UNIVERSIDADE DE ÉVORA

UMA ABORDAGEM DO ENSINO DO CÁLCULO COM ORDENS DE GRANDEZA

Uma dissertação apresentada por

Maria Domingas Sousa Reforço

para a obtenção do grau de
Mestre em Matemática Aplicada

Orientador: Prof. Doutor A. J. Franco de Oliveira

Mestrado em Matemática Aplicada
Biénio 1997/99

Agradecimentos

É com o maior prazer que exprimo a minha gratidão:

Ao Professor Doutor A. J. Franco de Oliveira por ter aceite orientar esta dissertação, pelas suas sugestões e críticas, por todo o apoio e estímulo que me deu e pela forma dedicada como acompanhou o meu trabalho.

À Universidade de Évora e aos docentes do mestrado em Matemática aplicada, biénio 1997–1999, pelos conhecimentos transmitidos.

Aos meus pais e irmã, pela confiança que em mim depositaram e pela paciência e incentivo que me concederam. Um agradecimento especial pela motivação que me incutiram na fase final deste trabalho.

Ao meu marido, José Manuel, pelo carinho, incentivo e positivismo que sempre me encorajava nos momentos mais difíceis.

A todos aqueles que de forma directa ou indirecta contribuíram para a realização deste trabalho, nomeadamente, ao Professor Doutor Luís Gonzaga, aos professores e amigos Luís e Inácia Cabanas, e aos colegas de mestrado Graça Carita e Romeu Silva.

Índice

Prefácio

As aulas de mestrado da cadeira de Análise Não-Standard, leccionada pelo Pr. A. J. Franco de Oliveira, foram o despertar do meu interesse pela Análise Não- - Standard.

Ao terminar a parte curricular do mestrado, falei com o Pr. A. J. Franco de Oliveira, e pedi-lhe que fosse ele a orientar a minha dissertação, pedido que foi aceite.

Em diálogo com o meu orientador acerca das possibilidades de desenvolver a dissertação em Análise Não-Standard, ficou assente que nela seria apresentada a preparação de alguns conteúdos, adequados a um curso de uma disciplina de Análise Não-Standard. Esta aplicação poderia verificar-se, por exemplo, na leccionação da disciplina de Análise Não-Standard a alunos da licenciatura em Matemática, do 1° ou 2° anos, ou equivalente, no caso das licenciaturas em Engenharia ou Física.

A dissertação teve como "fonte" principal o texto que serviu de base ao Curso de Análise Não-Standard leccionado por F. Diener e Gautheron, em Nice no ano de 1996.

É composta por sete capítulos. O capítulo I refere-se aos conjuntos finitos e infinitos, à distinção de objectos matemáticos em standard e não-standard e às regras que governam o seu uso. O capítulo II é consagrado ao conjunto dos números reais, e nele são descritas as principais propriedades dos referidos números, mas o próprio conceito de número real subjacente tem propriedades e elementos novos, relativamente à Matemática clássica, que advêm da distinção standard/não-standard no universo. De especial importância, em todo o cálculo, é o facto de esta distinção, standard/não standard, permitir introduzir diferentes ordens de grandeza dos números (noções como as de número infinitesimal, apreciável, infinitamente grande), e as regras para lidar com elas, nomeadamente as chamadas regras de Leibniz. No capítulo III estuda-se as sucessões. As derivadas fazem parte do capítulo IV, em que é definido o conceito de função de classe C^1, o que é a função inversa de uma função, e não é esquecida a observação de uma função à lupa; para finalizar o capítulo, são indicadas algumas aplicações das derivadas. As funções de várias variáveis são estudadas no capítulo V, considerando-se, na maior parte dos casos, funções de duas variáveis para melhor se visualizar as situações. O capítulo VI refere-se aos crescimentos limitados. Dele constam o enunciado e a demonstração do teorema dos acréscimos finitos, e as fórmulas de Taylor com uma e duas variáveis. Os integrais e as primitivas são estudados no último capítulo, onde se indica como calcular áreas aproximando-as por reuniões de rectângulos de áreas infinitesimais; estuda-se ainda a ligação entre o cálculo das primitivas e o das áreas.

Apresentada a estrutura do trabalho, seguem algumas considerações que focam a Análise Não-Standard em si mesma, na sua história, e na sua integração no universo da Matemática, e consequente interesse pela sua aplicação ao ensino.

A Análise Não-Standard é, antes de mais, um instrumento ao dispor dos matemáticos de qualquer denominação ou especialidade que, a cada área ou aplicação preferencial, pode trazer mais ou trazer menos, conforme o grau de investimento e utilização. Em certo sentido que pode ser precisado tecnicamente, a utilização dos infinitamente pequenos e dos infinitamente grandes actuais, em análise e noutras áreas matemáticas é, tão somente, uma mais sofisticada utilização de uma forma ou outra do Axioma da Escolha ou algo que se assemelha.

Desde finais do século XIX, a linguagem da Matemática é a linguagem da teoria dos conjuntos, e a teoria "mãe de todas as teorias" matemáticas é a teoria dos conjuntos, na versão de Zermelo-Fraenkel, incluindo o Axioma da Escolha, ZFC. Nesta teoria há um único conceito primitivo, o de pertença ($\in$). Por condescendência platonista, chamamos às variáveis x, y, z,...., X, Y, Z, ... conjuntos, e imaginamos que os axiomas ZFC descrevem um "universo" de conjuntos, suficientemente rico para que nele se representem como conjuntos os objectos matemáticos habituais (números, relações, funções, espaços, estruturas, etc.) e as demonstrações habituais nas diferentes disciplinas matemáticas.

Sempre que conveniente, a linguagem primitiva é enriquecida com novos símbolos, por meio de definições: constantes, isto é, símbolos que denotam conjuntos particulares como $\emptyset$, 2, $\mathbb{N}$, $\mathbb{R}^2$; símbolos relacionais ou predicativos, que denotam relações no universo, como as primitivas $=$, $\in$ e as relações de inclusão $\subset$, $\subseteq$; símbolos funcionais ou operacionais, que denotam operações no universo como o $\cup$, $\cap$, $\times$, etc.

É importante saber que a prática matemática corrente se pode formalizar em ZFC. Essa prática não vai sofrer qualquer alteração na Matemática Não-Standard, excepto no sentido de um duplo enriquecimento.

No final dos anos setenta surgiu uma axiomatização devido a E. Nelson, que toma como ponto de partida a teoria ZFC, na linguagem respectiva. À linguagem desta teoria (que é, em boa aproximação, a linguagem do discurso matemático tradicional), junta um novo predicado primitivo unário: "*standard*". Quer dizer, no universo dos objectos matemáticos (conjuntos) usuais (números, funções, espaços, etc.), distinguimos, por meio do novo conceito, os que são standard e os que o não são.

A teoria de Nelson é designada IST — é a *teoria dos conjuntos internos* (*Internal Set Theory*), pois, nela, os conjuntos são chamados conjuntos internos, os quais se dividem em conjuntos standard e conjuntos não-standard. Os objectos e conjuntos "ordinários" com que lidam os matemáticos no dia a dia são todos standard. Entretanto, a Análise Não-Standard tira partido de novas entidades — os objectos não-standard, que são, em certo sentido, objectos "ideais" relativamente aos standard (números infinitamente grandes e números infinitamente pequenos, etc.), e os chamados conceitos externos ou as suas extensões, as classes externas, que na literatura são vulgarmente chamados conjuntos externos, mas que, na realidade (pelo menos no formalismo de IST), não são conjuntos, estritamente falando. São classes externas, como, por exemplo, a classe dos naturais standard, a classe dos reais infinitesimais, etc. Tais classes são, em regra, classes relativamente "pequenas", pois que são classes contidas em conjuntos e definidas em termos do novo predicado standard — não é lícito aplicar a estas classes os axiomas e os teoremas clássicos (de ZFC).

IST é uma extensão conservativa de ZFC, podendo ser livremente utilizada para demonstrar resultados clássicos (normalmente, com simplificações notáveis nas demonstrações), e permitindo caracterizar noções clássicas (por exemplo, da Análise) de maneiras, em regra, mais simples e intuitivas.

Em IST, as extensões das estruturas clássicas são constituídas da mesma maneira que em ZFC, mas são mais ricas, pois nelas é permitido fazer mais distinções do que anteriormente (standard/não-standard, e outras definíveis a partir destes conceitos), tal como se estivessem sendo observadas com uma lupa cujo poder de resolução fosse agora maior. Por outro lado, o alcance da nova teoria parece maior, pois enriquece a antiga, ao nível conceptual e metodológico fundamental.

No trabalho não é exposta a teoria de IST. Para um curso introdutório de Análise Matemática, como se pretende ilustrar nesta dissertação, não é, talvez, necessário nem conveniente apresentar toda a potência demonstrativa dos axiomas de IST (todo o aparato e potência axiomática de IST), especialmente tudo aquilo que pressupõe um domínio aprofundado da lógica e da axiomática de ZFC. Não é de esperar tais conhecimentos prévios, em alunos que frequentam o primeiro ou o segundo ano de um curso de Análise Matemática. Por conseguinte, a abordagem feita no trabalho tem por base uma versão mais fraca da axiomática de Nelson, aproveitando desta apenas o suficiente para os aspectos do cálculo que interessa desenvolver num primeiro curso dessa natureza. Outros aspectos, que pressuporiam a totalidade dos axiomas de Nelson, poderão ser desenvolvidos num segundo curso. Aliás, como se pode verificar nos primeiros capítulos, a própria formulação dos princípios básicos da Análise Não-Standard não é em todos os casos tal como Nelson a axiomatizou. É o caso, por exemplo, do princípio da definição (princípio 1.5).

Começa a ser cada vez mais evidente, também, que os novos conceitos e os resultados que os envolvem, sem equivalentes clássicos, contribuem decisivamente para um mais fino estudo de fenómenos físicos e estatísticos, o que se confirma abundantemente na literatura mais recente.

Mesmo sem ter em conta estes recentes desenvolvimentos, outro facto importante tem sido observado: na prática Matemática corrente, apenas se faz um uso muito limitado dos princípios básicos da teoria dos conjuntos, a qual é, pois, um luxo! Ainda uma grande parte da Matemática corrente tem lugar numa extensão conservativa da Aritmética de 2.ª ordem (com axiomas de Compreensão e de Escolha apropriados). Sendo assim, o que as demonstrações não-standard de resultados clássicos têm permitido observar é um uso muito mais eficiente dos princípios clássicos, em comparação com as demonstrações clássicas (nos casos em que já eram conhecidas, ou vieram a ser posteriormente às não-standard).

A Análise Não-Standard é um tema aliciante e um caminho aberto ao desenvolvimento da Matemática e do ensino.

Seria muito interessante poder testar, numa situação concreta de sala de aula, o texto que aqui se apresenta. De uma tal experiência resultaria, certamente, uma melhoria e maior adequação ao fim proposto no próprio texto.

Capítulo 1

Conjuntos finitos e infinitos

1.1 Conjuntos finitos

Definição: Uma aplicação ou uma função do conjunto E para o conjunto F faz corresponder a todo o elemento de E um único elemento de F. Se todo o elemento de F tem, no máximo, um originário, é uma injecção. Se todo o elemento de F tem, pelo menos, um originário, é uma sobrejecção.

Estas definições podem transcrever-se através das seguintes fórmulas:

- Uma aplicação $f: E \to F$ é uma injecção sse
$$\forall\, x\,, x' \in E (x \neq x' \Rightarrow f(x) \neq f(x')).$$
- Uma aplicação $f: E \to F$ é uma sobrejecção sse
$$\forall\, y \in F\, \exists x \in E\ (y = f(x)).$$
- $f: E \to F$ é uma bijecção sse é ao mesmo tempo injectiva e sobrejectiva.

Definição: Um conjunto diz-se *infinito* sse existe uma bijecção do conjunto sobre uma das suas partes próprias.

Proposição 1.1

Todo o conjunto contendo um subconjunto infinito é infinito.

Dem.

Seja E infinito e $F \supseteq E$. Seja E' uma parte própria de E e seja $f: E \to E'$ uma bijecção, que existem por hipótese. O conjunto $F' = E' \cup (F - E)$ é uma parte própria de F e a aplicação g que é a identidade sobre $F - E$ e que é igual a f sobre E é uma bijecção de F sobre F'. □

Corolário 1.2

Os conjuntos $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$ *e* $\mathbb{C}$ *são infinitos.*

Definição: Um conjunto E é finito sse toda a injecção de E em si mesmo é uma sobrejecção.

Proposição 1.3

O conjunto vazio é finito e para todo o natural n*, o conjunto* $I_n := \{p \in \mathbb{N} : 0 \leq p \leq n\}$ *é finito.*

Dem.

O conjunto vazio é finito, porque a única aplicação de $\emptyset$ em $\emptyset$ é a aplicação vazia que é injectiva e sobrejectiva.

Por indução matemática: $I_0 = \{0\}$ é finito porque a única aplicação $f : I_0 \to I_0$ é a identidade. Suponhamos I_{n-1} finito e seja $f : I_n \to I_n$ uma aplicação injectiva; defina-se $g : I_n \to I_n$ por

$$g(p) = \begin{cases} p & \text{se } p \neq f(n) \text{ e } p \neq n \\ n & \text{se } p = f(n) \\ f(n) & \text{se } p = n \end{cases}$$

Se $f(n) \neq n$, esta aplicação g é uma *transposição*, isto é , uma simples troca dos inteiros n e $f(n)$, e os outros valores ficarão iguais. Senão, g é a identidade; em todos os casos, é claramente uma bijecção. A aplicação $\widehat{f} :\ g \circ f$ é injectiva e $\widehat{f}(n) = n$. Portanto, $\widehat{f}\big|_{I_{n-1}}$ é injectiva e, portanto, sobrejectiva por hipótese de indução. □

De facto, I_n é o protótipo dos conjuntos finitos, como indica o teorema seguinte, que se deve a Zermelo, e que admitiremos (a demonstração deste teorema utiliza o axioma da escolha, como pode ser consultado em A.J.F. Oliveira, *Teoria dos Conjuntos*, Esc. Ed. 1982) :

Teorema 1.4

Se $E \neq \emptyset$ *é um conjunto finito, então existe um único inteiro* $n \in \mathbb{N}$ *e uma bijecção* $f : I_n \to E$ *. Este número* n *é o número de elementos de* E *ou cardinal de* E*.*

Comentário: Este teorema pode ser compreendido da seguinte forma: se E é finito, *podem enumerar-se os seus elementos,* visto que a bijecção f associa a cada elemento de E um inteiro compreendido entre 0 e n. Isto conduz à ideia intuitiva (mas falsa) de que um conjunto finito é um *conjunto de que se pode fazer um inventário*, isto é, do qual se podem enumerar os elementos. Isto será verdade, com todo o rigor, se pudermos mostrar que todos os inteiros não negativos são *naturais*. Por exemplo, é indiscutível que o conjunto dos grãos de areia é um conjunto finito, mas poderemos jamais fazer o inventário ? O mesmo para o conjunto das moléculas de gás que estão dentro de um recipiente, ou o dos seres vivos que existem sobre a terra. Convém, portanto, distinguir entre a noção *intuitiva* de finito que está ligada aos números inteiros *acessíveis*, que se procura formalizar com o adjectivo "standard", e a noção *matemática* de finito (de que demos a definição) e se aplica, às vezes, aos conjuntos cujo cardinal é standard (conjunto dos dias da semana ou conjunto dos estudantes de anfíbios); e existem os conjuntos cujo cardinal é finito mas não standard (diz-se infinitamente grande), como o conjunto dos grãos de areia ou o dos consumidores numa economia, para todos os efeitos práticos, o seu número é inacessível.

1.2 Standard ou não-standard: os primeiros princípios

O método mais simples para permitir a utilização dos números com ordens de grandeza diferente é o de supor que todo o objecto matemático, conjunto, número, função, sucessão, ... é standard ou não é standard. Acrescenta-se, assim, um predicado ao nosso discurso, mas não se define este predicado de forma precisa, limitamo-nos a indicar as regras que governam o seu uso. Neste trabalho enunciaremos quatro. Vejamos os dois primeiros, chamados respectivamente *princípio da definição* e *princípio da idealização.* Veremos, posteriormente, os dois seguintes, o *princípio da sombra* (no capítulo 2) e o *princípio da transferência* (no capítulo 3).

Princípio 1.5 (Princípio da definição)

Se a definição de um objecto matemático, conjunto, número, função, ... pode escrever-se sem utilizar o adjectivo standard (e sem utilizar outros objectos não--standard), então este objecto é standard.

É em virtude deste princípio que os números 0, 1, 2, ... , π, e, ... são números standard, que os conjuntos $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{R}$, $[0,1]$, $[0, +\infty[$, ... são conjuntos standard, que as sucessões $x_n = 3n + 4$, $u_{n+1} = u_n + 2$, ... são sucessões standard ... . Pelo contrário, se w é um inteiro não-standard, a função $f(x) = x + w$ é uma função não-standard.

Deduz-se também deste princípio a seguinte consequência, muitas vezes útil:

Proposição 1.6

Seja f uma função standard. Sendo assim, necessariamente o conjunto $E = \operatorname{dom} f$ é standard, e a imagem por f de um elemento standard de E é um elemento standard (do conjunto de chegada de f).

Dem.

Seja $f : E \to F$ uma função e $x \in E$ um elemento standard. O domínio de f, E, é recuperável a partir de f enquanto conjunto de pares ordenados, logo é standard, pelo princípio da definição. O elemento $y = f(x)$ é definido de forma única a partir de objectos standard (x e f), e é, portanto, standard, também pelo princípio da definição.[1]

□

Nota-se, pelo contrário, que a imagem de um elemento não-standard por uma função standard pode ser standard ou não-standard, como o mostram as duas funções seguintes $f, g : \mathbb{N} \to \mathbb{N}$ definidas por: $f(n) = n$ e $g(x) = \begin{cases} 0 & \text{se } n \text{ é par} \\ 1 & \text{se } n \text{ é ímpar} \end{cases}$.

Princípio 1.7 (Princípio de idealização)

Todo o conjunto infinito contém elementos não-standard e todo o conjunto finito standard contém apenas elementos standard.

Deste princípio, resulta particularmente, o facto de que existem inteiros não-standard (visto que $\mathbb{N}$ é um conjunto infinito). Chamam-se a estes inteiros idealmente grandes (*ig*) ou infinitamente grandes, porque são, forçosamente, maiores que todos os inteiros standard: com efeito, se existir um w natural não-standard que seja menor do que um natural standard n, w será um elemento do conjunto $I_n := \{p \in \mathbb{N} : 0 \leq p \leq n\}$ que é um conjunto finito e standard (pelo princípio da definição) e, portanto, só pode ter elementos standard. Tem-se, portanto, o seguinte resultado:

Proposição 1.8

Todo o inteiro natural majorado por um inteiro standard é standard.

Dem.

Suponhamos $n \in \mathbb{N}$ standard. O conjunto $S_n = \{m \in \mathbb{N} : m < n\}$ é standard e finito, logo, todos os seus elementos são standard. □

---

[1] Se f é entendida como um triplo ordenado (f, E, F), então é claro que F também é standard. Em qualquer caso, se f é standard, $\operatorname{im} f$ = conjunto imagem de f, é standard, pelo referido princípio.

Exemplos:

Os conjuntos que se seguem são não-standard (supondo w ig):

- o conjunto dos múltiplos de w ;
- o conjunto I_w;
- o conjunto $\{w, w+1, w+2, ..., 2w\}$;
- o conjunto $\{w\}$;
- o conjunto dos números inferiores a w.

Os conjuntos que se seguem são standard (supondo n standard)

- o conjunto dos inteiros superiores a n ;
- $\{1, ..., 2^n\}$;
- o conjunto dos números pares;
- o conjunto das soluções de uma equação polinomial $P(x) = 0$ com coeficientes standard e de grau standard;
- o conjunto dos divisores de w estritamente inferiores a w, se w é um número primo *ig* : este conjunto é simplesmente $\{1\}$.

Mais exemplos:

• Em $\mathbb{N}$, existem outros números não-standard para além dos infinitamente grandes ?

Os infinitamente grandes em $\mathbb{N}$ são os números que são maiores que qualquer número standard, e é por isso que se chamam infinitamente grandes, logo, em $\mathbb{N}$ não existem outros números não-standard para além dos infinitamente grandes.

• Mostrar que o conjunto $\{1, 2, ..., 2001\} \subset \mathbb{N}$ não contém números não-standard e que é um conjunto standard.

O conjunto $\{1, 2, ..., 2001\}$ é standard pelo princípio da definição (princípio 1.5) e é finito, logo, só contém elementos standard pelo princípio de idealização (princípio 1.7).

• Mostrar que o conjunto dos números ímpares é um conjunto standard.

O conjunto dos números ímpares é o conjunto $\{n \in \mathbb{N} : \exists p \in \mathbb{N}\ n = 2p + 1\}$, que é bem definido a partir de objectos standard; logo, pelo princípio da definição, é standard.

• O conjunto dos números primos é finito? Existem números primos não-standard ?

O conjunto dos números primos não é (classicamente) finito, porque, para todo o número primo p, pode construir-se um número primo maior, que é o produto de todos os números primos inferiores ou iguais a p adicionando 1. Pelo princípio de idealização, existem números primos não-standard.

Capítulo 2

Ordens de grandeza em $\mathbb{R}$

Neste capítulo, consagrado ao corpo ordenado dos números reais, $\mathbb{R}$, serão descritas as suas principais propriedades.

2.1 Operações com os reais

Todos sabemos que o conjunto $\mathbb{R}$ está munido de duas operações (o comum dos mortais conhece quatro, mas o matemático só retém duas; as outras duas são simplesmente compreendidas como as inversas das primeiras), que são duas aplicações, representadas por $+$ e $\times$ respectivamente

$$+ : \mathbb{R} \times \mathbb{R} \to \mathbb{R}, \qquad \times : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$$

$$(a, b) \mapsto a + b, \qquad (a, b) \mapsto ab$$

e que têm as seguintes propriedades: para todos os reais a, b e c,

- $a + b = b + a$ e $ab = ba$ (comutatividade de $+$ e $\times$);
- $a + (b + c) = (a + b) + c$ e $a(bc) = (ab)c$ (associatividade de $+$ e $\times$);
- existe $0 \in \mathbb{R}$ tal que $a + 0 = a$ e existe $1 \in \mathbb{R}$ tal que $a1 = a$ (existência de elemento neutro);
- para todo $a \in \mathbb{R}$, existe b ($:= \ -a$) tal que $a + b = 0$ (existência de simétrico) e, se $a \neq 0$, existe $b(:= 1/a)$ tal que $a \times b = 1$ (existência de inverso);
- $a(b + c) = ab + ac$ (distributividade de $\times$ em relação a $+$).

Quando um conjunto possui duas operações que verificam as propriedades anteriores, diz-se que este é um *corpo comutativo*. O conjunto $\mathbb{R}$ é, portanto, um corpo comutativo. Em contrapartida, o conjunto $\mathbb{Z}$, munido das mesmas propriedades, não é um corpo comutativo, porque $1/a$ não existe geralmente em $\mathbb{Z}$. Quanto ao conjunto $\mathbb{Q}$ é, tal como $\mathbb{R}$, um corpo comutativo para estas duas operações.

2.2 Ordem em $\mathbb{R}$

Sabe-se que os números reais são ordenados de uma forma natural, o que não é o caso dos vectores de $\mathbb{R}^2$ ou dos complexos $\mathbb{C}$, por exemplo. Esta relação, representada por $\leq$, verifica as seguintes propriedades: para todo $a, b, c \in \mathbb{R}$,

- $a \leq a$ (reflexividade);
- $(a \leq b \text{ e } b \leq a) \Rightarrow a = b$ (anti-simetria);
- $(a \leq b \text{ e } b \leq c) \Rightarrow a \leq c$ (transitividade);
- tem-se sempre $a \leq b$ ou $b \leq a$ (ordem total).

Diz-se que esta relação de ordem parcial (três primeiras propriedades) é total (quarta propriedade) e que se trata de um *corpo ordenado*, isto é, que é *compatível com as duas operações*: com efeito, para todo $a, b, c \in \mathbb{R}$, tem-se:

- $(a \leq b) \Rightarrow a + c \leq b + c$
- $(a \leq b \text{ e } c > 0) \Rightarrow ac \leq bc$

Enfim, o conjunto dos números reais, tal como $\mathbb{Q}$, tem uma importante propriedade ligada à sua ordem: é um conjunto arquimediano, isto é, que verifica:

$$\forall x > 0,\ \forall A \geq 0,\ \exists n \in \mathbb{N}\ nx \geq A.$$

Com o "passo" x (tão pequeno quanto se queira), podemos ultrapassar todo o número A (por maior que seja), desde que se dê um número n suficientemente grande de passos.

2.3 Números para todas as escalas

Em $\mathbb{N}$ os números inteiros não-standard são todos *ig*, isto é, maiores do que os inteiros standard, em $\mathbb{R}$ os reais standard e não-standard estão mais "misturados". Podemos convencer-nos facilmente disto, observando que todo o intervalo $]a, b[$ não vazio de $\mathbb{R}$, por muito pequeno que seja, é um conjunto infinito e, portanto, contém forçosamente números não-standard, pelo princípio de idealização.

Distinguem-se, habitualmente, entre os números reais, diversas ordens de grandeza:

Definição:

- Um real x diz-se ilimitado ou infinitamente grande (*ig*) sse para todo o $n \in \mathbb{N}$ standard $|x| > n$.
- Um real x diz-se limitado (*lmt*) sse existe $n \in \mathbb{N}$ standard tal que $|x| \leq n$.
- Um real x diz-se infinitesimal ou infinitamente pequeno (*ip*) sse para todo

$n \in \mathbb{N}_+$ standard $|x| \leq \frac{1}{n}$.

- Um real x diz-se apreciável (*ap*) sse x não é nem infinitamente pequeno nem infinitamente grande.
- Dois reais x e y dizem-se infinitamente próximos sse $x - y$ é infinitamente pequeno, e escreve-se $x \simeq y$.

Observação: Note-se que estas ordens de grandeza "cortando" sobre o conjunto $\mathbb{R}$ diferentes regiões (os números infinitamente grandes a cada extremo, os números limitados que se decompõem em números infinitesimais e números apreciáveis), as fronteiras entre estas regiões não são pontos precisos. Por exemplo se x é um real limitado, $x + 1$ será também limitado (porque se $|x| < n$ e n é um inteiro standard, $n + 1$ também é (princípio da definição: $n + 1$ está definido a partir dos números standard n e 1 unicamente) e ele majora $x + 1$). E se $x > 0$ é infinitamente grande, também o é $x - 1$. Da mesma forma, se x é um real apreciável, $\frac{x}{2}$ é também um real apreciável.

Exemplos:

- Em $\mathbb{R}$, existem outros números não-standard para além dos infinitamente grandes e dos infinitesimais, por exemplo $2 + \varepsilon$, com ε infinitesimal diferente de zero, tem-se $2 + \varepsilon \simeq 2$, mas $2 + \varepsilon$ é um número real não-standard.
- O intervalo real $[3, 4]$ contém números não-standard porque é infinito, pois pelo princípio de idealização (princípio 1.7), todo o conjunto infinito contém elementos não-standard.

2.4 Operações e ordens de grandeza

Pelo princípio da definição, as duas operações $+$ e $\times$ são aplicações standard e o resultado obtido efectuando a soma ou o produto de dois números standard é também um número standard, sempre em virtude do mesmo princípio (este resultado só depende dos dois argumentos e da operação). Pode-se perguntar o que são números não-standard e mais precisamente se é possível conhecer a ordem de grandeza da soma ou do produto conhecendo a dos termos ou a dos factores. Com efeito é fácil verificar, como faremos em seguida para algumas delas, que a influência dos operadores sobre as ordens de grandeza obedecem a regras, chamadas *regras de Leibniz,* descritas pelas três tabelas que se seguem. Estas exprimem, quando é possível, as ordens de grandeza para a soma e a diferença ($\pm$), o produto ($\times$), e o quociente (/) de dois reais.

