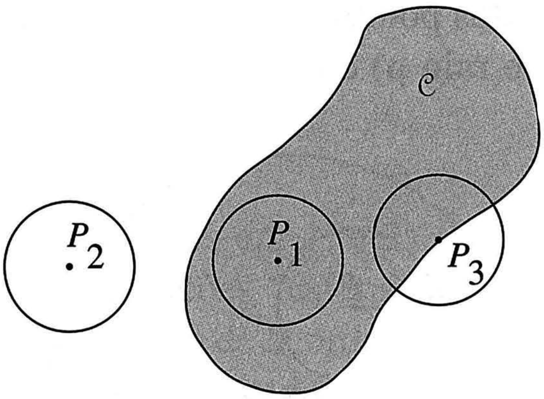
DEFINIÇÃO 5. *Convencionaremos chamar vizinhanças dum dado ponto P de* $\mathbf{R}^n$ *a todas as esferas abertas com centro em P; mais precisamente, chamaremos vizinhança* (δ) *de P à esfera aberta de centro P e raio* δ.

Consideremos agora em $\mathbf{R}^n$ um conjunto $\mathcal{C}$ de pontos e um ponto P qualquer; seguem-se algumas importantes DEFINIÇÕES:

I) – Diz-se que P é *interior* a $\mathcal{C}$ quando existe pelo menos uma vizinhança de P contida em $\mathcal{C}$, isto é, formada só por pontos de $\mathcal{C}$ (caso de P_1 na figura).

II) – Diz-se que P é *exterior* a $\mathcal{C}$ quando é interior ao complementar de $\mathcal{C}$, isto é, ao conjunto $\mathbf{R}^n \backslash \mathcal{C}$, formado pelos pontos que não pertencem a $\mathcal{C}$ (caso de P_2 na figura).

III) – Diz-se que P é ponto *fronteiro* de $\mathcal{C}$, quando não é interior nem exterior a $\mathcal{C}$, isto é, quando em qualquer vizinhança de P há sempre pelo menos um ponto de $\mathcal{C}$ e um ponto de $\mathbf{R}^n \backslash \mathcal{C}$ (caso de P_3 na figura).

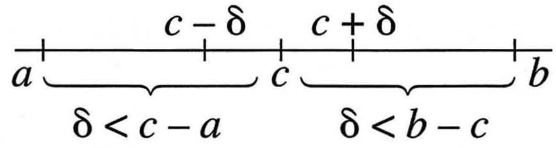


Chama-se *interior* de $\mathcal{C}$ o conjunto dos pontos interiores de $\mathcal{C}$; *exterior* de $\mathcal{C}$, o conjunto dos pontos exteriores a $\mathcal{C}$; *fronteira* de $\mathcal{C}$, o conjunto dos pontos fronteiros de $\mathcal{C}$.

Um conjunto diz-se *aberto*, quando coincide com o seu interior (portanto, se não for vazio, será formado só de pontos que lhe são interiores); diz-se *fechado*, quando contém a sua fronteira.

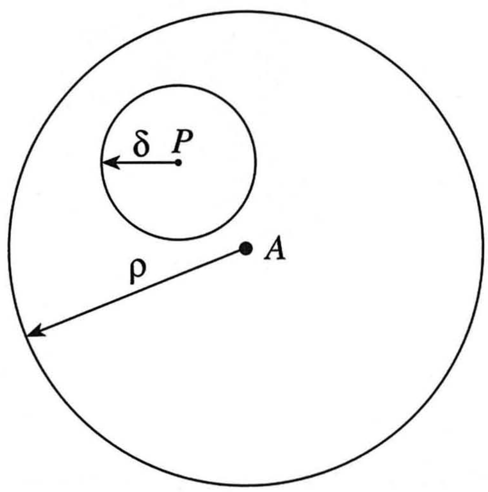
Exemplos – 1) Na recta numérica $\mathbf{R}$, um intervalo aberto de extremos a, b (conjunto dos pontos x tais que $a < x < b$) é um conjunto aberto, pois que todo o seu ponto c lhe é interior; para ter uma vizinhança

(δ) de c contida em $]a, b[$ basta tomar δ inferior ao menor dos números $c - a$ e $b - c$. Contudo, o intervalo $]a, b[$ não é fechado,



visto que não contém a sua fronteira, constituıda pelos dois extremos a, b. Por sua vez, o intervalo $[a, b]$ (conjunto dos pontos x de $\mathbf{R}$ tais que $a \leqslant x \leqslant b$) é um conjunto fechado, porque contém os seus pontos fronteiros a, b; mas não é um conjunto aberto, visto não ser formado só de pontos interiores. Finalmente, o intervalo $[a, b[$ (fechado à esquerda e aberto à direita) dá-nos o exemplo dum conjunto que não é aberto nem fechado.

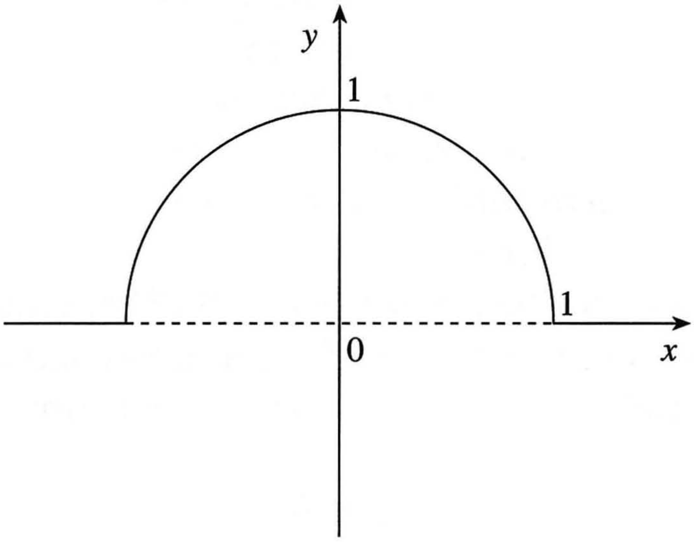
2) Fixado em $\mathbf{R}^2$ um ponto A qualquer e sendo ρ um número positivo, o conjunto dos pontos X de $\mathbf{R}^2$ tais que $|X - A| < \rho$ (círculo aberto do centro A e raio ρ) é um conjunto aberto, visto que, para



todo o seu ponto P, é possível obter uma vizinhança desse ponto, contida no conjunto (basta que o raio δ da vizinhança seja inferior à diferença $\rho - |P - A|$ entre o raio do círculo e a distância de P a A). Por sua vez, o círculo (fechado) de centro A e raio ρ é um conjunto fechado, pois contém a sua fronteira, constituída pela respectiva circunferência; mas não é, evidentemente, um conjunto aberto. Consideremos, por último, em $\mathbf{R}^2$ o conjunto dos pontos (x, y) tais que seja simultaneamente:

$$\begin{cases} x^2+y^2 \leqslant 1 \\ y>0 \end{cases}.$$

Este conjunto, constituído pela parte do círculo de centro O e raio 1 situada *acima* do eixo dos xx, não é aberto nem fechado, pois que contém os pontos fronteiros de ordenada positiva ($x^2+y^2=1$, $y>0$), mas não os pontos fronteiros de ordenada nula ($y=0$, $-1 \leqslant x \leqslant 1$).



3) As considerações feitas em 1) e 2) podem estender-se a $\mathbf{R}^n$, com n qualquer. As designações "esfera aberta" e "esfera fechada" atrás introduzidas estão assim de acordo com as definições de "conjunto aberto" e de "conjunto fechado" que foram dadas em seguida.

Note-se que um conjunto pode ser simultaneamente aberto e fechado: por exemplo, o espaço inteiro, $\mathbf{R}^n$, é um conjunto aberto, porque todos os seus pontos lhe são interiores, e fechado, porque contém a sua fronteira (neste caso vazia). Demonstra-se, porém, que em $\mathbf{R}^n$ só há dois conjuntos simultaneamente abertos e fechados: o *conjunto vazio* e o *espaço inteiro*.

Interessa-nos ainda, para o seguimento, registar a seguinte proposição:

PROPOSIÇÃO. *O interior de um qualquer conjunto $\mathcal{C}$ de pontos de* $\mathbf{R}^n$ *é sempre um conjunto aberto.*

Demonstração. Se $\mathcal{C}$ não tem nenhum ponto interior, o seu interior reduz-se ao conjunto vazio, que é aberto.

Suponhamos agora que $\mathcal{C}$ tem pelo menos um ponto interior e designemos por $\mathcal{C}^*$ o interior de $\mathcal{C}$; vamos provar que $\mathcal{C}^*$ é aberto. Com efeito, se for P um ponto qualquer de $\mathcal{C}^*$, portanto interior a $\mathcal{C}$, existirá, pelo menos, uma esfera aberta ϑ, com centro em P e contida em $\mathcal{C}$.

Mas qualquer ponto de ϑ é interior a ϑ e portanto a $\mathcal{C}$. Quer isto dizer que a vizinhança ϑ de P é formada só de pontos interiores a $\mathcal{C}$, e portanto está contida em $\mathcal{C}^*$. Assim, todo o ponto P de $\mathcal{C}^*$ lhe é interior, o que significa que $\mathcal{C}^*$ é aberto (q.e.d.).

Convém ainda notar que o *complementar dum conjunto aberto é um conjunto fechado* (*e vice-versa*) *e que a reunião de vários conjuntos abertos* (*em número finito ou infinito, com pontos comuns ou não*) *é ainda um conjunto aberto.*

Interessam-nos também as seguintes DEFINIÇÕES:

IV) – Dados dois conjuntos $\mathcal{A}$, $\mathcal{B}$ em $\mathbf{R}^n$ diz-se que $\mathcal{A}$, $\mathcal{B}$ são *separados*, quando cada um deles está contido no exterior do outro; caso contrário, dizem-se *ligados*.

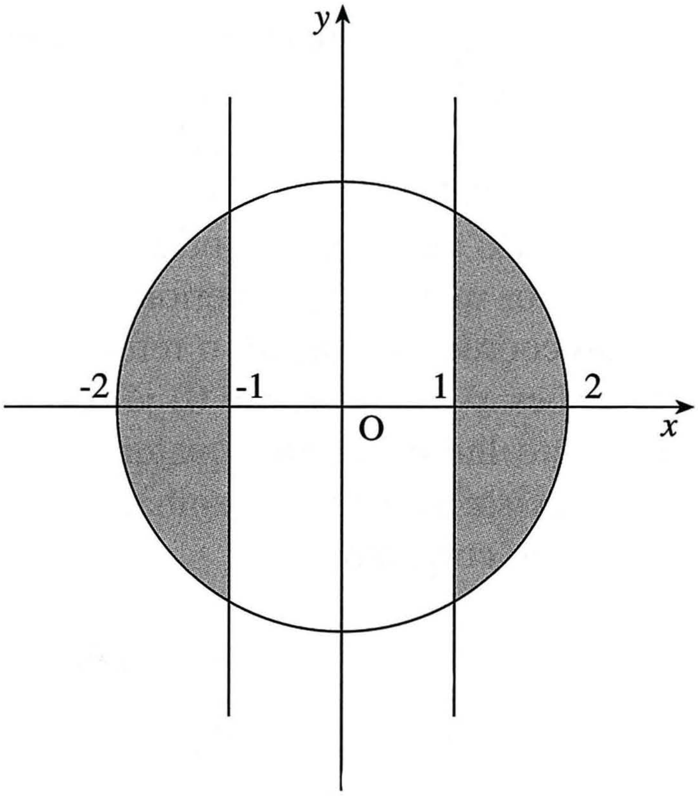
Por exemplo, em $\mathbf{R}^2$, dois círculos abertos sem ponto comum são conjuntos separados (mesmo que as respectivas circunferências sejam tangentes); mas um círculo aberto e um círculo fechado cujas circunferências se toquem já não são separados, porque nem todos os pontos do círculo fechado são exteriores ao círculo aberto (há um que está na fronteira deste).

V) – Um conjunto $\mathcal{C}$ de pontos de $\mathbf{R}^n$ diz-se *desconexo*, quando existem dois conjuntos $\mathcal{A}$, $\mathcal{B}$ separados de que $\mathcal{C}$ seja a reunião, isto é, tais que $\mathcal{C} = \mathcal{A} \cup \mathcal{B}$. Caso contrário, o conjunto $\mathcal{C}$ diz-se *conexo*.

Por exemplo, em $\mathbf{R}^3$ são conjuntos conexos: uma esfera, uma superfície esférica, um cubo, o espaço inteiro, etc. Mas consideremos agora o campo de existência da função de x, y:

$$z = \log(x^2 - 1) + \sqrt{4 - x^2 - y^2}\ ;$$

trata-se, como é fácil de ver, do conjunto dos pontos (x, y) de $\mathbf{R}^2$ que verificam simultaneamente as condições: $x^2 - 1 \geqslant 0$ e $x^2 + y^2 \leqslant 4$; é, pois, um conjunto desconexo, visto ser formado pelas duas partes do círculo de centro em O e raio 2, situadas nos semi-planos $x \geqslant 1$ e $x \leqslant -1$ (conjunto fechado).



Notemos, por último, que as anteriores definições de *interior, exterior, fronteira, conjunto aberto, conjunto fechado, conjunto conexo,* etc., foram todas formuladas partindo unicamente da noção de "ponto interior". Dum modo geral, chamam-se noções topológicas todas aquelas que derivam logicamente da noção de "ponto interior", tomada como única noção inicial.

Mas, por sua vez, o conceito de "ponto interior" foi dado a partir do conceito de "vizinhança" (que não é uma noção topológica).

Convencionámos nós chamar vizinhanças dum ponto P em $\mathbf{R}^n$ às esferas abertas com centro em P; mas é fácil ver que se tivéssemos dado essa designação às esferas fechadas com centro em P, a definição de "ponto interior" obtida seria equivalente à anterior. E ainda poderíamos adoptar muitas outras definições de vizinhança conducentes às mesmas noções topológicas, chamando vizinhança de A, por exemplo, aos cubos ou aos paralelipípedos centrados em A, etc. (as noções de cubo e de paralelipípedo são extensíveis a $\mathbf{R}^n$).

NOTA. Dados dois números naturais $m, n,$ com $m > n$, o espaço $\mathbf{R}^n$ pode ser identificado, de infinitos modos diversos, a um subconjunto de $\mathbf{R}^m$. Por exemplo, o espaço $\mathbf{R}$ pode, em $\mathbf{R}^3$, ser identificado a um dos eixos coordenados; o espaço $\mathbf{R}^2$ pode, em $\mathbf{R}^3$, ser identificado a um dos planos coordenados, etc., etc.

Porém, feita a identificação de $\mathbf{R}^n$, por um subconjunto de $\mathbf{R}^m$, há que atender ao seguinte: um ponto P, que no espaço $\mathbf{R}^n$ seja interior a um dado conjunto $\mathcal{C}$, pode já não ser interior a $\mathcal{C}$ em $\mathbf{R}^m$.

Por exemplo, no espaço $\mathbf{R}^2$, o conjunto dos pontos (x, y) tais que $x^2 + y^2 < 1$ (círculo aberto de centro na origem e raio 1) é *aberto,* isto é, formado de pontos que lhe são interiores. Mas, identificando $\mathbf{R}^2$ ao primeiro plano coordenado de $\mathbf{R}^3$, o referido conjunto (agora definido pelas condições $x^2 + y^2 < 1$, $z = 0$) já não é aberto em $\mathbf{R}^3$, pois que nenhum ponto lhe é interior: qualquer que seja o ponto desse conjunto, não existe nenhuma vizinhança do ponto (esfera nele centrada) contida no conjunto.

6. Noções de limite para as sucessões de pontos de $\mathbf{R}^n$

A noção de limite para as sucessões de pontos dum espaço numérico pluridimensional introduz-se de modo inteiramente análogo ao que foi seguido para as sucessões de números reais (pontos de $\mathbf{R}^2$):

DEFINIÇÃO 6. *Consideremos um espaço euclideano* $\mathbf{R}^p$. *Dada uma sucessão*

$$U_1, U_2, \ldots, U_n, \ldots$$

constituída por pontos de $\mathbf{R}^p$, *diz-se que tal sucessão tende para um ponto A de* $\mathbf{R}^p$ (*tem por limite A ou converge para A*) *e escreve-se:*

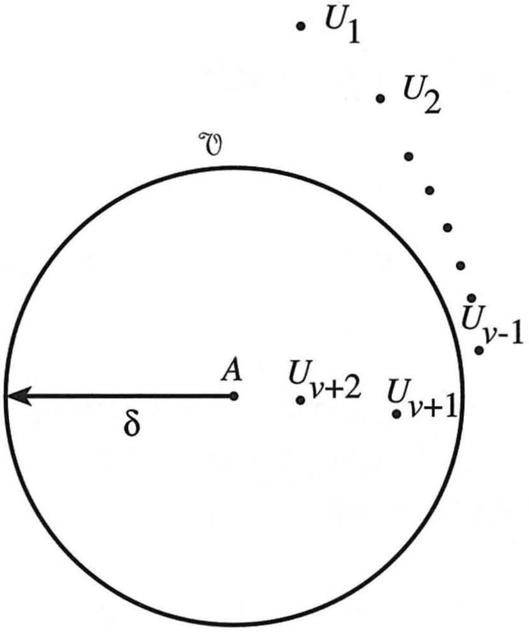
$$\lim_{n \to \infty} U_n = A \quad \text{ou} \quad U_n \to A\,,$$

quando, a todo o número $\delta > 0$, *corresponder um número natural* ν, *tal que se tenha* $|U_n - A| < \delta$ *para todo o* $n > \nu$. (Quando não houver lugar para equívoco, usaremos a notação "$\lim U_n$" em vez de "$\lim\limits_{n \to \infty} U_n$"; geralmente, o equívoco só pode dar-se quando na expressão do termo geral, figurarem outras variáveis além de n).

Desde logo se reconhece que se pode introduzir a mesma noção deste outro modo:

DEFINIÇÃO 6a. *Diz-se que o ponto variável* U_n *tende para o ponto fixo A, quando se tem* $\lim\limits_{n \to \infty} |U_n - A| = 0$, *isto é, quando a distância de* U_n *a A tende para zero, ao tender n para* ∞.

Por outro lado, observemos o seguinte: Dizer que o ponto variável U_n satisfaz à condição $|U_n - A| < \delta$ equivale a dizer que U_n pertence à vizinhança (δ) de A, (isto é, à esfera aberta de centro A e raio δ).



Então, é fácil ver que as anteriores definições resultam ainda equivalentes à seguinte:

DEFINIÇÃO 6b. *Diz-se que a sucessão U_n tende para A, quando, qualquer que seja a vizinhança ϑ de A, existir sempre uma ordem ν depois da qual todos os termos da sucessão estejam em ϑ.*

Posto isto, podemos estabelecer o seguinte:

TEOREMA 1. *Para que o ponto variável, U_n, tenha por limite o ponto fixo A, é necessário e suficiente que cada uma das coordenadas de U_n tenha por limite a coordenada homóloga de A.*

Demonstração. Pondo

$$U_n = (u_{1n}, u_{2n}, \dots, u_{pn}), \; A = (a_1, a_2, \dots, a_p),$$

virá (veja-se a DEFINIÇÃO 2)

$$(1) \qquad |U_n - A| = \sqrt{(u_{1n} - a_1)^2 + (u_{2n} - a_2)^2 + \cdots + (u_{pn} - a_p)^2}\,,$$

ou seja,

$$|U_n - A|^2 = (u_{1n} - a_1)^2 + (u_{2n} - a_2)^2 + \cdots + (u_{pn} - a_p)^2$$

e portanto:

$$|U_n - A|^2 \geqslant (u_{in} - a_i)^2, \quad 1 \leqslant i \leqslant p,$$

ou ainda:

$$(2) \qquad |u_{in} - a_i| \leqslant |U_n - A|, \quad 1 \leqslant i \leqslant p.$$

Posto isto:

a) – Demonstremos que a condição é suficiente.

Suponhamos que cada uma das coordenadas de U_n tende para a coordenada homóloga de A, isto é, que se tem:

$$a_1 = \lim_{n \to \infty} u_{1n}, \; a_2 = \lim_{n \to \infty} u_{2n}, \; \ldots, \; a_p = \lim_{n \to \infty} u_{pn};$$

vamos provar que, nesta hipótese, $\lim U_n = A$. Para isso, basta tomar os limites a ambos os membros de (1) e aplicar as regras estabelecidas em Matemáticas Gerais para limites de sucessões de números.

Ter-se-á:

$$\lim_{n \to \infty} |U_n - A| = 0,$$

o que, segundo a DEFINIÇÃO 6a, significa que:

$$\lim U_n = A.$$

b) Demonstremos que a condição é necessária. Suponhamos que $\lim\limits_{n \to \infty} U_n = A$, isto é, que $\lim\limits_{n \to \infty} |U_n - A| = 0$; trata-se de provar que, nesta hipótese, cada uma das coordenadas de U_n tende para a coordenada homóloga de A, isto é, que:

$$\lim_{n \to \infty} |u_{in} - a_i| = 0, \quad 1 \leqslant i \leqslant p,$$

mas isto é uma consequência imediata de (2) e da hipótese.

7. A noção de limite para as funções vectoriais duma variável vectorial

Tal como acontece para as funções reais duma variável real, a noção de limite para as funções vectoriais duma variável vectorial pode ser introduzida, pelo menos, de dois modos diversos:

Segundo a orientação de HEINE e segundo a orientação de CAUCHY.

Consideremos uma função

$$Y = F(X)$$

definida num subconjunto $\mathcal{C}$ de $\mathbf{R}^m$, com $Y \in \mathbf{R}^p$ (m, p números naturais quaisquer).

Designando por A e B, respectivamente, um ponto de $\mathbf{R}^m$ e um ponto de $\mathbf{R}^p$, a noção de limite segundo HEINE pode ser definida tal como segue:

DEFINIÇÃO 7. *Diz-se que $F(X)$ tende para B (ou tem por limite B) ao tender X para A, e escreve-se*

$$\lim_{X \to A} F(X) = B,$$

quando, a toda a sucessão $X_1, X_2, \ldots, X_n, \ldots$ de pontos de $\mathcal{C}$, distintos de A, tal que

$$\lim_{n \to \infty} X_n = A,$$

corresponder uma sucessão $F(X_1), F(X_2), \ldots, F(X_n), \ldots$ de pontos de $\mathbf{R}^p$ tal que

$$\lim_{n \to \infty} F(X_n) = B.$$

Segundo CAUCHY, a mesma noção é introduzida do seguinte modo:

DEFINIÇÃO 7a. *Diz-se que $F(X)$ tende para B (ou tem por limite B) ao tender X para A e escreve-se*

$$\lim_{X \to A} F(X) = B$$

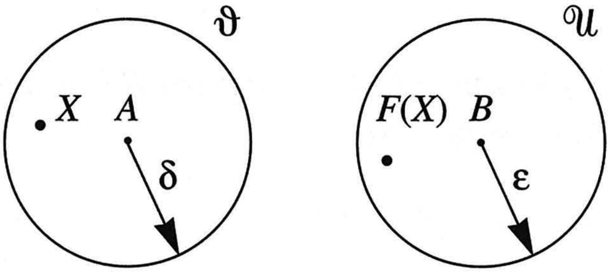
quando, a todo o número $\varepsilon > 0$, se possa associar um número $\delta > 0$, de tal modo que se tenha

$$|F(X)-B|<\varepsilon \quad \textit{sempre que} \quad |X-A|<\delta,$$

com $X \in \mathcal{C}, X \neq A$.

Tal como no caso das funções reais duma variável real, a equivalência das duas definições pode ser estabelecida com o emprego do axioma de Zermelo. Qualquer delas pressupõe que existe em $\mathcal{C}$ pelo menos uma sucessão de pontos distintos de A, convergente para este ponto (diz-se, então, que A é ponto limite de $\mathcal{C}$).

Recordemos, ainda, que a condição "$|X-A|<\delta$" se pode exprimir dizendo: "X pertence à vizinhança (δ) de A". Então, a DEFINIÇÃO 7a poderá ainda formular-se deste modo: "Diz-se que $F(X)$ tende para B ao tender X para A, quando, para toda a vizinhança $\mathcal{U}$ de B, existir pelo menos uma vizinhança ϑ de A, tal que, enquanto o ponto X percorre ϑ (permanecendo em $\mathcal{C}$ e sem passar por A), o ponto correspondente $F(X)$ não sai de $\mathcal{U}$".