$\pm$	ip	lmt	ap	ig
ip	ip	lmt	ap	ig
lmt	lmt	lmt	lmt	ig
ap	ap	lmt	lmt	ig
ig	ig	ig	ig	?

$\times$	ip	lmt	ap	ig
ip	ip	ip	ip	?
lmt	ip	lmt	lmt	?
ap	ip	lmt	ap	ig
ig	?	?	ig	ig

$\downarrow / \rightarrow$	ip	lmt	ap	ig
ip	?	?	ip	ip
lmt	?	?	lmt	ip
ap	ig	?	ap	ip
ig	ig	ig	ig	?

Estas regras retêm-se facilmente; note-se, contudo, que dois reais não têm sempre o mesmo sinal, a soma de dois apreciáveis pode bem ser infinitamente pequena, e a soma de dois infinitamente grandes pode ser limitada, os casos $ap + ap = lmt$ (não é sempre apreciável), e $ig + ig = ?$ Imagina-se facilmente porque não existe um resultado geral para $ip \times ig$ e ip/ip; é conveniente portanto, visto que um limitado pode ser infinitamente pequeno, não esquecer que não existem resultados gerais para $lmt \times ig$ e ig/lmt: é por isso que é útil distinguir os reais apreciáveis de entre os limitados.

Vejamos, através de dois exemplos, o tipo de raciocínio que convém fazer:

Mostremos que a soma de dois números ip é ip. Sejam α e β dois números ip. Seja $n \neq 0$ um número inteiro standard. Como n é standard, $2n$ também é (princípio da definição) e portanto, visto que α e β são ip, $|\alpha| < \frac{1}{2n}$ e $|\beta| < \frac{1}{2n}$. Logo $|\alpha+\beta| < |\alpha|+|\beta| < \frac{1}{2n} + \frac{1}{2n} = \frac{1}{n}$.

Mostremos que o produto de um número limitado por um número infinitamente pequeno é infinitamente pequeno. Sejam l um número lmt e α um número ip. Seja n um número inteiro standard. Sabe-se que existe um inteiro standard m tal que $|l| < m$. Como m é um inteiro standard, mn também é (princípio da definição) e portanto $|\alpha| < \frac{1}{mn}$. Resulta então que $|l\alpha| < m\frac{1}{mn} = \frac{1}{n}$.

Exemplos: Mostrar, utilizando as regras de Leibniz que:

• Se w é infinitamente grande então $a = \frac{w-1}{w+1} - 1$ é infinitesimal.

Como $a = \frac{w-1}{w+1} - 1 = 1 - \frac{2}{w+1} - 1 = -\frac{2}{w+1}$, w é ig logo $w+1$ que é a soma de um ap com um ig é ig, logo $\frac{2}{w+1}$ é infinitesimal porque é o quociente entre um ap e um ig.

• Se w é infinitamente grande então $b = \sqrt{w+1} - \sqrt{w}$ é infinitesimal.

Temos $b = \left(\sqrt{w+1} - \sqrt{w}\right) = \left(\sqrt{w+1} - \sqrt{w}\right)\left(\frac{\sqrt{w+1}+\sqrt{w}}{\sqrt{w+1}+\sqrt{w}}\right) =$

$= \frac{w+1-w}{\sqrt{w+1}+\sqrt{w}} = \frac{1}{\sqrt{w+1}+\sqrt{w}}$, como w é ig então $\sqrt{w+1}$ e $\sqrt{w}$ são infinitamente grandes, $\sqrt{w+1} + \sqrt{w}$ é ig, logo $\frac{1}{\sqrt{w+1}+\sqrt{w}}$ que é o quociente entre um ap e um ig é, portanto, infinitesimal.

Exemplos:

• Existem infinitesimais racionais: por exemplo, $\frac{1}{w+5}$, com w natural infinitamente grande, é racional e é infinitesimal. E existem infinitesimais irracionais: se $\varepsilon \neq 0$ é infinitesimal racional então $\varepsilon\sqrt{2}$ é irracional e é infinitesimal.

• Mostrar que se ε é infinitesimal então $\frac{3}{4} + \varepsilon$ pode ser irracional.

Seja ε irracional, por exemplo, $\varepsilon = \lambda\sqrt{3}$ com λ infinitesimal racional e diferente de zero; neste caso $\frac{3}{4} + \varepsilon$ é um número irracional.

• Se N é infinitamente grande então $\frac{1}{N}$ e $\frac{1}{N+1}$ são infinitesimais; assim, $\frac{1}{N} - \frac{1}{N+1} = \frac{N+1-N}{N(N+1)} = \frac{1}{N(N+1)}$ é infinitesimal.

• Seja n um número inteiro infinitamente grande. O número real $a = \frac{1}{n^2}$ é infinitesimal. O que poderemos dizer dos números na, $7an^2$ e an^3 ?

Como n é infinitamente grande tem-se, $na = n.\frac{1}{n^2} = \frac{1}{n}$ que, pelas regras de Leibniz, é infinitesimal. Verificamos que $7an^2 = 7.\frac{1}{n^2}.n^2 = 7$, que é um número apreciável. Finalmente, $an^3 = \frac{1}{n^2}.n^3 = n$, que é infinitamente grande.

• Mostrar que existe um número x tal que $\frac{x+1}{x-1}$ $(x \neq 1)$seja :

◇ igual a 2;

Vejamos que $\frac{x+1}{x-1} = 2 \Leftrightarrow \frac{x+1}{x-1} - 2 = 0 \Leftrightarrow \frac{-x+3}{x-1} = 0 \Leftrightarrow -x+3 = 0 \Leftrightarrow x = 3$ logo, com $x \neq 1$ $\frac{x+1}{x-1} = 2$ sse $x = 3$.

◇ infinitamente grande;

Seja w infinitamente grande. Então (com $x \neq 1$) $\frac{x+1}{x-1} = w \Leftrightarrow \frac{x+1}{x-1} - w = 0 \Leftrightarrow$ $\Leftrightarrow x - wx + 1 + w = 0 \Leftrightarrow x = \frac{w+1}{w-1}$, logo $x = \frac{w+1}{w-1} = 1 + \frac{2}{w-1} = 1 + \varepsilon$, onde $\varepsilon = \frac{2}{w-1}$ é infinitesimal. Assim, $\frac{x+1}{x-1}$ é infinitamente grande se $x = 1 + \varepsilon$ onde $\varepsilon = \frac{2}{w-1}$ com w infinitamente grande.

◇ infinitesimal não nulo.

Seja $\varepsilon \neq 0$ infinitesimal. Então (com $x \neq 1$) $\frac{x+1}{x-1} = \varepsilon \Leftrightarrow x = \frac{\varepsilon+1}{\varepsilon-1} = 1 + \frac{2}{\varepsilon-1} = 1 + w$ com $w = \frac{2}{\varepsilon-1}$ infinitamente grande. Portanto, $\frac{x+1}{x-1}$ é infinitesimal não nulo se $x = 1 + w$, onde $w = \frac{2}{\varepsilon-1}$, com $\varepsilon \neq 0$ infinitesimal.

Observação: É conveniente meditar um pouco sobre a aparente contradição (que não o é, felizmente!) entre a primeira das regras de Leibniz, segundo a qual a soma se dois *ip* é *ip* (portanto, em particular, se ε é *ip*, $2\varepsilon = \varepsilon + \varepsilon$ deve ser também, o mesmo que $3\varepsilon, 4\varepsilon, ...$) e o facto de $\mathbb{R}$ ser arquimediano, de onde resulta que para todo o real dado existe um inteiro n tal que $n\varepsilon$ o ultrapassa, por exemplo 1 (que não é *ip*). Não se pode deduzir como consequência do facto de que $\varepsilon, 2\varepsilon, 3\varepsilon, 4\varepsilon, ...$, são *ip* que $n\varepsilon$ é *ip* para todo $n \in \mathbb{N}$. Isto não será verdade, de facto, senão para n "passos" com n não muito grande; os n standard, em particular, e alguns outros podem ser, mas não certamente os *ig* superiores a $E(\frac{1}{\varepsilon}) + 1$, onde $E(x)$ designa a parte inteira de x.

Da mesma forma se escolhermos um "passo" x limitado e um "objectivo" a ultrapassar A *ig* o número n de "passos" necessários será seguramente *ig*.

Observação: Resultam facilmente da definição e das regras de Leibniz as propriedades de *reflexividade*, *simetria* e *transitividade* da relação de proximidade infinitesimal: por exemplo se $x \simeq y$ e $y \simeq z$, então $x - y$ e $y - z$ são infinitesimais, logo $(x-y) + (y-z) = x - z$ é infinitesimal, por uma regra de Leibniz, logo $x \simeq z$.

Proposição 2.1. (Colocação de factores em evidência)

Sejam N standard, $p_1, p_2, ..., p_N$ números positivos e $r_1, r_2, ..., r_N$ números da mesma ordem de grandeza (isto é, todos ip, ou todos limitados ou todos apreciáveis com o mesmo sinal, ou todos infinitamente grandes com o mesmo sinal). Então existe um real r da mesma ordem de grandeza que os r_i tal que $\sum p_i r_i = r \sum p_i$.

Dem.

Seja $S := \sum p_i$. Se todos os p_i são nulos, o resultado é trivial. Senão, $S > 0$, e pode escrever-se portanto $r := \frac{1}{S}\sum p_i r_i$. Mostremos que este número r é da mesma ordem de grandeza que os r_i. Seja r_- o menor dos números r_i e r_+ o maior. Tem-se:

$$r_- = \frac{r_-}{S}\sum p_i = \sum \frac{p_i}{S} r_- \leq \sum \frac{p_i}{S} r_i \leq \sum \frac{p_i}{s} r_+ = \frac{r_+}{S}\sum p_i = r_+$$

e portanto $r_- \leq r \leq r_+$, o que mostra que r é da mesma ordem de grandeza que r_- e r_+. □

2.5 O cálculo com os "quase iguais"

O sentido do símbolo " $\simeq$ " é o da quase igualdade de dois números: a é infinitamente próximo de b (representa-se $a \simeq b$) significa que os números a e b não são forçosamente iguais mas diferem no máximo de uma quantidade infinitesimal. A semelhança deste símbolo com o igual usual " $=$ " conduz naturalmente a de utilizar nos raciocínios seguintes as regras, aprendidas no ensino secundário, permitindo estabelecer equivalências entre igualdades, como as regras seguintes (a, b e c são números reais quaisquer):

- $a = b \Leftrightarrow a + c = b + c$
- $a = b \Leftrightarrow ac = bc$ (se $c \neq 0$)
- $ac = bc$ e $c \neq 0 \Rightarrow a = b$
- $ab = 0 \Leftrightarrow a = 0$ ou $b = 0$

É interessante colocar a seguinte questão: estas regras mantém-se verdadeiras substituindo " $=$ " por " $\simeq$ " (e " $\neq$ " por " $\not\simeq$ ") ? Vejamos um exemplo útil para responder a esta questão: se $\varepsilon \neq 0$ é infinitesimal, os números ε e ε^2 são infinitamente próximos (e são infinitamente próximos de zero) mas se os multiplicarmos pelo número infinitamente grande $1/\varepsilon$, obtêm-se dois números que não são mais infinitamente próximos:

$$\varepsilon \simeq \varepsilon^2 \text{ e } 1 = \varepsilon \times \frac{1}{\varepsilon} \not\simeq \varepsilon^2 \times \frac{1}{\varepsilon} = \varepsilon.$$

A segunda regra não é portanto sempre verdadeira se substituirmos " $=$ " por " $\simeq$ ". Tem-se no entanto a seguinte proposição:

Proposição 2.2

Para todos os números reais a,b e c, tem-se:

- $a \simeq b \Leftrightarrow a + c \simeq b + c$
- $a \simeq b$ *e c limitado* $\Rightarrow ac \simeq bc$
- $ac \simeq bc$ e $c \not\simeq 0 \Rightarrow a \simeq b$
- $ab \simeq 0 \Leftrightarrow a \simeq 0$ *ou* $b \simeq 0$

Dem.

A demonstração destas implicações ou equivalências é fácil mediante a observação seguinte: pode-se sempre substituir $a \simeq b$ por $(\exists \varepsilon \simeq 0,\ a = b + \varepsilon)$. Esta observação é *aliás uma regra de conduta que é vivamente recomendada a aplicar sistematicamente* nos raciocínios matemáticos onde o símbolo $\simeq$ aparece.

Mediante esta observação, as quatro regras aparecem simplesmente como aplicações da regras clássicas correspondentes para a igualdade (recordadas anteriormente) e as regras de Leibniz.

- $(a \simeq b)$ equivale a $(\exists \varepsilon \simeq 0,\ a = b + \varepsilon)$, portanto também a $(\exists \varepsilon \simeq 0,$ $a + c = b + c + \varepsilon)$ e portanto $(a + c \simeq b + c)$.
- $(a \simeq b$ e c limitado$)$ implica $(\exists \varepsilon \simeq 0,\ a = b + \varepsilon$ e c limitado$)$, logo $(\exists \varepsilon \simeq 0,\ ac = bc + \varepsilon c$ e c limitado$)$, seja $(\exists \varepsilon \simeq 0,\ ac = bc + \eta)$ e portanto $(ac \simeq bc)$.
- $(ac \simeq bc$ e $c \not\simeq 0)$ implica $(\exists \varepsilon \simeq 0,\ ac = bc + \varepsilon c$ e $c \not\simeq 0)$, o que pode escrever-se também $(\exists \varepsilon \simeq 0,\ ac = c(b + \frac{\varepsilon}{c})$ e $c \not\simeq 0)$, seja $(\exists \eta \simeq 0,\ a = b + \eta)$, logo $(a \simeq b)$.
- Se $(ab \simeq 0$ e a $\not\simeq 0)$ então $(\exists \varepsilon \simeq 0,\ ab = \varepsilon$ e $a \not\simeq 0)$ e portanto $(\exists \varepsilon \simeq 0, b = \frac{\varepsilon}{a}$ e $a \not\simeq 0)$, logo $(\exists \eta \not\simeq 0, b = \eta)$ e portanto $(b \simeq 0)$. □

2.6 A existência de uma sombra

Todas as propriedades de $\mathbb{R}$ que recordámos são também propriedades de $\mathbb{Q}$, que partilha com $\mathbb{R}$ o facto de ser um *corpo ordenado arquimediano*. Não será, portanto, surpresa estudar que $\mathbb{R}$ tem igualmente uma propriedade que $\mathbb{Q}$ não tem, e que faz o interesse e a especificidade daquele conjunto. Senão porque não fazer a análise em $\mathbb{Q}$?

Esta propriedade diz que $\mathbb{R}$ é um conjunto "completo" e será estudada com detalhe no capítulo sobre as sucessões: significa intuitivamente "que não há buracos". Será apresentada segundo a forma de um princípio, o terceiro.

Princípio 2.3 (Princípio da sombra)

Para todo o número real limitado x, existe um real standard do qual x é infinitamente próximo. Este real standard é único, representa-se por ^{o}x e chama-se a sombra de x ou a parte standard de x.

Observações:

- Este princípio permite "afinar " a ideia de que se pode construir a recta real: na sua parte limitada, todo o número ou é standard, ou é infinitamente próximo de um standard. Cada número standard é, portanto, rodeado (cercado) de um *halo* de números não-standard infinitamente próximos, todos estes *halos* justapõem-se uns aos outros para formar a parte limitada da recta real.

• A unicidade da sombra tem por consequência um resultado frequentemente utilizado, o *princípio de Carnot*, segundo o qual *dois reais standard infinitamente próximos são forçosamente iguais*[2].

• O princípio da sombra é falso em $\mathbb{Q}$: consideremos um número irracional x, por exemplo $x = \sqrt{2}$, e seja $\tilde{x}$ uma aproximação decimal de x tendo um número infinitamente grande de casas decimais. O número $\tilde{x}$ é um racional limitado (menor do que 2 por exemplo) que não pode ter sombra em $\mathbb{Q}$ visto que tem uma sombra única $^o\tilde{x} = \sqrt{2}$ que não pertence a $\mathbb{Q}$.

Proposição 2.4

Têm-se as seguintes regras: para todos os reais limitados x e y,

a) $^o(x+y) = {^o x} + {^o y}$;

b) $^o(xy) = {^o x}{^o y}$;

c) se $^o y \neq 0$, $^o(x/y) = {^o x}/{^o y}$;

d) $x < y \Rightarrow {^o x} \leq {^o y}$.

Dem.

Estas regras deduzem-se facilmente das regras de Leibniz e do princípio de Carnot:

a) Sejam x e y dois reais limitados. Existem α e β infinitesimais tais que $x = {^\circ x} + \alpha$ e $y = {^\circ y} + \beta$. Portanto $x + y = ({^\circ x} + \alpha) + ({^\circ y} + \beta) = ({^\circ x} + {^\circ y}) +$ $+ (\alpha + \beta)$. Como $\alpha + \beta$ é infinitesimal (regras de Leibniz), resulta que $^\circ(x + y)$ e $^\circ x + {^\circ y}$ são dois números standard infinitamente próximos, portanto são iguais.

b) Sejam x e y dois reais limitados. Existem α e β infinitesimais tais que $x = {^\circ x} + \alpha$ e $y = {^\circ y} + \beta$. Portanto $xy = ({^\circ x} + \alpha)({^\circ y} + \beta) = {^\circ x}{^\circ y} + (\alpha{^\circ y} + \beta{^\circ x} +$ $+ \alpha\beta)$. Como $\alpha{^\circ y}$, $\beta{^\circ x}$ e $\alpha\beta$ são infinitesimais, resulta então que $^\circ(xy)$ e $^\circ x{^\circ y}$ são dois números standard infinitamente próximos, portanto são iguais.

---

[2] Sejam a e b números reais standard. Se $x \simeq a$ e $x \simeq b$, então $a \simeq b$ (v. Observ. pág. 20) logo, $a = b$ (porque o único infinitesimal standard é zero). Deste modo, podemos concluir que dois números reais standard infinitamente próximos são iguais.

c) Depois das duas demonstrações anteriores, é suficiente mostrar a regra para $x = 1$. Seja $y \neq 0$ um real limitado. Existe β infinitesimal tal que $y = {}^\circ y + \beta$ com ${}^\circ y \neq 0$. Portanto,

$$\frac{1}{y} = \frac{1}{({}^\circ y + \beta)} = \frac{1}{{}^\circ y} - \frac{1}{{}^\circ y} + \frac{1}{{}^\circ y + \beta} = \frac{1}{{}^\circ y} - \frac{\beta}{{}^\circ y({}^\circ y + \beta)}.$$

Como ${}^\circ y({}^\circ y + \beta)$ é apreciável e β é infinitesimal, resulta que ${}^\circ(\frac{1}{y})$ e $\frac{1}{{}^\circ y}$ são dois números standard infinitamente próximos, portanto são iguais.

d) Sejam x e y dois reais limitados tais que $x < y$. Existem α e β infinitesimais tais que $x = {}^\circ x + \alpha$ e $y = {}^\circ y + \beta$. Portanto ${}^\circ x = x - \alpha < y - \alpha = {}^\circ y + \beta - \alpha$. Como $\beta - \alpha$ é infinitesimal, então sendo ${}^\circ x < {}^\circ y$, e ${}^\circ x$ e ${}^\circ y$ infinitamente próximas, portanto são iguais visto que são standard.

Exemplos:

• Mostrar que existem números racionais equivalentes a um dado número irracional.

Se x é um número irracional, $\varepsilon \neq 0$ um infinitesimal. Temos $x \simeq x + \varepsilon$, consideremos $x + \varepsilon$ o número cuja representação décimal é constituída pelos N primeiros algarismos do desenvolvimento décimal de x. Então $x + \varepsilon$ é um número racional.

• Existe um número x tal que a parte standard de $A_i(x)$ seja igual a 1 (um) quando:

$\diamond$ $A_1(x) = \frac{x+1}{x-1}$;

Temos $A_1(x) = \frac{x+1}{x-1} = 1 + \frac{2}{x-1} \simeq 1$ se $\frac{2}{x-1} \simeq 0$, logo $A_1(x) \simeq 1$ se x for infinitamente grande.

$\diamond$ $A_2(x) = 1 + x$;

Temos, $1 + x = 1 \Leftrightarrow x = 0$. Se $x = 0$, $A_2(x) = 1$.

$\diamond$ $A_3(x) = \frac{1}{x}$.

Verificamos que $\frac{1}{x} = 1 \Leftrightarrow x = 1$. Se $x = 1$ a parte standard de $A_3(x)$ é 1.

• Dois infinitesimais são sempre infinitamente próximos. Sejam ε e λ infinitesimais. Como $\varepsilon \simeq 0$ e como $\lambda \simeq 0$, então $\varepsilon \simeq 0 \simeq \lambda$ por simetria, logo $\varepsilon \simeq \lambda$ por tansitividade.

• Dois infinitamente grandes podem ser infinitamente próximos. Por exemplo, se w é um número infinitamente grande, então $\frac{1}{w} \simeq 0$, e podemos considerar os números infinitamente grandes w e $w + \frac{1}{w}$: $w \simeq w + \frac{1}{w}$, porque $\frac{1}{w} \simeq 0$.

• Sejam $x \leq x'$ dois números reais infinitamente próximos. Dizer quais os reais standard que podem existir em $[x, x']$ segundo os casos (x standard ou não, x' standard ou não).

Sejam x e x' standard, então $x = x'$ porque dois standard infinitamente próximos são iguais, logo o intervalo $[x, x']$ reduz-se a um ponto, tem portanto um número que é standard. Se x é standard e x' é não-standard então existe $\varepsilon \neq 0$ infinitesimal tal que $x' = x + \varepsilon$, logo $[x, x'] = [x, x + \varepsilon]$, o intervalo $[x, x']$ tem infinitos números reais mas só tem um número real standard que é x. Se x não é standard e x' é standard então existe $\varepsilon \neq 0$ infinitesimal tal que $x = x' - \varepsilon$, logo $[x, x'] = [x' - \varepsilon, x']$ que tem infinitos números reais mas só tem um número real standard que é x'.

Capítulo 3

Sucessões

Neste capítulo estudam-se as principais propriedades de números reais. O princípio da sombra tem um papel importante. Será igualmente a ocasião de encontrar o último dos quatro princípios, o princípio da transferência.

3.1 Noção de convergência

Definição: Chama-se *sucessão de números reais* a toda a aplicação de $\mathbb{N}$ em $\mathbb{R}$.

Sabe-se que existem essencialmente duas formas de definir uma sucessão: seja directamente como uma aplicação $u_n = f(n)$, por exemplo $u_n = \frac{n}{2n^2+3}$ ou $u_n = (-\frac{1}{2})^n$, ou por recorrência indicando o primeiro termo e a regra que permite calcular o termo de ordem n em função dos anteriores, por exemplo do termo de ordem $(n-1)$, como para

$$\begin{cases} u_0 = 1 \\ u_n = 2 + u_{n-1}. \end{cases}$$

Note-se que para calcular u_{100}, é necessário neste caso ter calculado u_1,u_2, u_3,..., o que é inútil visto que se pode conhecer a sucessão anterior como uma função explicita, ou seja $u_n = 2n + 1$.

Evidentemente, existem sucessões standard e sucessões não-standard; por exemplo $u_n = \frac{2n+3}{n+1}$ é uma sucessão standard mas, se $\varepsilon \neq 0$ é um infinitesimal, $u_n = \varepsilon^n$ ou $u_n = n + \frac{\varepsilon}{n}$ são sucessões não-standard.

Definição: Seja (u_n) uma sucessão standard e l um número standard. A sucessão (u_n) diz-se *convergente* de limite l se, para todo o n infinitamente grande, $u_n \simeq l$. Diz-se então que (u_n) tende para l ou que l é o limite da sucessão (u_n). Diz-se que a sucessão (u_n) tende para $+\infty$ (respectivamente $-\infty$) se, para todo o n infinitamente grande, u_n é infinitamente grande positivo (respectivamente infinitamente grande negativo).

Observações:

1. Quando uma sucessão (u_n) tende para l, diz-se que l é o limite da sucessão (u_n) (e representa-se $l = \lim\limits_{n\to\infty} u_n$) porque a sucessão só tem um limite: com efeito se para todo o n infinitamente grande se tem $u_n \simeq l$ e ao mesmo tempo $u_n \simeq l'$ então $l \simeq l'$; mas como l e l' são standard, $l = l'$ pelo princípio de Carnot.

2. Esta definição mostra que não se pode mudar a natureza da sucessão, isto é, o facto de uma sucessão ser convergente ou não, modificando um número finito dos seus termos de índice standard. Mais precisamente, duas sucessões standard que só diferem pelos seus valores em alguns índices standard (em número finito) são simultaneamente convergentes ou divergentes.

Exemplos:

1. A sucessão $u_n = \frac{2n+3}{n+1}$ tende para $l = 2$. Com efeito, para todo o $n > 0$, $u_n = \frac{2+\frac{3}{n}}{1+\frac{1}{n}}$ e portanto para n infinitamente grande, como $\frac{3}{n}$ e $\frac{1}{n}$ são infinitesimais, tem-se

$$^{\circ}(u_n) = {}^{\circ}\left(\frac{2+\frac{3}{n}}{1+\frac{1}{n}}\right) = \frac{^{\circ}(2+\frac{3}{n})}{^{\circ}(1+\frac{1}{n})} = \frac{2}{1} = 2.$$

2. A sucessão $u_n = n^3 - n^2$ tende para $+\infty$ porque, para n infinitamente grande $n^3 - n^2 = n^2(n-1)$ e o produto de dois inteiros infinitamente grandes positivos é um número infinitamente grande positivo.

3. A sucessão $u_n = (-1)^n$ não é convergente. Mais geral, se (u_n) e (v_n) são duas sucessões convergentes não tendo o mesmo limite então a sucessão (w_n) que é igual a u_n quando n é par e a v_n quando n é impar é uma sucessão divergente, isto é, não é convergente.

Definição: Um real standard l chama-se *ponto aderente* de uma sucessão standard (u_n) se existe um n infinitamente grande tal que $u_n \simeq l$.

Evidentemente que uma sucessão que tem vários pontos de aderência diferentes (eles não podem ser infinitamente próximos visto que são standard, a menos que sejam iguais), é necessariamente divergente.

3.2 Sucessões limitadas e o princípio de transferência

Definição: Diz-se que uma sucessão (u_n) é *majorada* (respectivamente *minorada*) se existe um número M (respectivamente um número m) tal que, para todo o n $u_n < M$ (respectivamente $m < u_n$). Uma sucessão que é majorada e minorada diz-se *limitada.*

Exemplos: A sucessão $u_n = \frac{1}{n}$ é majorada por 1 e minorada por 0, e portanto é limitada. Do mesmo modo, a sucessão $u_n = \frac{2n+3}{n+1}$ é majorada por 5 (por exemplo) e minorada por 0, portanto é limitada.

Pretendemos presentemente mostrar que toda a sucessão convergente é limitada (a recíproca é falsa) . Para isso vamos introduzir o último princípio, chamado princípio de transferência.

Este princípio tem duas formulações, logicamente equivalentes.

Princípio 3.1 (princípio de transferência)

- Primeira formulação: *se existir um objecto matemático (número, função, sucessão, conjunto, ...) verificando uma dada propriedade standard,*[3] *então existe um objecto standard verificando esta propriedade. Em particular, se aquele objecto é único, então ele é necessariamente standard.*

Este princípio pode enunciar-se também da seguinte forma: *se $P(x)$ é uma propriedade standard e existe x que satisfaz $P(x)$, então existe um x standard que satisfaz $P(x)$.*

- Segunda formulação: *para que uma propriedade standard seja verdadeira para todo o x é suficiente que seja verdadeira para todo o x standard.*

Este princípio pode enunciar-se também da seguinte forma: *seja $P(x)$ uma propriedade standard; se para todo o x standard, $P(x)$ é verdadeira, então para todo o x, $P(x)$ é verdadeira.*

Proposição 3.2

Toda a sucessão standard majorada é majorada por um standard.

[3] Se $P(x)$ é uma propriedade bem definida, sem o uso do adjectivo "standard" (nem dos conceitos derivados) a partir de objectos standard, então $P(x)$ é uma propriedade standard.

Dem.

Por hipótese existe um M tal que a propriedade "M é majorante da sucessão (u_n)" seja satisfeita. Esta é uma propriedade standard; portanto existe um M standard que verifica esta propriedade. □

Proposição 3.3

Toda a sucessão convergente é limitada.

Dem.

Por transferência, é suficiente verificar esta proposição para sucessões standard. Seja (u_n) uma sucessão standard convergente de limite l, limite este que é necessariamente standard, pelo princípio da transferência (primeira formulação). Então todos os termos de índice infinitamente grande são infinitamente próximos de l. Seja w um inteiro infinitamente grande e seja $k = \min_{n \leq w} u_n$ e $K = \max_{n \leq w} u_n$. Então $m = \min(k, l-1)$ e $M = \max(K, l+1)$ são, respectivamente, minorante e majorante para (u_n). □

3.3 Sucessões de Cauchy

Se quisermos calcular um limite através de um programa de "computador", é fácil escrever um programa que calcula u_0, u_1, u_2, ... e a questão à qual devemos responder é então a seguinte: "quando parar o cálculo?". Também na prática, considera-se frequentemente que se as primeiras casas decimais não evoluem mais após numerosas iterações, estas casas decimais são casas decimais exactas do limite, se este limite existir. Tal é um pouco a ideia das sucessões de Cauchy. Veremos que esta noção permite falar da convergência de uma sucessão sem fazer referência ao seu limite (o que se pode revelar bem útil quando não o conhecemos).

Definição: Diz-se que uma sucessão standard (u_n) *é de Cauchy* sse, para todo o p e q infinitamente grandes, $u_p \simeq u_q$.

Teorema 3.4

Em $\mathbb{R}$, toda a sucessão de Cauchy é convergente (a reciproca é verdadeira[4]*): diz-se que $\mathbb{R}$ é topologicamente completo. Por outro lado, o conjunto $\mathbb{Q}$ não é completo, isto é, existem sucessões de números racionais que são de Cauchy e que não convergem em $\mathbb{Q}$.*

Dem.

É suficiente, por transferência, demonstrar o teorema para as sucessões standard. Por hipótese, todos os termos da sucessão de índice infinitamente grande são infinitamente próximos uns dos outros. É portanto suficiente provar que são limitados, isto é, infinitamente próximos de um standard único. É claro que existe um índice N tal que

$$\forall n > N \ |u_n - u_N| < 1.$$

Com efeito é suficiente considerar N infinitamente grande. Pelo princípio de transferência, existe um N standard para o qual esta propriedade contínua verdadeira. Como a sucessão (u_n) é uma sucessão standard, os seus termos de índice standard são portanto limitados. Portanto u_N é um número standard e a propriedade anterior implica que para além do índice N todos os termos se mantém a uma distância deste inferior a 1, e portanto são limitados. A sua sombra comum é o limite procurado.

Contra exemplo em $\mathbb{Q}$: escolhe-se uma sucessão de números racionais que converge em $\mathbb{R}$ para um irracional (por exemplo a sucessão das aproximações decimais de $\sqrt{2}$). É uma sucessão de Cauchy visto que converge em $\mathbb{R}$ mas não converge em $\mathbb{Q}$ porque $\sqrt{2} \notin \mathbb{Q}$. □

Definição: Uma sucessão diz-se "*contractante*" sse existe $k \in]0,1[$ tal que $|u_{n+1} - u_n| < k|u_n - u_{n-1}|$ para todo o $n \geq 1$.

Exemplos:

- Uma progressão aritmética cujo primeiro termo é u_0 e de razão r, $u_n = u_0 + nr$, não é uma sucessão "*contractante*". A sucessão $u_n = n^2$ também não.

---

[4] Seja (u_n) standard convergente para a (necessariamente standard também). Queremos provar que (u_n) é de Cauchy, isto é, para quaisquer p e q infinitamente grandes, $u_p \simeq u_q$. Sejam p e q infinitamente grandes quaisquer. Então

$$|u_p - u_q| = |u_p - a + a - u_q| \leq |u_p - a| + |u_q + a| \simeq 0. \qquad \square$$

• Uma progressão geométrica cujo primeiro termo é u_0 e de razão q, $u_n = u_0 q^n$, é *"contractante"* de razão $k = q$ desde que $0 < q < 1$ e não *"contractante"* no caso contrário. Recorde a fórmula, para $q \neq 1$,

$$1 + q + q^2 + \cdots + q^n = \frac{1 - q^{n+1}}{1 - q}.$$

Proposição 3.5

Toda a sucessão standard "contractante" é convergente.

Dem.

Vamos mostrar que toda a sucessão "contractante" é de Cauchy. Para todo p, q infinitamente grandes, sem perda de generalidade, $(p > q)$, tem-se:

$$\begin{aligned}
\left|u_p - u_q\right| &\leq \left|u_p - u_{p-1}\right| + ... + \left|u_{q+1} - u_q\right| \\
&\leq (k^{p-q-1} + k^{p-q-2} + ... + k + 1)\left|u_{q+1} - u_q\right| \\
&\leq (k^{p-q-1} + k^{p-q-2} + ... + k + 1)k^q\left|u_1 - u_0\right| \\
&\leq \tfrac{1-k^{p-q}}{1-k} k^q \left|u_1 - u_0\right| \\
&\leq \tfrac{k^q - k^p}{1-k} \left|u_1 - u_0\right| \\
&\simeq 0
\end{aligned}$$

porque o último factor $(\left|u_1 - u_0\right|)$ é standard, e que, visto que $0 < k < 1$, $k^p \simeq 0$ e $k^q \simeq 0$, logo $\frac{k^q - k^p}{1-k} \simeq 0$, porque p e q são infinitamente grandes, por hipótese. □

3.4 Sucessões monótonas e limites superiores

O resultado fundamental deste parágrafo é o facto de que toda a sucessão monótona e limitada é convergente. Mas antes de mostrar este teorema, que é uma consequência do princípio da sombra, vamos introduzir as noções de supremo e de ínfimo de um conjunto.

Definição:

Seja A uma parte de $\mathbb{R}$. Um real s chama-se *supremo* de A (respectivamente *ínfimo*) se ele é o menor dos majorantes de A (respectivamente o maior dos minorantes de A). Representa-se $s = \sup\{x : x \in A\}$ ou simplesmente $s = sup\ A$ (resp. $s = \inf\{x : \ x \in A\}$ ou simplesmente $s = \inf A$).

É claro que, se A tem um supremo (ou ínfimo), ele é único e é standard pelo princípio de tranferência.

Definição: Seja N um inteiro infinitamente grande. Chama-se à sucessão $(x_n)_{n=0,...,N}$ uma *partição fina* do intervalo $[a, b]$ se é uma sucessão finita de números $a = x_0 < x_1 < ... < x_N = b$, tais que $x_i \simeq x_{i+1}$ para todo $i = 0, ..., N - 1$.

A demonstração do teorema seguinte utiliza a definição de partição fina de um intervalo.

Teorema 3.6

Toda a parte majorada não vazia em $\mathbb{R}$ tem um supremo. Toda a parte minorada não vazia em $\mathbb{R}$ tem um ínfimo.

Dem.

Seja A uma parte majorada não vazia de $\mathbb{R}$. Por transferência, basta fazer a demonstração para A standard. Supondo A standard e não vazio, pelo princípio da definição, o conjunto B dos majorantes de A também é standard e não vazio. Podemos portanto escolher um a standard pertencendo a A e um b standard pertencente a B. Seja N um natural infinitamente grande e seja $(x_n)_{n=0,...,N}$ uma partição fina do intervalo $[a, b]$. Representa-se por $\overline{x}$ o maior dos x_i tais que $[x_{i-1}, x_i]$ intersecta A: $\overline{x} = \max\limits_{i=1,...,N} \{x_i : [x_{i-1}, x_i] \cap A \neq \emptyset\}$. Como $\overline{x}$ pertence ao intervalo standard $[a, b]$, é limitado e portanto tem uma sombra $\overline{x}_0$. Verifiquemos que $\overline{x}_0$ é o supremo procurado. Tem-se, por construção, $x \leq \overline{x}$ para todo o $x \in A$ portanto $x \leq \overline{x}_0$ para todo o x pertencente a A. Portanto para todo o $x \in A$ standard, tem-se $x \leq \overline{x}_0$. Por transferência, $\overline{x}_0$ é portanto um majorante de A. É o menor dos majorantes porque por construção, $\overline{x} \in B$, portanto $\overline{x}_0 \leq x$ para todo $x \in B$ standard, onde, por transferência $\overline{x}_0 \leq x$ para todo o $x \in B$; é portanto o supremo.

A segunda parte do teorema deduz-se facilmente da primeira: se A é uma parte minorada de $\mathbb{R}$, $-A := \{-x : x \in A\}$ é uma parte majorada que tem, portanto, um supremo, digamos a, e $-a$ é o ínfimo de A. □

Observação:

Observamos que uma parte qualquer de $\mathbb{R}$ não tem necessariamente supremo: por exemplo $\mathbb{R}^+$ ou $\mathbb{N}$ não têm qualquer majorante, portanto, em particular supremo.

Notemos, também, que o supremo de uma parte de $\mathbb{R}$, quando existe, não pertence forçosamente a essa parte: por exemplo o número s é o supremo do intervalo $A =]-\infty, s[$ e não pertence a A.

Corolário 3.7

Toda a sucessão crescente majorada é convergente.

Toda a sucessão decrescente minorada é convergente.

Dem.

Por transferência podemos supor a sucessão (u_n) standard. Supomo-la crescente e majorada. Por hipótese, o conjunto $\{u_n : n \in \mathbb{N}\}$ é uma parte não vazia majorada de $\mathbb{R}$. Esta parte tem portanto um supremo l. Suponhamos por absurdo que existe um índice N infinitamente grande tal que $u_N < l$. Os dois números u_N e l não são infinitamente próximos, logo podemos encontrar um número standard k tal que $u_N < k < l$. O número k majora todos os termos da sucessão de índice standard visto que a sucessão é crescente. Por transferência, ele majora todos os termos da sucessão. Ora isto é absurdo visto que $k < l$ e l é o menor dos majorantes. □

Corolário 3.8

Uma sucessão crescente ou é convergente ou tende para $+\infty$.

Dem.

Seja (u_n) uma sucessão crescente, que podemos supor standard por transferência. Então ou u_n é infinitamente grande para todo o n infinitamente grande, e neste caso tende para infinito, ou existe um índice N infinitamente grande para o qual u_N é limitado. Neste caso pomos $l = {}^{o}(u_N)$. Como a sucessão é crescente, l é um majorante para todos os termos de índice standard da sucessão (porque estes são menores que u_N) portanto, por transferência, para todos os termos da sucessão. A sucessão é portanto majorada. Pelo teorema anterior a sucessão é convergente. □

Capítulo 4

Derivadas

Neste capítulo, consideram-se funções standard reais de variável real, $x \mapsto f(x)$, cujo domínio de definição $\mathcal{D}$ é um aberto de $\mathbb{R}$, isto é, uma parte tendo a seguinte propriedade: se $x \in \mathcal{D}$ e se $y \simeq x$ então $y \in \mathcal{D}$. Escolhemos definir a noção de função de classe C^1 e não a de função derivável porque quase todas as funções que encontramos e que são deriváveis são de classe C^1, e sobretudo porque esta noção é a noção mais operacional nos cálculos e nas aplicações.

4.1. Funções de classe C^1

Definição: Um real x diz-se *quase standard* num conjunto $\mathcal{D}$ sse tem uma sombra pertencente a $\mathcal{D}$. Se $\mathcal{D} = \mathbb{R}$, os pontos quase standard em $\mathcal{D}$ são simplesmente os pontos limitados (pelo princípio da sombra); se $\mathcal{D} \ = \ ]a, b[$, com a e b standard, os pontos quase standard em $\mathcal{D}$ são todos os pontos de $\mathcal{D}$ não infinitamente próximos de uma «extremidade».

Definição: Sejam f e f' duas funções standard definidas sobre um aberto standard $\mathcal{D}$ de $\mathbb{R}$. Diz-se que *f é de classe C^1 com derivada f'* sse para todo o x quase standard em $\mathcal{D}$ e todo $y \simeq x$, $y \neq x$,

$$\frac{f(y) - f(x)}{y - x} \simeq f'(x).$$

Observação: É útil assegurar que uma função f de classe C^1 não pode ter duas derivadas diferentes. Com efeito, se f' e $\widehat{f}'$ são derivadas de f, serão por definição infinitamente próximas, isto é $\forall\ x \in \mathcal{D}(f'(x) \simeq \widehat{f}'(x))$. Mas estas duas funções são standard, tomam valores standard nos pontos standard, portanto, pelo princípio de Carnot, devem coincidir em todos os pontos standard. Mas duas funções standard iguais nos pontos standard são necessariamente iguais, em todo o ponto standard ou não-standard, por transferência.

Observação (Notações diferenciais): As notações seguintes, introduzidas por Leibniz, são frequentemente cómodas. Designa-se por dx um número infinitamente pequeno não nulo (representando um pequeno crescimento da variável x) [Esta noção pode ter outros significados, conforme o contexto (por exemplo a de uma medida ou a de uma forma diferencial), mas neste capítulo não existe nenhum inconveniente em considerar dx como um número infinitamente pequeno representando um crescimento da variável x. Notemos que esta diversidade de sentido, que pode representar o uso desta notação delicada, faz também a sua riqueza]. Quando f é de classe C^1, com derivada f' tem-se a relação

$$\frac{f(x+dx)-f(x)}{dx} \simeq f'(x).$$

Se designarmos por $\triangle f$ a diferença $\triangle f = f(x+dx) - f(x)$, crescimento de f correspondendo a um crescimento dx da variável x, esta relação pode também escrever-se

$$\triangle f = f'(x)dx + \varepsilon dx$$

onde ε é um número infinitesimal. O diferencial de f, representado por df é, por definição, a quantidade $df = f'(x)dx$. Esta quantidade, tanto como $\triangle f$, depende não apenas da função f mas também do ponto x e do infinitesimal dx. Deverá precisar-se $df(x, dx)$. As duas quantidades $\triangle f$ e df não são iguais em geral (salvo se f é linear, por exemplo) como o mostra a figura 4.1. mas são muito próximas visto que, para que f seja de classe C^1, é necessário e suficiente que exista $\varepsilon \simeq 0$ tal que $\triangle f - df = \varepsilon dx$.

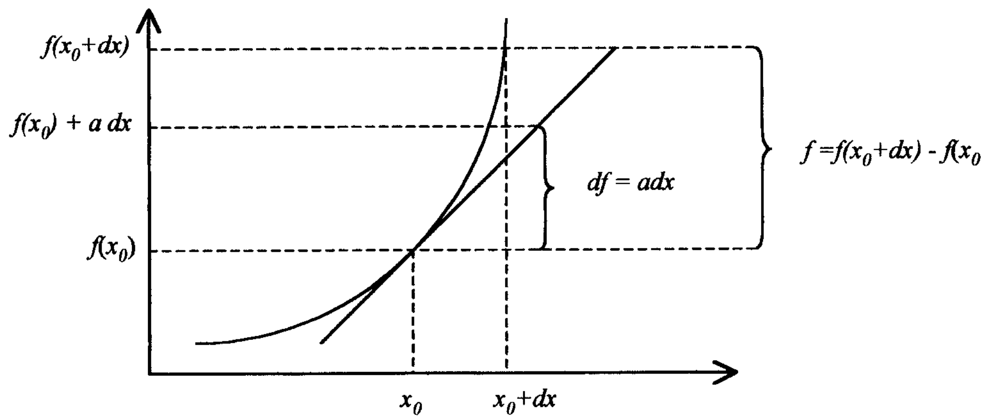


Figura 4.1: Notação diferencial: designa-se por $\triangle f = f(x_0 + dx) - f(x_0)$ o crescimento dx da variável x no ponto x_0 e representa-se por $df = adx$ o crescimento da função linear $y = ax$ tangente a f no ponto x_0 correspondente ao mesmo crescimento dx da variável. A fórmula anterior pode recriar-se $\triangle f = df + \varepsilon dx$, o que significa que o afastamento entre $\triangle f$ e df deve não só ser infinitamente pequeno mas mais precisamente infinitamente pequeno perante dx.

Exemplo:

• Mostremos que a função $f(x) = x^3$ é de classe C^1, de derivada $f'(x) = 3x^2$. Sejam x quase standard e $dx \simeq 0$, $dx \neq 0$. Tem-se, utilizando as regras de Leibniz (xdx e dx^2 são infinitamente pequenos):

$$\frac{(x+dx)^3 - x^3}{dx} = \frac{x^3 + 3x^2dx + 3x(dx)^2 + (dx)^3 - x^3}{dx} = 3x^2 + 3xdx + (dx)^2 \simeq 3x^2.$$

• Mostremos que a função $f(x) = \frac{1}{x}$ é de classe C^1, de derivada $f'(x) = -\frac{1}{x^2}$. Sejam x quase standard em $\mathbb{R}\backslash\{0\}$ (isto é x é apreciável) e $dx \simeq 0$, $dx \neq 0$. Tem-se:

$$\frac{\frac{1}{x+dx} - \frac{1}{x}}{dx} = \frac{\frac{x-x-dx}{x(x+dx)}}{dx} = \frac{-dx}{xdx(x+dx)} = \frac{-1}{x(x+dx)} \simeq -\frac{1}{x^2}.$$

• Mostremos que a função $f(x) = \sqrt{x}$ é de classe C^1, de derivada $f'(x) = \frac{1}{2\sqrt{x}}$. Sejam x quase standard em $]0, +\infty[$ e $dx \simeq 0$, $dx \neq 0$. Tem-se:

$$\begin{aligned}\frac{f(x+dx) - f(x)}{dx} &= \frac{\sqrt{x+dx} - \sqrt{x}}{dx} = \frac{x + dx - x}{dx(\sqrt{x+dx} + \sqrt{x})} \\ &= \frac{dx}{dx(\sqrt{x+dx} + \sqrt{x})} = \frac{1}{(\sqrt{x+dx} + \sqrt{x})} \simeq \frac{1}{2\sqrt{x}}.\end{aligned}$$

Definição: Uma função f diz-se S-contínua sse para todo o x, y no seu domínio,

$$x \simeq y \Rightarrow f(x) \simeq f(y).$$

Observação: Notar-se-á que toda a função standard f de classe C^1 verifica a seguinte propriedade, propriedade de *S-continuidade*,

$$x \simeq y \Rightarrow f(x) \simeq f(y)$$

em todos os pontos x e y quase standard do conjunto aberto $\mathcal{D}$ (porque se tem derivada é contínua, logo, como é standard, é S-contínua nos pontos quase-standard).

Com a definição de uma função de classe C^1, não é difícil mostrar as fórmulas de derivação usuais para somas, produtos e quocientes, que são muito úteis nos cálculos:

Proposição 4.1

Sejam f e g duas funções de classe C^1, de derivadas f' e g' respectivamente, e seja $\lambda \in \mathbb{R}$. Então:

- *a soma $f+g$ é de classe C^1, de derivada $(f+g)' = f' + g'$;*
- *o produto λf é de classe C^1, de derivada $(\lambda f)' = \lambda f'$;*
- *o produto fg é de classe C^1, de derivada $(fg)' = f'g + fg'$;*
- *se g não se anula, o quociente $\frac{f}{g}$ é de classe C^1, de derivada $\left(\frac{f}{g}\right)' = \frac{f'g - fg'}{g^2}$.*

Dem.

Por transferência, podemos supor que f, g e λ (e portanto $f+g$, λf, fg e $\frac{f}{g}$ (com $g \neq 0$)) são standard, $\lambda \in \mathbb{R}$. Podemos aplicar a definição anterior, seja x um ponto quase standard de $\mathcal{D}$ e seja $dx \simeq 0$. Tem-se por hipótese:

$$f(x+dx) - f(x) = f'(x)dx + \varepsilon dx, \text{ com } \varepsilon \simeq 0$$

e

$$g(x+dx) - g(x) = g'(x)dx + \overline{\varepsilon}\, dx, \text{ com } \overline{\varepsilon} \simeq 0.$$

Tem-se portanto:

$$\bullet\, (f+g)(x+dx) - (f+g)(x) \begin{aligned}[t] &= f(x+dx) + g(x+dx) - f(x) - g(x) \\ &= (f(x+dx) - f(x)) + (g(x+dx) - g(x)) \\ &= f'(x)dx + \varepsilon dx + g'(x)dx + \overline{\varepsilon}\, dx \\ &= (f'(x) + g'(x))dx + \widetilde{\varepsilon}\, dx \end{aligned}$$

com $\widetilde{\varepsilon} \simeq 0$. Logo $(f+g)' = f' + g'$.

$$\bullet\, (\lambda f)(x+dx) - (\lambda f)(x) \begin{aligned}[t] &= \lambda f(x+dx) - \lambda f(x) = \lambda(f(x+dx) - f(x)) \\ &= \lambda(f'(x)dx + \varepsilon dx) = \lambda f'(x)dx + \widehat{\varepsilon} dx \end{aligned}$$

com $\widehat{\varepsilon} \simeq 0$. Desde modo $(\lambda f)' = \lambda f'$ com λ constante.

$$\bullet\, (fg)(x+dx) - (fg)(x) \begin{aligned}[t] &= f(x+dx)g(x+dx) - f(x)g(x) \\ &= (f(x) + f'(x)dx + \varepsilon dx)(g(x) + g'(x)dx + \overline{\varepsilon}\, dx) + \\ &- f(x)g(x) = (f(x)g'(x) + f'(x)g(x))dx + \varepsilon^* dx \end{aligned}$$

com $\varepsilon^* \simeq 0$. Assim podemos concluir que $(fg)' = f'g + fg'$.

• Se g nunca se anula, temos

$$\begin{aligned}
\frac{f}{g}(x+dx)-\frac{f}{g}(x) &= \frac{f(x+dx)}{g(x+dx)}-\frac{f(x)}{g(x)}=\frac{f(x+dx)g(x)-f(x)g(x+dx)}{g(x)g(x+dx)}\\
&= \frac{(f(x)+f'(x)dx+\varepsilon dx)g(x)-f(x)(g(x)+g'(x)dx+\overline{\varepsilon}\,dx)}{g(x)(g(x)+g'(x)dx+\overline{\varepsilon}\,dx)}\\
&= \frac{g(x)f'(x)dx+\varepsilon g(x)dx-f(x)g'(x)dx-\overline{\varepsilon}\,f(x)dx}{(g(x))^2+g(x)g'(x)dx+g(x)\overline{\varepsilon}\,dx}\\
&= \frac{(g(x)f'(x)-f(x)g'(x))dx}{(g(x))^2}+\varepsilon' dx
\end{aligned}$$

com $\varepsilon' \simeq 0$. Logo podemos concluir que $\left(\frac{f}{g}\right)' = \frac{f'g-fg'}{g^2}$. □

Teorema 4.2 (Derivação das funções compostas (regra da cadeia))

Sejam $\mathcal{D}$ e $\mathcal{E}$ dois conjuntos abertos de $\mathbb{R}$ e sejam $f:\mathcal{D}\longrightarrow\mathcal{E}$ e $g:\mathcal{E}\longrightarrow\mathbb{R}$ duas funções de classe C^1, de derivadas f' e g' respectivamente. Então a função composta $g\circ f$ é de classe C^1, de derivada $(g\circ f)'(x)=g(f(x))f'(x)$.

Observação: Em notação diferencial, esta fórmula é quase evidente visto ter-se $d(g\circ f)=dg(y)=g'(y)dy=g'(y)f'(x)dx$ e $d(g\circ f)=(g\circ f)'(x)dx$.

Dem.

O teorema da derivação das funções compostas é um teorema standard que podemos portanto provar por transferência no caso em que as funções f e g são standard. Consideremos um ponto x quase standard em $\mathcal{D}$ e a sua imagem $y=f(x)$. O ponto y é quase standard em $\mathcal{E}$ porque a função f verifica a propriedade de S-continuidade e em virtude do facto de que a imagem por uma função standard de um ponto standard é standard (transferência). Por hipótese, temos, para todo o dx e dy infinitesimais,

$$f(x+dx)-f(x)=f'(x)dx+\varepsilon dx,\ \text{com}\ \varepsilon\simeq 0$$

e
$$g(y+dy)-g(y)=g'(y)dy+\overline{\varepsilon}dy,\ \text{com}\ \overline{\varepsilon}\simeq 0.$$

Utilizando sucessivamente a primeira hipótese, depois a segunda com $dy=(f'(x)+\varepsilon)dx$ (que é um número infinitesimal), temos:

$$\begin{aligned}
g(f(x+dx))-g(f(x)) &= g(f(x)+f'(x)dx+\varepsilon dx)-g(f(x))\\
&= g(f(x))+g'(f(x))(f'(x)+\varepsilon)dx+\\
&+\overline{\varepsilon}(f'(x)+\varepsilon)dx-g(f(x))=g'(f(x))f'(x)dx+\widetilde{\varepsilon}\,dx
\end{aligned}$$

onde $\tilde{\varepsilon} = g'(f(x))\varepsilon + \overline{\varepsilon}(f'(x) + \varepsilon)$ é infinitesimal. □

4.2 Inversa de uma função

Definição: Uma função $f : I = [a, b] \to J = f([a, b])$ é *invertível* se existe uma função $g : J \to I$ tal que $g \circ f = Id_I$ e $f \circ g = Id_J$, onde $Id_{\mathcal{D}}$ ($\mathcal{D} = I$ ou J) é a função *identidade*, isto é a função definida sobre $\mathcal{D}$ tal que $Id_{\mathcal{D}}(x) = x$.

Exemplo: A função ln é invertível de função inversa exp; a função Id é invertível e igual à sua função inversa; a função $f(x) = \frac{1}{2x+3}$ é invertível no seu domínio e tem por inversa a função $f(x) = \frac{1-3x}{2x}$.

Proposição 4.3

Uma função é invertível sse é bijectiva (isto é, injectiva e sobrejectiva).

Dem.

Suponhamos que $f : I \to J$ é invertível, e seja $g : J \to I$ uma inversa. Então $g \circ f = Id_I$ e $f \circ g = Id_J$. Mas Id_I é injectiva, consideremos x, $x' \in I$ arbitrários, tais que $f(x) = f(x')$. Então $g(f(x)) = g(f(x'))$ ou seja $g \circ f(x) = g \circ f(x')$, donde $x = x'$ por $g \circ f$ ser injectiva. Conclui-se então que f é injectiva. E Id_J é sobrejectiva, suponhamos que $f \circ g$ é sobrejectiva. Dado $y \in J$ arbitrário, existe $z \in J$ tal que $y = (f \circ g)\ (z) = f(g(z))$, mas $g(z) \in I$. Quer dizer, para todo o $y \in J$ existe $x = g(z) \in I$ tal que $f(x) = y$, o que mostra que f é sobrejectiva.