A DEFINIÇÃO 7 é geralmente mais cómoda para as aplicações. *Consideremos em particular o caso duma função real de m variáveis reais:*

$$y = f(X) = f(x_1, x_2, \ldots, x_m)$$

definida no conjunto $\mathcal{C}$. Seja ainda $A = (a_1, a_2, \ldots, a_m)$ um ponto limite de $\mathcal{C}$; a DEFINIÇÃO 7 poderá agora aplicar-se *mesmo quando B for, em vez dum número real, um dos símbolos* ∞, $+\infty$ ou $-\infty$. E, para designar o "limite de $f(x_1, x_2, \ldots, x_m)$, quando X tende para A" (se tal limite existe), poderá usar-se em vez da notação

$$\lim_{X \to A} f(x_1, x_2, \ldots, x_m)$$

esta outra:

$$\lim_{\substack{x_1 \to a_1 \\ x_2 \to a_2 \\ \cdots\cdots \\ x_m \to a_m}} f(x_1, x_2, \ldots, x_m)\,.$$

Exemplos – 1) Consideremos a função z de x, y, assim definida:

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

e seja (a, b) um ponto de $\mathbf{R}^2$ *distinto da origem*, isto é, $(a, b) \neq (0, 0)$.

Atribuindo ao par de variáveis x, y uma *qualquer* sucessão de valores (x_1, y_1), (x_2, y_2), ..., (x_n, y_n), ... que tenha por limite (a, b), virá, pelo TEOREMA 1,

$$\lim x_n = a, \quad \lim y_n = b,$$

e obter-se-á para a variável z uma sucessão de valores $z_1, z_2, \ldots, z_n, \ldots$ dados por

$$z_n = \frac{1 - 2\sqrt{x_n y_n}}{x_n^2 + y_n^2}.$$

Posto isto, a álgebra dos limites, estabelecida em Matemáticas Gerais para sucessões numéricas, permite-nos escrever:

$$\lim z_n = \frac{1 - 2\sqrt{(\lim x_n) \cdot (\lim y_n)}}{(\lim x_n)^2 + (\lim y_n)^2}$$

$$= \frac{1 - 2\sqrt{ab}}{a^2 + b^2} \quad (\text{visto que } a^2 + b^2 \neq 0).$$

Por conseguinte, ter-se-á, segundo a DEFINIÇÃO 7:

$$\lim_{(x, y) \to (a, b)} z = \frac{1 - 2\sqrt{ab}}{a^2 + b^2} \text{ sendo } (a, b) \neq (0, 0).$$

Discorrendo de modo análogo, vê-se que:

$$\lim_{(x, y) \to (0, 0)} z = +\infty.$$

2) Consideremos a função z de x, y, assim definida:

$$z = \frac{y-1}{x+2}, \quad \text{para } (x, y) \neq (-2, 1),$$

e investiguemos se tal função tende ou não para um limite quando (x, y) tende para $(-2, 1)$. Atribuindo a (x, y) uma sucessão de valores (x_n, y_n) distintos de $(-2, 1)$, tal que $(x_n, y_n) \to (-2, 1)$, resultará para z a sucessão de valores dada por

$$z_n = \frac{y_n - 1}{x_n + 2}.$$

Suponhamos que os pontos (x_n, y_n) são escolhidos sobre uma recta que passe por $(-2, 1)$, isto é, uma recta de equação $y - 1 = k(x+2)$; então virá:

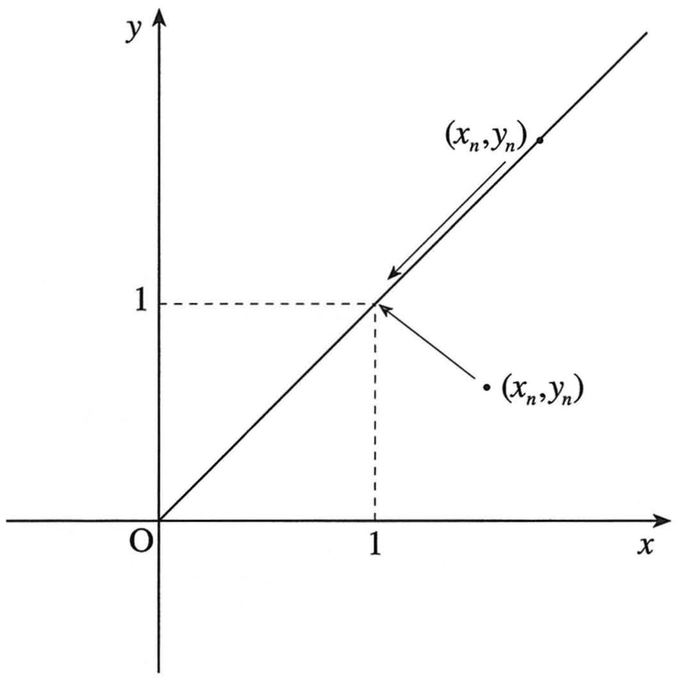
$$z_n = \frac{y_n - 1}{x_n + 2} = \frac{k(x_n + 2)}{x_n + 2} = k$$

e portanto $\lim z_n = k$. Mas como k, coeficiente angular dessa recta, pode tomar todos os valores desde $-\infty$ a $+\infty$, segue que o limite de z_n *depende da maneira como o ponto variável* (x_n, y_n) *tende para o ponto fixo* $(-2, 1)$. Por outras palavras: *não existe aquilo a que é lícito chamar "limite de z quando* (x, y) *tende para* $(-2, 1)$".

3) Seja a função:

$$\varphi(x, y) = \begin{cases} 0, & \text{se } x \neq y, \\ 1, & \text{se } (x, y) = (0, 0), \\ x^2, & \text{se } x = y \text{ e } (x, y) \neq (0, 0), \end{cases}$$

e consideremos por exemplo o ponto $(1,1)$. Se tomarmos uma sucessão de pontos (x_n, y_n) sobre a recta $x = y$ (sendo portanto $x_n = y_n$), de modo que (x_n, y_n) tenda para $(1,1)$, ter-se-á, atendendo à maneira como é definida a função $\varphi(x, y)$ para $x = y$: $\lim\limits_{n\to\infty} \varphi(x_n, y_n) = (\lim x_n)^2 = 1$.



Mas também se pode considerar uma sucessão de pontos (x_n, y_n), todos situados fora da recta $x=y$ e tal que $(x_n, y_n) \to (1,1)$; então será sempre (pelo menos a partir de certa ordem):

$$\varphi(x_n, y_n)=0, \text{ donde } \lim_{n\to\infty} \varphi(x_n, y_n)=0.$$

Como são diferentes os dois limites obtidos, *não existe o limite da função* $\varphi(x, y)$ ao tender (x, y) para $(1,1)$.

Consideremos agora o ponto $(0,0)$.

É fácil ver que, para toda a sucessão de pontos (x_n, y_n) *distintos* de $(0,0)$ e tal que:

$$(x_n, y_n) \to (0,0),$$

se tem $\varphi(x_n, y_n) \to 0$. (Neste caso é indispensável a condição de os pontos da sucessão serem distintos do ponto $(0,0)$ considerado). Teremos, portanto:

$$\lim_{\substack{x\to 0\\ y\to 0}} \varphi(x, y) = 0.$$

Os teoremas relativos aos limites da soma, do produto e do cociente de funções reais duma variável real estendem-se imediatamente ao caso das funções reais de m variáveis reais.

Consideremos duas funções reais:

$$y = f(x_1, x_2 \ldots, x_m)\ ,\ z = g\ (x_1, x_2 \ldots, x_m),$$

definidas num mesmo conjunto $\mathcal{C}$ e suponhamos que, ao tender X para um dado ponto de $\mathbf{R}^m$, tais funções tendem para *limites finitos*.

Então, a DEFINIÇÃO 7, por um lado, e a álgebra dos limites para sucessões numéricas, por outro lado, mostram-nos imediatamente que:

$$\lim_{X \to A} (y \pm z) = \lim_{X \to A} y \pm \lim_{X \to A} z$$

$$\lim_{X \to A} (yz) = (\lim_{X \to A} y) \cdot (\lim_{X \to A} z)\,.$$

Se, além disso, for $\lim\limits_{X \to A} z \neq 0$, ter-se-á, ainda,

$$\lim_{X \to A} \frac{y}{z} = (\lim_{X \to A} y) \,/\, (\lim_{X \to A} z)\,.$$

Análogas considerações, nos casos dos limites infinitos.

Vamos ver agora que, no estudo dos limites, o caso das funções *vectoriais* se reduz sempre ao caso das funções *reais*. Tem-se, com efeito:

TEOREMA 2. *Consideremos uma função vectorial $Y = F(X)$ definida num subconjunto $\mathcal{C}$ de $\mathbf{R}^m$ pelo sistema de funções reais*

$$\begin{cases} y_1 = f_1(x_1, x_2, \ldots, x_m) \\ y_2 = f_2(x_1, x_2, \ldots, x_m) \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ y_p = f_p(x_1, x_2, \ldots, x_m) \end{cases}$$

e consideremos dois pontos

$$A = (a_1, a_2, \ldots, a_m)\ ,\ B = (b_1, b_2, \ldots, b_p)\,,$$

respectivamente, de $\mathbf{R}^m$ e de $\mathbf{R}^p$. Condição necessária e suficiente *para que, ao tender X para A, Y tenda para B, é que cada uma das coordenadas de Y tenda para a coordenada homóloga de B, isto é, que se tenha:*

$$\lim_{X \to A} y_1 = b_1, \ \lim_{X \to A} y_2 = b_2, \ \ldots, \lim_{X \to A} y_p = b_p.$$

Aplicando a DEFINIÇÃO 7, facilmente se reconnece que este teorema é uma consequência imediata do TEOREMA 1.

8. A noção de continuidade para as funções vectoriais duma variável vectorial

Esta noção introduz-se imediatamente a partir da anterior noção de limite:

DEFINIÇÃO 8. *Diz-se que uma função $Y = F(X)$ é contínua num dado ponto A, quando se verificam as seguintes condições:*

1) – *a função é definida no ponto A;*
2) – *existe o limite de $F(X)$ ao tender X para A;*
3) – *tem-se:*

$$\lim_{X \to A} F(X) = F(A).$$

(A função dir-se-á *descontínua* em A, quando não for contínua em A, isto é, quando uma, pelo menos, das condições 1), 2) ou 3) não for verificada).

Mas tal noção poderá ainda introduzir-se directamente, tal como segue:

DEFINIÇÃO 8a. *Diz-se que $F(X)$ é contínua num ponto A do seu campo de existência, quando, por menor que seja o número $\varepsilon > 0$, for sempre possível associar-lhe um número $\delta > 0$, de modo que resulte $|F(X) - F(A)| < \varepsilon$ sempre que $|X - A| < \delta$.*

Subentende-se que X pertence ao campo de existência da função, mas já não é necessária a restrição $X \neq A$, visto que, para $X = A$, a condição $|F(X) - F(A)| < \varepsilon$ é sempre verificada. Pela mesma razão, deixa de ser necessário impor a A a restrição de ser ponto limite do campo de existência de $F(X)$.

No caso particular da função real de n variáveis reais,

$$y = F(X) = F(x_1, x_2, \ldots, x_n),$$

como há a possibilidade de a função assumir valores infinitos, a DEFINIÇÃO 8 tem de ser completada, substituindo nela a condição 1) por esta outra: "a função toma em A um valor *finito*".

Do TEOREMA 2 deduz-se imediatamente este outro:

TEOREMA 3. *Dada uma função vectorial $Y = F(X)$ definida num subconjunto de $\mathbf{R}^n$ pelas m funções reais*

$$\begin{cases} y_1 = f_1(x_1, x_2, \ldots, x_n) \\ y_2 = f_2(x_1, x_2, \ldots, x_n) \\ \cdots\cdots\cdots\cdots\cdots\cdots \\ y_m = f_m(x_1, x_2, \ldots, x_n) \end{cases}$$

condição necessária e suficiente para que $F(X)$ seja contínua num dado ponto A de $\mathbf{R}^n$ *é que cada uma das funções coordenadas $f_1(X)$, $f_2(X), \ldots, f_m(X)$ seja contínua em A.*

Exemplos – 1) A função real assim definida:

$$\varphi(x, y) = \begin{cases} \dfrac{y-1}{x+2}, & \text{se } x \neq -2, \\ 0, & \text{se } x = -2, \end{cases}$$

é contínua em todo o ponto (x_0, y_0) tal que $x_0 \neq -2$, pois tem nessa hipótese:

$$\lim_{(x, y) \to (x_0, y_0)} \varphi(x, y) = \frac{y_0 - 1}{x_0 + 2} \neq \infty;$$

mas em qualquer ponto $(-2, y_0)$ da recta $x = -2$ a função é descontínua porque: ou $y_0 \neq 1$ e então $\varphi(-2, y_0) = \infty$, ou $y_0 = 1$ e, como já vimos atrás, não existe o limite de $\varphi(x, y)$ quando $(x, y) \to (-2, 1)$.

2) Seja agora a função vectorial assim definida em $\mathbf{R}^2$,

$$\begin{cases} u = x^2 + y^2 \\ v = \lim\limits_{n \to \infty} (1 + x^2 + y^2)^{-n} \end{cases}.$$

A primeira função é manifestamente contínua em todo o ponto (x, y) de $\mathbf{R}^2$. Quanto à segunda, tem-se:

$$v = \begin{cases} 0, \text{ se } (x, y) \neq (0,0), \\ 1, \text{ se } (x, y) = (0,0), \end{cases}$$

portanto a função será contínua em todo o ponto distinto da origem; mas será descontínua na origem, visto que, ao tender (x, y) para $(0,0)$, v tende para 0, *não sendo esse o valor que a função toma em* $(0,0)$.

Finalmente, aplicando o TEOREMA 3, conclui-se que a função vectorial considerada é contínua em todos os pontos de $\mathbf{R}^2$, com excepção da origem.

9. Teoremas sobre funções contínuas

A álgebra dos limites permite-nos, desde logo, estabelecer as seguintes regras:

Se duas funções reais,

$$y = f(x_1, x_2, \ldots, x_n), \quad z = g(x_1, x_2, \ldots, x_n),$$

forem ambas contínuas num dado ponto A de $\mathbf{R}^n$, *também a soma ou diferença* $y \pm z$ *e o produto* $y \cdot z$ *são funções contínuas de* X *em* A; *se, além disso,* z *não for nula para* $X = A$, *então o cociente* y/z *será ainda uma função contínua de* X *em* A.

Em particular, *toda a função racional inteira é contínua em qualquer ponto*. Basta lembrar que se chama "função inteira" de x_1, x_2, …, x_n a toda a função que se pode obter como soma de um número finito de funções $\varphi(X)$ do tipo

$$\varphi(X) = \gamma\, x_1^{p_1}\, x_2^{p_2} \cdots\cdots x_n^{p_n},$$

sendo γ um número real qualquer (coeficiente do monómio) e $p_1, p_2, \ldots, p_n$, números inteiros *não negativos.*

Qualquer que seja o ponto $A = (a_1, a_2, \ldots, a_n)$, a álgebra dos limites habilita-nos a escrever:

$$\lim_{X \to A} \varphi(X) = \gamma(\lim_{X \to A} x_1)^{p_1} \ldots (\lim_{X \to A} x_n)^{p_n}.$$

Por outro lado, tem-se, em virtude do TEOREMA 1:

$$\lim_{X \to A} x_1 = a_1,\ \lim_{X \to A} x_2 = a_2,\ \ldots,\ \lim_{X \to A} x_n = a_n,$$

donde, finalmente,

$$\lim_{X \to A} \varphi(X) = \gamma\, a_1^{p_1}\, a_2^{p_2} \cdots\cdots a_n^{p_n} = \varphi(A),$$

o que significa, precisamente, ser $\varphi(X)$ contínua em A; portanto, qualquer função inteira, sendo a soma de um número finito de funções deste tipo, será também contínua no ponto arbitrário A.

Analogamente se reconhece que *toda a função racional fraccionária é contínua nos pontos que não anulam o denominador.*

Por exemplo, a função

$$f(x, y, z) = \frac{3x - y + 5z}{x^2 + 3y^2 + 2z^2 - 4}$$

é contínua em todos os pontos do espaço $\mathbf{R}^3$, excepto os que pertencem ao elipsóide $x^2 + 3y^2 + 2z^2 = 4$.

Para reconhecer se uma dada função é ou não contínua num dado ponto, é útil, muitas vezes, o teorema relativo às funções compostas que vamos demonstrar adiante. Antes disso, porém, convém esclarecer o conceito de função composta. Consideremos a função vectorial

$$Y = F(U) \tag{1}$$

definida num subconjunto $\mathcal{C}$ de $\mathbf{R}^m$, com $Y \in \mathbf{R}^p$, e seja

$$U = G(X)$$

uma segunda função vectorial, definida num subconjunto $\mathcal{D}$ de $\mathbf{R}^n$ e cujo contradomínio esteja contido em $\mathcal{C}$ (quer dizer, quando o ponto X percorre $\mathcal{D}$, o ponto U deve manter-se em $\mathcal{C}$).

É claro que, substituindo em (1), U por $G(X)$, se obtém uma nova função vectorial,

$$Y = F(G(X)) = \Phi(X),$$

definida em $\mathcal{D}$, com $Y \in \mathbf{R}^p$; diz-se, então, que *Y é função composta de X por intermédio de U.*

Adoptando a linguagem introduzida em Matemáticas Gerais, podemos também dizer que *G representa uma transformação de $\mathcal{D}$ para $\mathcal{C}$, F uma transformação* de $\mathcal{C}$ para $\mathbf{R}^p$ e finalmente Φ *uma transformação de $\mathcal{D}$ para $\mathbf{R}^p$* que *é o produto* de F por G:

$$\Phi = F.G\,.$$

Convém não perder de vista que a função vectorial $Y = F(U)$ equivale a um sistema de p funções reais de m variáveis reais

$$(2) \qquad \begin{cases} y_1 = f_1(u_1, u_2, \ldots, u_m) \\ y_2 = f_2(u_1, u_2, \ldots, u_m) \\ \dots\dots\dots\dots\dots\dots\dots\dots \\ y_p = f_p(u_1, u_2, \ldots, u_m) \end{cases}$$

definidas em $\mathcal{C}$, enquanto a função $U = G(X)$ equivale a um sistema de m funções reais de n variáveis reais

$$(3) \qquad \begin{cases} u_1 = g_1(x_1, x_2, \ldots, x_n) \\ u_2 = g_2(x_1, x_2, \ldots, x_n) \\ \dots\dots\dots\dots\dots\dots\dots\dots \\ u_m = g_m(x_1, x_2, \ldots, x_n) \end{cases}$$

definidas em $\mathfrak{D}$. É claro que para ter $y_1, y_2, \ldots, y_p$, directamente em função de $x_1, x_2, \ldots, x_n$, chegando ao sistema de funções reais representativo da função $Y = \Phi(X)$, bastará em (2) substituir $u_1, u_2, \ldots, u_m$, pelas respectivas expressões (3) em funções de $x_1, x_2, \ldots, x_n$.

Também se diz neste caso que se efectua sobre as funções (2) *a substituição ou mudança de variável definida por* (3).

(Recordemos ainda que, se as transformações F, G forem lineares, também o produto $F.G$ é uma transformação linear, que tem por matriz o produto da matriz de F pela matriz de G).

Por exemplo, da função de u, v

$$z = \sqrt{u^2 + v^2} \tag{4}$$

definida em $\mathbf{R}^2$, e, da transformação

$$\begin{cases} u = x + y \\ v = 2x - y \end{cases} \tag{5}$$

de $\mathbf{R}^2$ em $\mathbf{R}^2$, resulta uma função composta de (x, y) por intermédio de (u, v), que pode directamente escrever-se sob a forma

$$z = \sqrt{(x + y)^2 + (2x - y)^2} = \sqrt{5x^2 - 2xy + 2y^2}\,.$$

Efectuámos, portanto, em (4), *a substituição ou mudança de variáveis definida por* (5).

Posto isto, podemos estabececer o anunciado teorema relativo às funções compostas:

TEOREMA 4. *Dadas as funções $Y = F(U)$, $U = G(X)$, de modo que o domínio da primeira contenha o contradomínio da segunda*, se *$G(X)$ é contínua num certo ponto X_0 e $F(U)$ é contínua no ponto correspondente $U_0 = G(X_0)$, também a função composta $F(G(X))$ é contínua no ponto X_0.*

Demonstração. Suponhamos $G(X)$ contínua em X_0, $F(U)$ contínua em $U_0 = G(X_0)$, e seja ε um número positivo arbitrário. Segundo a

DEFINIÇÃO 8a, existirá, um número $\sigma > 0$ tal que se tenha

$$(6) \qquad |F(U) - F(U_0)| < \varepsilon \text{ para } |U - U_0| < \sigma;$$

por sua vez, segundo a mesma definição, existirá um número $\delta > 0$ tal que resulte

$$(7) \qquad |G(X) - G(X_0)| < \sigma \text{ sempre que } |X - X_0| < \delta.$$

Então, de (6) e (7) conclui-se que será $|F(G(X)) - F(G(X_0))| < \varepsilon$ todas as vezes que for $|X - X_0| < \delta$, o que, atendendo à arbitrariedade de ε, significa previamente que $F(G(X))$ é contínua para $X = X_0$, ficando completa a demonstração.

Como exemplo de aplicação consideremos a função

$$z = \log\,(x - \sqrt{1 + (x - y)^2}),$$

a qual pode considerar-se como resultante da seguinte cadeia de funções:

$$z = \log u, \quad u = v - \sqrt{w}\,, \quad v = x, \quad w = 1 + (x - y)^2.$$

Notando que a função $\log u$ é contínua para $u > 0$ e que $\sqrt{w}$ é contínua para $w > 0$, atendendo à álgebra dos limites e aplicando, finalmente, o TEOREMA 4 duas vezes sucessivas, conclui-se que z é função contínua de (x, y) em todo o ponto tal que $x^2 > 1 + (x - y)^2$.

10. Funções contínuas num conjunto. Teoremas de CANTOR e de WEIERSTRASS

No n.º precedente foi definido o conceito de *"função contínua num dado ponto"*. Agora convém deter um pouco a atenção sobre o conceito de *"função contínua num dado conjunto"*.

DEFINIÇÃO 9. *Seja $F(X)$ uma função vectorial definida num conjunto $\mathcal{C} \subset \mathbf{R}^n$. Dado um conjunto $\mathcal{M}$ contido em $\mathcal{C}$ (podendo em particular coincidir com $\mathcal{C}$), diz-se que $F(X)$ é contínua no conjunto $\mathcal{M}$, quando a restrição de $F(X)$ a $\mathcal{M}$ é contínua em cada um dos* *pontos de $\mathcal{M}$. (Recordemos que se chama restrição de $F(X)$ a $\mathcal{M}$ à função $F^*(X)$ cujo campo de existência é $\mathcal{M}$ e tal que $F^*(X)=F(X)$ em cada ponto $X \in \mathcal{M}$).*

Exemplos – 1) Seja a função z de x, y, assim definida em todo o espaço $\mathbf{R}^2$:

$$z = \begin{cases} x^2 + e^{x-2y}, & \text{para } x^2 + y^2 \leqslant 1, \\ 0, & \text{para } x^2 + y^2 > 1. \end{cases}$$

Em qualquer ponto (x_0, y_0) da circunferência $x^2 + y^2 = 1$ a função apresenta uma discontinuidade, pois que se verificam os seguintes factos: a) – o valor de z nesse ponto é, com certeza, >0; b) – fazendo tender (x, y) para (x_0, y_0) por uma sucessão de pontos (x_n, y_n), situados fora do círculo $x^2 + y^2 \leqslant 1$, z tende para 0.

Por conseguinte, a função *não* é contínua no seu domínio de existência. Todavia, se designarmos por $\mathcal{M}$ o círculo $x^2 + y^2 \leqslant 1$, a função será contínua no conjunto $\mathcal{M}$, pois que, fazendo tender (x, y) para um ponto (x_0, y_0) qualquer de $\mathcal{M}$, *passando unicamente por pontos de $\mathcal{M}$*, a função tende para $x_0^2 + e^{x_0 - 2y_0}$, valor que assume nesse ponto.

2) Seja a função:

$$\varphi(x, y, z) = \begin{cases} x^2 + y^2 + z^2, & \text{para } x - y + z = 0, \\ 0, & \text{para } x - y + z \neq 0. \end{cases}$$

A função não é contínua no seu domínio de existência, porque em qualquer ponto do plano $x - y + z = 0$, excepto a origem, não existe sequer limite da função. Porém, se a restringirmos ao plano $x - y + z = 0$, onde é representada pela expressão $x^2 + y^2 + z^2$, a função já será contínua.