Reciprocamente, se f é bijectiva então para todo o $y \in J$ existe, por sobrejectividade um $x \in I$ tal que $y = f(x)$ e o número x é único, por injectividade; isto permite portanto associar a todo o y um x isto de forma única e portanto de definir uma função g (sendo $g(y) := x$) que, por construção é a função inversa procurada. □

Observação: Se $G_f = \{(x, f(x)) : x \in [a, b]\}$ é o gráfico de f e $G_g = \{(x, g(x)) : x \in f([a, b])\}$ é o de g, onde f e g são inversas uma da outra, então G_f e G_g são duas curvas simétricas em relação à recta de equação $y = x$.

Exemplo: Inversas das funções trigonométricas

As funções trigonométricas não são invertíveis em $\mathbb{R}$ mas são invertíveis, pelo teorema precedente, sobre todo o intervalo $[a, b]$ onde são definidas e monótonas. Por convenção, dá-se um nome às inversas das funções sen, cos, tg, cotg sobre os intervalos $\left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$, $[0,\pi]$, $\left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$ e $[0,\pi]$ respectivamente. Estas funções chamam-se arcoseno, arcocoseno, arcotangente e arcocotangente e são respectivamente definidas nos intervalos $[-1,1]$, $[-1,1]$, $]-\infty, +\infty[$ e $]-\infty, +\infty[$. Pela observação interior, obtêm-se os gráficos destas funções por simples simetria por razão à primeira bissectriz.

Por exemplo:

$$\operatorname{arcsen}(-1) = \tfrac{3}{2}\pi, \quad \arccos\left(\tfrac{\sqrt{2}}{2}\right) = \tfrac{\pi}{4}, \quad \operatorname{arctg}(-\sqrt{3}) = \tfrac{2}{3}\pi.$$

Proposição 4.4

Seja $f : I \to I$ uma função invertivel, de função inversa f^{-1}. Se f é de classe C^1 e se $f'(x) \neq 0$ sobre I então f^{-1} é de classe C^1 e tem-se, para todo $y \in I$, imagem por f de x, $\left(f^{-1}\right)'(y) = \frac{1}{f'(x)}$.

Dem.

Se a função f^{-1} é de classe C^1, obtemos a fórmula imediatamente a partir do teorema 4.2. derivando a igualdade $f^{-1}(f(x)) = x$. Resta portanto mostrar que f^{-1} é de classe C^1.

Podemos supor, por transferência, que f (e portanto f^{-1}) são standard. Temos por hipótese, para todo o y quase standard em I e todo o $dx \simeq 0$:

$$f\left(f^{-1}(y) + dx\right) - f\left(f^{-1}(y)\right) = f'\left(f^{-1}(y)\right)dx + \varepsilon dx.$$

Em particular porque o número $dx = f^{-1}(y+dy) - f^{-1}(y)$ que é infinitesimal desde que dy seja infinitesimal, visto que f^{-1} é S-contínua. Onde,

$$f\left(f^{-1}(y+dy)\right) - f\left(f^{-1}(y)\right) = \left(f'\left(f^{-1}(y)\right)\right) + \varepsilon\big)\left(f^{-1}(y+dy) - f^{-1}(y)\right)$$

e portanto

$$dy = \left(f'(x) + \varepsilon\right)\left(f^{-1}(y+dy) - f^{-1}(y)\right).$$

Se $f'(x) \neq 0$, temos também $f'(x) + \varepsilon \neq 0$ porque $f'(x)$ é standard; portanto

$$f^{-1}(y+dy) - f^{-1}(y) = \frac{dy}{f'(x)+\varepsilon} = \frac{dy}{f'(x)} + \overline{\varepsilon}\, dy. \qquad \square$$

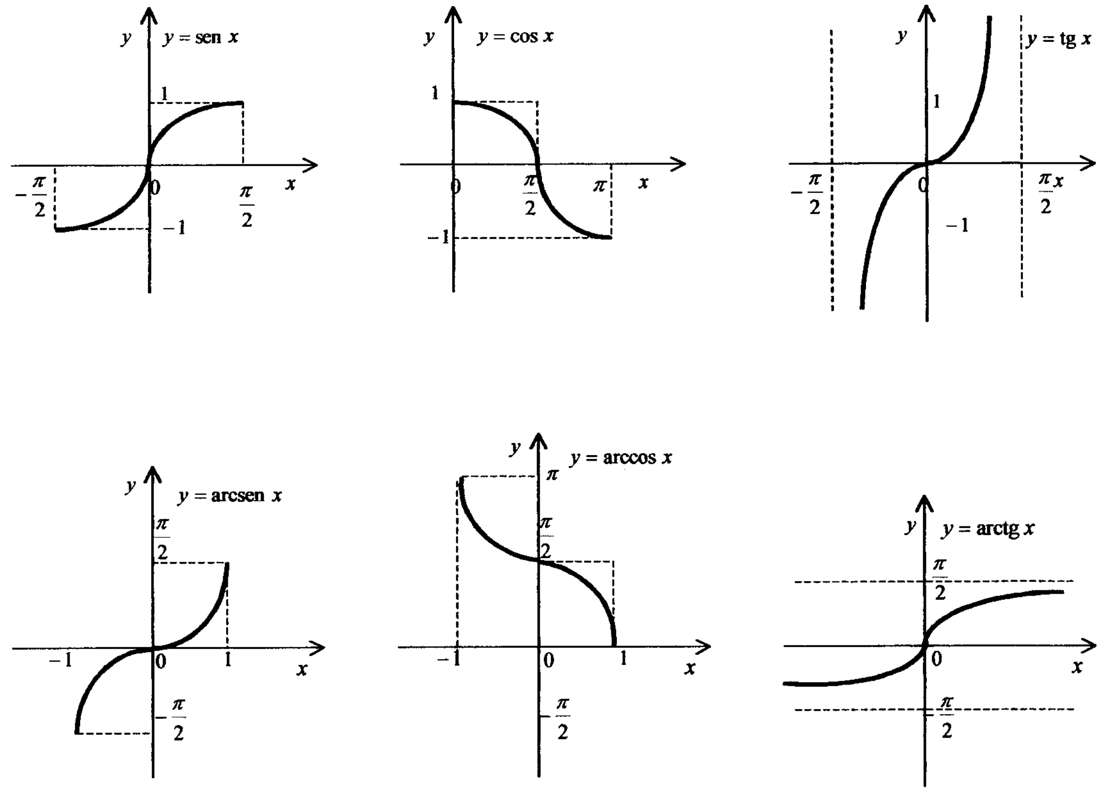


Figura 4.2: As funções trigonométricas convenientemente restringidas e as suas inversas.

Exemplos: Derivadas das funções inversas das funções trigonométricas: deduzem-se da proposição anterior as seguintes fórmulas:

$$(\text{arcsen})'(x) = \frac{1}{\sqrt{1-x^2}} \quad , \quad (\text{arccos})'(x) = -\frac{1}{\sqrt{1-x^2}}$$

$$(\text{arctg})'(x) = \frac{1}{1+x^2} \quad , \quad (\text{arccotg})'(x) = -\frac{1}{1+x^2}\ .$$

4.3 Observar uma função à lupa

Na medida em que a derivada é uma noção que se presta particularmente bem aos cálculos, tem-se frequentemente tendência de esquecer a ideia geométrica fundamental subjacente: esta ideia é que uma função de classe C^1 é uma função "*bem aproximada*" em cada ponto por uma função linear. O objecto deste parágrafo é precisar o que se entende por "*bem aproximada*".

Definição: Seja $(x_0, y_0) \in \mathbb{R}^2$ e seja $\varepsilon \simeq 0$ positivo. A uma homotetia da forma $(x, y) \mapsto \left(X = \frac{x-x_0}{\varepsilon}, Y = \frac{y-y_0}{\varepsilon}\right)$ chama-se lupa de centro (x_0, y_0) e de ampliação $\frac{1}{\varepsilon}$(ou também lupa-ε de centro (x_0, y_0)).

Definição: Se $x \mapsto f(x)$ é uma função tal que $f(x_0) = y_0$, chama-se *imagem de f* sob a lupa de centro $(x_0, f(x_0))$ e de ampliação $\frac{1}{\varepsilon}$, a função $X \mapsto Y = F(X)$ definida por $F(X) = \frac{f(x_0+\varepsilon X)-f(x_0)}{\varepsilon}$.

Observação: Notar-se-à que mesmo quando f é uma função standard, a sua imagem sobre uma lupa-ε é geralmente não-standard. Mas o que nos interessa, quando observamos uma função f através de uma lupa, não é a imagem F, mas a sua sombra, isto é, a função standard que é infinitamente próxima.

Definição: Seja $F : \mathbb{R} \to \mathbb{R}$ uma função, eventualmente não-standard, e seja a um real standard. Diz-se que F tem por sombra a função (standard) linear $Y = ax$ se se tem, para todo o X limitado, $F(X) \simeq aX$.

Exemplos:

• A imagem de uma função linear $f(x) = ax$ sobre uma lupa-ε de centro $(0, 0)$ é a própria função linear $F(X) = aX$.

• A imagem da função $f(x) = x^2$ sobre uma lupa-ε de centro $(0, 0)$ tem por sombra a função linear $F(X) = 0$.

• A imagem da função $f(x) = x^2$ sobre uma lupa-ε de centro $(1, 1)$ tem por sombra a função linear $F(X) = 2X$.

• A imagem da função $f(x) = |x|$ sobre uma lupa-ε de centro $(0, 0)$ não tem por sombra uma função linear visto que é a própria função $F(X) = |X|$.

Proposição 4.5

Seja um conjunto aberto standard $\mathcal{D} \subseteq \mathbb{R}$ e seja $f : \mathcal{D} \to \mathbb{R}$ uma função standard. A função f é de classe C^1 de derivada f', sse, para todo o ponto $(x_0, f(x_0))$ quase standard, a imagem de $(x_0, f(x_0))$ por toda a lupa-ε centrada em $(x_0, f(x_0))$ tem por sombra a função linear (standard) de declive $f'(x_0)$.

Dem.

Suponhamos, em primeiro lugar, f de classe C^1 de derivada f', isto é, suponhamos que para todo $(x_0, f(x_0))$ quase standard e para todo $dx \simeq 0$, tem-se:

$$f(x_0 + dx) - f(x_0) = f'(x_0)dx + \overline{\varepsilon}\, dx, \text{ com } \overline{\varepsilon} \simeq 0.$$

Representemos $a := {}^{\circ}(f'(x_0))$. Calculemos a imagem de f sobre uma lupa-ε centrada em $(x_0, f(x_0))$:

$$F(X) = \frac{f(x_0 + \varepsilon X) - f(x_0)}{\varepsilon} = \frac{a\varepsilon X + \overline{\varepsilon}\varepsilon X}{\varepsilon} = aX + \overline{\varepsilon} X.$$

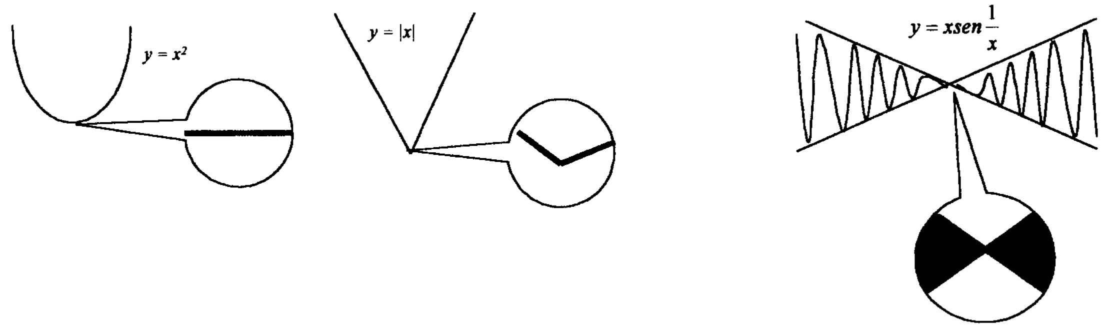


Figura 4.3: Exemplos de funções observadas à lupa.

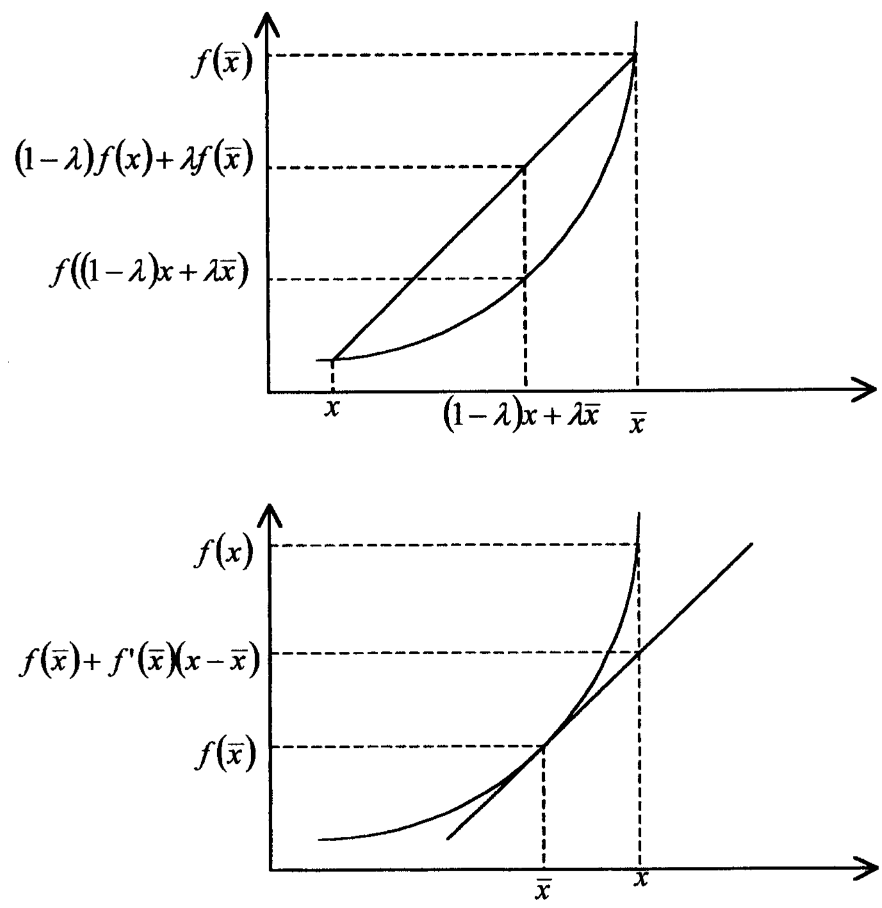


Figura 4.4 Uma função é convexa sse o seu gráfico está situado debaixo de toda a corda, ou (definição equivalente se a função é de classe C^1) sse o seu gráfico está situado acima de toda a tangente.

Portanto, para todo o X limitado $F(X) \simeq ax$ visto que $\overline{\varepsilon} X$ é infinitesimal.

Reciprocamente, suponhamos que a imagem de f sobre toda a lupa-ε centrada em $(x_0, f(x_0))$ tem por sombra a função linear de declive $a = {}^\circ(f'(x_0))$. Seja dx um infinitesimal qualquer e designemos por $F(X)$ a imagem de f sobre a lupa-dx centrada neste ponto. Tem-se para o limitado $X = 1$, $F(1) \simeq f'(x_0)$, se a sombra de F é a aplicação $X \to aX$. Donde,

$$F(1) = \frac{f(x_0 + dx) - f(x_0)}{dx} \simeq f'(x_0). \qquad \square$$

Exemplos: Se examinarmos as funções $f(x) = x^2$, $g(x) = |x|$ e

$$h(x) = \begin{cases} x \operatorname{sen}\frac{1}{x} & \text{se} \quad x \neq 0 \\ 0 & \text{se} \quad x = 0 \end{cases}$$

sob uma lupa-ε centrada em $(0,0)$, observamos (ver figura 4.3.) que a primeira tem por sombra uma função linear, o que não é o caso das duas seguintes (que não são de classe C^1).

4.4 Algumas aplicações das derivadas

4.4.1 Funções convexas, côncavas

Uma das noções matemáticas frequentemente utilizadas em Economia é a noção de convexidade. Vamos dar aqui duas caracterizações e alguns exemplos.

Definição: Diz-se que uma função $f :]a,b[\to \mathbb{R}$ é *convexa* (respectivamente *estritamente convexa*) se, para todos os pontos x, $\overline{x}$ em $]a,b[$, e todo o número real $\lambda \in [0,1]$, tem-se: $f((1-\lambda)x + \lambda\overline{x}) \leq (1-\lambda)f(x) + \lambda f(\overline{x})$ (respectivamente $f((1-\lambda)x + \lambda\overline{x}) < (1-\lambda)f(x) + \lambda f(\overline{x})$). Diz-se que uma função $f :]a,b[\to \mathbb{R}$ é *côncava* (respectivamente *estritamente côncava*) se $-f$ é convexa (respectivamente estritamente convexa).

Esta definição de convexidade interpreta-se facilmente em termos geométricos (ver figura 4.4): esta definição afirma que se $M = (x, f(x))$ e $N = (\overline{x}, f(\overline{x}))$ são dois pontos do gráfico de f, todos os pontos da corda MN estão situados acima (respectivamente abaixo para a função côncava) dos pontos do gráfico com a mesma abcissa. Ou seja, o gráfico de f está abaixo de todas as cordas.

Exemplos:

- As funções $f(x) = e^x$ e $f(x) = x^2$ são convexas em $\mathbb{R}$.

• As funções $f(x) = \ln(x)$ e $f(x) = \sqrt{x}$ são côncavas em $]0, +\infty[$.

• A função $f(x) = ax + b$ é a única função que é convexa e côncava ao mesmo tempo.

• As funções $f(x) = \text{sen } x$ e $f(x) = x^3$ não são nem convexas nem côncavas em $\mathbb{R}$.

Proposição 4.6

Seja $f :]a, b[\to \mathbb{R}$ uma função de classe C^1. A função f é convexa sse se verifica uma das duas caracterizações equivalentes seguintes:

• *para todos os x e $\overline{x}$, $a < x < \overline{x} < b$, tem-se $f(x) \geq f(\overline{x}) + f'(\overline{x})(x - \overline{x})$.*

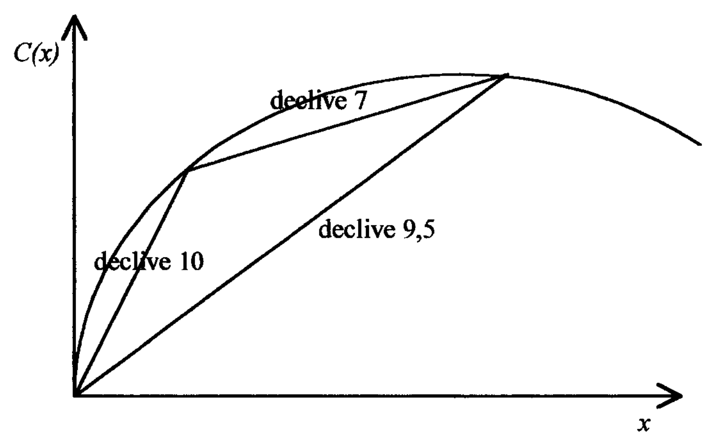
• *f' é crescente.*

A primeira caracterização compreende-se geometricamente (ver figura 4.4.) da seguinte forma: $y = f(\overline{x}) + f'(\overline{x})(x - \overline{x})$ é simplesmente a equação da tangente ao gráfico de f no ponto de abcissa $\overline{x}$. A desigualdade significa portanto que no ponto de abcissa x o gráfico de f está acima da sua tangente em $\overline{x}$. O gráfico de uma função convexa de classe C^1 está portanto sempre situado acima das suas tangentes.

4.4.2 Saber distinguir o "médio" do "marginal"

Vejamos um pequeno exemplo instrutivo: um patrão de uma empresa fabrica 100 000 unidades de um produto para um custo total de 1 000 000 de escudos. O preço de custo unitário é portanto de 10 escudos. Um cliente propôs-lhe uma encomenda suplementar de 20 000 unidades mas só aceita pagar 9 escudos por peça. O patrão da empresa deve recusar esta encomenda suplementar ?

O patrão estuda qual seria o custo total correspondente ao fabrico de 120 000 unidades. Suponhamos que este é elevado para 1 140 000 escudos, seja um preço unitário, desta vez, de 9,5 escudos. Deve ele recusar ?



Não, visto que o custo unitário de 20 000 unidades suplementares é, na realidade, de (1 140 000 – 1 000 000) /20 000, é de 7 escudos. Se ele confundir custo médio $(9,5$ escudos) e custo marginal (7 escudos), ele perderá $(9-7)\times 20000 = 40000$ escudos! O esquema da figura anterior mostra o custo médio $\left(C_M = \frac{\Delta C}{\Delta x}\right)$ que é aqui uma corda e o custo marginal que se pode assimilar, supondo que $\frac{1}{20\,000}$é infinitesimal, à tangente ao gráfico no ponto A, $C_m = C'(x)$.

4.4.3 Derivadas logarítmicas e elasticidade

Definição: Chama-se *derivada logarítmica* de uma função $f:\mathbb{R}\to\mathbb{R}$ à função $\frac{f'(x)}{f(x)}$ quando esta existe. É igualmente a derivada de $\ln f(x)$, daí o seu nome.

Chama-se *elasticidade* de uma função $f:\mathbb{R}\to\mathbb{R}$ à função $e(f)$ definida por $e(f)(x) = x\frac{f'(x)}{f(x)}$.

Observação: Para calcular a derivada logarítmica de uma função, é por vezes preferível calcular a derivada de $\ln f(x)$, em vez do quociente $\frac{f'(x)}{f(x)}$. Por exemplo

$$f(x) = \frac{x^{\alpha}(\ln x)^{\beta}e^{\gamma x}}{ax+b}$$

$$\ln f(x) = \alpha \ln x + \beta \ln(\ln x) + \gamma x - \ln(ax+b)$$

$$\frac{f'(x)}{f(x)} = \frac{\alpha}{x} + \frac{\beta}{x\ln x} + \gamma - \frac{a}{ax+b}\,.$$

Em termos económicos a elasticidade tem uma interpretação simples: é o quociente entre a variação relativa de f, $\frac{f(x)-f(x_0)}{f(x_0)}$, e a variação relativa de x, $\frac{x-x_0}{x_0}$. Com efeito,

$$\frac{\frac{f(x)-f(x_0)}{f(x_0)}}{\frac{x-x_0}{x_0}} = \frac{f(x)-f(x_0)}{x-x_0}\,\frac{x_0}{f(x_0)} \simeq \frac{f'(x_0)}{f(x_0)}x_0 = e(f)(x_0)$$

a aproximação utilizada só é válida se f e x_0 são standard e se $x \simeq x_0$ para $x \neq x_0$. Portanto se por exemplo f representa a quantidade procurada de um produto em função do seu preço x, dizer que esta função tem uma elasticidade de $-2,5\%$ significa que um aumento de preço de 1% conduz a uma diminuição da procura de $2,5\%$. Diz-se que uma tal função de procura f é muito elástica se for sensível às variações de preço, o que pode ser o caso, por exemplo, se existirem produtos de substituição. Diz-se que a função é pouco elástica se é pouco sensível às variações de preço.

Proposição 4.7

A elasticidade verifica as seguintes propriedades :

$$e(f) = e(-f); \quad e(fg) = e(f) + e(g); \quad e\left(\frac{f}{g}\right) = e(f) - e(g);$$

$$e(f^\alpha) = \alpha e(f); \quad e(f)(x) = \frac{\left(\ln|f(x)|\right)'}{\left(\ln|x|\right)'}.$$

Exemplos:

• Recordemos que $dx \neq 0$ designa um (pequeno) incremento da variável x. A uma função f de classe C^1 e a um incremento dx, podem associar-se duas quantidades: o diferencial de f, representado por df, que é o produto da derivada de f por dx ($df = f'(x)dx$), e a quantidade $\triangle f$ que é a variação de f resultante do incremento de x pela quantidade dx ($\triangle f = f(x+dx) - f(x)$). Quando dx é infinitesimal, estas duas quantidades são não só infinitamente próximas, mas mais precisamente o seu afastamento é infinitesimal em relação a dx, e tem-se a fórmula,

$$f(x_0 + dx) - f(x_0) = f'(x_0)dx + \varepsilon dx$$

ou ainda

$$\triangle f = df + \varepsilon dx,$$

onde ε é infinitesimal. Para as seguintes funções, o cálculo do incremento $\triangle f$ e do diferencial df no ponto x_0 indicado, para $dx = 0,01$ conduz à conclusão de que o afastamento entre estas duas quantidades é muito pequeno em relação a dx.

◇ $f : x \mapsto \frac{x}{1-x}$ ($x_0 = 0$ e $dx = 0,01$);

Temos $df = \frac{1}{(1-x)^2}dx$ logo, obtemos $df = 0,01$ e $\triangle f = 0,01 + \varepsilon \times 0,01$, mas ε é infinitesimal, $\triangle f \simeq 0,01$.

◇ $f : x \mapsto \sqrt{1+x}$ ($x_0 = 0$ e $dx = 0,01$);

Temos $df = \frac{1}{2\sqrt{1+x}}dx$, logo, $df = 0{,}005$ e $\triangle f = 0{,}005 + \varepsilon \times 0,01 \simeq 0{,}005$, porque ε é infinitesimal.

◇ $f : x \mapsto x\ln x - x$ ($x_0 = 1$ e $dx = 0,01$);

Temos $df = \left(\ln x + \frac{1}{x} - 1\right)dx$, assim obtemos $df = (\ln 1 + 1 - 1) \times 0,01 = 0$ e $\triangle f = 0 + \varepsilon \times 0,01 \simeq 0$, porque ε é infinitesimal.

◇ $f : x \mapsto e^{-x^2}$ ($x_0 = 1$ e $dx = 0,01$);

Temos $df = \left(-2xe^{-x^2}\right)dx$, logo $df = \frac{-2}{e} \times 0,01 = \frac{-2}{100e} \simeq -0,00736$ (5 casas decimais) e $\triangle f = \frac{-2}{100e} + \varepsilon \times 0,01 \simeq \frac{-2}{100e}$, porque ε é infinitesimal.

• Sejam f e g duas funções standard de classe C^k (diz-se que f é de classe C^k se f é de classe C^{k-1}, e se todas as suas derivadas de ordem $(k-1)$ são de classe C^1) definidas em $\mathbb{R}$. Mostrar (pela fórmula de Taylor) que, para $x \simeq 0$ mas $x \neq 0$, se

$$f(0) = f'(0) = f''(0) = f^{(k-1)}(0) = 0,$$

então

$$f^{(k)}(0) = {}^{o}\!\left(\frac{k!}{x^k} f(x)\right).$$

Como o resultado que se pretende mostrar é standard, é suficiente estabelecê-lo para k standard. A função $f^{(k)}$ é portanto standard. Para todo o $h \simeq 0$ e $h \neq 0$, existe $\theta \in \,]0, 1[$ tal que,

$$f(x+h) = f(x) + hf'(x) + \frac{h^2}{2!} f''(x) + \ldots + \frac{h^{k-1}}{(k-1)!} f^{(k-1)}(x) + \frac{h^k}{k!} f^{(k)}(x + \theta h).$$

Para $x = 0$, as primeiras $(k-1)$ derivadas são nulas, logo tem-se $f^{(k)}(\theta h) = \frac{k!}{h^k} f(h)$. As partes standard destes dois números são iguais, a parte standard de $f^{(k)}(\theta h)$ é $f^{(k)}(0)$ visto que $f^{(k)}$ é contínua; assim, obtém-se o resultado pretendido.

Capítulo 5

Funções de várias variáveis

5.1 Definição e representação gráfica

Definição: Seja $\mathcal{D}$ um subconjunto de $\mathbb{R}^p$. Uma função $f : \mathcal{D} \to \mathbb{R}$, $(x_1, ..., x_p) \mapsto f(x_1, ..., x_p)$ chama-se *função de p variáveis*. O gráfico $G(f)$ de uma tal função é um subconjunto de $\mathbb{R}^{p+1}$ definido por

$$G(f) = \{(x_1, ..., x_p, f(x_1, ..., x_p)),\ (x_1, ..., x_p) \in \mathcal{D}\}.$$

Exemplos:

Para $p = 1$, o gráfico $G(f)$ é, em geral, uma curva, para $p = 2$, é uma superfície (de dimensão 2) de $\mathbb{R}^3$. Para $p \geq 2$, é uma hiper-superficie (de dimensão p) de $\mathbb{R}^{p+1}$.

Em geral não é fácil traçar o gráfico de uma função $(x, y) \mapsto f(x, y)$ porque, mesmo se se conhecer a forma, a representação de uma superfície num espaço a três dimensões é um problema em si (que necessita, por exemplo, alguns conhecimentos em matéria de perspectiva). Um caso particular mais fácil é o das funções afins $(x, y) \mapsto ax + by + c$ porque o gráfico é então o plano de equação $z = ax + by + c$. Podemos representar por exemplo os gráficos das funções seguintes:

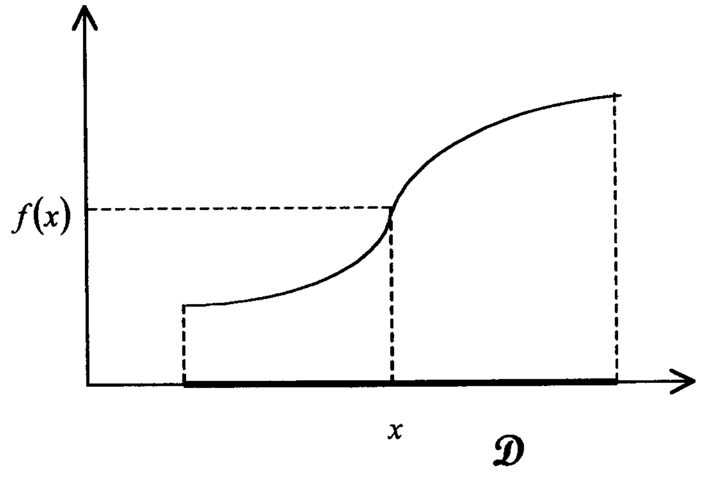
$$(x, y) \mapsto f(x, y) = c,\ (x, y) \mapsto p_1(x, y) = x,\ (x, y) \mapsto p_2(x, y) = y,$$

$$(x, y) \mapsto g(x, y) = x - y,\ \ (x, y) \mapsto h(x, y) = \ -x - 2y + 2.$$

Para isto, o mais simples é frequentemente determinar as intersecções dos planos procurados com os três eixos de coordenadas (isto é as rectas de equações $x = y = 0$, $x = z = 0$ e $y = z = 0$ respectivamente) ou com dois dos planos de coordenadas (isto é, dois dos planos de equação $x = 0$, $y = 0$ ou $z = 0$).

As funções afins de duas (ou p) variáveis podem ser calculadas em todos os pontos de $\mathbb{R}^2$ (ou de $\mathbb{R}^p$), o que não é o caso, por exemplo, da função

$(x, y) \mapsto f(x, y) = \ln(x - y)$ que só é definida se $x - y > 0$, isto é se (x, y) pertence ao semi-plano de equação $x - y > 0$.



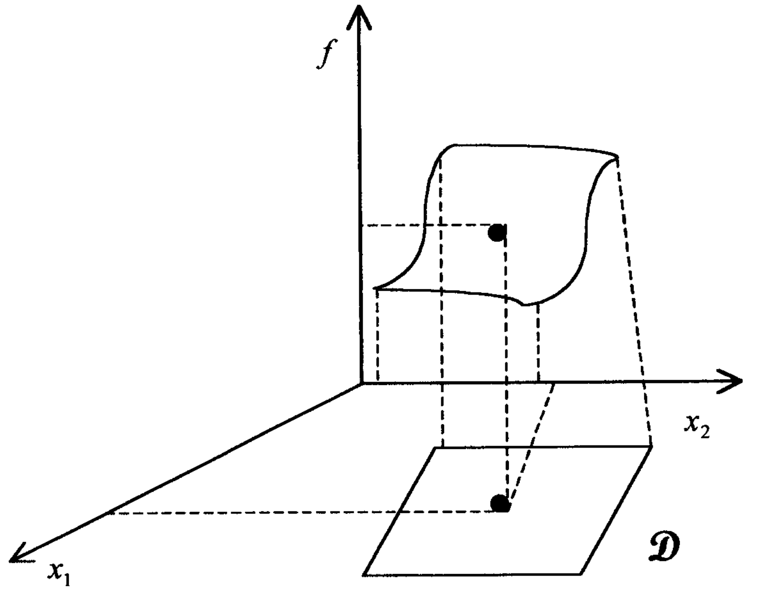


Figura 5.1. Gráfico de uma função de uma variável e gráfico de uma função de duas variáveis

Este semi-plano é o domínio de definição da função f, isto é o maior subconjunto de $\mathbb{R}^2$ sobre o qual a expressão de f pode ser calculada.

Em princípio o conhecimento de uma função contém a definição do seu domínio e em particular duas funções diferentes podem muito bem ser definidas por uma mesma expressão algébrica e deferirem apenas pelo domínio sobre o qual estão definidas.

Exemplos:

O domínio da definição da função $(x,y) \mapsto \sqrt{xy}$ é a reunião de dois quadrantes.

$\{(x,y) \in \mathbb{R}^2, x \geq 0 \text{ e } y \geq 0\} \cup \{(x,y) \in \mathbb{R}^2, x \leq 0, \text{e } y \leq 0\}.$

O domínio da definição da função $(x,y) \mapsto \frac{1}{1-x^2-y^2}$ é o círculo aberto de centro $(0,0)$ e de raio 1:

$$\{(x,y) \in \mathbb{R}^2 : x^2 + y^2 < 1\}.$$

Definição: Seja $f : \mathcal{D} \to \mathbb{R}$ uma função de p variáveis definida num domínio $\mathcal{D}$ de $\mathbb{R}^p$. Seja $X^* = (x_1^*, x_2^*, ..., x_p^*)$ um vector de $\mathcal{D}$. Chama-se *aplicação parcial* de f em X^* a cada uma das aplicações obtidas fixando todas as coordenadas $(x_1, x_2, ..., x_p)$ em $(x_1^*, x_2^*, ..., x_p^*)$ salvo uma. Por exemplo,

$$f_1^{X^*} : \mathcal{D}_1 \to \mathbb{R}, (x_1) \to f_1^{X^*}(x_1) = f(x_1, x_2^*, ..., x_p^*).$$

O domínio de definição $\mathcal{D}_1$ de $f_1^{X^*}$ é constituído pelos x_1 que são a primeira coordenada dos pontos de intersecção de $\mathcal{D}$ com a recta de $\mathbb{R}^p$ de equações $x_2 = x_2^*, ..., x_p = x_p^*$ e o gráfico de $f_1^{X^*}$ e a intersecção do gráfico de f com o plano "vertical" de equações $x_2 = x_2^*, ..., x_p = x_p^*$ em $\mathbb{R}^{p+1}$ (ver a figura 5.2).

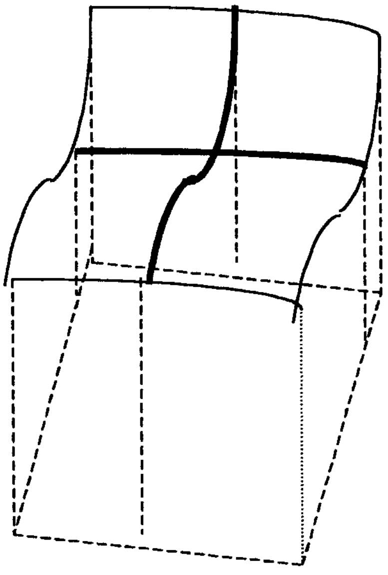


Figura 5.2 As duas aplicações parciais de uma função de duas variáveis.

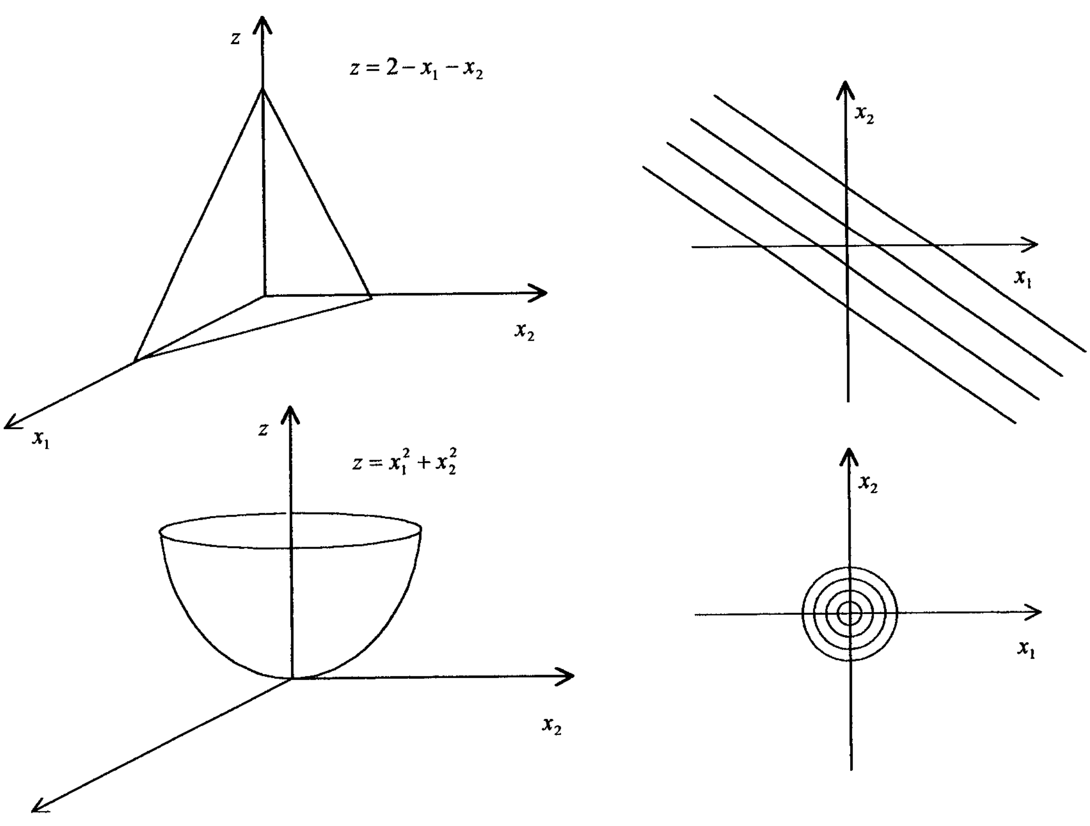
5.2 Curvas e superfícies de nível

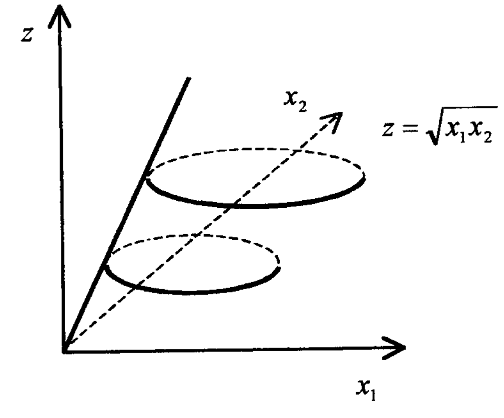
Uma outra forma de representar o gráfico de uma função de duas variáveis sobre uma folha de papel (de dimensão 2!) é traçar as suas curvas de nível, como se faz, por exemplo, para representar relevo sobre um mapa de um percurso.

Definição: Seja $\mathcal{D}$ um subconjunto de $\mathbb{R}^2$, $f : \mathcal{D} \to \mathbb{R}$ uma função de duas variáveis e c um número pertencente a $f(\mathcal{D})$. Chama-se *curva de nível* c de f à curva de $\mathcal{D} \subset \mathbb{R}^2$ de equação $f(x, y) = c$.

Mais geralmente:

Definição: Seja $\mathcal{D}$ um subconjunto de $\mathbb{R}^p$, $f : \mathcal{D} \to \mathbb{R}$ uma função com p variáveis e c um número pertencente a $f(\mathcal{D})$. Chama-se *superfície de nível* c de f à hipersuperfície de $\mathcal{D} \subset \mathbb{R}^p$ de equação $f\ (x_1, x_2, ..., x_p) = c$.





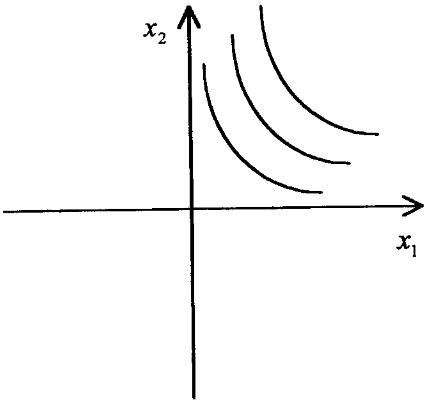


Figura 5.3: Gráficos e curvas de nível de algumas funções com duas variáveis. Observar-se-á que a curva de nível é a projecção, no plano das variáveis (x, y) da intersecção do gráfico $G(f)$ com o plano horizontal de cota c.

Observação: Ao contrário do conjunto $\mathbb{R}$ que é naturalmente ordenado (sabe-se o que significa $x \leq y$), o conjunto $\mathbb{R}^p$ para $p \geq 2$ não é naturalmente ordenado (o que é "o maior" dos vectores (2,5) e (3,4) ?). Um dos meios mais simples para definir uma "relação de ordem" em $\mathbb{R}^p$ é o de dar uma função $f : \mathbb{R}^p \to \mathbb{R}$: põe-se então simplesmente,

$$(x_1, x_2, ..., x_p) \leq (x_1', x_2', ..., x_p') \Leftrightarrow f(x_1, x_2, ..., x_p) \leq f(x_1', x_2', ..., x_p').$$

Este procedimento é frequentemente utilizado em Economia: por exemplo na teoria do consumidor, chama-se *cabaz dos bens* ao conjunto das quantidades de cada um dos bens que é atribuída por um consumidor: x_i é a quantidade do bem i que ele possui. O cabaz dos bens não é nada mais do que um vector de $\mathbb{R}^p$. Para formalizar a hipótese de que este consumidor é capaz de decidir, perante dois cabazes de bens, qual cabaz prefere ou se estes são equivalentes para ele, introduz-se uma *função de utilidade* $U : \mathbb{R}_+^p \to \mathbb{R}$ cuja relação de ordem associada será a relação de preferência do consumidor. As curvas de nível desta função têm então o nome de *curvas de indiferença.*

5.3 Diferencial e derivadas parciais

Neste parágrafo, vamos generalizar às funções de várias variáveis a noção de função de classe C^1, estudada no caso das funções com uma variável no capítulo 4. Por razões de simplificação, limitamo-nos muito frequentemente ao caso das funções de duas variáveis; a generalização a n variáveis será fácil.

Tal como para uma variável, diz-se que uma função é de classe C^1 na vizinhança de um dado ponto quando está *bem aproximada* por uma função linear na vizinhança desse ponto. A questão será precisar o que se entende por *bem aproximada*. É útil notar que, se com uma só variável, uma função linear $x \mapsto L(x)$ é assimilável a um número, seja a por exemplo, $L(x) = ax$, o declive do seu gráfico, com duas variáveis ter-se-á $(x, y) \mapsto L(x, y) = ax + by$, sendo conveniente, portanto, dar dois números a e b (e com n variáveis n números). A aproximação linear de uma função f de classe C^1 com duas variáveis na vizinhança de um ponto (x, y) será, portanto, uma aplicação linear de $\mathbb{R}^2$ em $\mathbb{R}$, chamada *diferencial de* f no ponto (x, y), seja $Df_{(x,y)}$ que poder-se-á escrever,

$$(h, k) \mapsto Df_{(x,y)}[h, k] = ah + bk$$

onde a e b são dois números que dependem do ponto (x, y) considerado.

Antes de dar uma definição precisa do diferencial Df e das suas propriedades, indiquemos primeiro como calcular os dois números a e b: trata-se do valor das derivadas parciais de f no ponto (x, y).

Definição: Sejam $\mathcal{D}$ um aberto de $\mathbb{R}^2$, $f : \mathcal{D} \to \mathbb{R}$ uma função standard e $M_0 = (x_0, y_0) \in \mathcal{D}$ um ponto standard. Fixa-se a variável y com o valor y_0 e considera-se a aplicação parcial $x \mapsto f_1^{X^*}(x) = f(x, y_0)$. Se $f_1^{X^*}$ é de classe C^1, de derivada $\left(f_1^{X^*}\right)'$, o número $a = \left(f_1^{X^*}\right)'(x_0)$ chama-se *derivada parcial de* f *em relação a* x no ponto (x_0, y_0) e representa-se por:

$$\frac{\partial f}{\partial x}(x_0, y_0) = f'_x(x_0, y_0) = a.$$

Define-se de forma análoga a *derivada parcial de* f *em relação a* y no ponto (x_0, y_0):

$$\frac{\partial f}{\partial y}(x_0, y_0) = f'_y(x_0, y_0) = b.$$

Observação: Para calcular os dois números a e b, pode-se portanto, voltando à definição de derivada de uma função de classe C^1 com uma variável, utilizar as seguintes fórmulas (dx e dy são números infinitesimais quaisquer diferentes de zero):

$$\frac{\partial f}{\partial x}(x_0, y_0) = {}^{o}[(f(x_0 + dx, y_0) - f(x_0, y_0))/dx]$$

e

$$\frac{\partial f}{\partial y}(x_0, y_0) = {}^{o}[(f(x_0, y_0 + dy) - f(x_0, y_0))/dy].$$

Mas, na prática, raramente é necessário voltar a estas definições porque pode-se utilizar directamente as derivadas conhecidas das funções usuais com uma variável: em particular não haverá aqui novas fórmulas de derivação a aprender.

Exemplo: Calculemos as derivadas parciais de $f(x,y) = x^3 + y^2 - 3xy$ no ponto $(1,2)$. As duas aplicações parciais de f neste ponto são respectivamente:

$$x \mapsto x^3 + 4 - 6x \ , \ y \mapsto 1 + y^2 - 3y.$$

Portanto as duas derivadas parciais são $\dfrac{\partial f}{\partial x}(x,2) = 3x^2 - 6$ e $\frac{\partial f}{\partial y}(1,y) = 2y - 3$ e $a = \frac{\partial f}{\partial x}(1,2) = -3$ e $b = \frac{\partial f}{\partial y}(1,2) = 1$.

Generaliza-se facilmente a definição anterior às funções com n variáveis:

Definição: Seja $(x_1,...,x_n) \mapsto f(x_1,...,x_n)$ uma função standard definida sobre um domínio $\mathcal{D}$ de $\mathbb{R}^n$ com valores em $\mathbb{R}$. Seja $M^0 = (x_1^0,...,x_n^0)$ um ponto standard de $\mathcal{D}$. Chama-se derivada parcial de ordem i de f no ponto M^0 à derivada (se existe) no ponto x_i^0 da aplicação parcial (função com uma só variável) $x_i \mapsto f(x_1^0,...,x_{i-1}^0, x_i, x_{i+1}^0,...,x_n^0)$. Este número representa-se por $\frac{\partial f}{\partial x_i}(x_1^0,...,x_n^0)$.

O gradiente[5] de f em $x_1^0,...,x_n^0$ é:

$$Grad(f)(x_1^0,...,x_n^0) = \left(\frac{\partial f}{\partial x_1}(x_1^0,...,x_n^0),...,\frac{\partial f}{\partial x_n}(x_1^0,...,x_n^0)\right).$$

Exemplo: Para $f(x,y) = 2x^4y^3 - xy^2 + 3y + 1$ e $g(x,y) = xe^y + y \operatorname{sen} x$, tem-se:

$$Grad(f) = (8x^3y^3 - y^2, 6x^4y^2 - 2xy + 3),$$
$$Grad(g) = (e^y + y\cos x, xe^y + \operatorname{sen} x).$$

Voltemos ao assunto inicial, que consiste em generalizar às funções de duas variáveis a noção de função de classe C^1.

---

[5] O gradiente de f é um "vector simbólico" cujas componentes são as n derivadas parciais de f. Podemos representar o gradiente de f por $Grad(f)$ ou ∇f:

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

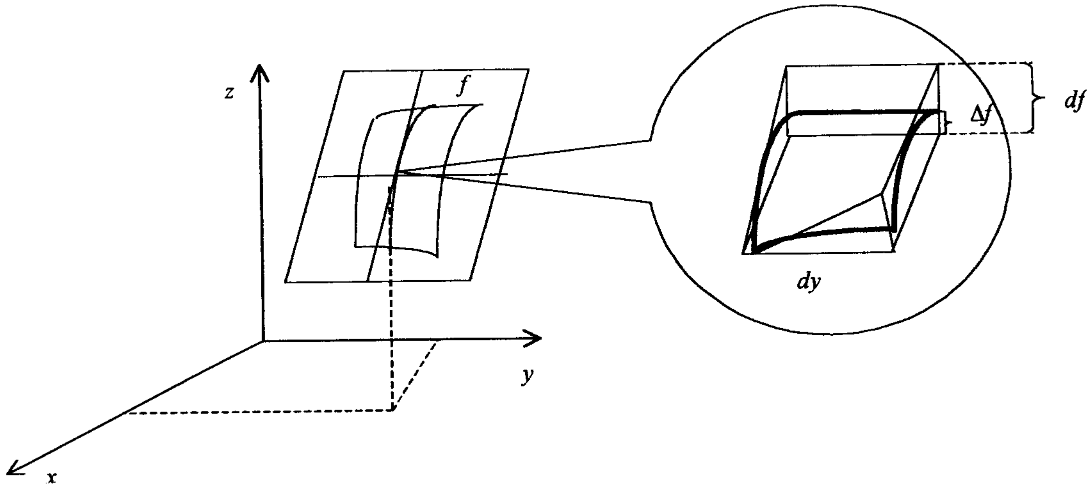


Figura 5.4: O crescimento $\triangle f$ do diferencial df para uma função $f(x,y)$ de duas variáveis.

Definição: Seja $\mathcal{D}$ um conjunto aberto standard de $\mathbb{R}^2$ e sejam $f:\mathcal{D}\to\mathbb{R}$, $f'_x:\mathcal{D}\to\mathbb{R}$ e $f'_y:\mathcal{D}\to\mathbb{R}$ funções standard. Diz-se que f é de *classe* C^1 de derivadas parciais f'_x e f'_y se para todo o (x,y) quase standard em $\mathcal{D}$ e para todos os dx e dy infinitesimais,

$$f(x+dx,y+dy)-f(x,y)=f'_x(x,y)dx+f'_y\,(x,y)dy+\varepsilon_x dx+\varepsilon_y dy,$$

com $\varepsilon_x\simeq 0$ e $\varepsilon_y\simeq 0$. Para (x,y) fixo, a aplicação linear

$$(dx,dy)\mapsto f'_x(x,y)dx+f'_y(x,y)dy$$

representa-se por $Df_{(x,y)}$ e chama-se diferencial de f no ponto (x,y).

Observação: Notações diferenciais: Se se designar por $\triangle f$ e df as quantidades:

$$\triangle f=f(x+dx,y+dy)-f(x,y)$$

e

$$df=adx+bdy=f'_x(x,y)dx+f'_y(x,y)dy$$

a fórmula anterior pode escrever-se simplesmente,

$$\triangle f=df+\varepsilon_x dx+\varepsilon_y dy.$$

Exemplo:

A função $f(x,y)=x^2+xy+y^2$ é de classe C^1, de derivadas parciais, $f'_x(x,y)=2x+y$ e $f'_y\,(x,y)=x+2y$.

5.4 Observar uma função à lupa

Como no caso das funções com uma variável, vamos agora observar os gráficos das funções de duas variáveis à lupa, e veremos que as funções de classe C^1 são precisamente as funções cujos gráficos têm por sombra, sobre toda a lupa-ε, um plano (que não é mais do que o plano tangente ao gráfico).

Definição: Seja $(x_0, y_0, z_0) \in \mathbb{R}^3$ e seja $\varepsilon \simeq 0$ positivo. Uma homotetia da forma

$$(x, y, z) \mapsto \left(X = \frac{x - x_0}{\varepsilon}, Y = \frac{y - y_0}{\varepsilon}, Z = \frac{z - z_0}{\varepsilon}\right)$$

chama-se uma *lupa de centro* (x_0, y_0, z_0) e de ampliação $\frac{1}{\varepsilon}$ [ou, também, uma ε-lupa de centro (x_0, y_0, z_0)].

Definição: Se $(x, y) \mapsto f(x, y)$ é uma função tal que $f(x_0, y_0) = z_0$, chama-se *imagem de* f sobre a lupa de centro $(x_0, y_0, f(x_0, y_0))$ e de ampliação $\frac{1}{\varepsilon}$, a função $(X, Y) \mapsto Z = F(X, Y)$ definida por:

$$F(X, Y) = \frac{f(x_0 + \varepsilon X, y_0 + \varepsilon Y) - f(x_0, y_0)}{\varepsilon}.$$

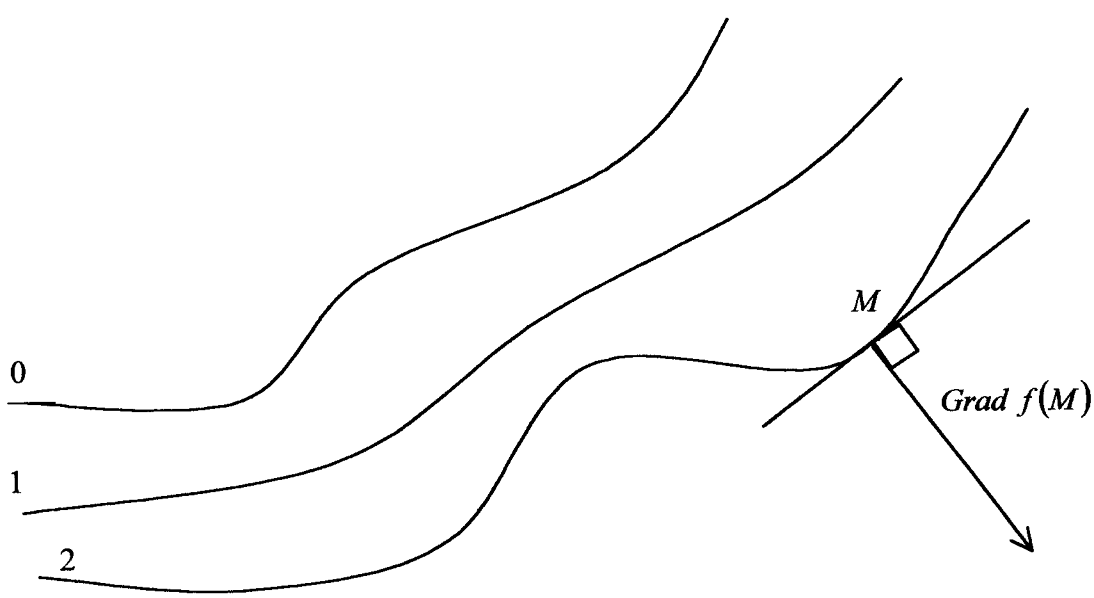


Figura 5.5 Em cada ponto, o vector gradiente é perpendicular à curva de nível da função que passa por este ponto, dirigida no sentido dos níveis crescentes.