Um outro conceito de grande importância em Análise Matemática é o de "continuidade uniforme":

DEFINIÇÃO 10. *Diz-se que a função $F(X)$ é uniformemente contínua num dado conjunto $\mathcal{M}$ contido no seu campo de existência* *quando, por menor que seja o número* $\varepsilon > 0$*, exista sempre um correspondente número* $\delta > 0$*, de modo que se tenha* $|F(X) - F(X')| < \varepsilon$ *para todo o par de pontos* X, X' *de* $\mathcal{M}$ *cuja distância seja inferior a* δ*, isto é, tais que* $|X - X'| < \delta$.

Confrontando esta definição com as DEFINIÇÕES 8a e 9, tem-se à primeira vista a impressão de que "função uniformemente contínua em $\mathcal{M}$" e "função contínua em $\mathcal{M}$" são uma e a mesma coisa. Note-se, porém, que, na DEFINIÇÃO 8a, o número δ em questão dependerá, em geral, não só do número ε dado, *mas também do ponto A de que se trata*; o que há de essencialmente novo na DEFINIÇÃO 10 é o facto de o número δ depender unicamente de ε. E, na verdade, conhecem-se exemplos muito simples de funções que, sendo contínuas num conjunto $\mathcal{M}$, não resultam uniformemente contínuas nesse conjunto.

Há, porém, um caso importante em que a simples continuidade arrasta a continuidade uniforme. Para o apresentar, convirá ainda introduzir a seguinte:

DEFINIÇÃO 11. *Um conjunto* $\mathcal{M}$ *de pontos de* $\mathbf{R}^n$ *diz-se limitado quando for finito o extremo superior das distâncias dos seus pontos à origem; por outros termos: quando existir, pelo menos, uma esfera que o contenha. Caso contrário, diz-se ilimitado.*

Por exemplo, em $\mathbf{R}^3$, um cubo, um elipsóide, etc., são conjuntos limitados, enquanto uma recta, um plano, um parabolóide, etc., são conjuntos ilimitados.

Posto isto, enunciaremos, sem demonstração, o

TEOREMA 5 (de CANTOR). *Toda a função* $F(X)$ *que seja contínua num subconjunto limitado e fechado de* $\mathbf{R}^n$ *é uniformemente contínua nesse conjunto.* (Veja-se no n.º 5 a definição de conjunto fechado).

Consideremos agora, em particular, o caso das funções reais. É fácil ver que, a tais funções, se estendem naturalmente os conceitos de extremo superior, máximo, etc., definidos para funções reais de *uma* variável real.

Dada uma função real $y = f(X)$ definida num conjunto $\mathcal{C} \subset \mathbf{R}^n$, chama-se *extremo superior* de $f(X)$ num dado conjunto $\mathcal{M} \subset \mathcal{C}$, e representa-se por

$$\sup_{X \in \mathcal{M}} f(X)$$

o extremo superior[1] do conjunto dos valores que toma y quando X percorre $\mathcal{M}$. Se existir, pelo menos, um ponto X_0 de $\mathcal{M}$ tal que $f(X_0)$ iguale o extremo superior de $f(X)$ em $\mathcal{M}$, então (e só então) este poderá chamar-se o *máximo* de $f(X)$ em $\mathcal{M}$.

Análogas definições para "extremo inferior" e "mínimo". A função dir-se-á *limitada em* $\mathcal{M}$, quando forem finitos os seus extremos superiores e inferior em $\mathcal{M}$, ou, o que é equivalente, quando existir, pelo menos, um número k tal que

$$|f(X)| < k, \text{ qualquer que seja } X \in \mathcal{M}.$$

O TEOREMA DE WEIERSTRASS subsiste para as funções reais de n variáveis reais com o seguinte aspecto.

TEOREMA 6 (de WEIERSTRASS). *Toda a função real $f(X)$, contínua num conjunto limitado e fechado $\mathcal{M} \subset \mathbf{R}^n$, tem nesse conjunto um máximo e um mínimo; isto é, existirão, pelo menos, um ponto A e um ponto B em $\mathcal{M}$, tais que*

$$f(A) = \sup_{X \in \mathcal{M}} f(X) \quad , \quad f(B) = \inf_{X \in \mathcal{M}} f(X) .$$

Também não demonstraremos aqui este teorema. Observe-se no entanto que dele se deduz ainda este facto:

---

(1) – Recordemos que, dado um conjunto H de números reais, se chama *majorante* de H a todo o número que goze da propriedade de não ser inferior a nenhum elemento de H. Se um tal número existe, H diz-se *limitado superiormente* e chama-se *extremo superior* de H (em símbolos: sup H) o menor dos seus majorantes. Este pode pertencer ou não a H: por exemplo, o conjunto dos números *menores* que 1, tem 1 por extremo superior, número este que não pertence ao conjunto. Se o conjunto H não é limitado superiormente, convenciona-se dizer que o seu extremo superior é $+\infty$.

Toda a função real $f(X)$, contínua num conjunto limitado e *fechado $\mathcal{M} \subset \mathbf{R}^n$, é limitada em $\mathcal{M}$.*

Com efeito, o máximo e o mínimo de $f(X)$ em $\mathcal{M}$ serão finitos por serem valores dessa *função que é contínua em $\mathcal{M}$.*

11. Infinitésimos com X. Conceito de ordem

Consideremos uma função real de n variáveis reais

$$y = \varphi(X) = \varphi(x_1, x_2, \ldots, x_n)$$

definida num conjunto $\mathcal{C} \subset \mathbf{R}^n$, e seja A um ponto limite de $\mathcal{C}$; diz-se que $\varphi(X)$ é um *infinitésimo* ao tender X para A (ou *um infinitésimo com $X - A$*), quando se tem

$$\lim_{X - A \to \vec{0}} \varphi(X) = 0\,,$$

onde $\vec{0}$ designa a *origem* ou *vector nulo* de $\mathbf{R}^n$:

$$\vec{0} = (0, 0, \ldots, 0)\,.$$

(Em particular, quando A é a origem, vem $X - A = X$, e então falaremos simplesmente dum infinitésimo com X).

DEFINIÇÕES 12 e 13. *Dados dois infinitésimos com X, $\varphi(X)$ e $\psi(X)$, diz-se que tais infinitésimos são da mesma ordem, quando o cociente $\dfrac{\varphi(X)}{\psi(X)}$ tender para um limite finito e diferente de* 0, *ao tender X para $\vec{0}$*[1]. *Diz-se que o infinitésimo $\varphi(X)$ é de ordem superior à de $\psi(X)$, quando se tiver*

$$\lim_{X \to \vec{0}} \frac{\varphi(X)}{\psi(X)} = 0.$$

(1) – Alguns autores utilizam um conceito de "infinitésimos da mesma ordem" mais geral do que este aqui definido, o qual porém é suficiente para as necessidades do nosso curso.

Sejam, por exemplo, as funções de duas variáveis

$$\varphi(x, y) = x\sqrt{y}, \quad \psi(x, y) = \sqrt{x^2 + y^2} = |(x, y)| .$$

Visto que $|x| \leqslant \sqrt{x^2 + y^2}$, será sempre

$$\left|\frac{x}{\sqrt{x^2 + y^2}}\right| \leqslant 1 ;$$

ora como:

$$\lim_{\substack{x \to 0 \\ y \to 0}} \sqrt{y} = 0 ,$$

tem-se:

$$\lim_{\substack{x \to 0 \\ y \to 0}} \left(\frac{x}{\sqrt{x^2 + y^2}} \sqrt{y}\right) = 0$$

pois que o produto duma variável limitada por um infinitésimo é ainda um infinitésimo; portanto $x\sqrt{y}$ é um infinitésimo com (x, y) de ordem superior à de $\sqrt{x^2 + y^2}$.

Note-se, porém, que, dados dois infinitésimos com X, $\varphi(X)$, $\psi(X)$, pode acontecer que *não exista* o limite de $\dfrac{\varphi(X)}{\psi(X)}$ ao tender X para $\vec{0}$; diremos neste caso que se trata de *infinitésimos incomparáveis*.

DEFINIÇÃO 14. *Sejam μ um número real > 0 e $\varphi(X)$ um infinitésimo com X, definido num subconjunto de $\mathbf{R}^n$.*

Diz-se que o infinitésimo $\varphi(X)$ é de ordem μ, quando for da mesma ordem que o infinitésimo $|X|^{\mu}$ (ao tender X para $\vec{0}$); isto é, quando existir o

$$\lim_{X \to \vec{0}} \frac{\varphi(X)}{|X|^{\mu}} ,$$

sendo este finito e diferente de 0.

Pode acontecer (e acontece em muito casos da prática) que a um dado infinitésimo $\varphi(X)$ não corresponda nenhum número real aferidor da sua ordem. Assim, por exemplo, a função $x y^{1/2}$ é um infinitésimo com (x, y) a cuja ordem não corresponde nenhum número real, como se pode verificar.

12. Derivadas parciais

Comecemos por considerar o caso duma função real de duas variáveis reais. Seja a função:

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

definida num subconjunto $\mathcal{C}$ de $\mathbf{R}^2$. Dado um ponto (a, b) *interior* a $\mathcal{C}$, existirá (por definição de ponto interior), pelo menos, um número $\delta > 0$ tal que a vizinhança (δ) de (a, b) esteja contida em $\mathcal{C}$.

Ora bem, se atribuirmos a y o valor b, a função $f(x, y)$ converte-se na função $f(x, b)$ *só de* x, que será definida, pelo menos, para os valores de x tais que $a - \delta < x < a + \delta$ (pois que, então, a distância de (x, b) a (a, b) é $< \delta$). Ora, pode acontecer que a função $f(x, b)$ de x admita derivada no ponto a; em tal hipótese, essa derivada chamar-se-á a *derivada parcial da função* $f(x, y)$ *em ordem a* x *no ponto* (a, b) e designar-se-á indiferentemente por qualquer das notações

$$f'_x(a, b), \frac{\partial f}{\partial x}(a, b), \left(\frac{\partial z}{\partial x}\right)_{(a, b)}, D_x f(a, b).$$

Recordando a definição de derivada para funções duma só variável, dada em Matemáticas Gerais, teremos, pois,

$$f'_x(a, b) = \lim_{h \to 0} \frac{f(a + h, b) - f(a, b)}{h}.$$

De modo perfeitamente análogo, *se definirá derivada parcial de* $f(x, y)$ *em ordem a* y: será o limite da razão incremental

$$\frac{f(a, b + k) - f(a, b)}{k}$$

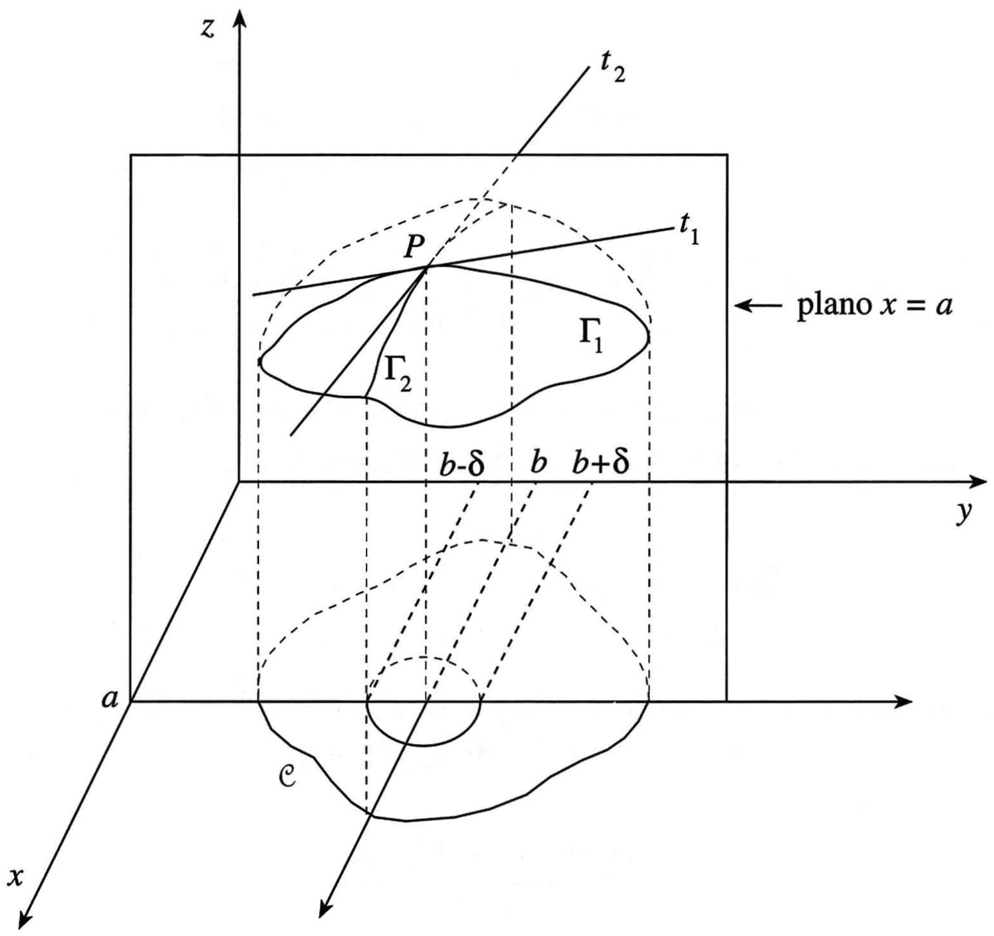
quando $k \to 0$, se este limite existir; e designar-se-á por qualquer das notações:

$$f'_y(a,\, b),\ \frac{\partial f}{\partial y}(a,\, b),\ \left(\frac{\partial z}{\partial y}\right)_{(a,\, b)},\ D_y f(a,\, b)\,.$$

Não será difícil, agora, encontrar um significado geométrico para as derivadas parciais $f'_x(a, b)$, $f'_y(a, b)$.

Supondo a função $f(x, y)$ contínua em $\mathcal{C}$, sendo este um conjunto limitado por uma linha fechada (por ex., uma circunferência), o gráfico de $f(x, y)$ será aquilo a que podemos chamar uma *superfície*, Σ; então, no plano $x = a$, a curva Γ_1 de equação $z = f(a, y)$ é a intersecção desse plano com a superfície Σ.

Portanto, pelo que sabemos de Matemáticas Gerais, se designarmos por P o ponto de coordenadas a, b, $f(a, b)$, a derivada parcial $f'_y(a, b)$ será a tangente trigonométrica do ângulo que a recta t_1, tangente a Γ_1 em P, faz com a recta $x = a$ do plano xy (orientada no sentido do eixo dos yy).



Analogamente, representando por Γ_2 a intersecção do plano $y=b$ com a superfície Σ, a derivada parcial $f'_x(a, b)$ será a tangente trigonométrica do ângulo que a recta t_2, tangente a Γ_2 em P, faz com a recta $y=b$ do plano xy (orientada no mesmo sentido que o eixo dos xx).

Pode acontecer que a função $f(x, y)$ admita derivada parcial em ordem a x em *todo* o ponto (x, y) *interior* a $\mathcal{C}$. A derivada parcial de $f(x, y)$ em ordem a x no ponto genérico (x, y) será então uma nova função das *duas variáveis* x, y, definida no interior de $\mathcal{C}$ e que poderá designar-se por qualquer das notações,

$$f'_x(x, y), \quad \frac{\partial f}{\partial x}, \quad \frac{\partial z}{\partial x}, \quad D_x f(x, y), \text{ etc.}$$

Analogamente, se $f(x, y)$ admitir derivada parcial em ordem a y em todos os pontos interiores a $\mathcal{C}$, ficará definido no interior de $\mathcal{C}$, uma nova função de x, y, que se representará por qualquer dos símbolos

$$f'_y(x, y), \quad \frac{\partial f}{\partial y}, \quad \frac{\partial z}{\partial y} \quad D_y f(x, y), \text{ etc.}$$

Na prática, quando nos é dada uma função elementar $f(x, y)$ que admita derivadas parciais no interior de algum conjunto, para calcular a derivada de $f(x, y)$ em relação a uma das variáveis, o que há a fazer é aplicar as regras de derivação em ordem a essa variável, *considerando a outra como constante.*

Exemplo – Seja a função

$$z = \operatorname{sen}(x - 3y) + \log(1 - x^2 - y^2),$$

definida no círculo de centro na origem e raio 1; ter-se-á, no interior desse círculo:

$$\frac{\partial z}{\partial x} = \cos(x - 3y) - \frac{2x}{1 - x^2 - y^2},$$

$$\frac{\partial z}{\partial y} = -3\cos(x - 3y) - \frac{2y}{1 - x^2 - y^2}.$$

Mais geralmente, dada uma função real,

$$y = f(x_1, x_2, \ldots, x_n),$$

definida num subconjunto $\mathcal{C}$ de $\mathbf{R}^n$, com $n>1$ e sendo $A=(a_1, a_2, \ldots, a_n)$ um ponto *interior* de $\mathcal{C}$, pode acontecer que, ao fazer as substituições

$$x_1 = a_1, \ldots, x_{i-1} = a_{i-1}, x_{i+1} = a_{i+1}, \ldots, x_n = a_n,$$

se obtenha uma função só de x_i,

$$y = f(a_1 \ldots, a_{i-1}, x_i, a_{i+1}, \ldots, a_n),$$

que admita derivada no ponto a_i; chamar-lhe-emos, então, a *derivada parcial de y em ordem a* x_i e podemos representá-la por qualquer dos símbolos

$$f'_{x_i}(A), \frac{\partial f}{\partial x_i}(A), \left(\frac{\partial y}{\partial x_i}\right)_A, D_{x_i} f(A)$$

(é claro que i pode ser $= 1, 2, \ldots, n$).

Se a função considerada admitir derivada parcial em ordem a x_i em qualquer ponto interior de $\mathcal{C}$, essa derivada parcial será uma função das n variáveis $x_1, x_2, \ldots, x_n$, que poderá representar-se por um dos símbolos

$$f'_{x_i}(X), \frac{\partial f}{\partial x_i}, \frac{\partial y}{\partial x_i}, D_{x_i} f(X), \text{etc.}$$

13. Conceitos de diferença finita e de diferencial para funções de mais de uma variável

O conceito de diferença finita tem grande interesse em Matemáticas aplicadas. Podemos introduzi-lo do seguinte modo:

DEFINIÇÃO 15. *Seja* $f(X)$ *uma função real definida num conjunto* $\mathcal{C} \subset \mathbf{R}^n$ *e seja* $H = (h_1, h_2, \ldots, h_n)$ *um elemento qualquer de* $\mathbf{R}^n$.

Chama-se diferença finita de $f(X)$, correspondente a H, no ponto X, a diferença

$$f(X+H)-f(X)=f(x_1+h_1, x_2+h_2, \ldots, x_n+h_n)-f(x_1, x_2, \ldots, x_n)$$

que resulta para $f(X)$ quando se passa do ponto X para o ponto $X+H$.

É claro que a diferença $f(X+H)-f(X)$ será uma função de X definida no conjunto $\mathcal{C}^*$ dos pontos X tais que $X \in \mathcal{C}$, $X+H \in \mathcal{C}$. (O conjunto $\mathcal{C}^*$ aproxima-se tanto mais de $\mathcal{C}$ quanto menor for $|H|$; em particular, pode ser vazio).

Para designar a diferença finita de $f(X)$ correspondente a H usaremos a notação

$$\Delta_H f(X)$$

ou, simplesmente, $\Delta f(X)$, quando estiver subentendido o acréscimo H de que se trata. Ter-se-á, pois, por definição:

$$\Delta_H f(X)=f(X+H)-f(X) \quad (\text{para } X \in \mathcal{C}^*).$$

TEOREMA 7. *Se a função real $f(X)$ admite derivadas parciais, $f'_{x_i}(X)$, $f'_{x_2}(X)$, $\ldots$, $f'_{x_n}(X)$, numa vizinhança ϑ dum ponto A* de $\mathbf{R}^n$ *e se essas derivadas, como funções de X, são contínuas no ponto A, então pode escrever-se*:

$$f(A+H)-f(A)=f'_{x_1}(A)\,h_1+f'_{x_2}(A)\,h_2+\ldots, f'_{x_n}(A)\,h_n+\omega(H)$$

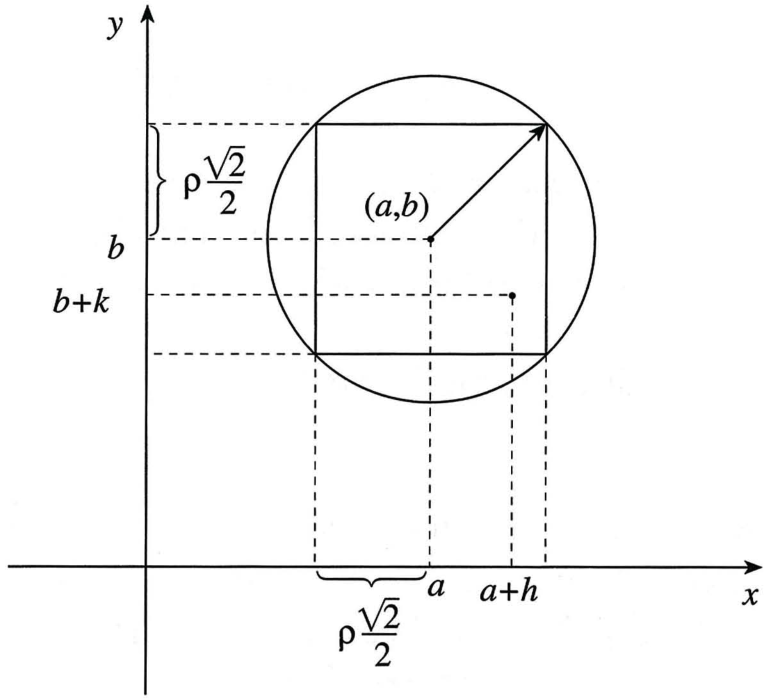
para todo o H tal que $A+H \in \vartheta$, sendo $\omega(H)$ *um infinitésimo com H de ordem superior à primeira, isto é, tendo-se*

$$\lim_{H \to \vec{0}} \frac{\omega(H)}{|H|}=0.$$

Além disso, a função $f(X)$ será contínua em A.

Bastará fazer a demonstração para o caso das funções de duas variáveis reais, pois que, no caso geral, a demonstração é análoga.

Dada uma função real $f(x, y)$ das duas variáveis reais, x, y, suponhamos que existe um ponto (a, b) de $\mathbf{R}^2$ e um número $\rho > 0$, tal que $f(x, y)$ admite na vizinhança (ρ) de (a, b) derivadas parciais $f'_x(x, y)$,



$f'_y(x, y)$, contínuas no ponto (a, b). (Podemos, então, supor ρ já escolhido de modo que as derivadas parciais sejam finitas na referida vizinhança, pois que, quando $(x, y) \to (a, b)$, elas tendem para limites finitos). Se tomarmos $|h| < \rho\frac{\sqrt{2}}{2}$, $|k| < \rho\frac{\sqrt{2}}{2}$, será $\sqrt{h^2 + k^2} < \rho$ e, portanto, o ponto $(a + h, b + k)$ estará na vizinhança (ρ) de (a, b). Ponhamos, então,

$$(1) \qquad \Delta f = f(a + h, b + k) - \mathrm{f}\,(a, b)$$

$$= [\,f(a + h, b + k) - f(a, b + k)] + [\,f(a, b + k) - f(a, b)]$$

para valores de h, k tais que $|h| < \rho\frac{\sqrt{2}}{2}$, $|k| < \rho\frac{\sqrt{2}}{2}$.

Ora, $f(x, b + k)$, como função de x, é regular no intervalo $[a, a + h]$, pois que admite aí *derivada finita* $f'_x(x, b + k)$; podemos, então, aplicar-lhe o teorema dos acréscimos finitos, o que dá

$$(2) \qquad f(a + h, b + k) - f(a, b + k) = hf'_x(a + \Theta\, h, b + k),$$

sendo Θ um número (dependente de h) tal que $0 < \Theta < 1$. É claro que será $\lim\limits_{h \to 0} \Theta\, h = 0$; então, quando (h, k) tende para $(0,0)$ o ponto $(a + \Theta\, h, b + k)$ tenderá para (a, b) e, por sua vez, $f'_x(a + \Theta\, h, b + k)$ irá tender para $f'_x(a, b)$, visto que, por hipótese, a função $f'_x(x, y)$ de x, y é contínua em (a, b). Deste modo, se pusermos

$$f'_x(a + \Theta\, h, b + k) = f'_x(a, b) + \varepsilon_1,$$

será ε_1 um infinitésimo com (h, k), tendo-se, além disso, por substituição em (2)

$$f(a + h, b + k) - f(a, b + k) = h\,[f'_x(a, b) + \varepsilon_1].$$

Analogamente se reconhece que

$$f(a, b + k) - f(a, b) = k\,[f'_x(a, b) + \varepsilon_2],$$

sendo ε_2 um infinitésimo com (h, k). (Para esta conclusão bastaria até entrar com a definição de derivada). Finalmente, por substituição em (1) tem-se:

$$\Delta f = h\, f'_x(a, b) + k\, f'_y(a, b) + h\varepsilon_1 + k\varepsilon_2. \tag{3}$$

Ora, por ser $|h| \leqslant \sqrt{h^2 + k^2}$, $|k| \leqslant \sqrt{h^2 + k^2}$, vem:

$$\left|\frac{h}{\sqrt{h^2 + k^2}}\right| \leqslant 1, \quad \left|\frac{k}{\sqrt{h^2 + k^2}}\right| \leqslant 1,$$

donde:

$$\left|\frac{h\varepsilon_1 + k\varepsilon_2}{\sqrt{h^2 + k^2}}\right| \leqslant |\varepsilon_1| \left|\frac{h}{\sqrt{h^2 + k^2}}\right| + |\varepsilon_2| \left|\frac{k}{\sqrt{h^2 + k^2}}\right| \leqslant |\varepsilon_1| + |\varepsilon_2|$$

e, como ε_1, ε_2 tendem para 0 quando $(h, k) \to (0,0)$, o mesmo acontece com o 1.º membro, o que significa (DEFINIÇÕES 13 e 14) que $h\varepsilon_1 + k\varepsilon_2$ é um infinitésimo com (h, k) de ordem superior à de $\sqrt{h^2 + k^2}$, ou seja, de ordem superior à 1.ª.

Finalmente, a continuidade de $f(x, y)$ no ponto (a, b) decorre imediatamente de (3), notando que $\Delta f \to 0$ quando $(h, k) \to (0,0)$.

DEFINIÇÃO 16. *Diremos que a função real* $f(X)$ *é continuamente derivável num dado conjunto aberto em* $\mathbf{R}^n$, *quando admite derivadas parciais* $f'_{x_1}(X)$, $f'_{x_2}(X)$, $\ldots$, $f'_{x_n}(X)$, *todas contínuas nesse conjunto.*

DEFINIÇÃO 17. *Na hipótese de a função real* $f(X)$ *ser continuamente derivável no interior de* $\mathcal{C} \subset \mathbf{R}^n$, *e sendo* $H = (h_1, h_2, \ldots, h_n)$ *um elemento qualquer de* $\mathbf{R}^n$ *chama-se diferencial de* $f(X)$ *correspondente a* H *no ponto* X, *a função linear de* $h_1, h_2, \ldots, h_n$,

$$f'_{x_1}(X)\, h_1 + f'_{x_2}(X)\, h_2 + \cdots + f'_{x_n}(X)\, h_n.$$

É claro que os coeficientes deste polinómio em $h_1, h_2, \ldots, h_n$ são funções de X definidas no *interior de* $\mathcal{C}$. Para designar o diferencial de $f(X)$ correspondente ao acréscimo H, usaremos a notação de $d_H f(X)$ ou, simplesmente, $df(X)$, quando estiver subentendido o acréscimo de que se trata. Ter-se-á pois, por definição,

$$d_H f(X) = \frac{\partial f}{\partial x_1} h_1 + \frac{\partial f}{\partial x_2} h_2 + \cdots + \frac{\partial f}{\partial x_n} h_n .$$

O TEOREMA 7 indica-nos que, para valores pequenos de $|H|$, *o diferencial* $df(X)$ nos dá uma boa aproximação *da diferença finita* $\Delta f(X)$.

Mais precisamente, aquele teorema conduz-nos ao seguinte

COROLÁRIO. *Se a função real* $f(X)$ *é continuamente derivável no interior de C, tem-se, em relação ao acréscimo* H:

$$\Delta f(X) = df(X) + \omega \quad (para\ X \in \operatorname{int} \mathcal{C},\ X+H \in \mathcal{C}),$$

sendo ω *um infinitésimo com* H *de ordem superior à* 1.ª; por outras palavras: *a diferença finita e o diferencial de* f *em cada ponto* X *onde ambos são definidos, diferem por um infinitésimo de ordem superior à de* $|H|$.

Como exemplo, consideremos o caso duma função inteira do 1.º grau:

$$f(X) \equiv c_0 + c_1 x_1 + c_2 x_2 + \cdots + c_n x_n$$

sendo c_0, c_1, c_2, ..., c_n constantes (números reais quaisquer). Neste caso, é fácil ver que se tem

$$\Delta_H f(X) = d_H f(X) = c_1 h_1 + c_2 h_2 + \cdots + c_n h_n,$$

sendo, pois, constantes os coeficientes diferenciais; mas também é este o único caso em que tal sucede.

Mais particularmente ainda, suponhamos que todos os coeficientes c_1, c_2 ..., c_n, são nulos, excepto um, $c_i = 1$; então, o referido polinómio reduz-se, neste caso, a

$$f(X) \equiv x_i,$$

tendo-se $d_H f(X) = d_H x_i = h_i$. Em conclusão:

Os diferenciais das funções de X, que se reduzem às variáveis independentes x_1, x_2, ..., x_n, *coincidem sempre com os respectivos acréscimos* h_1, h_2, ..., h_n:

$$dx_1 = h_1, \ldots, dx_i = h_i, \ldots, dx_n = h_n.$$

Podemos, pois, escrever a expressão do diferencial de $f(X)$, no caso geral, sob a forma

$$\boxed{df(X) = \frac{\partial f}{\partial x_1} dx_1 + \frac{\partial f}{\partial x_2} dx_2 + \cdots + \frac{\partial f}{\partial x_n} dx_n}$$

que tem a vantagem de deixar de obrigar a uma indicação explícita do acréscimo H. É com este aspecto que os diferenciais se apresentam sistematicamente na prática.

Exemplo – Formar o diferencial da função de x, y:

$$z = \operatorname{sen} x + \sqrt{x - y}.$$

Visto que

$$\frac{\partial z}{\partial x} = \cos x + \frac{1}{2}(x-y)^{-\frac{1}{2}}, \qquad \frac{\partial z}{\partial y} = -\frac{1}{2}(x-y)^{-\frac{1}{2}},$$

ter-se-á:

$$dz = \left(\cos x + \frac{1}{2\sqrt{x-y}}\right) dx - \frac{1}{2\sqrt{x-y}}\, dy .$$

Note-se que o domínio de existência da função é o semiplano $x \geqslant y$, ao passo que os coeficientes do diferencial se consideram definidos no interior desse domínio, ou seja, para $x > y$.

Aplicação do conceito de diferencial às ciências da Natureza.

Vimos como o acréscimo Δf duma função continuamente derivável f, relativo aos acréscimos $h_1, h_2, \ldots, h_n$, das variáveis independentes, se decompõe na soma

$$\Delta f = df + \omega ,$$

sendo ω um infinitésimo com $(h_1, h_2, \ldots, h_n)$ de ordem superior à primeira. Daqui resulta que, para valores de $h_1, h_2, \ldots, h_n$ suficientemente pequenos, a diferença finita Δf pode ser substituída, *sem erro sensível*, pelo diferencial df.

Esta circunstância é aproveitada correntemente nas aplicações de Matemática (à Física, à Engenharia, etc.).

Em muitas questões de ordem prática, quer em cálculos quer em raciocínios, é uso tomar o diferencial df pela diferença finita Δf, como se fosse precisamente

$$\Delta f = df.$$

Nisto consistem os *métodos abreviados de cálculo e de raciocínio,* a que já fizemos alusão em Matemáticas Gerais, a propósito das funções reais duma só variável real. É claro que, procedendo assim, se comete geralmente um erro ω, o qual, porém, *é uma fracção de* $|H|$ *tanto mais pequena quanto menores forem os acréscimos*. Nestas condições, será muitas vezes cómodo, na prática, tomar df por Δf, mas convém não perder de vista que se trata dum processo destituído de rigor, que, a ser usado sem precauções, pode conduzir a resultados pouco satisfatórios.

Vimos também que se escreve frequentemente dx_1, dx_2, ... em vez de h_1, h_2, ... (diferenciais das variáveis independentes). Na prática, os símbolos dx_1, dx_2, ... costumam ser interpretados como *acréscimos muito pequenos* das variáveis x_1, x_2, ... Nos primórdios do Cálculo Infinitesimal, segundo a orientação de LEIBNIZ, os símbolos dx_1, dx_2, ... representavam acréscimos *infinitamente pequenos* de x_1, x_2, ... , mas, nesses tempos, dava-se a designação de "infinitamente pequenos" a quantidades que, sem serem nulas, seriam inferiores, em valor absoluto, a qualquer quantidade finita dada (>0).

Assim, entre 0 e os números reais maiores que 0 (*quantidades finitas*) estariam os infinitamente pequenos.

Porém, a análise lógica dos fundamentos da Matemática, levada a cabo no século passado, rejeitou esta concepção dos infinitamente pequenos, substituindo-a pela que se adopta hoje em todos os cursos de Cálculo Infinitesimal.

14. Regra de derivação das funções compostas

Comecemos por considerar uma função real,

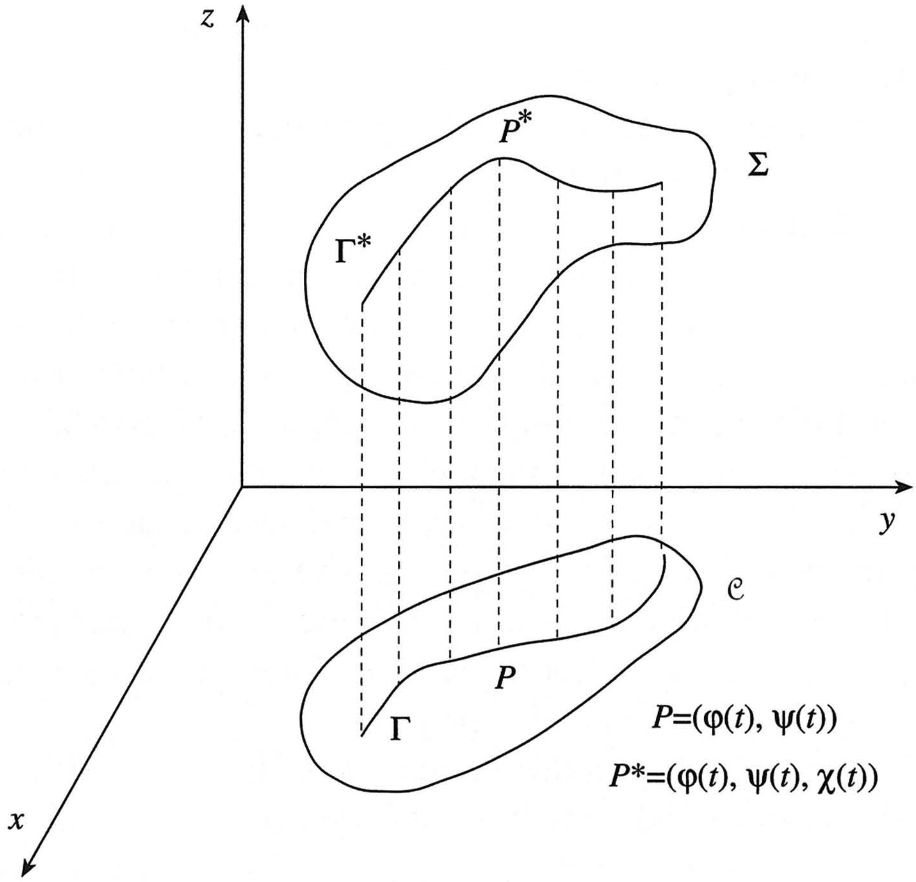
$$z = f(x, y), \tag{1}$$

das variáveis reais x, y *que seja continuamente derivável* num conjunto *aberto* $\mathcal{C}$ de pontos de $\mathbf{R}^2$.

Sejam, por outro lado,

$$x = \varphi(t) \quad , \quad y = \psi(t) \tag{2}$$

funções reais da variável t, que admitam derivadas *finitas* (de 1.ª ordem) num intervalo $\mathcal{I} \subset \mathbf{R}$ e tais que, ao variar t em $\mathcal{I}$ o ponto (x, y) se mantenha sobre o conjunto $\mathcal{C}$. (Geometricamente, sucede que, quando t percorre $\mathcal{I}$, o ponto P de coordenadas $(\varphi(t), \psi(t))$ no plano xy descreve uma linha Γ contida em $\mathcal{C}$). Associando as expressões



(1), (2), a variável z aparece como função composta de t, por intermédio de x, y, funções que representaremos por $\chi(t)$, isto é,

$$z = f(\varphi(t),\ \psi(t)) = \chi(t),$$

sendo esta (TEOREMA 4) uma função contínua de t em $\mathcal{J}$. (Geometricamente, quando t percorre $\mathcal{J}$, o ponto P^* de coordenadas $\varphi(t)$, $\psi(t)$, $\chi(t)$ descreve uma curva Γ^* sobre a superfície Σ de equação $z = f(x, y)$, sendo Γ a projecção de Γ^* sobre o plano $x\,y$).

Pois bem, vamos agora demonstrar que, *em tais condições, a função resultante* $z = \chi(t)$, *admite derivada em ordem a* t *em qualquer ponto de* $\mathcal{J}$, *tendo-se, precisamente*:

$$\frac{dz}{dt} = \frac{\partial z}{\partial x} \cdot \frac{dx}{dt} + \frac{\partial z}{\partial y} \cdot \frac{dy}{dt},$$

ou seja, com notações diversas

$$\chi'(t) = f'_x(\varphi(t),\ \psi(t))\ \varphi'(t) + f'_y(\varphi(t),\ \psi(t))\,\psi'(t).$$

Com efeito, seja t_0 um ponto qualquer de $\mathcal{J}$ e ponhamos $x_0 = \varphi(t_0)$, $y_0 = \psi(t_0)$. Então, se for Δt um acréscimo de t tal que $t_0 + \Delta t \in \mathcal{J}$, virão para x, y acréscimos h, k e para z um acréscimo Δz, tendo-se, por um lado,

$$\lim_{\Delta t \to 0} \frac{h}{\Delta t} = \varphi'(t_0)\,, \quad \lim_{\Delta t \to 0} \frac{k}{\Delta t} = \psi'(t_0)\,,$$

com $\varphi'(t_0) \neq \infty$, $\psi'(t_0) \neq \infty$ (em virtude da hipótese) e, por outro lado,

$$\Delta z = f'_x(x_0, y_0)\, h + f'_y(x_0, y_0)\, k + \omega$$

com $\lim\limits_{(h,k) \to (0,0)} \dfrac{\omega}{\sqrt{h^2 + k^2}} = 0$ (em virtude da hipótese e do TEOREMA 7).

Virá, portanto:

$$\text{(3)} \qquad \frac{\Delta z}{\Delta t} = f'_x(x_0, y_0)\, \frac{h}{\Delta t} + f'_y(x_0, y_0)\, \frac{k}{\Delta t} + \frac{\omega}{\Delta t}\,.$$

Mas, *ao tender Δt para zero,* também h, k tendem para zero e

$$\frac{\sqrt{h^2 + k^2}}{\Delta t} = \sqrt{\left(\frac{h}{\Delta t}\right)^2 + \left(\frac{k}{\Delta t}\right)^2} \to \sqrt{[\varphi'(t_0)]^2 + [\psi'(t_0)]^2} \neq \infty$$

donde:

$$\frac{\omega}{\Delta t} = \frac{\omega}{\sqrt{h^2 + k^2}} \times \frac{\sqrt{h^2 + k^2}}{\Delta t} \to 0 \,.\, \sqrt{[\varphi'(t_0)]^2 + [\psi'(t_0)]^2} = 0.$$

Portanto, conclui-se de (3) que, ao tender Δt para zero, $\dfrac{\Delta z}{\Delta t}$ tende para um limite finito, sendo este igual

$$\chi'(t_0) = f'_x(x_0, y_0)\, \varphi'(t_0) + f'_y(x_0, y_0)\, \psi'(t_0)\,.$$

Raciocinando de modo análogo, podemos, mais geralmente, estabelecer a seguinte *regra de derivação das funções compostas*:

TEOREMA 8. *Seja* $y = f(x_1, x_2, \ldots, x_n)$ *uma função real das variáveis reais* $x_1, x_2, \ldots, x_n$, *continuamente derivável num conjunto aberto* $\mathcal{C} \subset \mathbf{R}^n$. *Se, por sua vez, se tiver*

$$x_1 = \varphi_1(t),\ x_2 = \varphi_2(t),\ \ldots,\ x_n = \varphi_n(t),$$

sendo $\varphi_1(t), \varphi_2(t), \ldots, \varphi_n(t)$, *funções que admitem derivada finita num intervalo* $\mathcal{J} \subset \mathbf{R}$ *e tais que, ao variar* t *em* $\mathcal{J}$, *o ponto* $(x_1, x_2, \ldots, x_n)$ *permanece em* $\mathcal{C}$, *então* y *é uma função de* t *que admite em cada ponto de* $\mathcal{J}$ *uma derivada total dada pela fórmula:*

$$\frac{dy}{dt} = \frac{\partial y}{\partial x_1} \cdot \frac{dx_1}{dt} + \frac{\partial y}{\partial x_2} \cdot \frac{dx_2}{dt} + \cdots + \frac{\partial y}{\partial x_n} \cdot \frac{dx_n}{dt},$$

na qual em cada uma das expressões das derivadas parciais

$$\frac{\partial y}{\partial x_1} = f'_{x_1}(X),\ \frac{\partial y}{\partial x_2} = f'_{x_2}(X), \ldots, \frac{\partial y}{\partial x_n} = f'_{x_n}(X)$$

se supõe feita a substituição de $x_1, x_2, \ldots, x_n$, *pelas respectivas expressões em função de* t.

Como exemplo, procuremos a expressão da derivada em ordem a x de

$$y = u^v,$$

supondo u, v funções de x: $u = f(x)$, $v = g(x)$; ter-se-á:

$$\frac{dy}{dx} = \frac{\partial y}{\partial u}\frac{du}{dx} + \frac{\partial y}{\partial v}\frac{dv}{dx}$$

$$= v u^{v-1}\frac{du}{dx} + u^v \log u \frac{dv}{dx},$$

como de resto já sabemos do curso de Matemáticas Gerais, por outra via.

NOTA IMPORTANTE. Por vezes, na prática, a regra de derivação das funções compostas é aplicada em circunstâncias especiais, que embaraçam o principiante induzindo-o em equívocos. Suponhamos, por exemplo, que se trata duma função

$$z = f(x, y) \tag{4}$$

das duas variáveis x, y, e que se fez apenas a substituição $y = \varphi(x)$, que converte z em *função exclusiva de* x:

$$z = \chi(x) = f(x, \varphi(x)).$$

Para aplicar o TEOREMA 8, tal como foi enunciado, bastará considerar a mudança de variáveis:

$$x = t, \quad y = \varphi(t),$$

que, efectuada em (4), conduz ao resultado

$$z = f(t, \varphi(t)) = \chi(t).$$

Então, verificada a hipótese do teorema, podemos escrever:

$$\chi'(t) = \frac{dz}{dt} = \frac{\partial z}{\partial x} \cdot \frac{dx}{dt} + \frac{\partial z}{\partial y} \cdot \frac{dy}{dt}$$

$$= \frac{\partial z}{\partial x} + \frac{\partial z}{\partial y} \cdot \frac{dy}{dt};$$

ou seja, visto que $x = t$,

$$\frac{dz}{dx} = \frac{\partial z}{\partial x} + \frac{\partial z}{\partial y} \cdot \frac{dy}{dx}. \tag{5}$$

Na prática, dispensa-se geralmente a intervenção da variável auxiliar t, concebendo a mudança de variáveis sob a forma

$$x = x\ ,\ \ y = \varphi(x).$$

Simplesmente, a fórmula (5) *tem aspecto paradoxal*, porquanto no primeiro membro figura a *derivada total* $\dfrac{dz}{dx}$ de z em ordem a x, enquanto no segundo figura a *derivada parcial* $\dfrac{\partial z}{\partial x}$.

Mas trata-se apenas de uma deficiência de notação, que não terá importância de maior para quem esteja elucidado sobre os significados dos símbolos. Com $\dfrac{dz}{dx}$ pretende-se designar a derivada de z como *nova função* de x, isto é:

$$\frac{dz}{dx} = \chi'(x) = \frac{d}{dx} f(x, \varphi(x))\,;$$

ao passo que $\dfrac{\partial z}{\partial x}$ representa a derivada parcial em ordem a x da *primitiva função* z de x, y, isto é:

$$\frac{\partial z}{\partial x} = f'_x(x, y) = \frac{\partial f}{\partial x}(x, y)\,.$$

Quando se quiser evitar confusão, bastará substituir na fórmula (5) as notações

$$\frac{\partial z}{\partial x}\,,\ \frac{\partial z}{\partial y}\,,$$

respectivamente, pelas notações

$$\frac{\partial f}{\partial x}\,,\ \frac{\partial f}{\partial y}$$

que são, na verdade, preferíveis às primeiras do ponto de vista da coerência lógica.

15. Diferencial duma função composta. Invariância do diferencial

Consideremos agora, mais geralmente, uma função real

$$y = f(x_1, x_2, \ldots, x_n)$$

continuamente derivável num conjunto aberto $\mathcal{C} \subset \mathbf{R}^n$ e suponhamos as variáveis $x_1, x_2, \ldots, x_n$, expressas como funções de outras variáveis reais $t_1, t_2, \ldots, t_p$,

$$(1) \qquad \begin{cases} x_1 = \varphi_1(t_1, t_2, \ldots, t_p) \\ x_2 = \varphi_2(t_1, t_2, \ldots, t_p) \\ \cdots\cdots\cdots\cdots\cdots\cdots \\ x_n = \varphi_n(t_1, t_2, \ldots, t_p) \end{cases}$$

sendo estas funções *continuamente deriváveis num conjunto aberto* $\mathcal{D} \subset \mathbf{R}^p$ (de modo que o ponto X de coordenadas $x_1, x_2, \ldots, x_n$, esteja em $\mathcal{C}$ quando o ponto T de coordenadas $t_1, t_2, \ldots, t_p$, está em $\mathcal{D}$). Então, y será função composta de $t_1, t_2, \ldots, t_p$ por intermédio de $x_1, x_2, \ldots, x_n$; seja

$$y = \psi(t_1, t_2, \ldots, t_p)$$

esta nova função, definida em $\mathcal{D}$.

Posto isto, supondo fixadas todas as variáveis $t_1, t_2, \ldots, t_p$, excepto uma, t_i, as variáveis $x_1, x_2, \ldots, x_n$, resultarão funções só de t_i, com derivadas finitas (em certos intervalos) que coincidem com as derivadas parciais

$$\frac{\partial x_1}{\partial t_i}, \frac{\partial x_2}{\partial t_i}, \ldots, \frac{\partial x_n}{\partial t_i}.$$

Por sua vez, a variável y resultará, assim, função só de t_i, por intermédio de $x_1, x_2, \ldots, x_n$, e a sua derivada em ordem a t_i coincidirá com a derivada parcial $\dfrac{\partial y}{\partial t_i}$. Podemos, pois, aplicar ao cálculo desta derivada o TEOREMA 8, que nos dá

$$(2) \qquad \frac{\partial y}{\partial t_i} = \frac{\partial y}{\partial x_1} \cdot \frac{\partial x_1}{\partial t_i} + \frac{\partial y}{\partial x_2} \cdot \frac{\partial x_2}{\partial t_i} + \cdots + \frac{\partial y}{\partial x_n} \cdot \frac{\partial x_n}{\partial t_i}.$$

Temos, assim, a regra de derivação das funções compostas generalizada ao caso em que o número das novas variáveis independentes é superior a 1.