Definição: Seja $F : \mathbb{R}^2 \to \mathbb{R}$ uma função, eventualmente não-standard, e sejam a e b dois números reais standard. Diz-se que F *tem por sombra a função (standard) linear* $Z = aX + bY$ se se tiver, para todo (X, Y) limitado, $F(X,Y) \simeq aX + bY$.

Exemplos:

• a imagem de uma função linear $f(x,y) = ax + by$ sobre uma lupa-ε de centro $(0,0,0)$ é a função linear $F(X,Y) = aX + bY$.

• A imagem da função $f(x) = x^2 + y^2$ sobre uma lupa-ε de centro $(0,0,0)$ tem por sombra a função linear $L(X,Y) = 0$.

• A imagem da função $f(x) = x^2 + y^2$ sobre uma lupa-ε de centro $(1,1,2)$ tem por sombra a função linear $L(X,Y) = 2X + 2Y$.

Proposição 5.1

Seja $\mathcal{D}$ um conjunto aberto standard de $\mathbb{R}^2$ e sejam f, f'_x, $f'_y : \mathcal{D} \to \mathbb{R}$ funções standard. A função f é de classe C^1, de derivadas parciais f'_x e f'_y sse para todo o (x_0, y_0) quase standard em $\mathcal{D}$, a sua imagem F sobre toda a lupa-ε *centrada em $(x_0, y_0, f(x_0, y_0))$ tem por sombra a função linear (standard) $L(X,Y) = aX + bY$, onde $a = {}^o(f'_x(x_0,y_0))$ e $b = {}^o(f'_y(x_0,y_0))$.*

Dem.

Sob uma lupa-ε, a função f pode escrever-se:

$$F(X,Y) = \frac{f(x_0 + \varepsilon X, y_0 + \varepsilon Y) - f(x_0, y_0)}{\varepsilon}.$$

Como f é de classe C^1 e é standard, e para todo o X e todo o Y limitados, $\varepsilon X \simeq 0$ e $\varepsilon Y \simeq 0$ a razão

$$\frac{f(x_0 + \varepsilon X, y_0 + \varepsilon Y) - f(x_0, y_0)}{\varepsilon}$$

é quase standard. Portanto, para todo o X e todo o Y limitados

$$F(X,Y) \simeq f'_x(x_0,y_0)X + f'_y(x_0,y_0)Y,$$

donde,

$${}^o(F(X,Y)) = f'_x(x_0,y_0)X + f'_y(x_0,y_0)Y,$$

porque, por hipótese, f'_x e f'_y são números standard.

Reciprocamente, para todo o ε infinitesimal diferente de zero, a razão

$$\frac{f(x_0+\varepsilon X, y_0+\varepsilon Y)-f(x_0,y_0)}{\varepsilon}$$

é igual a $F(X,Y)$, onde $F(X,Y)$ é a imagem de f sob a lupa-ε. Como, por hipótese, a sombra F é uma função linear (standard) tem-se

$$^{o}(F(X,Y)) = L(X,Y) = aX + bY$$

em que $a = {}^{o}(f'_x(x_0,y_0))$ e $b = {}^{o}\big(f'_y(x_0,y_0)\big)$. □

Definição: Seja $\mathcal{D}$ um conjunto aberto standard de $\mathbb{R}^2$ e seja $f : \mathcal{D} \to \mathbb{R}$ de classe C^1, de derivadas parciais f'_x e f'_y. Chama-se plano tangente ao gráfico de f no ponto $(x_0, y_0, z_0 = f(x_0,y_0))$ ao plano de equação $z = a(x-x_0)+b(y-y_0)+z_0$, onde $a = f'_x\,(x_0,y_0)$ e $b = f'_y(x_0,y_0)$.

Exemplo:

A equação do plano tangente ao gráfico de $f(x,y) = x^2+xy+y^2$ no ponto $(2,\,-1)$ é $z = 3x-3$ (derivando a função f em ordem a x, temos $f'_x(x,y) = 2x+y$ logo, $f'_x(2,\,-1) = 3$. E derivando a função f em ordem a y, temos $f'_y(x,y) = x+2y$ logo, $f'_y(2,\,-1) = 0$, assim, $z = 3(x-2)+0(y-1)+3 = 3x-3$)

Vejamos agora um outro resultado, muito útil, que se pode igualmente obter examinando uma função de duas variáveis à lupa:

Proposição 5.2

Seja f uma função de classe C^1 com duas variáveis. Seja (x_0,y_0) um ponto onde onde o vector $Grad f$ não é o vector nulo. Então a tangente da curva de nível de f que passa por este ponto tem por vector director $\left(\frac{\partial f}{\partial y}(x_0,y_0),\,-\frac{\partial f}{\partial x}(x_0,y_0)\right)$ e por vector normal $\left(\frac{\partial f}{\partial x}(x_0,y_0),\frac{\partial f}{\partial y}(x_0,y_0)\right)$. Portanto, quando não é nulo, o vector gradiente de f no ponto (x_0,y_0) é perpendicular à curva de nível que passa por este ponto, e está dirigido no sentido dos níveis crescentes ("sentido do maior declive ascendente").

Dem.

Podemos supor, por transferência, que a função f e o ponto (x_0,y_0) são standard. Observamos a curva de equação $f(x,y) = f(x_0,y_0)$ à lupa-ε $(x,y) \mapsto \left(X = \frac{x-x_0}{\varepsilon}, Y = \frac{y-y_0}{\varepsilon}\right)$. A curva imagem tem por equação $F(X,Y) = 0$

onde $F(X,Y) = f(x_0 + \varepsilon X, y_0 + \varepsilon Y) - f(x_0, y_0)$. Temos que $F(X,Y)$ tem por sombra a função linear $(X,Y) \mapsto \frac{\partial f}{\partial x}(x_0,y_0)X + \frac{\partial f}{\partial y}(x_0,y_0)Y$. Logo o resultado. □

5.5 Mostrar que uma função é de classe C^1

Para mostrar que uma função dada é de classe C^1, pensamos em primeiro lugar na definição dada anteriormente:

Para todo o (x,y) quase standard em $\mathcal{D}$ e para todos os dx e dy infinitesimais,

$$f(x+dx, y+dy) - f(x,y) = f'_x(x,y)dx + f'_y\,(x,y)dy + \varepsilon_x dx + \varepsilon_y dy,$$

com $\varepsilon_x \simeq 0$ e $\varepsilon_y \simeq 0$.

Mas, como para as funções com uma só variável, é raramente necessário, na prática, recorrer a esta definição. Habitualmente, demonstramos que uma função é de classe C^1 mostrando que esta se exprime como uma *combinação* (soma, produto, composição, etc ...) de funções *usuais*, isto é, de funções para as quais se mostrou que são de classe C^1. Estas funções *usuais* são especialmente os polinómios, as fracções racionais, a exponencial, o logaritmo, as funções trigonométricas e trigonométricas inversas, as funções hiperbólicas e hiperbólicas inversas, as potências (em particular a raíz quadrada) e o valor absoluto. Para cada uma destas funções, conhece-se o maior domínio onde são de classe C^1. A maior parte das funções que se encontram podem exprimir-se a partir de algumas funções usuais através das operações descritas pela proposição seguinte. Esta proposição permite então mostrar que as funções são de classe C^1.

Proposição 5.3

• *As duas projecções $(x,y) \mapsto x$ e $(x,y) \mapsto y$ são de classe C^1.*

• *Sejam $(x,y) \mapsto f(x,y)$ e $(x,y) \mapsto g(x,y)$ duas funções de classe C^1, com derivadas parciais f'_x , f'_y e g'_x , g'_y respectivamente, e seja $\lambda \in \mathbb{R}$. Então:*

– a soma $(x,y) \mapsto (f+g)(x,y) = f(x,y) + g(x,y)$ é de classe C^1, de derivadas parciais $(f+g)'_x = f'_x + g'_x$ e $(f+g)'_y = f'_y + g'_y$.

– o produto $(x,y) \mapsto (fg)(x,y) = f(x,y) \cdot g(x,y)$ é de classe C^1, de derivadas parciais $(fg)'_x = f'_x g + f g'_x$ e $(fg)'_y = f'_y\, g + f g'_y$.

– se $(x,y) \mapsto g(x,y)$ não se anula, o quociente $(x,y) \mapsto \frac{f(x,y)}{g(x,y)}$ é de classe C^1, de derivadas parciais $\left(\frac{f}{g}\right)'_x = \frac{f'_x g - f g'_x}{g^2}$ e $\left(\frac{f}{g}\right)'_y = \frac{f'_y g - f g'_y}{g^2}$.

• *Se* $(x,y) \mapsto f(x,y)$ *e* $z \mapsto h(z)$ *são de classe* C^1*, então* $h \circ f$ *é de classe* C^1*, de derivadas parciais* $(h \circ f)'_x = (h' \circ f)f'_x$ *e* $(h \circ f)'_y = (h' \circ f)f'_y$*.*

• *Se* $t \mapsto h(t)$*,* $t \mapsto k(t)$ *e* $(x,y) \mapsto f(x,y)$ *são três funções de classe* C^1*, então* $t \mapsto f(h(t),k(t))$ *é de classe* C^1*, de derivadas:*

$$f'_x(h(t),k(t))h'(t) + f'_y(h(t),k(t))k'(t).$$

Vejamos um exemplo desta proposição:

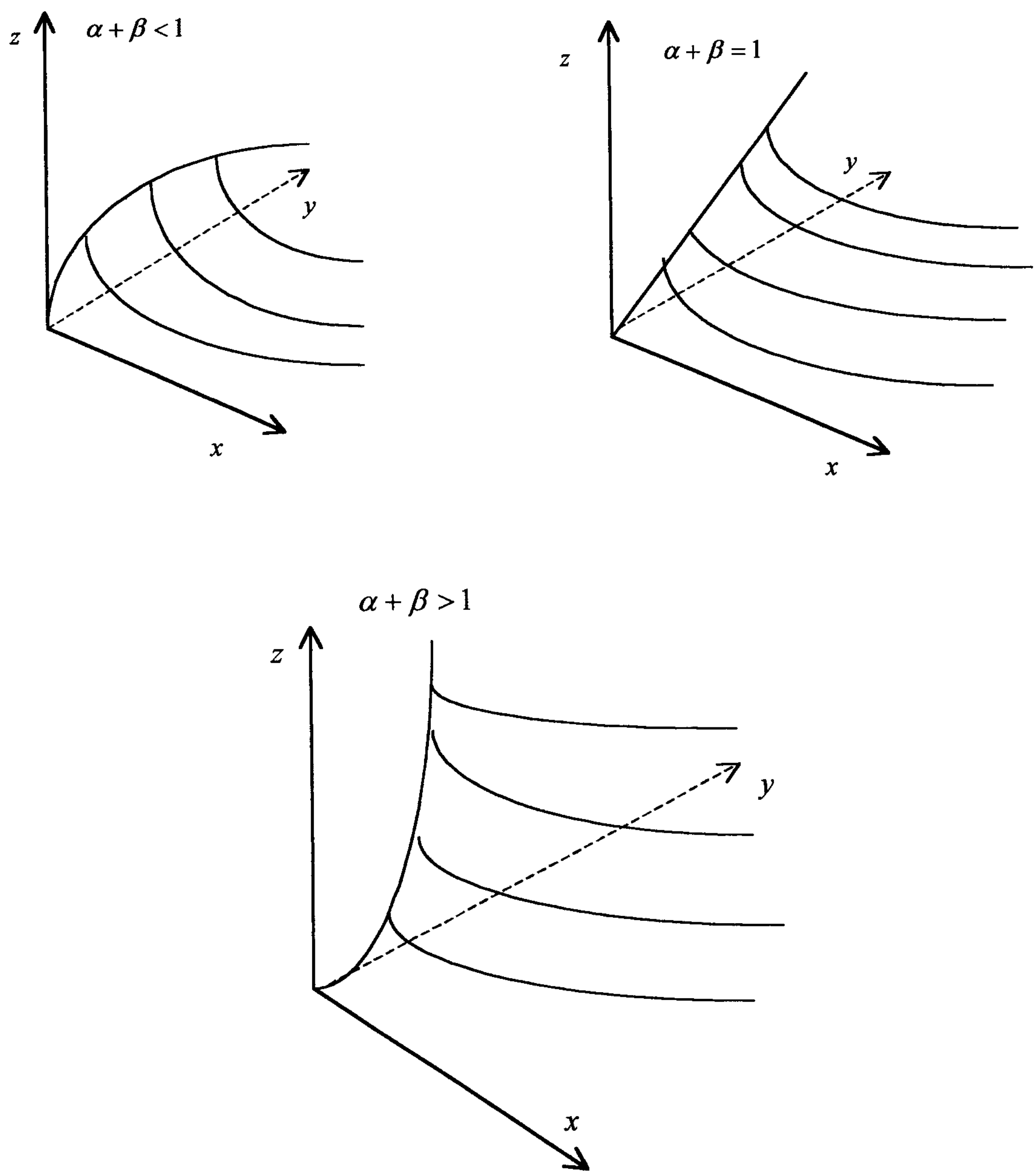


Figura 5.6: Gráfico de uma função de Cobb-Douglas, segundo os valores de α e β.

Exemplo: Encontrar o maior domínio $\mathcal{D}$ de $\mathbb{R}^2$ sobre o qual a função $f(x,y) = y\ln x - \sqrt{1-x^2-y^2}$ é de classe C^1.

Para que $\ln x$ esteja definido (e seja de classe C^1), é necessário que $x > 0$ e para que $-\sqrt{1-x^2-y^2}$ seja de classe C^1, é necessário que $1-x^2-y^2 > 0$. Considere-se portanto $\mathcal{D} = \{(x,y) \in \mathbb{R}^2,\ 1-x^2-y^2 > 0 \text{ e } x > 0\}$ que é um "semicírculo aberto". Para verificarmos que f é de classe C^1 sobre $\mathcal{D}$, observamos que f é a diferença de duas funções

$$(x,y) \mapsto y \ln x \text{ e } (x,y) \mapsto -\sqrt{1-x^2-y^2}$$

É suficiente assegurar que cada uma é de classe C^1. A primeira é o produto de duas funções $(x,y) \mapsto y$ (projecção sobre a segunda variável) e $(x,y) \mapsto \ln x$ (composta da primeira projecção e da função usual $x \mapsto \ln x$). A segunda é a composta de $(x,y) \mapsto 1-x^2-y^2$ e $u \mapsto -\sqrt{u}$ que são duas funções *usuais*.

5.6 Funções homogéneas

Definição: Seja $\mathcal{D} \subseteq \mathbb{R}^p$. Uma função $f : \mathcal{D} \to \mathbb{R}$, $(x_1,...,x_p) \mapsto f\,(x_1,...,x_p)$ diz-se *homogénea de grau* n se verifica, para todo $\lambda > 0$ e para todo $(x_1,x_2,...,x_p) \in \mathcal{D}$,

$$f(\lambda x_1, \lambda x_2, ..., \lambda x_p) = \lambda^n f(x_1, x_2, ..., x_p).$$

Vejamos alguns exemplos:

Exemplos:

- Os polinómios seguintes são homogéneos, respectivamente de grau $1, 2$ e 3 : $ax + by,\ ax^2 + bxy + cy^2,\ ax^3 + bx^2y + cxy^2 + dy^3$.
- As funções $\sqrt{x^2+y^2}$ e $\frac{1}{\sqrt{x^2+y^2}}$ são respectivamente homogéneas de grau 1 e -1.
- Se $f : \mathbb{R} \to \mathbb{R}$, então $\varphi(x,y) = f(x/y)$ é homogénea de grau 0.
- As funções e^{x+y}, $\ln(x^2+y^2)$ e $x^3 + x^3y^3 + y^3$ não são homogéneas.
- Se α e β são dois números reais, a função de Cobb-Douglas $(x,y) \mapsto (x^\alpha y^\beta)$ é uma função homogénea de grau $\alpha + \beta$ (ver figura 5.6.).

Nos modelos económicos, são frequentemente escolhidas funções de produção homogéneas. Diz-se que uma tal função é de rendimento de escala crescente, (respectivamente constante, respectivamente decrescente) quando o seu grau de homogeneidade é estritamente superior (respectivamente igual, respectivamente inferior) a 1. Assim uma função do tipo de Cobb-Douglas, isto é, da forma $f(x,y) = x^\alpha y^\beta$, tem rendimentos crescentes se $\alpha + \beta > 1$.

A homogeneidade de uma função traduz-se igualmente numa propriedade geométrica do seu gráfico: este é invariante para homotetias de centro $(0,...,0)$. Assim se se designar por $\|X\| = \|(x_1, x_2, ..., x_n)\|$ a norma euclideana $\left(\|X\| = \sqrt{x_1^2 + x_2^2 + ... + x_n^2}\right)$ e por $S = \{X \in \mathbb{R}^n : \|X\| = 1\}$ a esfera unitária, o conhecimento de uma função homogénea f em todo o ponto X da esfera S determina completamente f. Com efeito se $Y \in \mathbb{R}^n$ é um elemento qualquer não nulo, o vector $\frac{Y}{\|Y\|}$ pertence a S e portanto para calcular a imagem de Y por f, é suficiente conhecer a imagem deste vector de S:

$$f(Y) = f\left(\|Y\| \frac{Y}{\|Y\|}\right) = \|Y\|^\alpha f\left(\frac{Y}{\|Y\|}\right).$$

Proposição 5.4

Se $f : \mathbb{R}^n \to \mathbb{R}$ é de classe C^1 e homogénea de grau α, as suas derivadas parciais $\frac{\partial f}{\partial x_i}$ são funções homogéneas de grau $\alpha - 1$.

Dem.

Seja $f : \mathbb{R}^n \to \mathbb{R}$ uma função homogénea de grau α então verifica, para todo o $\lambda > 0$ e para todo o $(x_1, ..., x_n) \in \mathbb{R}^n$,

$$f(\lambda x_1, \lambda x_2, ..., \lambda x_n) = \lambda^n f(x_1, x_2, ..., x_n).$$

Como f é de classe C^1, por hipótese, temos

$$Grad\, f(\lambda x_1, \lambda x_2, ..., \lambda x_n) = \left(\frac{\partial f}{\partial x_1}(\lambda x_1, \lambda x_2, ..., \lambda x_n), ..., \frac{\partial f}{\partial x_n}(\lambda x_1, \lambda x_2, ..., \lambda x_n)\right)$$

logo, para todo o $i = 1, ..., n$, $\frac{\partial f}{\partial x_i}(\lambda x_1, \lambda x_2, ..., \lambda x_n)$ é uma função homogénea de grau $\alpha - 1$. □

A proposição seguinte chama-se *fórmula de Euler*:

Proposição 5.5

Se $f : \mathbb{R}^n \to \mathbb{R}$ é de classe C^1 e homogénea de grau α, então tem-se:

$$x_1 \frac{\partial f}{\partial x_1} + ... + x_n \frac{\partial f}{\partial x_n} = \alpha f.$$

Dem.

A demonstração resulta do lema seguinte, que se mostra directamente a partir da definição da propriedade de classe C^1 para funções com uma variável:

Lema 5.6

Sejam $x_1(t)$, $x_2(t)$, ... , $x_n(t)$ *n funções de classe* C^1 *de* $\mathbb{R}$ *em* $\mathbb{R}$ *e* $f: \mathbb{R}^n \to \mathbb{R}$ *uma função de classe* C^1. *Então a função* $g: t \mapsto f(\, x_1(t), x_2(t), ..., \, x_n(t))$ *é de classe* C^1, *com derivada*

$$g'(t) = \frac{\partial f}{\partial x_1} x_1'(t) + ... + \frac{\partial f}{\partial x_n} x_n'(t).$$

Dem. (do lema)

A função $g(t) = f(\, x_1(t), x_2(t), ..., \, x_n(t))$ é função da função f e das funções $x_1(t)$, $x_2(t)$, ... , $x_n(t)$, que são por hipótese funções de classe C^1, logo, pela proposição 5.3, g é de classe C^1 e tem como derivada a função:

$$g'(t) = \frac{\partial f}{\partial x_1} x_1'(t) + ... + \frac{\partial f}{\partial x_n} x_n'(t). \qquad \square$$

Para mostrar a fórmula de Euler, calculamos de duas formas diferentes a derivada da função $\lambda \mapsto g(\lambda)$ definida por $g(\lambda) = f(\lambda x_1, ..., \lambda x_n)$. Obtém-se :

$$x_1 \frac{\partial f}{\partial x_1} + ... + x_n \frac{\partial f}{\partial x_n} = \alpha \lambda^{\alpha - 1} f.$$

Logo a fórmula, pondo $\lambda = 1$. $\square$

Para uma função de produção com rendimentos de escala constante ($\alpha = 1$), se supusermos que o preço de uma unidade de produto $f(x_1, ..., x_n)$ é igual a 1 e cada factor x_i é exactamente recompensado pela sua produtividade marginal $\frac{\partial f}{\partial x_i}$, deve ter-se $\sum x_i \frac{\partial f}{\partial x_i} = f$, o que não é outra coisa senão a fórmula de Euler: chama-se neste contexto o *teorema do enfraquecimento dos produtos*.

Exemplo: Recorde que para a função f de duas variáveis de classe C^1, tem-se a fórmula,

$$f(x_0 + dx, y_0 + dy) - f(x_0, y_0) = f_x'(x_0, y_0)dx + f_y'(x_0, y_0)dy + \varepsilon_1 dx + \varepsilon_2 dy$$

onde ε_1 e ε_2 são infinitesimais e (x_0, y_0) é um ponto quase standard no domínio de f. Calcular para as seguintes funções, o crescimento $\triangle f$ e o diferencial df no ponto (x_0, y_0) indicado e considerando $dx = 0,01$ e $dy = 0,025$ e verificar que a diferença entre estas duas quantidades não é só pequena mas é muito pequena perante dx e dy.

a) $f : (x, y) \mapsto x^2 y \ , \ (x_0, y_0) = (1, 1)$.

$$df = f'_x(1,1)dx + f'_y(1,1)dy = 2 \times 0,01 + 0,025 = 0,045$$
$$\triangle f = f'_x(1,1)dx + f'_y(1,1)dy + \varepsilon_1 dx + \ \varepsilon_2 dy =$$
$$= 0,045 + 0,01\varepsilon_1 + 0,025\,\varepsilon_2 \simeq 0,045$$

porque ε_1 e ε_2 são infinitesimais.

b) $f : (x, y) \mapsto x^3 - 2y^3 + 3xy \ , \ (x_0, y_0) = (1, 2)$.

$$df = f'_x(1,2)dx + f'_y(1,2)dy = 9 \times 0,01 - 21 \times 0,025 = \ -0,435$$
$$\triangle f = f'_x(1,2)dx + f'_y(1,2)dy + \varepsilon_3 dx + \ \varepsilon_4 dy =$$
$$= \ -0,435 + 0,01\varepsilon_3 + 0,025\,\varepsilon_4 \simeq \ -0,435$$

porque ε_3 e ε_4 são infinitesimais.

c) $f : (x, y) \mapsto e^x + y \ , \ (x_0, y_0) = (0, 0)$.

$$df = f'_x(0,0)dx + f'_y(0,0)dy = 0,035$$
$$\triangle f = f'_x(0,0)dx + f'_y(0,0)dy + \varepsilon_3 dx + \ \varepsilon_4 dy =$$
$$= 0,035 + 0,01\varepsilon_5 + 0,025\,\varepsilon_6 \simeq 0,035$$

porque ε_5 e ε_6 são infinitesimais.

Nas alíneas a), b) e c) podemos concluir que em cada caso $df \simeq \triangle f$.

Capítulo 6

Crescimentos limitados

A derivada de uma função permite calcular uma aproximação linear, isto é, uma aproximação por um polinómio de grau 1. Este polinómio é de desenvolvimento limitado de ordem 1. Neste capítulo vamos ver como calcular *desenvolvimentos limitados* de ordem mais elevada (ver definição adiante), utilizando em particular as derivadas de segunda, terceira, ... , e como utilizar este tipo de desenvolvimento para melhor conhecer as funções.

6.1 O teorema dos acréscimos finitos

O teorema dos acréscimos finitos, também chamado teorema dos valores intermédios ou teorema de Bolzano, possui como caso particular o teorema de Rolle.

Teorema 6.1 (Teorema de Rolle)

Seja $f : \mathcal{D} \to \mathbb{R}$ *uma função de classe* C^1 *e sejam* $a \in \mathcal{D}$ *e* $b \in \mathcal{D}$ *tais que* $[a, b] \subset \mathcal{D}$ *e* $f(a) = f(b)$. *Então existe* $c \in \left]a, b\right[$ *tal que* $f'(c) = 0$.

A demonstração deste teorema utiliza o teorema de *Weierstrass*.

Teorema 6.2 (teorema de Weierstrass)

Toda a função real contínua f *num domínio* $\mathcal{D}$ *fechado e limitado tem um maximizante e um minimizante*[6] *pertencentes a* $\mathcal{D}$.

Dem. (teorema de Weierstrass)

Por transferência, é suficiente demonstrar o teorema para o caso em que f (e portanto $\mathcal{D}$) é standard. Para simplificar, suponhamos que $\mathcal{D}$ é um intervalo fechado $[a, b]$ (e portanto limitado!). Seja N infinitamente grande e $a = x_0 < x_1 < ... < x_N = b$ uma partição fina do intervalo $[a, b]$ (v. página 33). Consideremos $y_i = f(x_i)$, $i = 0, ..., N$ e sejam $y_p = \max_{i=0,...,N} y_i$ e x_p tais que

[6] Seja $f : \mathcal{D} \to \mathbb{R}$, diz-se que x' é um maximizante em $\mathcal{D}$ sse para todo o x pertencente a $\mathcal{D}$, $f(x') \geq f(x)$. Diz-se que x^* é um minimizante em $\mathcal{D}$ sse para todo o x pertencente a $\mathcal{D}$, $f(x^*) \leq f(x)$.

$f(x_p) = y_p$. Como x_p pertence ao intervalo (fechado e limitado) $[a,b]$, tem uma sombra ${}^\circ x_p$ que pertence também a $[a,b]$. Representemos por $\overline{x}$ esta sombra e mostremos que este número é o maximizante procurado. Tem-se por construção, para todo $n = 0, ..., N$, $f(x_n) \leq f(x_p)$. Trata-se de verificar que, para todo o $x \in [a,b]$, se tem $f(x) \leq f(\overline{x})$. Por transferência, é suficiente verificar para os x standard. Seja $x \in [a,b]$ standard. Existe $n \in \{0, ..., N\}$ tal que $x \simeq x_n$. Como f é contínua, $f(x) \simeq f(x_n) \leq f(\overline{x})$, portanto $f(x) \lesssim f(\overline{x})$. Mas como estes dois números são standard, isto conduz a que $f(x) \leq f(\overline{x})$. □

Dem. (Teorema de Rolle)

Se $f(x) = f(a)$ para todo o $x \in [a,b]$, o teorema é evidente (porque a derivada de uma função constante é nula), caso contrário a função f toma pelo menos num ponto $x_0 \in [a,b]$ um valor diferente de $f(a)$, estritamente superior ou estritamente inferior a $f(a)$. Vejamos o primeiro caso, o segundo é análogo. A função f é de classe C^1, então em particularmente é contínua em $[a,b]$ e pode-se portanto aplicar o teorema de Weierstrass. Portanto existe $c \in [a,b]$ tal que $f(c) = \max\limits_{x \in [a,b]} f(x)$. Como $f(c) \geq f(x_0) > f(a) = f(b)$, c não pode ser nem a nem b. Portanto, $c \in \,]a,b[$. Para terminar a demonstração, vamos mostrar, que neste ponto, se tem $f'(c) = 0$. Isto resulta do lema seguinte:

Lema 6.3

Se $f : \,]a,b[\,\to \mathbb{R}$ é uma função de classe C^1, que tem um extremo num ponto $c \in \,]a,b[$, então $f'(c) = 0$.