Por exemplo, seja:

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

uma função de x, y com derivadas parciais $\frac{\partial z}{\partial x}$ e $\frac{\partial z}{\partial y}$ contínuas em todo o plano e consideremos a mudança de variáveis

$$\begin{cases} x = \rho \cos \Theta \\ y = \rho \operatorname{sen} \Theta \end{cases}$$

Ter-se-á, então:

$$\begin{cases} \dfrac{\partial z}{\partial \rho} = \dfrac{\partial z}{\partial x} \cdot \dfrac{\partial x}{\partial \rho} + \dfrac{\partial z}{\partial y} \cdot \dfrac{\partial y}{\partial \rho} \\[2ex] \dfrac{\partial z}{\partial \Theta} = \dfrac{\partial z}{\partial x} \cdot \dfrac{\partial x}{\partial \Theta} + \dfrac{\partial z}{\partial y} \cdot \dfrac{\partial y}{\partial \Theta} \end{cases}$$

ou seja:

$$\frac{\partial z}{\partial \rho} = f'_x(x, y) \cos \Theta + f'_y(x, y) \operatorname{sen} \Theta$$

$$\frac{\partial z}{\partial \Theta} = - f'_x(x, y)\, \rho \operatorname{sen} \Theta + f'_y(x, y)\, \rho \cos \Theta$$

Restará, portanto, substituir, em $f'_x(x, y)$, $f'_y(x, y)$, as variáveis x, y (entre parêntesis) respectivamente por $\rho \cos \Theta$ e $\rho \operatorname{sen} \Theta$.

Tornando ao caso geral, observemos, agora, que, atendendo à hipótese e aos teoremas sobre funções contínuas, se deduz de (2) que as derivadas $\frac{\partial y}{\partial t_1}, \frac{\partial y}{\partial t_2}, \ldots, \frac{\partial y}{\partial t_p}$, são contínuas em $\mathcal{D}$; portanto, (DEFINIÇÕES 16 e 17), y será, como função de T, continuamente derivável em $\mathcal{D}$, tendo-se

$$d\psi(T) = \frac{\partial y}{\partial t_1} dt_1 + \frac{\partial y}{\partial t_2} dt_2 + \cdots + \frac{\partial y}{\partial t_p} dt_p$$

ou, abreviadamente,

$$d\psi = \sum_{i=1}^{p} \frac{\partial y}{\partial t_i} dt_i$$

donde atendendo a (2):

$$(3) \qquad d\psi = \sum_{i=1}^{p} \left(\frac{\partial y}{\partial x_1} \cdot \frac{\partial x_1}{\partial t_i} + \frac{\partial y}{\partial x_2} \cdot \frac{\partial x_2}{\partial t_i} + \cdots + \frac{\partial y}{\partial x_n} \cdot \frac{\partial x_n}{\partial t_i} \right) dt_i =$$

$$= \frac{\partial y}{\partial x_1} \sum_{i=1}^{p} \frac{\partial x_1}{\partial t_i} dt_i + \cdots + \frac{\partial y}{\partial x_n} \sum_{i=1}^{p} \frac{\partial x_n}{\partial t_i} dt_i .$$

Mas, pela definição de diferencial, tem-se:

$$\sum_{i=1}^{p} \frac{\partial x_1}{\partial t_i} dt_i = \frac{\partial x_1}{\partial t_1} dt_1 + \cdots + \frac{\partial x_1}{\partial t_p} dt_p = dx_1$$

e, dum modo geral,

$$(4) \qquad \sum_{i=1}^{p} \frac{\partial x_k}{\partial t_i} dt_i = dx_k = d\varphi_k(T) \quad (k = 1, 2, \ldots, n) .$$

Entrando com estes valores em (3) resulta

$$(5) \qquad d\psi = \frac{\partial y}{\partial x_1} dx_1 + \frac{\partial y}{\partial x_2} dx_2 + \cdots + \frac{\partial y}{\partial x_n} dx_n \,;$$

isto é, poderemos, escrever

$$d\psi(t_1, t_2, \ldots, t_p) = df(x_1, x_2, \ldots, x_n),$$

desde que, no segundo membro, se tome

$$x_k = \varphi_k(t_1, t_2, \ldots, t_p), \; k = 1, 2, \ldots, n,$$

interpretando $dx_1, dx_2, \ldots, dx_n$, como os diferenciais de $x_1, x_2, \ldots, x_n$, em ordem a $t_1, t_2, \ldots, t_p$, dados por (4).

Em conclusão:

O diferencial de y em ordem a $t_1, t_2, \ldots, t_p$, *pode obter-se formando primeiro o diferencial de y em ordem às variáveis* $x_1, x_2, \ldots, x_n$ (*supostas independentes*) *e exprimindo em seguida* $dx_1, dx_2, \ldots, dx_n$, *como diferenciais de* $x_1, x_2, \ldots, x_n$, *em ordem a* $t_1, t_2, \ldots, t_p$.

Este resultado é conhecido por *princípio da invariância do diferencial.*

O seu significado será melhor apreendido quando tratarmos de diferenciais de ordem superior à primeira, a respeito dos quais este princípio já não se verifica.

16. Cálculo prático dos diferenciais

Aplicando directamente a DEFINIÇÃO 17, o cálculo do diferencial duma função de mais de uma variável reduz-se sempre ao cálculo das suas derivadas parciais. Mas, na prática, pode ainda com vantagem tirar-se partido do princípio de invariância há pouco demonstrado. Com efeito, resulta desse princípio que as regras de diferenciação para funções de mais de uma variável são perfeitamente análogas às regras de derivação para as funções de uma só variável. Assim, designando por *u, v* duas quaisquer funções continuamente deriváveis, das variáveis $x_1, x_2, \ldots, x_n$, ter-se-á:

1) – $d(u \pm v) = du \pm dv$;

2) – $d(a\, u) = a\, du$, sendo a constante;

3) – $d(u\, v) = u\, dv + v\, du$;

4) – $d\left(\dfrac{u}{v}\right) = \dfrac{v\, du - u\, dv}{v^2}$;

5) – $d \log u = \dfrac{du}{u}$;

6) – $d\, u^v = v\, u^{v-1}\, du + u^v \log u\, dv$;

7) – d sen u = cos $u\, du$;

8) – d cos u = – sen $u\, du$;

9) – d tg u = sec^2 $u\, du$;

10) – d arcsen $u = \dfrac{du}{\sqrt{1-u^2}}$;

11) – d arctg $u = \dfrac{du}{1+u^2}$.

Deve ainda observar-se que o *diferencial duma constante é zero*.

Para justificar, por exemplo, a regra 3), ponhamos $y = u\,v$. Como função das duas variáveis u, v, o diferencial de y é de facto $dy = v\,du + u\,dv$, pois que $\dfrac{\partial y}{\partial u} = v$, $\dfrac{\partial y}{\partial v} = u$. Por outro lado, o princípio da invariância do diferencial diz-nos que o diferencial de y em ordem a $x_1, x_2, \dots, x_n$, é dado ainda por $v\,du + u\,dv$, desde que se interpretem du, dv, respectivamente, como os diferenciais de u e v em ordem a $x_1, x_2, \dots, x_n$.

Exemplo – Tem-se:

$$d \log \sqrt{1+x^2 y} = \frac{1}{2}\, d \log (1+x^2 y) =$$

$$= \frac{d(1+x^2 y)}{2(1+x^2 y)} = \frac{2xy\,dx + x^2\,dy}{2(1+x^2 y)} =$$

$$= \frac{xy}{(1+x^2 y)}\,dx + \frac{x^2}{2(1+x^2 y)}\,dy.$$

É claro que se terá, por definição de diferencial,

$$\frac{\partial}{\partial x} \log \sqrt{1+x^2 y} = \frac{xy}{(1+x^2 y)}$$

$$\frac{\partial}{\partial y} \log \sqrt{1+x^2 y} = \frac{x^2}{2(1+x^2 y)}.$$

Poderíamos ter construído directamente o diferencial a partir das derivadas parciais. Mas, geralmente, é preferível o emprego das regras de diferenciação, quando se tenha adquirido prática no manejo dessas regras.

17. Conceito de derivada total para as funções vectoriais duma variável vectorial

Consideremos uma função vectorial, $Y = F(X)$, definida num conjunto aberto $\mathcal{C} \subset \mathbf{R}^m$, por meio das m funções reais

$$(1) \qquad y_i = f_i(x_1, x_2, \dots, x_n), \quad i = 1, 2, \dots, m.$$

Suponhamos que estas m funções são continuamente deriváveis em $\mathcal{C}$. Então, aplicando o TEOREMA 7, podemos afirmar que, quando se passa do ponto X para o ponto $X + H = (x_1 + h_1, x_2 + h_2, \dots, x_n + h_n)$, a função Y sofre um acréscimo $\Delta Y = (\Delta y_1, \Delta y_2, \dots, \Delta y_m)$ tal que:

$$\Delta y_1 = \frac{\partial f_1}{\partial x_1} h_1 + \frac{\partial f_1}{\partial x_2} h_2 + \cdots + \frac{\partial f_1}{\partial x_n} h_n + \omega_1$$

$$\Delta y_2 = \frac{\partial f_2}{\partial x_1} h_1 + \frac{\partial f_2}{\partial x_2} h_2 + \cdots + \frac{\partial f_2}{\partial x_n} h_n + \omega_2$$

$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$

$$\Delta y_m = \frac{\partial f_m}{\partial x_1} h_1 + \frac{\partial f_m}{\partial x_2} h_2 + \cdots + \frac{\partial f_m}{\partial x_n} h_n + \omega_m,$$

sendo $\omega_1, \omega_2, \dots, \omega_m$ infinitésimos com H de ordem superior à primeira, isto é, tendo-se

$$\lim_{H \to \vec{0}} \frac{\omega_i}{|H|} = 0, \text{ para } i = 1, 2, \dots, m.$$

Dizem-nos as anteriores expressões que, desprezando $\omega_1, \omega_2, \dots, \omega_m$ (o que não provocaria erro sensível quando $h_1, h_2, \dots, h_n$, fossem bastante pequenos), o acréscimo ΔY resultaria do acréscimo H, aplicando a este uma *transformação linear de matriz:*

$$(2) \qquad \begin{bmatrix} \dfrac{\partial f_1}{\partial x_1} & \dfrac{\partial f_1}{\partial x_2} & \cdots\cdots & \dfrac{\partial f_1}{\partial x_n} \\ \dfrac{\partial f_2}{\partial x_1} & \dfrac{\partial f_2}{\partial x_2} & \cdots\cdots & \dfrac{\partial f_2}{\partial x_n} \\ \cdots & \cdots & \cdots & \cdots \\ \dfrac{\partial f_m}{\partial x_1} & \dfrac{\partial f_m}{\partial x_2} & \cdots\cdots & \dfrac{\partial f_m}{\partial x_n} \end{bmatrix}$$

ou, abreviadamente:

$$\left[\frac{\partial f_i}{\partial x_k}\right],$$

subentendendo-se que, neste símbolo, i é o índice de linha e k o índice de coluna do elemento genérico $\dfrac{\partial f_i}{\partial x_k}$ da matriz. Pois bem:

DEFINIÇÃO 18. *A matriz* (2) *é chamada matriz jacobiana da transformação* $Y=F(X)$ (*ou do sistema de funções* $f_1, f_2 \ldots, f_m$ *em ordem a* $x_1, x_2, \ldots, x_n$). *Também se lhe dá* (*modernamente*) *o nome de derivada de* $F(X)$ *em ordem a* X, *sendo representada por qualquer das notações...*

$$\frac{dY}{dX}, F'(X), \frac{dF}{dX}.$$

(Não esquecer que esta definição é dada na hipótese de as componentes de $F(X)$ serem continuamente deriváveis em $\mathcal{C}$).

É claro que os elementos da matriz (2) *são funções de* X *definidas em* $\mathcal{C}$; mas se fizermos em todas elas a substituição de X por uma determinada constante $A=(a_1, a_2, \ldots, a_n) \in \mathcal{C}$, obtém-se uma matriz que tem por elementos números. Para designar esta matriz numérica, podemos ainda usar o símbolo (2), escrevendo A em índice. Abreviadamente:

$$\left[\frac{\partial f_i}{\partial x_k}\right]_A = \left[\frac{\partial f_i}{\partial x_k}(A)\right].$$

A esta matriz chamaremos, ainda, *matriz jacobiana* ou *derivada* de $F(X)$ no ponto A, e poderemos designá-la por qualquer dos símbolos

$$\left(\frac{dY}{dX}\right)_A, \; F'(A), \; \frac{dF}{dX}(A).$$

No caso particular em que $m = n$, tem-se, ainda, a

DEFINIÇÃO 19. *Chama-se jacobiano ou determinante funcional das funções* $f_1, f_2, \dots, f_n$ *em ordem às variáveis* $x_1, x_2, \dots, x_n$ *(ou de* $F(X)$ *em ordem a* X*), e representa-se por*

$$J\begin{pmatrix} f_1 \, f_2 \cdots f_n \\ x_1 \, x_2 \cdots x_n \end{pmatrix},$$

o determinante da matriz jacobiana $\dfrac{dF}{dX}$*, isto é, o determinante*

$$\left|\frac{\partial f_i}{\partial x_k}\right| = \begin{vmatrix} \dfrac{\partial f_1}{\partial x_1} & \dfrac{\partial f_1}{\partial x_2} & \cdots\cdots & \dfrac{\partial f_1}{\partial x_n} \\ \dfrac{\partial f_2}{\partial x_1} & \dfrac{\partial f_2}{\partial x_2} & \cdots\cdots & \dfrac{\partial f_2}{\partial x_n} \\ \cdots & \cdots & \cdots & \cdots \\ \dfrac{\partial f_n}{\partial x_1} & \dfrac{\partial f_n}{\partial x_2} & \cdots\cdots & \dfrac{\partial f_n}{\partial x_n} \end{vmatrix}.$$

Deste modo, o jacobiano duma transformação terá por valores números (e não matrizes). Para designar o valor do jacobiano num determinado ponto A, bastará representar esse ponto em índice, tal como se faz para a matriz jacobiana.

Exemplo – Seja a transformação

$$\begin{cases} x = \rho \cos \Theta \\ y = \rho \operatorname{sen} \Theta \end{cases}$$

que faz corresponder a cada par (ρ, Θ) de números reais um outro par (x, y) de números reais (transformação de $\mathbf{R}^2$ sobre si mesmo). A matriz jacobiana desta transformação será:

$$\frac{d(x, y)}{d(\rho, \Theta)} = \begin{bmatrix} \dfrac{\partial x}{\partial \rho} & \dfrac{\partial x}{\partial \Theta} \\ \dfrac{\partial y}{\partial \rho} & \dfrac{\partial y}{\partial \Theta} \end{bmatrix} = \begin{bmatrix} \cos\Theta & -\rho \operatorname{sen}\Theta \\ \operatorname{sen}\Theta & \rho\cos\Theta \end{bmatrix}.$$

O respectivo jacobiano é o determinante desta matriz, ou seja,

$$J\begin{pmatrix} x & y \\ \rho & \Theta \end{pmatrix} = \begin{vmatrix} \cos\Theta & -\rho \operatorname{sen}\Theta \\ \operatorname{sen}\Theta & \rho\cos\Theta \end{vmatrix} = \rho\cos^2\Theta + \rho \operatorname{sen}^2\Theta = \rho\,.$$

Em particular, o valor deste jacobiano no ponto $\rho = 1$, $\Theta = \pi$ será

$$J\begin{pmatrix} x & y \\ \rho & \Theta \end{pmatrix}_{(1,\,\pi)} = 1.$$

A designação de "derivada" atribuída à matriz jacobiana de $F(X)$ e as correlativas notações $\dfrac{dY}{dX}$, $F'(X)$, etc., só há poucos anos foram introduzidas em Matemática. A sua justificação encontra-se em numerosos factos, entre os quais o que vamos apontar.

Consideremos, de novo, a função vectorial $Y = F(X)$ definida no conjunto aberto $\mathcal{C} \subset \mathbf{R}^n$ pelas m funções reais

$$y_i = f_i(x_1, x_2, \ldots, x_n), \quad i = 1, 2, \ldots, m\,,$$

e seja $X = G(T)$ uma função da nova variável T, definida num conjunto aberto $\mathcal{D} \subset \mathbf{R}^p$ pelas n funções reais

$$x_j = g_j(t_1, t_2, \ldots, t_p), \quad j = 1, 2, \ldots, n\,,$$

de modo que $\mathcal{C}$ contenha o contradomínio de $G(T)$. Então, Y é função composta de T por intermédio de X:

$$Y = F(G(T)) = \Phi(T)\,.$$

Introduzamos as seguintes hipóteses suplementares: as funções $f_i(X)$ são continuamente deriváveis no conjunto $\mathcal{C}$ e as funções $g_k(T)$ são continuamente deriváveis no conjunto $\mathcal{D}$. Então, atendendo ao que foi estabelecido no número 14, podemos concluir que também as componentes da função $Y=\Phi(T)$ são continuamente deriváveis em D, tendo-se:

$$\frac{\partial y_i}{\partial t_k}=\frac{\partial y_i}{\partial x_1}\cdot\frac{\partial x_1}{\partial t_k}+\frac{\partial y_i}{\partial x_2}\cdot\frac{\partial x_2}{\partial t_k}+\cdots+\frac{\partial y_i}{\partial x_n}\cdot\frac{\partial x_n}{\partial t_k}$$

$$i=1,2,\ldots,m;\quad k=1,2,\ldots,p.$$

Mas notemos que o segundo membro desta igualdade é o produto interno do vector

$$\left(\frac{\partial y_i}{\partial x_1},\frac{\partial y_i}{\partial x_2},\ldots,\frac{\partial y_i}{\partial x_n}\right),$$

linha i da matriz $\left[\dfrac{\partial y_i}{\partial x_j}\right]$, pelo vector

$$\left(\frac{\partial x_1}{\partial t_k},\frac{\partial x_2}{\partial t_k},\ldots,\frac{\partial x_n}{\partial t_k}\right),$$

coluna k de matriz $\left[\dfrac{\partial x_j}{\partial t_k}\right]$; quer dizer: $\left[\dfrac{\partial y_i}{\partial t_k}\right]$ é igual ao elemento da linha i e da coluna k da matriz produto $\left[\dfrac{\partial y_i}{\partial x_j}\right]\cdot\left[\dfrac{\partial x_j}{\partial t_k}\right]$.

Podemos, portanto, escrever

$$\left[\frac{\partial y_i}{\partial t_k}\right]=\left[\frac{\partial y_i}{\partial x_j}\right]\cdot\left[\frac{\partial x_j}{\partial t_k}\right],$$

ou seja:

$$\boxed{\frac{dY}{dT}=\frac{dY}{dX}\cdot\frac{dX}{dT}.}$$

Temos pois, que graças ao novo conceito de derivada, sob a forma de matriz, a regra de derivação das funções compostas, tal como se tinha apresentado para as funções reais de variável real, se estende exactamente ao caso das funções vectoriais de variável vectorial.

18. Casos particulares. Noção de gradiente

As considerações do número precedente referem-se a uma função vectorial

$$Y = F(X)$$

definida num conjunto $\mathcal{C} \subset \mathbf{R}^n$ e com os valores em $\mathbf{R}^m$ (isto é, $Y \in \mathbf{R}^m$), sendo *m, n* números naturais quaisquer. Mas convém dedicar especial atenção aos dois casos particulares seguintes:

a) – 1.º caso: $n=1$, *m qualquer*. Trata-se, então, duma função vectorial (ou pontual) duma variável real x:

$$Y = F(x)$$

representável por um sistema de m funções reais de x:

$$\begin{cases} y_1 = f_1(x) \\ y_2 = f_2(x) \\ \cdots\cdots\cdots\cdots \\ y_m = f_m(x) \end{cases}.$$

Se tais funções forem continuamente deriváveis num dado intervalo aberto, a derivada de Y em ordem a x em cada ponto a desse intervalo será dada pela matriz coluna, cujos elementos são as derivadas das componentes de Y em ordem a x no ponto a, isto é:

$$F'(a) = \begin{bmatrix} \dfrac{dy_1}{dx} \\ \dfrac{dy_2}{dx} \\ \cdots\cdots \\ \dfrac{dy_m}{dx} \end{bmatrix}_a = \begin{bmatrix} f_1'(a) \\ f_2'(a) \\ \cdots\cdots \\ f_m'(a) \end{bmatrix}.$$

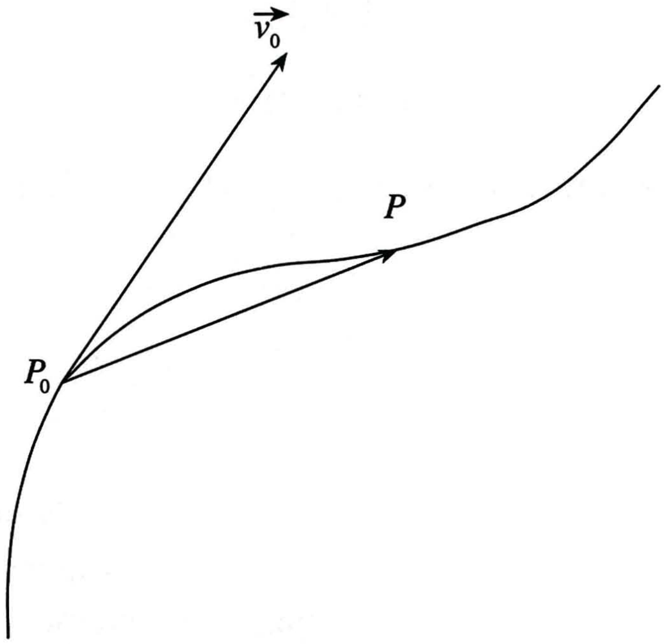
Neste caso, a derivada pode identificar-se com o vector cujas componentes são precisamente $f_1'(a)$, $f_2'(a)$, ..., $f_m'(a)$, sendo, então, susceptível da seguinte definição directa:

$$F'(a) = \lim_{x \to a} \frac{F(x) - F(a)}{x - a}.$$

Por força do TEOREMA 2, este limite, se existe, é o vector que tem por componentes

$$\lim_{x \to a} \frac{f_1(x) - f_1(a)}{x - a}, \ \lim_{x \to a} \frac{f_2(x) - f_2(a)}{x - a}, \ \ldots, \ \lim_{x \to a} \frac{f_m(x) - f_m(a)}{x - a},$$

ou seja, o vector $[f_1'(a), f_2'(a), \ldots, f_m'(a)]$, e assim fica estabelecida a equivalência das duas definições, sob a hipótese considerada. (É preciso notar, porém, que a segunda definição é válida sob hipóteses menos restritivas).



Tal conceito aparece sob variadíssimas vestes em questões de Mecânica, Electricidade, etc. Por exemplo, o movimento dum ponto no espaço é dado por uma função pontual

$$P = \Phi(t)$$

da variável real t (variável tempo), definida num intervalo $[a, b]$: a cada instante t situado neste intervalo corresponde uma determinada posição de P, de modo que, quando t percorre $[a, b]$, o ponto móvel P descreve uma linha no espaço (trajectória do movimento). Fixado um referencial cartesiano, as coordenadas x, y, z de P serão funções de t definidas em $[a, b]$:

$$x = \varphi_1(t), \quad y = \varphi_2(t), \quad z = \varphi_3(t).$$

A função pontual $\Phi(t)$ fica, pois, definida por este sistema de funções reais. Para que se trate, efectivamente, dum movimento, é necessário que a função $\Phi(t)$ seja contínua em $[a, b]$, o que, segundo o TEOREMA 3, equivale à continuidade simultânea das funções $\varphi_1(t)$, $\varphi_2(t)$, $\varphi_3(t)$, em $[a, b]$. Posto isto, seja t_0 um valor de t situado em $[a, b]$, e ponhamos $P_0 = \Phi(t_0)$. A razão incremental

$$\frac{P - P_0}{t - t_0} = \frac{\Phi(t) - \Phi(t_0)}{t - t_0}$$

será o cociente do vector $P - P_0$ pelo escalar (número real) $t - t_0$ e designa-se por *velocidade vectorial média do movimento no intervalo* $[t_0, t]$. Suponhamos, agora, que, ao tender t para t_0, aquela razão tende para um limite determinado; então, esse limite, que é a derivada de P em ordem a t em t_0:

$$\left(\frac{dP}{dt}\right)_{t_0} = \lim_{t \to t_0} \frac{P - P_0}{t - t_0},$$

será, por definição, a *velocidade vectorial do movimento no instante* t_0. Trata-se, pois, do vector que tem por componentes $\varphi_1'(t_0)$, $\varphi_2'(t_0)$, $\varphi_3'(t_0)$.

Se a função pontual $\Phi(t)$ admite derivada em cada ponto t de $[a, b]$, ficará definida neste intervalo a função vectorial

$$\vec{v} = \frac{dP}{dt} = \Phi'(t)$$

que dá a velocidade de P em cada instante. Pode acontecer que, por sua vez, esta função seja derivável em $[a, b]$; então, a função derivada

$$\frac{d^2P}{dt^2} = \Phi''(t)$$

dá-nos a chamada *aceleração vectorial de P no instante variável t*; as suas componentes serão as funções $\varphi_1''(t)$, $\varphi_2''(t)$, $\varphi_3''(t)$.

b) – 2.º caso: $m = 1$, n qualquer. Trata-se, então, duma só função real de n variáveis reais

$$y = f(x_1, x_2, \ldots, x_n).$$

Se a função $f(X)$ é continuamente derivável num conjunto aberto $\mathcal{C} \subset \mathbf{R}^n$, a sua derivada em ordem a X em cada ponto $A = (a_1, a_2, \ldots, a_n) \in \mathcal{C}$ é a matriz linha que tem por elementos as derivadas parciais da função em ordem às variáveis $x_1, x_2, \ldots, x_n$ no ponto A:

$$\left(\frac{dy}{dX}\right)_A = \left[\frac{\partial f}{\partial x_1}\ \frac{\partial f}{\partial x_2} \cdots \frac{\partial f}{\partial x_n}\right]_A.$$

Uma tal matriz pode identificar-se a um vector de $\mathbf{R}^n$, ao qual se dá o nome de *gradiente da função $f(X)$ no ponto A* e se representa pela notação

$$(\text{grad } f)_A.$$

Se em vez do ponto fixo A, se considerar o ponto genérico X, o gradiente de $f(X)$ em X será designado simplesmente pela notação: grad f.

Ter-se-á, pois, por definição,

$$\text{grad } f = \left(\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, \ldots, \frac{\partial f}{\partial x_n}\right).$$

O conceito de gradiente também se apresenta com variadas formas em muitas questões da Física, como veremos a propósito do conceito de potencial.

19. Derivadas direccionais

Continuemos a considerar uma função real

$$y = f(x_1, x_2, \dots, x_n)$$

definida num conjunto $\mathcal{C} \subset \mathbf{R}^n$ e seja $A = (a_1, a_2, \dots, a_n)$ um ponto interior de $\mathcal{C}$. Consideremos, por outro lado, um vector unitário U qualquer de $\mathbf{R}^n$, isto é, um vector

$$U = (u_1, u_2, \dots, u_n),$$

tal que:

$$|U| = \sqrt{u_1^2 + u_2^2 + \cdots + u_n^2} = 1 .$$

Por ser A interior a $\mathcal{C}$, existirá um número $\delta > 0$ tal que os pontos cuja distância a A é inferior a δ pertençam todos a $\mathcal{C}$; então, se for t um número real de módulo $< \delta$, será $|tU| = |t| \cdot 1 < \delta$ e, portanto, o ponto $A + tU$ estará em $\mathcal{C}$. Daqui resulta que, substituindo X por $A + tU$ em $f(X)$ se obtém uma função de t,

$$\varphi(t) = f(A + tU) = f(a_1 + tu_1, a_2 + tu_2, \dots, a_n + tu_n),$$

definida para $|t| < \delta$, ou seja, no intervalo $]-\delta, \delta[$.

Pode acontecer que a nova função $\varphi(t)$ admita derivada (em ordem a t) no ponto 0; ter-se-á, então,

$$\varphi'(0) = \lim_{t \to 0} \frac{f(A + tU) - f(A)}{t} .$$

Pois bem, a este limite, quando existe, dá-se o nome de *derivada da função $f(X)$ segundo a direcção do vector U no ponto A.*

Em particular, pode U coincidir com um dos vectores fundamentais de $\mathbf{R}^n$: $E_1, E_2, \dots, E_n$. (Em Matemáticas Gerais convencionámos representar por E_i o vector cuja coordenada de ordem i é igual a 1, sendo as restantes coordenadas todas nulas).

Ter-se-á:

$$f(A+tE_i)=f(a_1,\ \dots,\ a_{i-1},\ a_i+t,\ a_{i+1},\ \dots,\ a_n);$$

portanto, o

$$\lim_{t\to 0}\frac{f(A+tE_i)-f(A)}{t}$$

se existe, é, por definição, a derivada parcial $f'_{x_i}(A)$. Em resumo:

A derivada de $f(X)$ segundo o vector E_i no ponto A, se existe, é a derivada parcial de $f(X)$ em ordem a x_i em A.

Posto isto, demonstraremos o seguinte:

TEOREMA 9. *Se $f(X)$ é continuamente derivável no interior de $\mathcal{C}$ e admite derivadas em todas as direcções em cada ponto A interior a $\mathcal{C}$, a derivada de $f(X)$ segundo um dado vector unitário U em A será, então, igual a*

$$(\text{grad}\, f)_A \,|\, U,$$

isto é, será igual ao produto interno do gradiente de f em A pelo vector U.

Suponhamos $f(X)$ continuamente derivável no interior de $\mathcal{C}$ e seja A um ponto interior a $\mathcal{C}$.

Então, pondo $X=A+tU$ e $\varphi(t)=f(A+tU)$, teremos $x_1=a_1+tu_1$, $x_2=a_2+tu_2$, …, $x_n=a_n+tu_n$, e, aplicando o TEOREMA 8:

$$\varphi'(0)=\left(\frac{\partial f}{\partial x_1}\right)_A u_1+\left(\frac{\partial f}{\partial x_2}\right)_A u_2+\cdots+\left(\frac{\partial f}{\partial x_n}\right)_A u_n\,.$$

Ora, o segundo membro é precisamente o produto interno de $(\text{grad}\, f)_A$ pelo vector U, o que prova o teorema.

NOTA. Alguns autores definem a derivada de $f(X)$ segundo a direcção de U em A como sendo a derivada lateral

$$\varphi'(0^+)=\lim_{t\to 0^+}\frac{f(A+tU)-f(A)}{t}\,.$$

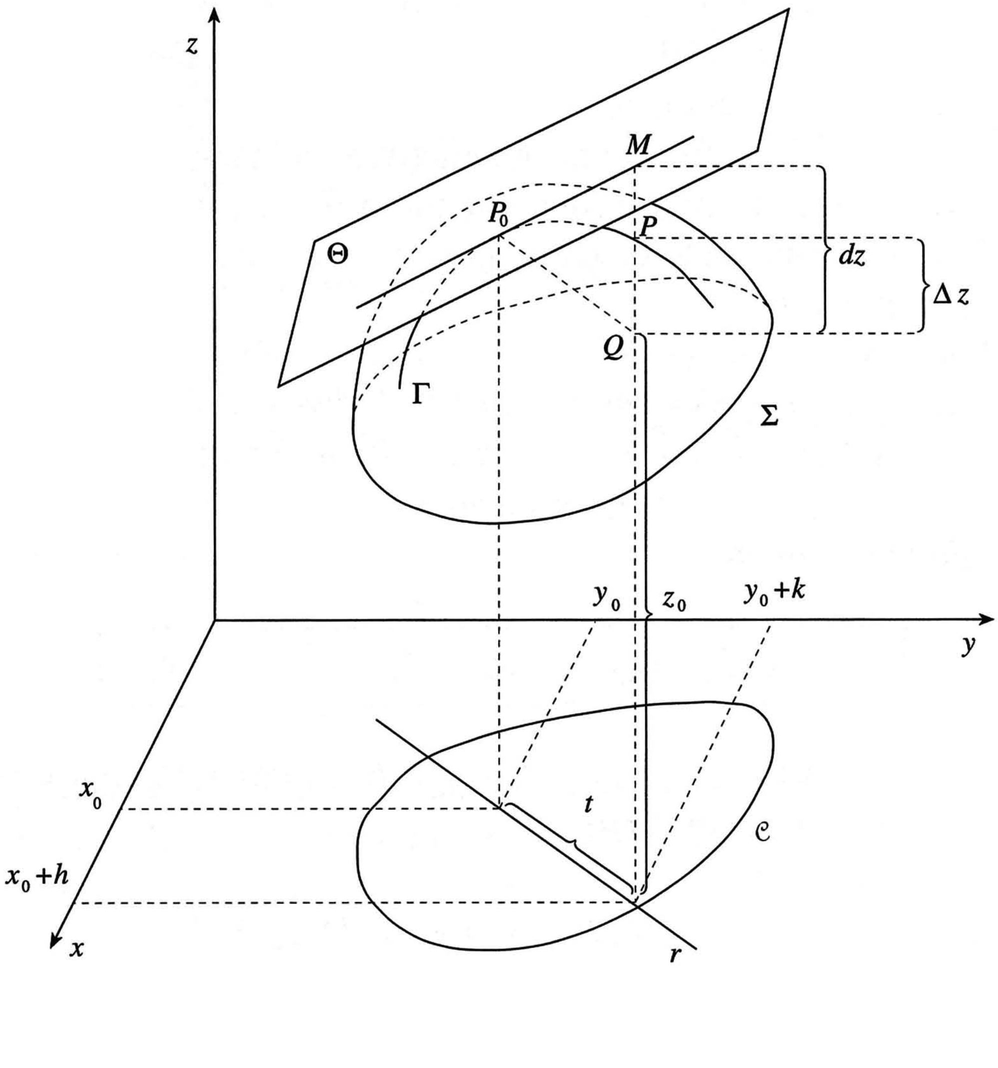
Mas parece-nos mais coerente chamar a este limite, quando existe, a "derivada de $f(X)$ segundo a direcção e o sentido de U em A".

20. Plano tangente a uma superfície. Interpretação geométrica do conceito de diferencial

Consideremos uma função real $z = f(x, y)$ definida e contínua num conjunto $\mathcal{C} \subset \mathbf{R}^2$ que supomos limitado por uma linha fechada. O gráfico de $f(x, y)$ será, então, uma superfície, Σ. Fixado um ponto (x_0, y_0) no interior de $\mathcal{C}$, ponhamos $x = x_0 + \alpha t$, $y = y_0 + \beta t$, com $\alpha^2 + \beta^2 = 1$. É claro que, quando t varia de $-\infty$ a $+\infty$, o ponto (x, y) descreve uma recta r que passa por (x_0, y_0). Por sua vez, a variável z será função composta de t por intermédio de x e y:

$$z = f(x_0 + \alpha t, y_0 + \beta t) = \varphi(t)$$

e, quando t varia numa vizinhança de 0, o ponto (x, y, z) descreve sobre a superfície Σ uma linha Γ que passa pelo ponto P_0 de coordenadas x_0, y_0, $z_0 = f(x_0, y_0)$.



Suponhamos, agora, que $f(x, y)$ é continuamente derivável no interior de $\mathcal{C}$. Então, segundo o TEOREMA 8, a função $\varphi(t)$ admitirá derivada em ordem a t no ponto 0 (derivada direccional de $f(x, y)$):

$$(1) \qquad \varphi'(0) = f'_x(x_0, y_0)\,\alpha + f'_y(x_0, y_0)\,\beta,$$

sendo fácil ver que $\varphi'(0)$ nos dá a tangente trigonométrica do ângulo que a recta P_0M, tangente a Γ em P_0, forma com a recta r orientada no sentido do vector unitário $\vec{u} = (\alpha, \beta)$. A recta P_0M terá, então, por equações paramétricas:

$$z - z_0 = \varphi'(0)\, t, \quad x - x_0 = \alpha t, \quad y - y_0 = \beta t.$$

Substituindo, na primeira destas equações, $\varphi'(0)$ pela sua expressão dada em (1), e atendendo às duas últimas equações, virá

$$z - z_0 = f'_x(x_0, y_0)(x - x_0) + f'_y(x_0, y_0)(y - y_0),$$

que é, como sabemos, a equação dum plano Θ, passante por P_0. Quer isto dizer que, qualquer que seja a direcção da recta r no plano x, y, a tangente à correspondente curva Γ no ponto P_0 está contida no plano Θ. Será, pois, Θ o lugar geométrico das tangentes às curvas Γ assim obtidas, quando r roda em torno de (x_0, y_0) no plano $x\,y$. Exprimiremos este facto dizendo que o plano Θ é *tangente à superfície* Σ *no ponto* P_0.

Portanto:

Se a função $f(x, y)$ *é continuamente derivável no interior de* $\mathcal{C}$, *a respectiva superfície representativa admite plano tangente em cada ponto* (x_0, y_0, z_0) *tal que* (x_0, y_0) *seja interior a* $\mathcal{C}$ *e* $z_0 = f(x_0, y_0)$. *A equação desse plano é*:

$$\boxed{z - z_0 = f'_x(x_0, y_0)(x_0 - x_0) + f'_y(x_0, y_0)(y - y_0).}$$

Ponhamos, agora, $x - x_0 = h$, $y - y_0 = k$; a estes acréscimos de x e y, corresponderá o acréscimo

$$\Delta z = f(x_0 + h, y_0 + k) - f(x_0, y_0)$$

da função considerada. O ponto P de coordenadas x_0+h, y_0+k, $z_0+\Delta z$, pertencerá, ainda, à superfície Σ; o ponto M do plano tangente Θ com a mesma abcissa e a mesma ordenada terá por cota z_0+dz, sendo:

$$dz = f'_x(x_0, y_0)h + f'_y(x_0, y_0)k,$$

diferencial de $f(x, y)$ no ponto (x_0, y_0) (correspondente aos acréscimos h, k). Ter-se-á, pois:

$$\Delta z = dz + \omega, \quad \lim_{\substack{h \to 0 \\ k \to 0}} \frac{\omega}{\sqrt{h^2 + k^2}} = 0.$$

É fácil agora ver que, na figura, os infinitésimos Δz, dz e ω são, em valor absoluto, representados, respectivamente, pelos cumprimentos $|PQ|$, $|MQ|$ e $|MP|$. Em conclusão:

Se $f(x, y)$ é continuamente derivável no interior de $\mathcal{C}$, o gráfico Σ de $f(x, y)$ para valores de (x, y) próximos de (x_0, y_0) pode, em primeira aproximação, ser substituído pelo plano tangente a Σ em (x_0, y_0, z_0); o erro cometido é um infinitésimo com (h, k) de ordem superior à primeira.

Note-se, porém, que esta interpretação geométrica é dada apenas para funções de duas variáveis.

21. Derivadas parciais de segunda ordem

Consideremos uma função real $z=f(x, y)$ que admite derivadas parciais, $\frac{\partial z}{\partial x}$, $\frac{\partial z}{\partial y}$, em todos os pontos dum dado conjunto aberto $\mathcal{A} \subset \mathbf{R}^2$.

Então, as duas derivadas

$$\frac{\partial z}{\partial x} = f'_x(x, y), \quad \frac{\partial z}{\partial y} = f'_y(x, y)$$

serão ainda funções de x, y, definidas no conjunto $\mathcal{A}$, e pode acontecer que tais funções admitam, por sua vez, derivadas parciais em ordem a x e a y no mesmo conjunto $\mathcal{A}$. Em tal hipótese, as novas derivadas dir-se-ão *derivadas parciais de segunda ordem* (ou simplesmente *segundas derivadas parciais*) de $f(x, y)$. A de $\frac{\partial z}{\partial x}$ em ordem a x designar-se-á por qualquer das notações

$$\frac{\partial^2 z}{\partial x^2}, \quad \frac{\partial^2 f}{\partial x^2}, \quad f''_{x^2}(x, y), \quad D_x D_x f(x, y), \text{ etc.};$$

e a de $\frac{\partial z}{\partial x}$ em ordem a y por uma das notações

$$\frac{\partial^2 z}{\partial x\, \partial y}, \quad \frac{\partial^2 f}{\partial x \partial y}, \quad f''_{xy}(x, y), \quad D_y D_x f(x, y), \text{ etc.}$$

Por sua vez, para as derivadas parciais de $\frac{\partial z}{\partial y}$ em ordem a x e a y teremos, respectivamente, as notações

$$\frac{\partial^2 z}{\partial y \partial x}, \quad \frac{\partial^2 f}{\partial y \partial x}, \quad f''_{yx}(x, y), \quad D_x D_y f(x, y),$$

$$\frac{\partial^2 z}{\partial y^2}, \quad \frac{\partial^2 f}{\partial y^2}, \quad f''_{y^2}(x, y), \quad D_y D_y f(x, y).$$

As derivadas parciais, $\frac{\partial z}{\partial x}, \frac{\partial z}{\partial y}$, dir-se-ão, agora, quando haja perigo de confusão, *derivadas parciais de* 1.ª *ordem* (ou *primeiras derivadas parciais*) da função $z = f(x, y)$.

Para designar o valor duma derivada parcial de 2.ª ordem num dado ponto (a, b), usaremos notações análogas às que foram adoptadas para as derivadas de 1.ª ordem. Assim, por exemplo, qualquer dos símbolos

$$\left(\frac{\partial^2 z}{\partial x\, \partial y}\right)_{(a,\, b)}, \; f''_{xy}(a, b), \quad \frac{\partial^2 f}{\partial x \partial y}(a, b), \quad D_y D_x f(a, b),$$

designará "a segunda derivada de $f(x, y)$ em ordem a x e a y no ponto (a, b)".

22. Permutabilidade das operações de derivação

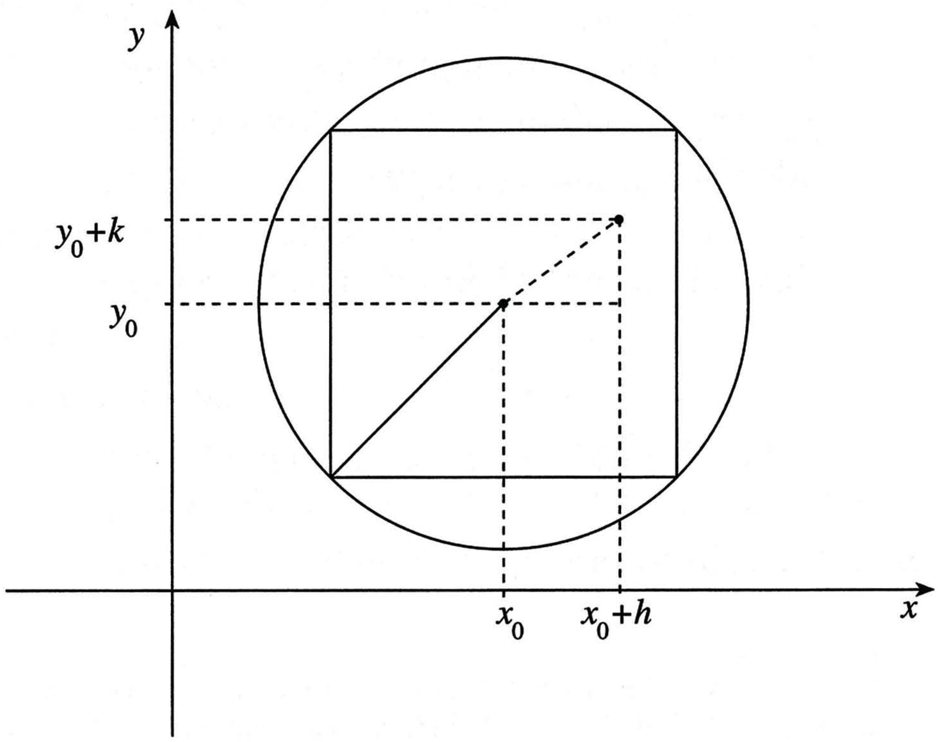
As derivadas $f''_{xy}(x, y)$, $f''_{yx}(x, y)$, quando existem, são denominadas *segundas derivadas mistas* ou *rectangulares* de $f(x, y)$; a primeira obtém-se aplicando sucessivamente os operadores D_x, D_y a $f(x, y)$, enquanto a segunda se obtém aplicando a $f(x, y)$ os mesmos operadores em ordem inversa. Serão estas operações permutáveis? Eis o que vai ser, em parte, esclarecido pelo seguinte:

TEOREMA 10 (de SCHWARZ). *Se as duas derivadas rectangulares $f''_{xy}(x, y)$, $f''_{yx}(x, y)$ são definidas numa vizinhança do ponto (x_0, y_0) e são contínuas nesse ponto, então assumem o mesmo valor em (x_0, y_0), isto é, tem-se:*

$$f''_{xy}(x_0, y_0) = f''_{yx}(x_0, y_0).$$

Demonstração. Seja ρ um número > 0 e suponhamos que as derivadas $f''_{xy}(x, y)$, $f''_{yx}(x, y)$ existem na vizinhança (ρ) do ponto (x_0, y_0), sendo contínuas nesse ponto. Então, desde que tomemos $|h| < \rho/\sqrt{2}$, $|k| < \rho/\sqrt{2}$, ter-se-á $\sqrt{h^2+k^2} < \rho$ e, portanto, o ponto $(x_0 + h, y_0 + k)$ pertencerá à vizinhança (ρ) de (x_0, y_0). Posto isto, consideremos a expressão

$$\Delta = f(x_0 + h, y_0 + k) - f(x_0 + h, y_0) - f(x_0, y_0 + k) + f(x_0, y_0)$$



para $|h| < \rho / \sqrt{2}$, $|k| < \rho / \sqrt{2}$. Se pusermos

$$\varphi(x) \equiv f(x, y_0 + k) - f(x, y_0),$$

virá:

$$\Delta = \varphi(x_0 + h) - \varphi(x_0).$$

Então, aplicando o teorema dos acréscimos finitos (o que é lícito(1), como se pode reconhecer), teremos:

$$\begin{aligned} \Delta &= h\varphi'(x_0 + \Theta_1 h) = \\ &= h[f'_x(x_0 + \Theta_1 h, y_0 + k) - f'_x(x_0 + \Theta_1 h, y_0)], \end{aligned}$$

com $0 < \Theta_1 < 1$. Por outro lado, pondo $\psi(y) \equiv f'_x(x_0 + \Theta_1 h, y)$, será:

$$\Delta = h[\psi(y_0 + k) - \psi(y_0)]$$

e, tornando a aplicar o mesmo teorema, obtém-se

$$\begin{aligned} \Delta &= hk\psi'(y_0 + \Theta_2 k) = \\ &= hkf''_{xy}(x_0 + \Theta_1 h, y_0 + \Theta_2 k), \quad 0 < \Theta_2 < 1. \end{aligned}$$

Façamos, agora, intervir a continuidade de $f''_{xy}(x, y)$ em (x_0, y_0). Quando h e k tendem para 0, também $\Theta_1 h$ e $\Theta_2 k$ tendem para zero, e $f''_{xy}(x_0 + \Theta_1 h, y_0 + \Theta_2 k)$ tenderá para $f''_{xy}(x_0, y_0)$ (em virtude da suposta continuidade em (x_0, y_0)). Virá, pois:

$$f''_{xy}(x_0 + \Theta_1 h, y_0 + \Theta_2 k) = f''_{xy}(x_0, y_0) + \varepsilon_1,$$

ou seja:

$$\Delta = hk[f''_{xy}(x_0, y_0) + \varepsilon_1],$$

sendo ε_1 um infinitésimo com (h, k).

---

(1) – Como na demonstração do TEOREMA 7, podemos supor já o número ρ escolhido de modo que as funções $f''_{xy}(x, y)$, $f''_{yx}(x, y)$ sejam finitas na vizinhança (ρ) de (x_0, y_0).

Notando, agora, que a expressão de Δ é simétrica em x, y, poderemos escrever, invertendo os papéis de x e de y:

$$\Delta = kh[f''_{yx}(x_0, y_0) + \varepsilon_2],$$

sendo ε_2 um infinitésimo com (h, k).

Ter-se-á, pois, para valores de h, k não nulos:

$$f''_{xy}(x_0, y_0) + \varepsilon_1 = f''_{yx}(x_0, y_0) + \varepsilon_2$$

donde, tomando limites quando $(h, k) \to (0, 0)$:

$$f''_{xy}(x_0, y_0) = f''_{yx}(x_0, y_0) \qquad \text{(q.e.d.)}.$$

Em particular, podemos afirmar que:

Se a função $f(x, y)$ admite derivadas $f''_{xy}(x, y)$, $f''_{yx}(x, y)$ contínuas no conjunto aberto $\mathcal{A}$, então será $f''_{xy}(x, y) = f''_{yx}(x, y)$ para todo o ponto de $\mathcal{A}$.