Dem.

Por transferência, podemos supor f e c standard. Suponhamos que c é um maximizante (o caso do minimizante é similar), isto é, suponhamos que, para todo o $x \in \,]a,b[$, tem-se $f(x) \leq f(c)$. Portanto a razão $\frac{f(x)-f(c)}{x-c}$ é negativa se $x > c$ e positiva se $x < c$. Como por definição $f'(c) = {}^\circ\left(\frac{f(x)-f(c)}{x-c}\right)$ quando $x \simeq c$, $f'(c)$ é um número standard infinitamente próximo de um número positivo e de um número negativo. Então só pode ser nulo. □

Teorema 6.4 (teorema dos acréscimos finitos)

Seja $f : \mathcal{D} \to \mathbb{R}$ uma função de classe C^1 e sejam $a \in \mathcal{D}$ e $b \in \mathcal{D}$ tais que $[a,b] \subset \mathcal{D}$. Então existe $c \in \,]a,b[$ tal que

$$f(b) - f(a) = (b - a)f'(c)$$

ou melhor (enunciado equivalente sendo $h = b - a$) existe $t \in \left]0, 1\right[$ tal que

$$f(a + h) - f(a) = hf'(a + th).$$

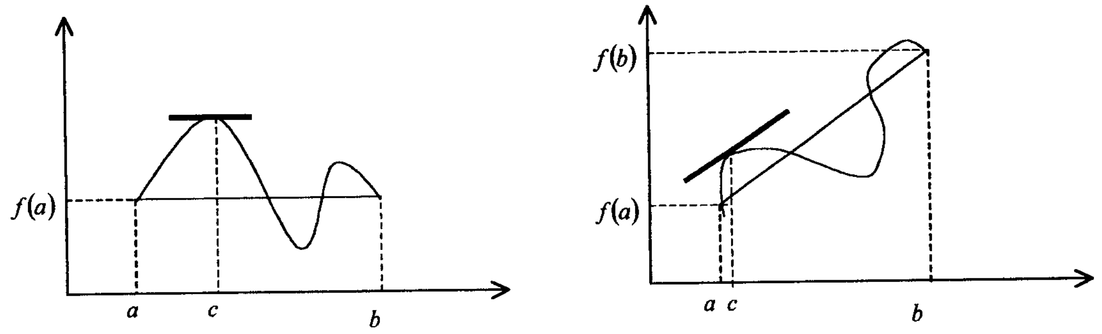


Figura 6.1 Ilustração do teorema de Rolle (à esquerda) e da igualdade dos acréscimos finitos (à direita).

Observação: No nome desta fórmula (acréscimos finitos), o adjectivo "finito" significa não infinitesimal (dizemos agora apreciável): a quantidade $f(a + h) - f(a)$ é um acréscimo finito de f no ponto a, enquanto a quantidade $f(a + dx) - f(a)$, utilizada na definição de derivada, é um acréscimo infinitesimal de f no ponto a.

Corolário 6.5

Seja $f : \mathcal{D} \to \mathbb{R}$ uma função de classe C^1 e sejam $a \in \mathcal{D}$ e $b \in \mathcal{D}$ tais que $[a, b] \subset \mathcal{D}$. Então

- *Se para todo o $x \in [a, b]$, $f'(x) \geq 0$, então f é crescente.*
- *Se para todo o $x \in [a, b]$, $f'(x) = 0$, então f é constante.*
- *Se para todo o $x \in [a, b]$, $f'(x) \leq 0$, então f é decrescente.*

6.2 A fórmula de Taylor para as funções com uma variável

Antes de introduzir a fórmula de Taylor que permite calcular os principais exemplos de desenvolvimentos limitados, é conveniente precisar o que se chamam "derivadas de ordem superior" de uma função.

Definição: Diz-se que uma função $f : \mathcal{D} \subset \mathbb{R} \to \mathbb{R}$ é de classe C^2 sse

- f é de classe C^1, de derivada f';

- f' é de classe C^1.

A derivada de f' chama-se segunda derivada de f e representa-se por f'' ou $f^{(2)}$.

Por indução matemática, diz-se que f é de classe C^n se f é de classe $C^{(n-1)}$ de $(n-1)$ iésima derivada $f^{(n-1)}(x)$ e se $f^{(n-1)}$ é de classe C^1. Uma função é de classe C^∞, ou simplesmente "suave" se é de classe C^n para todo o n. Os polinómios e a função e^x são exemplos de funções "suaves".

Exemplo: O exemplo seguinte é útil para compreender a ligação que existe entre os valores em zero das derivadas sucessivas de um polinómio e os coeficientes deste polinómio.

Seja $P(x) = a_0 + a_1 x + ... + a_n x^n$ um polinómio de grau n. Calculemos as derivadas sucessivas, $P'(x), P''(x), ...$. Observemos que $P(0) = a_0$, $P'(0) = a_1$, $P''(0) = (2!)a_2$, e mais geral, $P^{(k)}(0) = (k!)a_k$, para todo o $k \leq n$ e $P^{(k)}(0) = 0$ para $k > n$. Se conhecermos as n primeiras derivadas em $x = 0$ de um polinómio de grau n, podemos portanto calcular o valor deste polinómio em todo o ponto $x \neq 0$ através da seguinte fórmula:

$$P(x) = P(0) + \frac{P'(0)}{1!}x + \frac{P''(0)}{2!}x^2 + ... + \frac{P^{(n)}(0)}{n!}x^n.$$

Teorema 6.6 (fórmula de Taylor)

Seja $f : \mathcal{D} \to \mathbb{R}$ de classe C^{n+1} e sejam $a \in \mathcal{D}$ e $b \in \mathcal{D}$ tais que $[a,b] \subseteq \mathcal{D}$. Então existe $c \in \left]a,b\right[$ tal que :

$$f(b) = f(a) + (b-a)f'(a) + \frac{(b-a)^2}{2!}f''(a) + ... +$$

$$+ \frac{(b-a)^n}{n!}f^{(n)}(a) + \frac{(b-a)^{n+1}}{(n+1)!}f^{(n+1)}(c)$$

ou melhor (enunciado equivalente considerando $h := b-a$):
Existe $t \in \left]0,1\right[$ tal que:

$$f(a+h) = f(a) + hf'(a) + ... + \frac{h^n}{n!}f^{(n)}(a) + \frac{h^{n+1}}{(n+1)!}f^{(n+1)}(a+th).$$

Observação: Na expressão anterior, os n primeiros termos formam um polinómio de grau, no máximo, n, chamado *polinómio de Taylor de grau n de f em $x = a$*. O afastamento entre a função no ponto $a + h$ e o seu polinómio de Taylor é a quantidade:

$$R(h) = \frac{h^{n+1}}{(n+1)!} f^{(n+1)}(a+th)$$

que se chama *resto de ordem* $n+1$.

O desenvolvimento de Taylor de uma função f no ponto particular $a=0$ chama-se, por vezes, *desenvolvimento de Mac Laurin*.

Dem.

Consideremos $\varphi(x) = f(b) - f(x) - (b-x)f'(x) - .... - \frac{(b-x)^n}{n!} f^n(x) - \frac{(b-x)^{n+1}}{(n+1)!} A$, onde A é uma constante escolhida de tal forma que $\varphi(a) = 0$. Trata-se de mostrar a existência de um $c \in \left]a,b\right[$ tal que $A = f^{(n+1)}(c)$. Vamos para isto aplicar o teorema de Rolle à função φ. Tem-se, por definição, $\varphi(a) = 0$ e claramente $\varphi(b) = 0$. Por outro lado, a função φ é derivável sobre $[a,b]$com derivadas nos extremos a e b, e de classe C^1 sobre $\left]a,b\right[$ pelas hipóteses consideradas sobre f. Portanto, existe um c $\in \left]a,b\right[$ tal que $f'(c) = 0$. Ora o cálculo da derivada $\varphi'(x)$ mostra que $\varphi'(x) = \frac{(b-x)^n}{n!}\left(f^{(n+1)}(x) - A\right)$. Mas, como $b \neq c$, a igualdade $\varphi'(c) = 0$ implica que $A = f^{(n+1)}(c)$. □

Exemplos: Verifica-se facilmente que os polinómios de Taylor de grau n, $2n+1$ e $2n$, respectivamente, das funções exponencial, seno e coseno são os seguintes:

$$P_{\exp}(h) = 1 + \frac{h}{1!} + \frac{h^2}{2!} + ... + \frac{h^n}{n!};$$

$$P_{\text{sen}}(h) = h - \frac{h^3}{3!} + ... + (-1)^n \frac{h^{2n+1}}{(2n+1)!};$$

$$P_{\cos}(h) = h - \frac{h^2}{2!} + ... + (-1)^{n+1} \frac{h^{2n}}{(2n)!}.$$

6.3 Desenvolvimentos limitados

Definição: Diz-se que um polinómio standard $p: \mathbb{R} \to \mathbb{R}$ de grau n é um *desenvolvimento limitado de ordem* n da função standard $f: \left]a,b\right[\to \mathbb{R}$ no ponto standard $x_0 \in \left]a,b\right[$ se existe, para todo o $x \simeq x_0$, um número ε infinitesimal tal que:

$$f(x) = p(x) + (x - x_0)^n \varepsilon.$$

Um tal desenvolvimento limitado $p(x)$ existe, é único.

Em geral, para calcular o desenvolvimento limitado de uma função, exprime-se esta função como soma, produto, quociente ou composição de funções, e calcula-se o desenvolvimento procurado a partir de desenvolvimentos "conhecidos".

Vejamos uma lista de alguns desenvolvimentos que é útil conhecer:

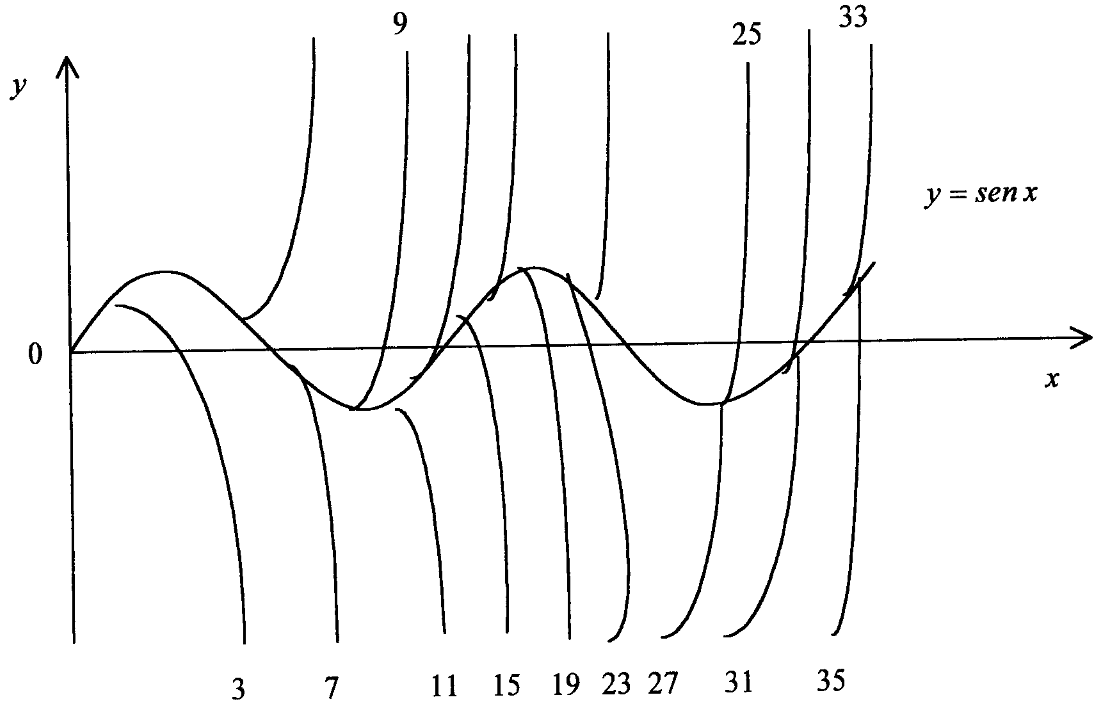


Figura 6.2 Gráficos da função seno e dos seus polinómios de Taylor em $x = 0$ de grau 3 a 35.

- $e^x = 1 + \frac{x}{1!} + \frac{x^2}{2!} + ... + \frac{x^n}{n!} + x^n\varepsilon$;
- $\cos(x) = 1 - \frac{x^2}{2!} + ... + (-1)^n \frac{x^{2n}}{(2n)!} + x^{2n}\varepsilon$;
- $\text{sen}(x) = x - \frac{x^3}{3!} + ... + (-1)^n \frac{x^{2n+1}}{(2n+1)!} + x^{2n+1}\varepsilon$;
- $\ln(1+x) = x - \frac{x^2}{2} + \frac{x^3}{3} + ... + (-1)^{n-1}\frac{x^n}{n} + x^n\varepsilon$;
- $\frac{1}{1+x} = 1 - x + x^2 - ... + (-1)^n x^n + x^n\varepsilon$;
- $\sqrt{1+x} = 1 + \frac{x}{2} - \frac{x^2}{8} + ... + (-1)^{n-1}\frac{(2n-2)!}{2^{2n-1}(n-1)!n!}x^n + x^n\varepsilon$;
- $\frac{1}{\sqrt{1+x}} = 1 - \frac{x}{2} + \frac{3x^2}{8} + ... + (-1)^n \frac{(2n)!}{2^{2n}(n!)^2}x^n + x^n\varepsilon$;
- Os três desenvolvimentos anteriores são casos particulares do seguinte:

$$(1+x)^\alpha = 1 + \frac{\alpha}{1!} + \frac{\alpha(\alpha-1)}{2!}x^2 + ... + \frac{\alpha(\alpha-1)...(\alpha-n+1)}{n!}x^n + x^n\varepsilon.$$

O cálculo algébrico com os desenvolvimentos limitados (soma, produto, quociente e composição) não representa problema particular se utilizarmos a fórmula exacta, compreendendo o "resto". O polinómio quociente obtém-se através de uma divisão segundo as potências crescentes.

Se uma função de classe C^n tem um desenvolvimento limitado de ordem n, este corresponde à fórmula de Taylor de ordem n (por unicidade) e obtém-se, neste caso, o desenvolvimento limitado de ordem $n-1$ da sua derivada f' derivando simplesmente termo a termo o de f.

Se uma função integrável tem um desenvolvimento limitado de ordem n, a sua primitiva tem um desenvolvimento limitado de ordem $n+1$ obtido do de f integrando termo a termo.

6.4 Derivadas parciais de ordem superior

Definição: Seja $f : \mathcal{D} \subseteq \mathbb{R}^2 \to \mathbb{R}$ uma função de classe C^1. Se as duas derivadas parciais de f, $\frac{\partial f}{\partial x}(x,y)$ e $\frac{\partial f}{\partial y}(x,y)$ são de classe C^1, diz-se que f é de classe C^2. As duas derivadas parciais de $\frac{\partial f}{\partial x}$ e de $\frac{\partial f}{\partial y}$ são, por definição, as quatro segundas derivadas de f (ou derivadas parciais de ordem dois) e representam-se respectivamente,

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

Por indução matemática, diz-se que f é de classe C^n se f é de classe $C^{(n-1)}$, e se todas as suas derivadas parciais de ordem $(n-1)$ são de classe C^1. Uma função é de classe C^∞, ou simplesmente "suave", se é de classe C^n para todo o n. Os polinómios ou a função $(x,y) \to e^{x+y}$ são exemplos de funções "suaves".

Notação: Recordemos que o vector gradiente é o vector formado pelas primeiras derivadas parciais de uma função f, e representa-se por:

$$Grad\ f(x,y) = \left(\frac{\partial f}{\partial x}(x,y), \frac{\partial f}{\partial y}(x,y)\right) = \nabla f(x,y).$$

Chama-se *matriz hessiana de f* à matriz 2×2 formada pelas segundas derivadas parciais de uma função f, e representa-se por:

$$Hess\ f(x,y) = \begin{pmatrix} \frac{\partial^2 f}{\partial x^2} & \frac{\partial^2 f}{\partial y \partial x} \\ \frac{\partial^2 f}{\partial x \partial y} & \frac{\partial^2 f}{\partial y^2} \end{pmatrix}.$$

Chama-se *determinante hessiano de f* ao determinante da matriz hessiana:

$$hess\ f(x,y) = \begin{vmatrix} \frac{\partial^2 f}{\partial x^2} & \frac{\partial^2 f}{\partial y \partial x} \\ \frac{\partial^2 f}{\partial x \partial y} & \frac{\partial^2 f}{\partial y^2} \end{vmatrix} = \frac{\partial^2 f}{\partial x^2} \cdot \frac{\partial^2 f}{\partial y^2} - \frac{\partial^2 f}{\partial x \partial y} \cdot \frac{\partial^2 f}{\partial y \partial x}$$

Exemplo: Seja $f(x,y) = e^{2x} \operatorname{sen} y$, tem-se:

$$Grad\ f(x,y) = \left(2e^{2x} \operatorname{sen} y\ ,\ e^{2x} \cos y\right)$$

e $$Hess\ f(x,y) = \begin{pmatrix} 4e^{2x}\ \text{sen}\, y & 2e^{2x}\cos y \\ 2e^{2x}\cos y & -\,e^{2x}\ \text{sen}\, y \end{pmatrix}.$$

A demonstração do resultado seguinte é um pouco técnica, mas proporciona a utilização do teorema dos acréscimos finitos.

Teorema 6.7 (lema de Schwarz)

Uma função $f(x,y)$ de classe C^2 sobre $\mathcal{D} \subseteq \mathbb{R}^2$ verifica, em todo o ponto (x,y) de $\mathcal{D}$, $\frac{\partial^2 f}{\partial x \partial y}(x,y) = \frac{\partial^2 f}{\partial y \partial x}(x,y)$.

Dem.

Por transferência, podemos supor f standard e só mostrar a igualdade das duas segundas derivadas "cruzadas" de f num ponto standard (x_0, y_0) do domínio da função.

Definem-se duas funções, F e G por:

$$F(x) = f(x, y_0 + k) - f(x, y_0) \quad , \quad G(y) = f(x_0 + h, y) - f(x_0, y).$$

Estas funções são de classe C^2 por hipótese. Recorde então que a quantidade,

$$f(x_0 + h, y_0 + k) - f(x_0 + h, y_0) - f(x_0, y_0 + k) + f(x_0, y_0)$$

é igual a $F(x_0 + h) - F(x_0)$ e $G(y_0 + k) - G(y_0)$. Pode-se, portanto, avaliar de duas formas, utilizando duas vezes o teorema dos acréscimos finitos (t e t_1 são elementos de $[0,1]$, assim como θ e θ_1):

$$F(x_0 + h) - F(x_0) = hF'(x_0 + th) = h(f'_x(x_0 + th, y_0 + k) - f'_x(x_0 + th, y_0)) =$$

$$= hk f''_{xy}(x_0 + th, y_0 + t_1 k)$$

$$G(y_0 + k) - G(y_0) = kG'(y_0 + \theta k) = k\big(f'_y(x_0 + h, y_0 + \theta k) - f'_y(x_0, y_0 + \theta k)\big) =$$

$$= kh f''_{yx}(x_0 + \theta_1 h, y_0 + \theta k).$$

Logo,

$$f''_{xy}(x_0 + th, y_0 + t_1 k) = f''_{yx}(x_0 + \theta_1 h, y_0 + \theta k).$$

Como f é de classe C^2, por hipótese, ter-se-á

$$f''_{xy}(x_0 + th, y_0 + t_1 k) \simeq f''_{xy}(x_0, y_0)$$

e

$$f''_{yx}(x_0 + \theta_1 h, y_0 + \theta k) \simeq f''_{yx}(x_0, y_0).$$

Com efeito, quando uma função é de classe C^2 as suas primeiras e segundas derivadas parciais são contínuas, e, como estabelecemos anteriormente, isto traduz-se precisamente pelas aproximações anteriores provando que h e k são infinitesimais. Resulta que:

$$f''_{xy}(x_0, y_0) \simeq f''_{yx}(x_0, y_0),$$

logo, a igualdade procurada, visto que as duas quantidades infinitamente próximas são standard (portanto iguais). □

6.5 Fórmula de Taylor para uma função de duas variáveis

Neste parágrafo restringimo-nos a estabelecer a fórmula de Taylor de ordem dois. Para tornar menos pesada a escrita, designamos por M_0 um ponto do domínio da função de coordenadas (x_0, y_0) e por M um ponto do domínio da função de coordenadas (x, y); um ponto qualquer do segmento M_0M pode, então, escrever-se M_t, onde $t \in [0,1]$, e as suas coordenadas são $(x_0 + t(x - x_0), y_0 + t(y - y_0))$.

Teorema 6.8

Seja f de classe C^2 sobre um aberto $\mathcal{D}$ e seja $M_0 \in \mathcal{D}$ e $M \in \mathcal{D}$ tal que o segmento M_0M está contido em $\mathcal{D}$. Então existe $t \in [0,1]$ tal que:

$$f(M) = f(M_0) + \frac{\partial f}{\partial x}(M_0)(x - x_0) + \frac{\partial f}{\partial y}(M_0)(y - y_0) + \frac{1}{2}\frac{\partial^2 f}{\partial x^2}(M_t)(x - x_0)^2 +$$

$$+ \frac{\partial^2 f}{\partial x \partial y}(M_t)(x - x_0)(y - y_0) + \frac{1}{2}\frac{\partial^2 f}{\partial y^2}(M_t)(y - y_0)^2.$$

Exemplo: Sejam f e g duas funções de classe C^1 em $\mathcal{D}$ e $[a, b] \subseteq \mathcal{D}$. Suponhamos que $g(a) \neq g(b)$ e que f' e g' não têm zeros comuns em $[a, b]$. Então existe $\xi \in]a, b[$ tal que $\frac{f(b)-f(a)}{g(b)-g(a)} = \frac{f'(\xi)}{g'(\xi)}$. Determinar ξ nos seguintes casos:

1) $f(x) = x^2 + 2$, $g(x) = x^3 - 1$, $[a, b] = [1, 2]$

$$\frac{f(2)-f(1)}{g(2)-g(1)}=\frac{f'(\xi)}{g'(\xi)}\Leftrightarrow\frac{3}{7}=\frac{2\xi}{3\xi^2}\Leftrightarrow\xi=0\vee\xi=\frac{14}{9}$$

logo $\xi=\frac{14}{9}$ porque $\xi\in\,]1,2[$.

2) $f(x)=\operatorname{sen}x$, $g(x)=\cos x$, $[a,b]=[0,\frac{\pi}{2}]$

$$\frac{f(\frac{\pi}{2})-f(0)}{g(\frac{\pi}{2})-g(0)}=\frac{f'(\xi)}{g'(\xi)}\Leftrightarrow\ -1=\ -\operatorname{cotg}\xi\Leftrightarrow\xi=\frac{\pi}{4}.$$

Capítulo 7

Integrais e primitivas

Neste capítulo indicamos como calcular áreas aproximando-as por reuniões de rectângulos de áreas infinitesimais, o que permite precisar o que se chama uma *função integrável*. Depois recorda-se a noção de primitiva de uma função e, através do *teorema fundamental do cálculo*, estuda-se a ligação que existe entre o cálculo das primitivas e o das áreas.

7.1 Cálculo de uma área através das somas de Riemann

Seja $f : I \to \mathbb{R}$ uma função, suposta, num primeiro tempo, não negativa, e seja R a região, compreendida entre o gráfico de f e um segmento $[a,b] \subseteq I$ do eixo dos xx (figura 7.1). Em seguida será calculada a área de uma tal região em função de f, a, e b.

Perante um problema para o qual não se vê imediatamente a solução, o matemático procede frequentemente da seguinte forma: define a noção que procura construir, neste caso a área, e faz a lista das propriedades que lhe desejar atribuir. É a técnica do caderno de encargos. Depois procura construir um objecto que satisfaça estas propriedades.

Definição: Seja I um intervalo de $\mathbb{R}$ e seja $f : I \to \mathbb{R}^+$. Chama-se *função de área de* f a uma função $A_f(a,b)$ definida para todo o par ordenado $(a,b) \in I^2$ com valores em $\mathbb{R}_0^+$ que verifica as propriedades seguintes:

1. *Propriedade da adição ou da aditividade*:
$$a \leq c \leq b \Rightarrow A_f(a,b) = A_f(a,c) + A_f(c,b)$$

2. *Propriedade do rectângulo:* se m e M são respectivamente o máximo e o mínimo de f sobre $[a,b]$: $m\ (b-a) \leq A_f(a,b) \leq M(b-a)$.

3. *Propriedade do sinal*: para todo o $a, b \in I$, $A_f(b,a) = -A_f(a,b)$ (e portanto $A_f(a,a) = 0$).

As duas primeiras propriedades são naturais. A terceira pode-se compreender pelo facto de que se deseja que a propriedade da adição seja verdadeira para todos os a,b,c e não só para os $a \leq c \leq b$.

Observação: A definição anterior estende-se facilmente ao caso de uma função f de qualquer sinal mediante a convenção de sinal seguinte: se f é menor ou igual a zero sobre $[a',b']$, $\int_{a'}^{b'} f(x)dx := -\int_{a'}^{b'} |f(x)|dx$. Isto quer dizer que se conta negativamente as áreas situadas abaixo do eixo dos xx.

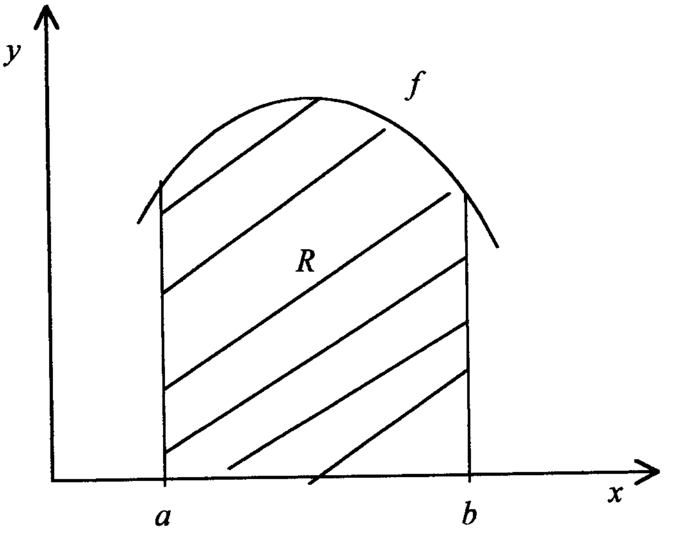


Figura 7.1 Área da região R.