23. Derivadas parciais de ordens sucessivas para funções de duas ou mais variáveis

Continuemos a considerar uma função real $f(x, y)$ de duas variáveis reais, definida num conjunto aberto $\mathcal{A} \subset \mathbf{R}^2$. Suponhamos que as operações D_x, D_y podem aplicar-se várias vezes, consecutivamente, sobre a função $f(x, y)$ e sobre as funções assim obtidas. As duas operações D_x, D_y, quando aplicadas uma a seguir à outra, poderão permutar-se, desde que conduzam, em ambos os casos, a funções contínuas em $\mathcal{A}$, (TEOREMA 10). Suponhamos que se efectuaram m derivações em ordem a x e n derivações em ordem a y; então, verificada a referida hipótese sobre a continuidade das derivadas, a função obtida será sempre a mesma qualquer que seja a ordem pela qual forem executadas as derivações, e poderá, assim, designar-se pelo símbolo único

$$\frac{\partial^{m+n} f}{\partial x^m \, \partial y^n}$$

que indica as m derivações em ordem a x precedendo as n derivações em ordem a y.

Por exemplo, as derivadas de 3.ª ordem

$$\frac{\partial^3 f}{\partial x^2 \partial y}, \quad \frac{\partial^3 f}{\partial x\, \partial y\, \partial x}, \quad \frac{\partial^3 f}{\partial y\, \partial x^2},$$

se existirem e forem contínuas, serão as três coincidentes e poderão, assim, designar-se pelo primeiro símbolo.

Consideremos, agora, mais geralmente, uma função real,

$$y = f(x_1, x_2, \ldots, x_n),$$

definida num conjunto aberto $A \subset \mathbf{R}^n$. Fixadas todas as variáveis independentes menos duas, x_i e x_k, com $i \neq k$, a variável y torna-se função exclusiva das duas variáveis x_i, x_k, função a que poderemos aplicar o TEOREMA DE SCHWARZ.

Deste modo, chegaremos facilmente à conclusão de que, se existirem as derivadas

$$\frac{\partial^2 f}{\partial x_i\, \partial x_k}, \quad \frac{\partial^2 f}{\partial x_k\, \partial x_i},$$

e forem (como funções de x_1, x_2, ..., x_n) contínuas no conjunto A, os seus valores serão iguais em todo o ponto de A.

Suponhamos que, sobre a função considerada, se efectuam k_1 derivações em ordem a x_1, k_2 derivações em ordem a x_2, ..., k_n derivações em ordem a x_n, e que todas as derivadas assim obtidas são contínuas em A qualquer que seja a ordem pela qual se efectuem aquelas derivações. Deste modo, a função final será uma derivada parcial de ordem $k_1 + k_2 + \cdots + k_n$, da função f e poderá sempre designar-se pela notação

$$\frac{\partial^{k_1 + k_2 + \cdots + k_n} f}{\partial x_1^{k_1}\, \partial x_2^{k_2} \ldots\ldots \partial x_n^{k_n}}.$$

Em particular, pode algum dos números $k_1, k_2, \ldots, k_n$, ser igual a zero: então, o correspondente factor simbólico do denominador terá o expoente 0 e poderá ser omitido.

NOTA IMPORTANTE. A regra de derivação das funções compostas pode ainda aplicar-se ao cálculo das derivadas de ordem superior. Para fixar ideias, limitemo-nos ao caso duma função de duas variáveis,

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

que admite derivadas parciais contínuas, até uma certa ordem μ, num aberto $\mathcal{A}$ de $\mathbf{R}^2$, sendo x, y funções duma só variável:

$$x = \varphi(t)\,, \quad y = \psi(t)\,,$$

que admitam derivadas contínuas até à mesma ordem μ num intervalo $\mathcal{J} \subset \mathbf{R}$ (e tais que $(x, y) \in \mathcal{A}$, quando $t \in \mathcal{J}$).

Então, virá:

$$\text{(1)} \qquad \frac{dz}{dt} = \frac{\partial f}{\partial x}\frac{dx}{dt} + \frac{\partial f}{\partial y}\frac{dy}{dt}\,.$$

É claro que $\dfrac{\partial f}{\partial x}, \dfrac{\partial f}{\partial y}$ serão ainda funções compostas de t por intermédio de x, y, de modo que se for $\mu \geqslant 2$, teremos:

$$\text{(2)} \qquad \begin{aligned} \frac{d}{dt}\left(\frac{\partial f}{\partial x}\right) &= \frac{\partial^2 f}{\partial x^2}\frac{dx}{dt} + \frac{\partial^2 f}{\partial x\,\partial y}\frac{dy}{dt} \\ \frac{d}{dt}\left(\frac{\partial f}{\partial y}\right) &= \frac{\partial^2 f}{\partial y\,\partial x}\frac{dx}{dt} + \frac{\partial^2 f}{\partial y^2}\frac{dy}{dt}\,. \end{aligned}$$

Derivando ambos os membros de (1) em ordem a t, virá, então,

$$\frac{d^2 z}{dt^2} = \frac{d}{dt}\left(\frac{\partial f}{\partial x}\right)\cdot\frac{dx}{dt} + \frac{d}{dt}\left(\frac{\partial f}{\partial y}\right)\cdot\frac{dy}{dt} + \frac{\partial f}{\partial x}\cdot\frac{d^2 x}{dt^2} + \frac{\partial f}{\partial y}\cdot\frac{d^2 y}{dt^2}\,,$$

donde, atendendo a (2):

$$\frac{d^2z}{dt^2} = \frac{\partial^2 f}{\partial x^2}\left(\frac{dx}{dt}\right)^2 + 2\,\frac{\partial^2 f}{\partial x\,\partial y}\cdot\frac{dx}{dt}\cdot\frac{dy}{dt} + \frac{\partial^2 f}{\partial y^2}\left(\frac{dy}{dt}\right)^2 +$$

$$+\frac{\partial f}{\partial x}\cdot\frac{d^2x}{dt^2} + \frac{\partial f}{\partial y}\cdot\frac{d^2y}{dt^2}\,.$$

E, analogamente, para as derivadas de ordem superior.

Em particular, a mudança de variáveis pode ser do tipo

$$x = x \quad , \quad y = \varphi(x)$$

já considerado na nota do n.° 14. Então, ter-se-á, como é fácil reconhecer,

$$\frac{d^2z}{dx^2} = \frac{\partial^2 f}{\partial x^2} + 2\,\frac{\partial^2 f}{\partial x\,\partial y}\cdot\frac{dy}{dx} + \frac{\partial^2 f}{\partial y^2}\left(\frac{dy}{dx}\right)^2 + \frac{\partial f}{\partial y}\cdot\frac{d^2y}{dx^2}\,,$$

resultado este a que se poderia ter chegado directamente.

Veremos, adiante, o partido que, nestes cálculos, se pode tirar dos diferenciais de ordem superior.

24. Polinómios homogéneos. Potência duma soma de *n* parcelas

Convém aqui abrir um parêntese para poder prosseguir o nosso estudo.

Já sabemos que se chama *monómio* em $x_1, x_2, \ldots, x_n$, toda a expressão do tipo

$$\gamma x_1^{k_1}\, x_2^{k_2} \cdots x_n^{k_n}\,,$$

sendo γ um número *real* qualquer e $k_1, k_2, \ldots, k_n$, números *inteiros não negativos* (alguns deles podem ser nulos). *Grau do monómio* é a soma $k_1 + k_2 + \cdots + k_n$ dos expoentes das suas variáveis.

Polinómio em $x_1, x_2, \ldots, x_n$, é toda a expressão que, ou se reduz a um monómio, ou é formada por vários monómios em $x_1, x_2, \ldots, x_n$, (em número finito) ligados entre si por sinais de adição algébrica. *Grau do polinómio* é o maior dos graus dos seu termos, isto é, dos monómios que o constituem.

Pode acontecer que todos os termos dum polinómio tenham um mesmo grau p: o polinómio diz-se, então, *homogéneo* (de grau p). Por exemplo o polinómio em x, y, z,

$$x^3 y z^2 - 7\,x^2 y^4 - 5\,x^6,$$

é homogéneo de grau 6.

Todo o polinómio pode ser dividido nas suas *secções homogéneas,* escrevendo em primeiro lugar o termo de grau 0 (independente), depois a soma dos termos de grau 1 (secção homogénea de grau 1) e assim sucessivamente até chegar aos termos de grau máximo. Por exemplo, o polinómio em $x, y,$

$$x^2 y - x^2 + 2x - 3y + xy - 2 + y^3 + 2y^2$$

pode ser dividido nas suas secções homogéneas, por ordem crescente, tal como segue:

$$-2 + (2x - 3y) + (-x^2 + xy + 2y^2) + (x^2 y + y^3).$$

Posto isto, convém fazer uma outra observação prévia. A conhecida fórmula do binómio, que dá o desenvolvimento da potência duma soma de duas parcelas (de expoente natural) pode ser generalizada ao caso duma soma de mais de duas parcelas. Com efeito, vamos demonstrar que, dados n números $x_1, x_2, \ldots, x_n$ (reais ou complexos) e sendo m um número natural qualquer, a potência de expoente m da soma $x_1 + x_2 + \cdots + x_n$ é dada pela seguinte fórmula chamada FÓRMULA DO POLINÓMIO DE LEIBNIZ:

$$(x_1 + x_2 + \cdots + x_n)^m = \sum_{k_1 + \ldots + k_n = m} \frac{m!}{k_1!\,k_2!\cdots k_n!}\,x_1^{k_1} x_2^{k_2} \cdots x_n^{k_n}$$

em que a soma é estendida a todos os sistemas $(k_1, k_2, \ldots, k_n)$ de números *inteiros não negativos* cuja soma é m. (Não esquecer que, por convenção, $0! = 1$).

A fórmula é manifestamente verdadeira para $n = 2$, pois que, então, coincide com a *fórmula do binómio*. Basta agora seguir o *método de indução matemática*: demonstremos que, sendo a fórmula verdadeira para as somas de n parcelas, também será verdadeira para as somas de $n+1$ parcelas.

Com efeito, dada uma soma de $n+1$ parcelas $x_1, x_2, \ldots, x_{n+1}$, podemos escrevê-la sob a forma de uma soma de duas parcelas,

$$S_{n+1} = (x_1 + x_2 + \cdots + x_n) + x_{n+1} = S_n + x_{n+1},$$

à qual se pode aplicar a fórmula do binómio

$$(S_n + x_{n+1})^m = \sum_{0 \le k_{n+1} \le m} \frac{m!}{(m - k_{n+1})!\, k_{n+1}!} S_n^{m-k_{n+1}} x_{n+1}^{k_{n+1}}$$

mas, supondo a FÓRMULA DE LEIBNIZ já demonstrada para o caso da soma de n parcelas, virá, atendendo a que $S_n = x_1 + x_2 + \cdots + x_n$, sendo o expoente a diferença, $m - k_{n+1}$:

$$S_n^{m-k_{n+1}} = \sum_{k_1 + \cdots + k_n = m - k_{n+1}} \frac{(m - k_{n+1})!}{k_1!\, k_2!\, \ldots\, k_n!} x_1^{k_1} x_2^{k_2} \cdots x_n^{k_n},$$

expressão esta que, introduzida na fórmula precedente, conduz logo ao resultado que se pretendia demonstrar:

$$S_{n+1}^m = \sum_{k_1 + \cdots + k_n = m} \frac{m!}{k_1!\, k_2 \ldots k_n!\, k_{n+1}!} x_1^{k_1} x_2^{k_2} \ldots x_{n+1}^{k_{n+1}}.$$

A FÓRMULA DE LEIBNIZ mostra-nos, em particular, que:

A *potência m* da soma $x_1 + x_2 + \cdots + x_n$ *é um polinómio homogéneo de grau m em* $x_1, x_2, \ldots, x_n$.

Tem-se, por exemplo:

$$(x_1 + x_2 + x_3)^3 = x_1^3 + x_2^3 + x_3^3 + 3\,(x_1^2 x_2 + x_1^2 x_3 + x_2^2 x_1 + x_2^2 x_3 + x_3^2 x_1 + x_3^2 x_2) + 6\, x_1 x_2 x_3.$$

25. As derivações parciais consideradas como operadores lineares

O conceito de operador linear, definido em Matemáticas Gerais para o caso de transformações entre espaços cartesianos, é susceptível de largas extensões. Comecemos por introduzir a

DEFINIÇÃO 20. *Uma função real* $f(x_1, x_2, \ldots, x_n)$ *diz-se continuamente derivável até uma dada ordem* μ *num conjunto aberto* $A \subset \mathbf{R}^n$, *quando nesse conjunto admite todas as derivadas parciais de ordens inferiores ou iguais a* μ, *contínuas em* A *(podendo ainda, é claro, admitir em* A *derivadas contínuas de ordem superior a* μ).

Consideremos, em primeiro lugar, o caso das funções de duas variáveis, $f(x, y)$, $g(x, y)$, ... continuamente deriváveis até à ordem μ num aberto $A \subset \mathbf{R}^2$ e designemos por $\mathfrak{F}_\mu$ a família dessas funções. Se for $\mu \geqslant 1$, os símbolos de derivações parcial

$$\frac{\partial}{\partial x} \text{ ou } D_x, \frac{\partial}{\partial y} \text{ ou } D_y$$

representam operadores (ou operações) lineares sobre as funções da família $\mathfrak{F}_\mu$, pois que se tem(1)

$$\frac{\partial}{\partial x}(f+g) = \frac{\partial}{\partial x} f + \frac{\partial}{\partial x} g, \quad \frac{\partial}{\partial x}(af) = a \frac{\partial}{\partial x} f$$

$$\frac{\partial}{\partial y}(f+g) = \frac{\partial}{\partial y} f + \frac{\partial}{\partial y} g, \quad \frac{\partial}{\partial y}(af) = a \frac{\partial}{\partial y} f,$$

quaisquer que sejam $f, g \in \mathfrak{F}_\mu$ e sendo a um número real qualquer.

Dum modo geral, dados dois operadores lineares L_1, L_2 sobre funções da família $\mathfrak{F}_\mu$, podemos definir a soma $L_1 + L_2$ como é uso em casos tais, $(L_1 + L_2) f = L_1 f + L_2 f$, qualquer que seja $f \in \mathfrak{F}_\mu$.

Por sua vez, o significado do produto $L_1 \cdot L_2$ subordina-se ao conceito geral de produto de operadores $(L_1 \cdot L_2) f = L_1 (L_2 f)$, para $f \in \mathfrak{F}_\mu$.

---

(1) – É claro que estes operadores transformam funções pertencentes a $\mathfrak{F}_\mu$ em funções pertencentes a $\mathfrak{F}_{\mu-1}$. Em particular, $\mathfrak{F}_0$ será a família das funções contínuas em A.

De acordo com estas convenções, o símbolo

$$a\frac{\partial}{\partial x}+b\frac{\partial}{\partial y}$$

onde a e b designam números reais quaisquer, terá um significado definido pela fórmula

$$\left(a\frac{\partial}{\partial x}+b\frac{\partial}{\partial y}\right)f=a\frac{\partial f}{\partial x}+b\frac{\partial f}{\partial y}\,,\ \text{para } f\in\mathfrak{F}_\mu;$$

isto é, *aplicar o operador* $a\dfrac{\partial}{\partial x}+b\dfrac{\partial}{\partial y}$ *a uma função* $f(x,y)$ *consiste em*:

1) – *derivar a função* f *em ordem a* x *e multiplicar o resultado por* a;
2) – *derivar a função* f *em ordem a* y *e multiplicar o resultado por* b;
3) – *somar os dois resultados assim obtidos.*

Ponhamos $L=a\dfrac{\partial}{\partial x}+b\dfrac{\partial}{\partial y}$. Ter-se-á, por definição:

$$L^2=L\cdot L,\ L^3=L\cdot L^2,\ \ldots,\ L^n=L\cdot L^{n-1},\ \ldots,$$

e ainda $L^1=L$, $L^0=I$ (operador identidade).

Suponhamos $\mu>1$. Então, qualquer que seja $f\in\mathfrak{F}_\mu$, virá, atendendo às anteriores definições de soma e de produto:

$$\begin{aligned}L^2f&=\left(a\frac{\partial}{\partial x}+b\frac{\partial}{\partial y}\right)\cdot\left(a\frac{\partial}{\partial x}+b\frac{\partial}{\partial y}\right)f\\&=\left(a\frac{\partial}{\partial x}+b\frac{\partial}{\partial y}\right)\left(a\frac{\partial f}{\partial x}+b\frac{\partial f}{\partial y}\right)\\&=a\frac{\partial}{\partial x}\left(a\frac{\partial f}{\partial x}+b\frac{\partial f}{\partial y}\right)+b\frac{\partial}{\partial y}\left(a\frac{\partial f}{\partial x}+b\frac{\partial f}{\partial y}\right)\end{aligned}$$

ou ainda, em virtude da linearidade de $\dfrac{\partial}{\partial x}$, $\dfrac{\partial}{\partial y}$ e tendo em vista o TEOREMA DE SCHWARZ:

$$L^2 f = a^2 \frac{\partial^2 f}{\partial x^2} + 2\,ab\,\frac{\partial^2 f}{\partial x \partial y} + b^2 \frac{\partial^2 f}{\partial y^2}.$$

Será, portanto:

$$L^2 = \left(a\frac{\partial}{\partial x} + b\frac{\partial}{\partial y}\right) = a^2\left(\frac{\partial}{\partial x}\right)^2 + 2\left(a\frac{\partial}{\partial x}\right)\left(b\frac{\partial}{\partial y}\right) + b^2\left(\frac{\partial}{\partial y}\right)^2.$$

Este resultado sugere-nos que a fórmula do binómio de Newton possa aplicar-se a operadores do tipo considerado; isto é, que se tenha qualquer que seja $m = 1, 2, \ldots,$

$$\left(a\frac{\partial}{\partial x} + b\frac{\partial}{\partial y}\right)^m = \sum_{p=0}^{m} \frac{m!}{p!\,(m-p)!}\left(a\frac{\partial}{\partial x}\right)^{m-p}\left(b\frac{\partial}{\partial y}\right)^p =$$

(1)

$$= a^m \frac{\partial^m}{\partial x^m} + ma^{m-1} b\,\frac{\partial^m}{\partial x^{m-1}\partial y} + \binom{m}{2} a^{m-2} b^2 \frac{\partial^m}{\partial x^{m-2}\partial y^2} + \cdots + b^m \frac{\partial^m}{\partial y^m}.$$

Pois bem, esta conjectura é confirmada *raciocinando por indução sobre m.*

Com efeito, demonstra-se que, se a fórmula é verdadeira para o expoente *m,* sê-lo-á também para o expoente $m + 1$ (exactamente como se faz em Álgebra clássica, para estabelecer o resultado correspondente).

A fórmula (1), *é, pois, válida qualquer que seja o número natural m.*

As condições precedentes generalizam-se, *mutatis mutandis,* ao caso dos operadores

$$\frac{\partial}{\partial x_1}, \frac{\partial}{\partial x_2}, \ldots, \frac{\partial}{\partial x_n}$$

aplicáveis às funções $f(x_1, x_2, \ldots, x_n)$ continuamente deriváveis num aberto $A \subset \mathbf{R}^n$. Sendo $h_1, h_2, \ldots, h_n$, constantes reais quaisquer, ter-se-á, por definição,

$$\left(h_1 \frac{\partial}{\partial x_1} + h_2 \frac{\partial}{\partial x_2} + \cdots + h_n \frac{\partial}{\partial x_n}\right) f = h_1 \frac{\partial f}{\partial x_1} + h_2 \frac{\partial f}{\partial x_2} + \cdots + h_n \frac{\partial f}{\partial x_n}.$$

Ora, designando por L o operador linear assim definido, é fácil reconhecer que a FÓRMULA DE LEIBNIZ subsiste para o desenvolvimento da potência L^m deste operador, isto é, que

$$L^m = \left(h_1 \frac{\partial}{\partial x_1} + h_2 \frac{\partial}{\partial x_2} + \cdots + h_n \frac{\partial}{\partial x_n}\right)^m =$$

$$= \sum_{k_1 + \cdots + k_n = m} \frac{m!}{k_1!\, k_2! \ldots k_n!} h_1^{k_1} h_2^{k_2} \ldots h_n^{k_n} \frac{\partial^m}{\partial x_1^{k_1} \partial x_2^{k_2} \ldots \partial x_n^{k_n}}$$

em que, para abreviar, se pôs:

$$\frac{\partial^m}{\partial x_1^{k_1} \partial x_2^{k_2} \ldots \partial x_n^{k_n}} = \left(\frac{\partial}{\partial x_1}\right)^{k_1} \cdot \left(\frac{\partial}{\partial x_2}\right)^{k_2} \cdot \ldots \cdot \left(\frac{\partial}{\partial x_n}\right)^{k_n}.$$

Mas é preciso não perder de vista que o operador L^m só pode ser aplicado a funções continuamente deriváveis até à ordem m no conjunto considerado.

26. Diferenças finitas e diferenciais de ordem superior para funções de mais de uma variável

Consideremos uma função real $f(X) = f(x_1, x_2, \ldots, x_n)$ definida num conjunto $\mathcal{C} \subset \mathbf{R}^n$. Já no n.º 13 precisámos o que se entende por diferença finita de $f(X)$ correspondente ao acréscimo $H = (h_1, h_2, \ldots, h_n)$:

$$\Delta_H f(X) = f(X + H) - f(X),$$

que é uma nova função de X definida num conjunto $\mathcal{C}_1 \subset \mathcal{C}$. Notemos, agora, que o operador Δ_H pode ainda ser aplicado à função obtida, o que equivale a aplicar o quadrado do mesmo operador à função inicial, $f(X)$; ter-se-á:

$$\Delta_H^2 f(X) = \Delta_H (\Delta_H f(X)) = \Delta_H (f(X + \mathrm{H}) - f(X)) =$$
$$= f(X + 2H) - 2\, f(X + H) + f(X).$$

Obtém-se, deste modo, uma outra função de X, que se *chama segunda diferença de* $f(X)$ (no ponto X) *relativa ao acréscimo* H. Por sua vez, a *terceira diferença relativa* a H será:

$$\begin{aligned}\Delta_H^3 f(X) &= \Delta_H \Delta_H^2 f(X)\\ &= f(X+3H) - 3f(X+2H) + 3f(X+H) - f(X).\end{aligned}$$

Dum modo geral, a *diferença de ordem* m *de* $f(X)$, relativa ao mesmo acréscimo, será definida por recorrência:

$$\Delta_H^m f(X) = \Delta_H(\Delta_H^{m-1} f(X)), \text{ para } m = 1, 2, \ldots,$$

supondo, é claro, $\Delta_H^0 = 1$ (operador idêntico); esta nova função de X resultará, pois, de aplicar m vezes sucessivas o operador Δ_H a $f(X)$.

A m-ésima diferença de $f(X)$ pode ainda ser designada pela notação $\Delta^m f(X)$, quando esteja subentendido o acréscimo H de X. (Trata-se, manifestamente, duma função de X definida no conjunto $\mathcal{C}_m$ dos pontos de X de $\mathcal{C}$ tais que os pontos acrescidos $X + H$, $X + 2H, \ldots, X + mH$, pertençam ainda a $\mathcal{C}$; em particular, $\mathcal{C}_m$ pode ser vazio.

Suponhamos agora que $f(X)$ *é continuamente derivável no interior de* $\mathcal{C}$ *até uma certa ordem* $\mu \geqslant 1$.

Como sabemos, o diferencial de $f(X)$ relativo a H é dado por

$$d_H f(X) = h_1 f'_{x_1}(X) + h_2 f'_{x_2}(X) + \cdots + h_n f'_{x_n}(X),$$

ou seja, segundo a convenção do n.º precedente:

$$d_H f(X) = \left(h_1 \frac{\partial}{\partial x_1} + h_2 \frac{\partial}{\partial x_2} + \cdots + h_n \frac{\partial}{\partial x_n}\right) f(X).$$

O operador d_H é, pois, idêntico ao que está aplicado a f no segundo membro da anterior igualdade, isto é:

$$d_H = h_1 \frac{\partial}{\partial x_1} + h_2 \frac{\partial}{\partial x_2} + \cdots + h_n \frac{\partial}{\partial x_n}.$$

Ora, já vimos que a potência m deste operador pode ser desenvolvida segundo a FÓRMULA DE LEIBNIZ.

Portanto, se aplicarmos m vezes sucessivas o operador d_H a $f(X)$ (sendo $m \leqslant \mu$), obtém-se o seguinte resultado:

$$d_H^m f(X) = \sum_{k_1+\cdots+k_n=m} \frac{m!}{k_1! \dots k_n!} h_1^{k_1} \dots h_n^{k_n} \frac{\partial^m f(X)}{\partial x_1^{k_1} \cdots \partial x_n^{k_n}},$$

que é um polinómio homogéneo de grau m em $h_1, h_2, \dots, h_n$, cujos coeficientes são funções de X definidas no interior de $\mathcal{C}$ ($m \leqslant \mu$).

DEFINIÇÃO 21. *O anterior polinómio em $h_1, h_2, \dots, h_n$, que resulta de aplicar m vezes sucessivas o operador d_H à função $f(X)$, é denominado diferencial de ordem m de $f(X)$ relativo ao acréscimo H (no ponto X).*

Para $m = 2, 3, \dots$ ter-se-á o *segundo diferencial*, e *terceiro diferencial*, etc. Para $m = 1$ tem-se, simplesmente, o diferencial de $f(X)$ que, para melhor se distinguir dos de ordem superior, se chamará, por vezes, *primeiro diferencial de $f(X)$*.

Observe-se, em particular, que *o conhecimento do m-ésimo diferencial de $f(X)$ equivale ao conhecimento de todas as derivadas parciais de ordem m de $f(X)$.*

Se pusermos agora, como foi justificado no n.º 13:

$$dx_1 = h_1, dx_2 = h_2, \dots, dx_n = h_n,$$

a expressão do m-ésimo diferencial de $f(X)$ toma novo aspecto, que dispensa a indicação explícita do acréscimo H. E é com tal aspecto que aparecerá sistematicamente na prática.

Por exemplo, para uma função $z = f(x, y)$ de duas variáveis reais, ter-se-á

$$d^3 z = \left(dx \frac{\partial}{\partial x} + dy \frac{\partial}{\partial y} \right)^3 f(x, y)$$

$$= \frac{\partial^3 f}{\partial x^3} dx^3 + 3 \frac{\partial^3 f}{\partial x^2 \partial y} dx^2 dy + 3 \frac{\partial^3 f}{\partial x \partial y^2} dx\, dy^2 + \frac{\partial^3 f}{\partial y^3} dy^3$$

onde, para simplicidade de notação, se escreveu:

$$dx^2 \text{ por } (dx)^2,\ dy^3 \text{ por } (dy)^3, \text{ etc.}$$

Fique, pois, bem explícito, que nesta maneira de escrever, os expoentes não afectam as variáveis x, y, ..., mas sim os respectivos diferenciais dx, dy, ... E não se confundam os símbolos dz^2, dz^3, ... (quadrado de dz, cubo de dz, etc.) *com os símbolos d^2z, d^3z, ... (segundo diferencial de z, terceiro diferencial de z,* etc.).

Assim, por exemplo, se for $z = \frac{y}{x}$, ter-se-á:

$$dz^2 = \left(\frac{\partial z}{\partial x} dx + \frac{\partial z}{\partial y} dy\right)^2 = \frac{y^2}{x^4} dx^2 - 2\frac{y}{x^3} dx\,dy + \frac{1}{x^2} dy^2$$

e

$$d^2z = \frac{\partial^2 z}{\partial x^2} dx^2 + 2\frac{\partial^2 z}{\partial x \partial y} dx\,dy + \frac{\partial^2 z}{\partial y^2} dy^2$$

$$= \frac{2y}{x^3} dx^2 - \frac{2}{x^2} dx\,dy.$$

Os conceitos de diferença e diferencial de ordem m estão relacionados entre si pelo seguinte teorema que não demonstraremos: *A diferença $\Delta_H^m f(X)$ e o diferencial $d_H^m f(X)$ nos pontos X onde ambos são definidos, diferem por um infinitésimo com H de ordem superior a m.*

27. Cálculo prático dos diferenciais de ordem superior. Não invariância destes diferenciais

Consideremos, em particular, o caso das funções inteiras do 1.º grau:

$$\varphi(x_1, x_2, \ldots, x_n) \equiv c_0 + c_1x_1 + c_2x_2 + \cdots + c_nx_n;$$

tem-se:

$$d\varphi(X) = c_1 dx_1 + c_2 dx_2 + \cdots + c_n dx_n$$

e como, neste caso, as *primeiras* derivadas parciais $c_1, c_2, \ldots, c_n$, são constantes, as de ordem superior resultam todas nulas, sendo portanto:

$$d^2\varphi(X) \equiv 0, \quad d^3\varphi(X) \equiv 0, \text{ etc.};$$

de resto, prova-se que este facto é exclusivo das funções lineares. Mais particularmente, ainda, os coeficientes $c_0, c_1, \ldots, c_n$, podem ser todos nulos, excepto um, $c_i = 1$; neste caso, a função $\varphi(X)$ reduz-se à variável independente x_i:

$$\varphi(x_1, x_2, \ldots, x_n) \equiv x_i,$$

e assim, do que precede, podemos concluir que

$$d^m x_i = 0, \text{ para } m > 1 \ (i = 1, 2, \ldots, n).$$

Isto é: *os diferenciais de ordem superior das variáveis independentes* $x_1, x_2, \ldots, x_n$, *são todos nulos*; o que equivale a dizer: *na diferenciação, os diferenciais* $dx_1, dx_2, \ldots, dx_n$, *comportam-se como constantes.*

Daqui e da definição recorrente

$$d^m f = d(d^{m-1} f), \ m = 2, 3, \ldots.$$

deduz-se a seguinte:

REGRA PRÁTICA – *Para obter os diferenciais de ordem superior de uma dada função* $f(x, y, \ldots)$, *basta aplicar as regras formais de diferenciação* (n.º 16) *primeiro à função dada, depois à expressão obtida e assim sucessivamente até à ordem a que se pretende chegar, considerando em todas estas diferenciações como constantes os diferenciais* $dx, dy, \ldots$, *das variáveis independentes.*

Por exemplo, se for $z = e^x \cos y$, ter-se-á, *considerando* x, y, *como variáveis independentes*:

$$dz = e^x \cos y\, dx - e^x \operatorname{sen} y\, dy,$$

$$d^2z = d(e^x \cos y)\, dx - d(e^x \operatorname{sen} y)\, dy$$
$$= (e^x \cos y\, dx - e^x \operatorname{sen} y\, dy)\, dx - (e^x \operatorname{sen} y\, dx + e^x \cos y\, dy)\, dy$$
$$= e^x \cos y\, dx^2 - 2\, e^x \operatorname{sen} y\, dx\, dy - e^x \cos y\, dy^2 .$$

É claro que se poderia chegar aos mesmos resultados aplicando directamente a definição.

Porém, quando se trata de *diferenciais de funções compostas*, a regra prática enunciada exige precauções especiais. Para fixar ideias, limitemo-nos ao caso duma função de duas variáveis,

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

continuamente derivável até uma certa ordem $\mu > 1$ num aberto $\mathcal{A}$ de $\mathbf{R}^2$, e suponhamos feita a mudança de variáveis

$$(1) \qquad x = \varphi(u, v), \quad y = \psi(u, v),$$

sendo φ, ψ, funções continuamente deriváveis até à ordem μ num aberto $\mathcal{B}$ de $\mathbf{R}^2$ (e tais que o ponto (x, y) esteja em $\mathcal{A}$ quando (u, v) está em $\mathcal{B}$). Então, z resulta função composta de u, v, por intermédio de x, y:

$$z = f(\varphi(u, v), \psi(u, v)) = \chi\,(u, v)$$

e já vimos no n.º 15 que se tem, à face de (1):

$$df = d\chi, \quad \text{isto é,} \quad \frac{\partial z}{\partial x} dx + \frac{\partial z}{\partial y} dy = \frac{\partial z}{\partial u} du + \frac{\partial z}{\partial v} dv .$$

Nisto consiste o que chamámos *princípio da invariância do (primeiro) diferencial*. Porém, vamos ver que tal princípio não subsiste para os diferenciais de ordem superior, isto é, vamos ver que, em geral, *não se tem* no caso considerado

$$d^n f = d^n \chi \quad \text{para} \quad n > 1 .$$

Basta verificá-lo para $n = 2$. Tem-se, como sabemos:

$$d^2 f = \frac{\partial^2 z}{\partial x^2}\, dx^2 + 2\, \frac{\partial^2 z}{\partial x\, \partial y}\, dx\, dy + \frac{\partial^2 z}{\partial y^2}\, dy^2$$

$$d^2 \chi = \frac{\partial^2 z}{\partial u^2}\, du^2 + 2\, \frac{\partial^2 z}{\partial u\, \partial v}\, du\, dv + \frac{\partial^2 z}{\partial v^2}\, dv^2$$

Ora:

$$(2) \qquad d\chi = df = \frac{\partial z}{\partial x}\, dx + \frac{\partial z}{\partial y}\, dy .$$

Sendo:

$$(3) \qquad \begin{cases} dx = d\varphi = \dfrac{\partial x}{\partial u}\, du + \dfrac{\partial x}{\partial v}\, dv \\[2ex] dy = d\psi = \dfrac{\partial y}{\partial u}\, du + \dfrac{\partial y}{\partial v}\, dv \end{cases}$$

e

$$(4) \qquad \begin{cases} d^2 x = \dfrac{\partial^2 x}{\partial u^2}\, du^2 + 2\, \dfrac{\partial^2 x}{\partial u \partial y}\, du\, dv + \dfrac{\partial^2 x}{\partial y^2}\, dv^2 \\[2ex] d^2 y = \dfrac{\partial^2 y}{\partial u^2}\, du^2 + 2\, \dfrac{\partial^2 y}{\partial u \partial v}\, du\, dv + \dfrac{\partial^2 y}{\partial v^2}\, dv^2 . \end{cases}$$

Portanto, diferenciando a expressão (2) de $d\chi$ *em ordem a* u e v, virá:

$$d^2 \chi = d\left(\frac{\partial z}{\partial x}\right) dx + d\left(\frac{\partial z}{\partial y}\right) dy + \frac{\partial z}{\partial x}\, d(dx) + \frac{\partial z}{\partial y}\, d(dy)$$

$$= \left(\frac{\partial^2 z}{\partial x^2}\, dx^2 + 2\, \frac{\partial^2 z}{\partial x \partial y}\, dx\, dy + \frac{\partial^2 z}{\partial y^2}\, dy^2\right) + \frac{\partial z}{\partial x}\, d^2 x + \frac{\partial z}{\partial y}\, d^2 y ,$$

ou seja:

$$\boxed{d^2 \chi = d^2 f + \frac{\partial z}{\partial x}\, d^2 x + \frac{\partial z}{\partial y}\, d^2 y}$$

sendo os diferenciais dx, dy, dados por (3) e d^2x, d^2y, dados por (4). Como, *em geral*, se terá neste caso $d^2x \not\equiv 0$, $d^2y \not\equiv 0$, será, *em geral*, $d^2\chi \not\equiv d^2f$, como tínhamos afirmado.

Estas considerações estendem-se imediatamente ao caso duma função $z = f(x_1, x_2, \ldots, x_n)$, na qual se tenha efectuado uma mudança de variáveis do tipo

$$x_i = \varphi_i(t_1, t_2, \ldots, t_p), \quad i = 1, 2, \ldots, n,$$

(sem ser, necessariamente, $n = p$); neste caso a variável z é função composta de $t_1, t_2, \ldots, t_p$: $z = \chi\,(t_1, t_2, \ldots, t_p)$ e ter-se-á:

$$\boxed{d^2\chi = d^2f + \frac{\partial z}{\partial x_1}\, d^2x_1 + \frac{\partial z}{\partial x_2}\, d^2x_2 + \cdots + \frac{\partial z}{\partial x_n}\, d^2x_n}$$

com $d^2x_i = d^2\varphi_i$ (2.º diferencial de x_i em ordem aos tt); e, analogamente, para os diferenciais de ordem $n > 2$.

Como vimos, há um caso em que $d^2x_i \equiv 0$: *aquele em que* x_i *é função linear de* $t_1, t_2, \ldots, t_p$ *(mas só nesse caso assim acontece).*

De todas estas considerações há que reter a seguinte:

ADVERTÊNCIA – *Ao aplicar a precedente regra prática deve sempre ter-se o cuidado prévio de registar quais são as variáveis independentes, isto é, as variáveis em ordem às quais se fazem as diferenciações, porque, exceptuando o referido caso particular, só os diferenciais dessas variáveis se comportam como constantes.*

Retomemos o anterior exemplo $z = e^x \cos y$. Se as variáveis independentes forem precisamente x, y, o diferencial d^2z é o que foi calculado atrás; mas se x, y, forem funções de outras variáveis $u, v, \ldots$, *em ordem às quais se faz a diferenciação*, há que juntar novos termos contendo os diferenciais, d^2x, d^2y: $d^2z = e^x \cos y\, dx^2 - 2\, e^x \operatorname{sen} y\, dx\, dy -$ $- e^x \cos y\, dy^2 + e^x \cos y\, d^2x - e^x \operatorname{sen} y\, d^2y$.

Ter-se-á $d^2x \equiv 0$, $d^2y \equiv 0$, se, e só se, x, y, forem funções *lineares* das novas variáveis $u, v, \ldots$.

NOTA. O cálculo dos diferenciais de ordem superior constitui um dos processos práticos para obter todas as derivadas até uma dada ordem, especialmente quando se trata de funções compostas. Consideremos, novamente, o caso da função $z = f(x, y)$ em que se fez a mudança de variáveis $x = \varphi(u, v)$, $y = \psi(u, v)$. Se quisermos, por exemplo, conhecer as segundas derivadas parciais de z em ordem a u e a v, bastará utilizar o segundo diferencial,

$$d^2z = \left(\frac{\partial^2 f}{\partial x^2}\, dx^2 + 2\, \frac{\partial^2 f}{\partial x \partial y}\, dx\, dy + \frac{\partial^2 f}{\partial y^2}\, dy^2\right) + \frac{\partial f}{\partial x}\, d^2x + \frac{\partial f}{\partial y}\, d^2y$$

e substituir dx, dy, d^2x, d^2y pelas respectivas expressões em função de du e dv: os coeficientes de du^2, $du\, dv$ e dv^2 na expressão obtida para d^2z dão-nos as derivadas parciais procuradas.

Por exemplo, tem-se (confrontar com a nota do n.º 23):

$$\frac{\partial^2 z}{\partial u^2} = \frac{\partial^2 f}{\partial x^2}\left(\frac{\partial x}{\partial u}\right)^2 + 2\, \frac{\partial^2 f}{\partial x \partial y}\left(\frac{\partial x}{\partial u}\right)\left(\frac{\partial y}{\partial u}\right) + \frac{\partial^2 f}{\partial y^2}\left(\frac{\partial y}{\partial u}\right)^2 +$$

$$+ \frac{\partial f}{\partial x}\, \frac{\partial^2 x}{\partial u^2} + \frac{\partial f}{\partial y}\, \frac{\partial^2 y}{\partial u^2}\,.$$

28. Fórmula de TAYLOR para as funções de mais de uma variável

Consideremos uma função real $f(X) = f(x_1, x_2, \ldots, x_n)$, das n variáveis reais $x_1, x_2, \ldots, x_n$, e um ponto $A = (a_1, a_2, \ldots, a_n)$ do seu campo de existência. Façamos ainda a seguinte hipótese:

(α) *A função $f(X)$ é continuamente derivável até uma certa ordem $m \geqslant 1$ numa vizinhança $\mathcal{V}$ do ponto A.*

Seja, então, ρ o raio de $\mathcal{V}$ e seja $H = (h_1, h_2, \ldots, h_n)$ um elemento arbitrário de $\mathbf{R}^n$ tal que:

$$|H| < \rho\,;$$

se pusermos $X = A + tH$, ter-se-á, manifestamente, $|X - A| = |t| \cdot |H|$, e portanto será $|X - A| < \rho$ desde que $|t| \leqslant 1$; isto é:

$$X \in \mathcal{V} \text{ quando } -1 \leqslant t \leqslant 1.$$

Ponhamos agora:

(1) $$\varphi(t) = f(A + tH) = f(a_1 + th_1, a_2 + th_2, \ldots, a_n + th_n)$$

para $t \in [-1, 1]$. Atendendo à hipótese e a que

(2) $$x_1 = a_1 + th_1, x_2 = a_2 + th_2, \ldots, x_n = a_n + th_n,$$

virá, aplicando o TEOREMA 8:

$$\varphi'(t) = \frac{\partial f}{\partial x_1} \cdot \frac{dx_1}{dt} + \frac{\partial f}{\partial x_2} \cdot \frac{dx_2}{dt} + \cdots + \frac{\partial f}{\partial x_n} \frac{dx_n}{dt}$$

$$= \frac{\partial f}{\partial x_1} h_1 + \frac{\partial f}{\partial x_2} h_2 + \cdots + \frac{\partial f}{\partial x_n} h_n,$$

ou seja, simbolicamente:

(3) $$\varphi'(t) = \left(h_1 \frac{\partial}{\partial x_1} + h_2 \frac{\partial}{\partial x_2} + \cdots + h_n \frac{\partial}{\partial x_n}\right) f(X),$$

para $X = A + tH$.

Quer dizer: obtém-se $\varphi'(t)$ aplicando a $f(X)$ o operador

$$d_H = h_1 \frac{\partial}{\partial x_1} + h_2 \frac{\partial}{\partial x_2} + \cdots + h_n \frac{\partial}{\partial x_n}$$

e fazendo *depois* a mudança de variáveis (2). Mas o segundo membro é ainda função composta de t, por intermédio de X, e a sua derivada em ordem a t obter-se-á pelo mesmo processo; isto é:

$$\varphi''(t) = d_H[d_H f(X)] = d_H^2 f(X),$$

para $X = A + tH$; e assim sucessivamente. Dum modo geral, teremos:

(4) $$\varphi^{(p)}(t) = d_H^p f(X) = \left(h_1 \frac{\partial}{\partial x_1} + h_2 \frac{\partial}{\partial x_2} + \cdots + h_n \frac{\partial}{\partial x_n}\right)^p f(X),$$

para $X = A + tH$ e supondo, é claro, $p \leqslant m$ (vide hipótese). Uma primeira conclusão a tirar daqui é a seguinte:

A função $\varphi(t)$ *admite derivadas contínuas (e portanto finitas) até à ordem m no intervalo* $[-1, 1]$. Ter-se-á, pois, segundo a FÓRMULA DE MAC-LAURIN, nas condições em que foi estabelecida em Matemáticas Gerais,

$$\varphi(t) = \varphi(0) + t\varphi'(0) + \frac{t^2}{2!}\varphi''(0) + \cdots + \frac{t^{m-1}}{(m-1)!}\varphi^{(m-1)}(0) +$$

$$+ \frac{t^m}{m!}\varphi^{(n)}(\Theta t), \quad \text{para } |t| \leqslant 1 \text{ e com } 0 < \Theta < 1$$

(Θ dependente de t). Em particular, fazendo $t = 1$, virá

$$(5) \quad \varphi(1) = \varphi(0) + \varphi'(0) + \frac{1}{2!}\varphi''(0) + \cdots + \frac{1}{(m-1)!}\varphi^{(m-1)}(0) +$$

$$+ \frac{1}{m!}\varphi^{(m)}(\Theta).$$

Mas, para $t = 1$, vem $X = A + tH = A + H$, e para $t = 0$, vem $X = A$. Portanto, virá, atendendo a (1),

$$\varphi(1) = f(A + H)\ , \ \varphi(0) = f(A)$$

e, para $0 < p < m$, segundo (4):

$$\varphi^{(p)}(0) = (d_H^p f)(A),$$

em que o símbolo A escrito no lugar de X está a indicar que, depois de aplicado ao operador diferencial, se deve substituir X por A, ou seja, x_1 por a_1, x_2 por a_2, $\ldots$, x_n por a_n: ter-se-á, portanto, segundo a fórmula de Leibniz,

$$(6) \quad (d_H^p f)(A) = \sum_{k1 + \cdots + kn = p} \frac{p!}{k_1 \ldots k_n!} h_1^{k1} \ldots h_n^{kn} \frac{\partial^p f}{\partial x_1^{k1} \ldots x_n^{kn}}(A).$$

Analogamente, ter-se-á

$$\varphi^m(\Theta) = (d_H^m f)(A + \Theta H).$$

E assim, por substituição em (5), virá finalmente a FÓRMULA DE TAYLOR GENERALIZADA:

$$f(A+H) = f(A) + (d_H f)(A) + \frac{1}{2!}(d_H^2 f)(A) + \cdots$$

$$\cdots + \frac{1}{(m-1)!}(d_H^{m-1} f)(A) + \frac{1}{m!}(d_H^m f)(A+\Theta H), \quad 0 < \Theta < 1,$$

a qual é válida para todo o acréscimo H tal que $|H| < \rho$, *supondo verificada a hipótese* (α) *e tendo em conta* (6).

Como exemplo, consideremos o caso duma função $f(x, y)$ de duas variáveis, continuamente derivável até à ordem 2 na vizinhança (ρ) do ponto (a, b). Virá, então:

$$f(a+h, b+k) = f(a, b) + [hf'_x(a, b) + kf'_y(a, b)] +$$

$$+ \frac{1}{2!}[h^2 f''_{x^2}(a+\Theta h, b+\Theta k) + 2hk f''_{xy}(a+\Theta h, b+\Theta k) +$$

$$+ k^2 f''_{y^2}(a+\Theta h, b+\Theta k)]$$

com $0 < \Theta < 1$, sob a condição de ser $h^2 + k^2 < \rho^2$.

Verificada a hipótese (α) e sendo $0 < \Theta < 1$, pode ainda demonstrar-se que:

$$(d_H^m f)(A+\Theta H) = (d_H^m f)(A) + \omega_m,$$

sendo ω_m um infinitésimo com H de ordem superior a m (ver DEFINIÇÕES 13 e 14). Daqui resulta o seguinte *aspecto particular da* FÓRMULA DE TAYLOR:

$$f(A+H) = \sum_{p=0} \frac{1}{p!}(d_H^p f)(A) + \omega_m, \quad \lim_{H \to \vec{0}} \frac{\omega_m}{|H|^m} = 0.$$

Se notarmos que, segundo (6), o termo geral do somatório aqui indicado é um polinómio homogéneo de grau p, a fórmula anterior diz-nos que:

Verificada a hipótese (α), *a função* $f(X)$ *será um polinómio em* $X - A$ *mais um infinitésimo com* $X - A$ *de ordem superior a m.*

Assim, para o exemplo de há pouco, ter-se-á:

$$f(a+h,\, b+k) = f(a,\, b) + f'_x(a,\, b)\,h + f'_y(a,\, b)\,k +$$

$$+ \frac{1}{2}\, f''_{x^2}(a,\, b)\,h^2 + f''_{xy}(a,\, b)\,h\,k + \frac{1}{2}\, f''_{y^2}(a,\, b)\,k^2 + \omega_2\,,$$

com $\lim\limits_{\substack{h \to 0 \\ k \to 0}} \dfrac{\omega_2}{h^2 + k^2} = 0$. Neste caso, a função aparece representada por um polinómio do 2.º grau em *h, k,* à parte o infinitésimo ω_2 de ordem superior à segunda.

ADVERTÊNCIA FINAL

Estes apontamentos referem-se a uma parte do nosso Curso de Cálculo Infinitesimal no Instituto Superior de Agronomia.

Indicamos, em seguida, as obras que mais frequentemente consultámos ao preparar as nossas lições:

G. SANSONE – *Lezioni di Analisi Matematica,* vol. I e II, 5.ª edição. Milano, Padova 1943.

G. VALIRON – *Théorie des Fonctions,* 2.ª edição. Masson, Paris 1948.

P. APPELL – *Cours de Mathématiques Générales – Analyse Mathématique,* 6.ª edição redigida por VALIRON. Masson, Paris 1948.

M. PICONE – *Lezioni di Analisi Infinitesimale*, vol. I. Catania 1923.

R. COURANT – *Differential and Integral Calculus*, 2.ª edição. Nordermann, New York 1940.

E. G. PHILLIPS – *A course of Analysis*. University Press, Cambridge 1946.

Porém, do ponto de vista didáctico, estas lições representam trabalho original.

Por isso mesmo pedimos, a quem porventura deseje fazer uso da orientação aqui seguida, o favor de citar a presente edição.

Aos alunos Srs. José Cardoso Soveral Dias e José Crespo Ascenso agradecemos a colaboração valiosa prestada na revisão e na reprodução destes apontamentos.

A nota do n.º 2 corresponde a um pedido de esclarecimento do aluno Sr. Tomás Moreira. A prática vai-nos mostrando que, mesmo em cursos técnicos, a Matemática não pode deixar de ser considerada com o rigor que lhe compete. Devem, sem dúvida, ser omitidas certas demonstrações, certos "raffinements" da teoria; mas deve-se, por outro lado, evitar cair em simplificações deturpadoras.

Lisboa, Março de 1953

J. Sebastião e Silva