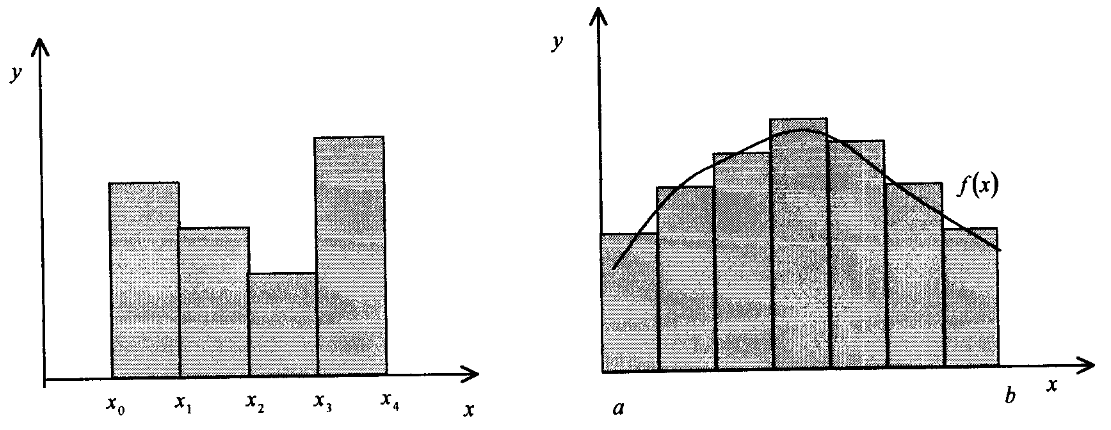


Figura 7.2 Área de uma função em escada e de uma função próxima de uma função em escada.

7.1.1 Caso das funções em escada

Para estas funções é fácil calcular o número $A_f(a,b)$.

Definição: Chama-se *partição* ou *subdivisão* de um intervalo $[a,b]$ a toda a sucessão (x_i) de $n+1$ elementos de I tal que $a = x_0 < x_1 < ... < x_n = b$. Quando, para todo o $i \in \{0,...,n-1\}$, $x_i \simeq x_{i+1}$, diz-se que a partição é infinitamente fina.

Definição: Uma função $f:\left[a,b\right]\to\mathbb{R}$ é uma *função em escada* se existe uma partição (x_i) de $\left[a,b\right]$ tal que f seja constante em cada intervalo $\left]x_i,x_{i+1}\right[$. Os valores que f toma nos pontos x_i não são especificados.

Definição: Para uma função em escada f que toma os valores c_i nos intervalos $\left]x_i,x_{i+1}\right[$ da partição, põe-se:

$$A_f(a,b)=\sum_{i=1}^{n-1}c_i(x_{i+1}-x_i).$$

É uma função de área que se representa por $\int_a^b f(x)dx$ e se chama *integral de* f sobre $\left[a,b\right]$ (figura7.2).

7.1.2 Caso de uma função qualquer

No caso geral, a ideia é aproximar a função (suposta standard) por uma função em escada (não-standard) definida sobre uma patirção infinitamente fina de tal forma que a sua área seja infinitamente próxima da área procurada. É suficiente então considerar a sombra.

Definição: Uma função $f:\left[a,b\right]\to\mathbb{R}$ standard diz-se *integrável no sentido de Riemann* (ou simplesmente integrável) se existe uma partição infinitamente fina (x_i) e duas funções em escada $g(x)$ e $h(x)$ sobre esta partição, tais que, para todo o $x\in I$, $g(x)\leq f(x)\leq h(x)$ e $\int_a^b g(x)dx\simeq\int_a^b h(x)dx$. Define-se então o *integral de* f *sobre* $\left[a,b\right]$ por:

$$\int_a^b f(x)dx={}^{\circ}\left(\sum_{i=0}^{n-1}f(\xi_i)(x_{i+1}-x_i)\right),$$

com $\xi_i\in\left[x_i,x_{i+1}\right]$ para $i=0,...,n-1$.

A soma, da qual é considerada a sombra, chama-se soma de Riemann de f. Esta soma é limitada quando f é integrável (e portanto tem uma sombra) visto que a função f é standard e limitada (por g e h), portanto limitada por números standard, e portanto com valores limitados. Assim tem-se:

$$\left|\sum_{i=0}^{n-1}f(\xi_i)(x_{i+1}-x_i)\right|\leq|\max_{i\in\{1,...,n-1\}}f(\xi_i)|\sum_{i=0}^{n-1}(x_{i+1}-x_i)=$$

$$= \left| \max_{i \in \{1,\dots,n-1\}} f(\xi_i) \right| (b-a).$$

Por outro lado, esta soma não depende da escolha dos pontos ξ_i nos intervalos $[x_i, x_{i+1}]$ visto que as desigualdades $g(\xi_i) \leq f(\xi_i) \leq h(\xi_i)$ conduzem a $\int_a^b g(x)dx \leq \sum f(\xi_i)(x_{i+1} - x_i) \leq \int_a^b h(x)dx$ e portanto,

$$\int_a^b f(x)dx = {}^{o}\left(\int_a^b g(x)dx\right) = {}^{o}\left(\int_a^b h(x)dx\right).$$

A proposição seguinte mostra que, além disso, o integral de uma função integrável também não depende da partição (x_i).

Proposição 7.1

Se f é uma função standard integrável, então o integral definido anteriormente (pela sombra da soma de Riemann) é uma função de área e é a única função de área associada a f sobre $[a,b]$.

Dem.

Mostremos as três propriedades que caracterizam uma função de área:

1. Sejam a, b e c números reais tais que, $a \leq c \leq b$, e suponhamos que f é integrável em qualquer dos intervalos $[a,c]$ e$[c,b]$. Temos, por definição, $A_f(a,c) = \int_a^c f(x)dx$ e $A_f(c,b) = \int_c^b f(x)dx$, assim,

$$\int_a^c f(x)dx + \int_c^b f(x)dx = \int_a^b f(x)dx$$

então, f é integrável em $[a,b]$ e temos, portanto,

$$A_f(a,b) = A_f(a,c) + A_f(c,b).$$

2. Sejam f uma função integrável em $[a,b], m$ e M, respectivamente, o mínimo e o máximo de f no intervalo $[a,b]$, portanto, $m \leq f(x) \leq M$, logo,

$$\int_a^b m dx \leq \int_a^b f(x)dx \leq \int_a^b M dx.$$

Portanto,

$$m(b-a) \leq A_f(a,b) \leq M(b-a).$$

3. Pela propriedade da adição (propriedade 1), temos,

$$\int_a^a f(x)dx = \int_a^b f(x)dx + \int_b^a f(x)dx,$$

mas, $\int_a^a f(x)dx = 0$, logo,

$$0 = \int_a^b f(x)dx + \int_b^a f(x)dx$$

então

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

Verifiquemos, agora, a unicidade: suponhamos que $A_f^1(a,b)$ e $A_f^2(a,b)$ são duas funções de área e suponhamos que g e h são as duas funções consideradas em escada sobre (x_i) que enquadram f. Tem-se, pela propriedade do rectângulo, para $j=1$ e $j=2$,

$$\int_{x_i}^{x_{i+1}} g(x)dx \leq A_f^j(x_i, x_{i+1}) \leq \int_{x_i}^{x_{i+1}} h(x)dx$$

sobre cada intervalo $[x_i, x_{i+1}]$ e, portanto, também sobre o intervalo $[a,b]$, pela propriedade da adição. Portanto $A_f^1(a,b) \simeq A_f^2(a,b)$, visto que os integrais de g e h são infinitamente próximos, o que conduz à igualdade visto que estas duas quantidades são standard. □

Proposição 7.2

Toda a função contínua $f : [a,b] \to \mathbb{R}$ é integrável.

Dem.

É utilizado de novo, nesta demonstração, o teorema do valor extremo ou teorema de Weierstrass (teorema 6.2). Este teorema afirma que se $f : [a,b] \to \mathbb{R}$ é uma função contínua, então existe $x_M \in [a,b]$, tal que, para todo o $x \in [a,b]$, $f(x) \leq f(x_M)$ e existe $x_m \in [a,b]$, tal que, para todo o $x \in [a,b]$, $f(x) \geq f(x_m)$. Os pontos x_M e x_m chamam-se respectivamente maximizante e minimizante de f em $[a,b]$.

Por transferência, é suficiente mostrar o teorema para uma função f standard. Escolhemos uma partição infinitamente fina (x_i) qualquer e definimos, para todo o $x \in [x_i, x_{i+1}]$, $g(x) = m_i = \inf_{x \in [x_i, x_{i+1}]} f(x)$ e $h(x) = M_i = \sup_{x \in [x_i, x_{i+1}]} f(x)$. As quantidades m_i e M_i existem, em virtude do teorema de Weierstrass porque f é contínua; mais $m_i \simeq M_i$ visto que f é S-contínua[7] e $x_i \simeq x_{i+1}$. Tem-se evidentemente $g(x) \leq f(x) \leq h(x)$ por construção e, portanto,

[7] Recordemos que a S-continuidade em $[a,b]$ se exprime simplesmente pelo facto de que para todos os $x, y \in [a,b]$, $x \simeq y \Rightarrow f(x) \simeq f(y)$. Utilizaremos de novo a S-continuidade para formular o teorema fundamental no parágrafo seguinte.

$$\int_a^b h(x)dx - \int_a^b g(x)dx = \sum_{i=0}^{n-1}(M_i - m_i)(x_{i+1} - x_i) \leq$$

$$\leq \varepsilon \sum_{i=0}^{n-1}(x_{i+1} - x_i) = \varepsilon(b-a) \simeq 0$$

onde ε é o maior dos números $M_i - m_i$, que é infinitesimal visto que cada um destes números o é. Então

$$\int_a^b g(x)dx \simeq \int_a^b h(x)dx,$$

e como $g(x) \leq f(x) \leq h(x)$ em $[a,b]$ e, g e h são funções em escada, f é integrável, pela definição. □

Proposição 7.3

Toda a função $f : [a,b] \to \mathbb{R}$ *monótona é integrável.*

Dem.

Por transferência, podemos supor f standard. Se $f(a) = f(b)$, f é constante em $[a,b]$ e portanto integrável. Suponhamos, por exemplo, que $f(a) < f(b)$ (f crescente, para f decrescente a demonstração é análoga).

Sejam $\varepsilon > 0$ e $a = x_0 < ... < x_n = b$ uma partição de $[a,b]$ que verifica,

$$0 < x_i - x_{i-1} < \frac{\varepsilon}{f(b) - f(a)} \quad (\text{para } i = 1, ..., n).$$

Como f é crescente tem-se:

$$m_i = \inf_{x \in [x_{i-1}, x_i]} f(x) = f(x_{i-1}) \text{ e } M_i = \sup_{x \in [x_{i-1}, x_i]} f(x) = f(x_i),$$

portanto,

$$\sum_{i=1}^{n}(M_i - m_i)(x_i - x_{i-1}) = \sum_{i=1}^{n}[f(x_i) - f(x_{i-1})](x_i - x_{i-1}) \leq$$

$$\leq \sum_{i=1}^{n}[f(x_i) - f(x_{i-1})]\frac{\varepsilon}{f(b) - f(a)} = (f(b) - f(a)) \cdot \frac{\varepsilon}{f(b) - f(a)} = \varepsilon$$

logo f é integrável em $[a,b]$. □

7.2 O teorema fundamental do cálculo

Teorema 7.4

Seja f uma função contínua num intervalo $[a,b]$ e seja $F(x)$ a função definida por $x \mapsto F(x) = \int_a^x f(\xi)d\xi$. Então F é uma função de classe C^1 sobre $[a,b]$, de derivada $F'(x) = f(x)$.

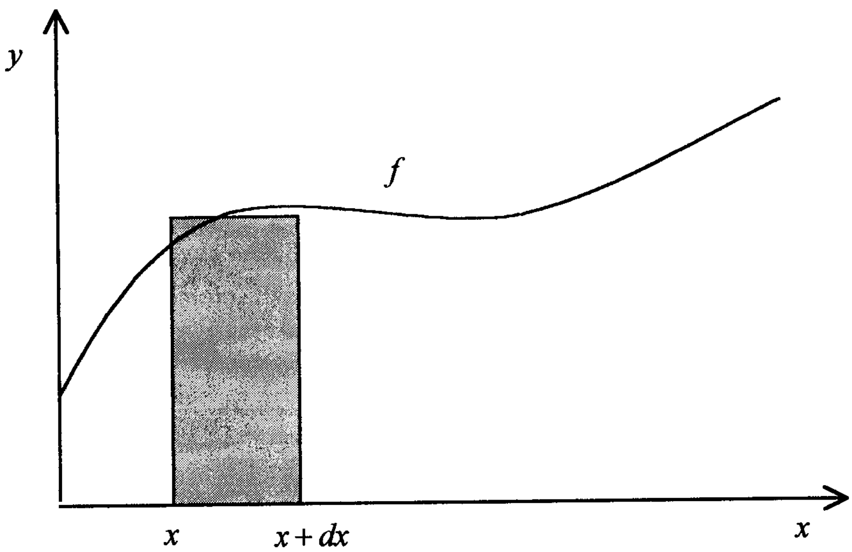


Figura 7.3 A ideia da demonstração do teorema fundamental é que a região situada abaixo do gráfico de f e acima do segmento infinitesimal $[x, x + dx]$ pode ser aproximada por um rectângulo quando f é contínua (porque f é "quase constante" entre x e $x + dx$).

Dem.

Por transferência, podemos supor f, a, b e, portanto, F standard. É suficiente, assim, verificar que em todo o ponto $x \in]a,b[$ e para todo o $dx \simeq 0$ com $dx \neq 0$ se tem:

$$\frac{F(x+dx) - F(x)}{dx} = f(x) + \varepsilon$$

com ε infinitesimal. Como f é contínua, tem-se $f(\xi) - f(x) \simeq 0$ em todo o ponto ξ de $[x, x + dx]$. Ora a quantidade $F(x + dx) - F(x)$ é precisamente a área da região situada abaixo do gráfico de f e acima do segmento infinitesimal $[x, x + dx]$. A ideia da demonstração é de notar que, quando f é contínua, esta área pode ser "aproximada" pela de um rectângulo (ver figura 7.3). Mais precisamente tem-se:

$$F(x+dx) - F(x) = \int_x^{x+dx} f(\xi)d\xi = \int_x^{x+dx} (f(\xi) - f(x))d\xi + \int_x^{x+dx} f(x)d\xi.$$

Portanto, se α é um majorante infinitamente pequeno de $f(\xi) - f(x)$ no intervalo $[x, x + dx]$ (um tal majorante existe porque a função $x \mapsto f(\xi) - f(x)$ é contínua e, portanto limitada neste intervalo), existe um ponto x_M no intervalo $[x, x + dx]$ tal que $f(x) \leq f(x_M)$ em todo o ponto x do intervalo; ora deve ter-se $f(x_M) \simeq 0$; é suficiente, portanto, considerar-se $\alpha = f(x_M)$, e tem-se:

$$\left|F(x + dx) - F(x) - \int_x^{x+dx} f(x) d\xi\right| \leq \int_x^{x+dx} \alpha d\xi.$$

Portanto, se se considerar

$$\varepsilon := \frac{F(x + dx) - F(x)}{dx} - f(x)$$

ter-se-á $\varepsilon \simeq 0$ visto que $\varepsilon \leq \alpha$. Isto mostra que $F'(x) = f(x)$ porque f e F são standard por hipótese. □

A principal consequência deste teorema é a seguinte: a função F, definida no enunciado, é uma primitiva de f (visto que a sua derivada é f) e é mesmo a primitiva de f que vale 0 em a.

Corolário 7.5

Se f é uma função contínua em $[a, b]$, então f tem uma primitiva F e

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

Dem.

Por transferência, podemos supor f standard. Seja $\varphi(x) = \int_a^x f(t) dt$ uma primitiva de f em $[a, b]$. Sendo F também uma primitiva de f nesse intervalo, existe uma constante c tal que $F(x) = \varphi(x) + c$ em cada ponto $x \in [a, b]$. Assim verificam-se as igualdades:

$$\begin{aligned} F(b) - F(a) &= (\varphi(b) + c) - (\varphi(a) + c) = \varphi(b) - \varphi(a) = \\ &= \int_a^b f(t) dt - \int_a^a f(t) dt = \int_a^b f(t) dt. \end{aligned}$$

□

O corolário seguinte chama-se *teorema do valor médio (forma integral)*:

Corolário 7.6

Para toda a função $f : [a, b] \to \mathbb{R}$ contínua, existe $c \in \,]a, b[$ tal que

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

Dem.

Por transferência, podemos supor f e c standard. Se F é uma primitiva de f, então, pelo teorema dos acréscimos finitos, existe $c \in \left]a,b\right[$ tal que

$$F(b) - F(a) = (b-a)F'(c)$$

logo,

$$\int_a^b f(x)dx = f(c)(b-a). \qquad \square$$

Notação: Designa-se frequentemente uma primitiva de f por $\int f(x)\,dx$. A esta notação (onde não é preciso os limites de integração) chama-se *integral indefinido*, ao contrário do integral $\int_a^b f(x)\,dx$ (sombra de uma soma de Riemann) chamado *integral definido*, que é um número.

7.3 Cálculo de primitivas

Existem diversos métodos para calcular primitivas (e por consequência as áreas):

1. Utilizando a decomposição e a linearidade:

$$\int (\lambda f + \mu g)(x)dx = \lambda \int f(x)dx + \mu \int g(x)dx.$$

Exemplo: Mostrar que $\int_{-1}^{1} \frac{4}{1+x^2} = 2\pi$.

$$\int_{-1}^{1} \frac{4}{1+x^2} = 4\int_{-1}^{1} \frac{1}{1+x^2} = 4([\text{arctan}x]_{-1}^{1}) = 4\Big(\frac{\pi}{4} - \Big(-\frac{\pi}{4}\Big)\Big) = 2\pi.$$

2. Integrando por partes: $\int u(x)v'(x)dx = u(x)v(x) - \int v(x)u'(x)dx + C$, em que C é uma constante.

Exemplo: Mostrar que $\int xe^x dx = (x-1)e^x + C$, onde C é uma constante. Integrando por partes temos, considerando $u'(x) = e^x$ e $v(x) = x$,

$$\int xe^x dx = xe^x - \int e^x d = xe^x - e^x + C = (x-1)e^x + C.$$

3. Utilizando a mudança de variável: para $\varphi : [c,d] \to [a,b]$ $(a = \varphi(c)$ e $b = \varphi(d))$ monótona e derivável $(u = \varphi(x))$, tem-se:

$$\int_a^b f(u)du = \int_c^d f(\varphi(x))\varphi'(x)dx \text{ e } \int f(u)du = \int f(\varphi(x))\varphi'(x)dx.$$

Exemplo: Mostrar esta fórmula utilizando a definição de integral (definido) pelas somas de Riemann.

É suficiente mostrar esta fórmula para as funções f e φ standard. Seja (u_i) uma partição infinitamente fina do intervalo $[a, b]$ e $x_i := \varphi^{-1}(u_i)$ a partição correspondente (φ é uma bijecção) do intervalo $[c, d]$. A partição é também uma partição finita porque φ é contínua. Tem-se:

$$\int_a^b f(u)du = {}^\circ\left(\sum_{i=0}^{n-1} f(v_i)(u_{i+1} - u_i)\right), \text{com } v_i \in [u_i, u_{i+1}].$$

portanto

$$\int_a^b f(u)du = {}^\circ\left(\sum_{i=0}^{n-1} f(\varphi(\xi_i))(\varphi(x_{i+1}) - \varphi(x_i))\right), \text{com } \xi_i \in [x_i, x_{i+1}].$$

Mas como φ é derivável e $x_i \simeq x_{i+1}$, tem-se :

$$\varphi(x_{i+1}) - \varphi(x_i) = (x_{i+1} - x_i)\varphi'(x_i) + \phi(x_{i+1} - x_i)$$

logo,

$$\int_a^b f(u)du = {}^\circ\left(\sum_{i=0}^{n-1} f(\varphi(x_i)\varphi'(x_i))(x_{i+1} - x_i) + \sum_{i=0}^{n-1} f(\varphi(x_i))(x_{i+1} - x_i)\phi\right).$$

Esta última soma, $\sum_{i=0}^{n-1} f(\varphi(x_i))(x_{i+1} - x_i)\phi$, é infinitesimal, e a primeira, $\sum_{i=0}^{n-1} f(\varphi(x_i)\varphi'(x_i))(x_{i+1} - x_i)$, é infinitamente próxima de $\int_c^d f(\varphi(x))\varphi'(x)dx$, donde a fórmula pretendida.

4. Se a função f é uma fracção racional, decompõe-se em elementos simples:

Exemplos:

(a) Caso em que o denominador tem raízes reais simples: como $\frac{5x+17}{x^2-3x-10} = \frac{-1}{x+2} + \frac{6}{x-5}$, pode-se calcular:

$$\int \frac{5x+17}{x^2-3x-10}dx = \int \frac{-1}{x+2}dx + \int \frac{6}{x-5}dx = -\ln|x+2| + 6\ln|x-5|.$$

(b) Caso em que o denominador tem uma raíz real múltipla: $\frac{x^2}{(x+1)^3} = \frac{1}{(x+1)^3} - \frac{2}{(x+1)^2} + \frac{1}{x+1}$, logo,

$$\int \frac{x^2}{(x+1)^3}dx = \int \frac{1}{(x+1)^3}dx - \int \frac{2}{(x+1)^2}dx + \int \frac{1}{x+1}dx$$

e, portanto,

$$\int \frac{x^2}{(x+1)^3}dx = -\frac{1}{2}(x+1)^{-2} + 2(x+1)^{-1} + \ln(x+1) =$$

$$= -\frac{1}{2(x+1)^2} + \frac{2}{x+1} + \ln(x+1).$$

(c) Caso em que o denominador não tem raízes reais: como $\frac{2}{x^2+x+1} = \frac{2}{\frac{3}{4}(t^2+1)}$, com $t = \frac{4}{3}\left(x + \frac{1}{2}\right)$, tem-se:

$$\int \frac{2}{x^2+x+1}dx = \int \frac{2}{(t^2+1)}dt = 2\text{arctan } x.$$

Exemplo: Mostrar que $\int_{-1}^{+1} \frac{1}{(1+x^2)(1+(z-x)^2)}dx = \frac{\pi}{z^2+4}$($z$ é um parâmetro).

Decompondo a fracção em:

$$\frac{1}{(1+x^2)\left(1+(z-x)^2\right)} = \frac{ax+b}{1+x^2} + \frac{a'x+b'}{1+(z-x)^2}$$

Consideremos $a = \frac{2}{z(z^2+4)}$, $b = \frac{1}{z^2+4}$, $a' = \frac{-2}{z(z^2+4)}$ e $b' = \frac{3}{z^2+4}$. Por outro lado tem-se:

$$\frac{ax+b}{1+x^2} + \frac{a'x+b'}{1+(z-x)^2} =$$

$$= \left(\frac{a}{2}\right)\frac{2x}{1+x^2} + \frac{b}{1+x^2} + \left(\frac{a'}{2}\right)\frac{(-2)(z-x)}{1+(z-x)^2} + \frac{a'z+b'}{1+(z-x)^2}.$$

Portanto o integral é igual a:

$$\left[\frac{a}{2}\ln\left(1+x^2\right)\right]_{-1}^{+1} + [b \text{ arctan } x]_{-1}^{+1} + \left[\frac{a'}{2}\ln\left(1+(z-x)^2\right)\right]_{-1}^{+1} +$$

$$+ \left[(a'z+b')\text{arctan }(z-x)\right]_{-1}^{+1}$$

e, portanto, os dois termos em ln são nulos (funções pares),

$$\int_{-1}^{+1} \frac{1}{(1+x^2)(1+(z-x)^2)}dx = \frac{1}{z^2+4}\left(\frac{\pi}{2}\right) + \left(\frac{-2}{z^2+4} + \frac{3}{z^2+4}\right)\left(\frac{\pi}{2}\right) = \frac{\pi}{z^2+4}.$$

Exemplos:

- Mostrar que

$$\int \frac{4x^2+x+1}{x^3-x}dx = \ln\left[\left|\frac{(x-1)^3}{x}\right|(x+1)^2\right] + C$$

(onde C é uma constante).
Decompondo a fracção em:

$$\frac{4x^2+x+1}{x^3-x} = \frac{a}{x} + \frac{b}{(x-1)} + \frac{c}{(x+1)}$$

Obtém-se então $a = -1, b = 3$ e $c = 2$. Então,

$$\frac{4x^2+x+1}{x^3-x} = \frac{-1}{x} + \frac{3}{(x-1)} + \frac{2}{(x+1)}.$$

Portanto:

$$\int \frac{4x^2+x+1}{x^3-x}dx = \int \left(-\frac{1}{x} + \frac{3}{x-1} + \frac{2}{x+1}\right)dx =$$

$$= \int -\frac{1}{x}dx + \int \frac{3}{x-1}dx + \int \frac{2}{x+1}dx =$$

$$= -\ln|x| + 3\ln|x-1| + 2\ln|x+1| + C =$$

$$= \ln\left[\left|\frac{(x-1)^3}{x}\right|(x+1)^2\right] + C.$$

- Mostrar que $\int \frac{2x^3+5x^2+6x+2}{x(x+1)^3}dx = \ln(x^2) + \frac{1}{x+1} - \frac{1}{2}\left(\frac{1}{(x+1)^2}\right) + C$

(onde C é uma constante). Decompondo a fracção em:

$$\frac{2x^3+5x^2+6x+2}{x(x+1)^3} = \frac{a}{x} + \frac{b_1}{x+1} + \frac{b_2}{(x+1)^2} + \frac{b_3}{(x+1)^3}$$

Obtém-se então $a = 2, b_1 = 0$ $b_2 = -1$ e $b_3 = 1$. Então,

$$\frac{2x^3+5x^2+6x+2}{x(x+1)^3} = \frac{2}{x} + \frac{-1}{(x+1)^2} + \frac{1}{(x+1)^3}$$

Logo,

$$\int \frac{2x^3+5x^2+6x+2}{x(x+1)^3}dx = \int \left(\frac{2}{x} - \frac{1}{(x+1)^2} + \frac{1}{(x+1)^3}\right)dx =$$

$$= \int \frac{2}{x}dx + \int -\frac{1}{(x+1)^2}dx + \int \frac{1}{(x+1)^3}dx =$$

$$= \ln x^2 + \frac{1}{x+1} - \frac{1}{2}\left(\frac{1}{(x+1)^2}\right) + C\,.$$

• Mostrar que

$$\int \tfrac{x+2}{x^3-1}dx = \ln|x-1| - \tfrac{1}{2}\ln\sqrt{x^2+x+1} - \tfrac{\sqrt{3}}{3}\arctan\left[\tfrac{2\sqrt{3}}{3}(x+\tfrac{1}{2})\right] + C$$

(onde C é uma constante).

Decompondo a fracção em:

$$\frac{x+2}{x^3-1} = \frac{a}{x-1} + \frac{bx+c}{x^2+x+1}$$

Obtém-se então $a = 1$, $b = -1$ e $c = -1$. Então,

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

Logo,

$$\int \frac{x+2}{x^3-1}dx = \int \left(\frac{1}{x-1} - \frac{x+1}{x^2+x+1}\right)dx =$$

$$= \int \frac{1}{x-1}dx - \int \frac{x+1}{x^2+x+1}dx =$$

$$= \ln|x-1| - \int \frac{x+1}{x^2+x+1}dx + C =$$

$$= \ln|x-1| - \left(\frac{1}{2}\left(\int \frac{2x+1}{x^2+x+1}dx + \int \frac{1}{x^2+x+1}\right)dx\right) + C =$$

$$= \ln|x-1| - \tfrac{1}{2}\left(\ln\sqrt{x^2+x+1} + \tfrac{2\sqrt{3}}{3}\arctan\left[\tfrac{2\sqrt{3}}{3}(x+\tfrac{1}{2})\right)\right) + C =$$

$$= \ln|x-1| - \frac{1}{2}\ln\sqrt{x^2+x+1} - \frac{\sqrt{3}}{3}\arctan\left[\frac{2\sqrt{3}}{3}\left(x+\frac{1}{2}\right)\right] + C\,.$$

Conclusão

Com este trabalho espero contribuir um pouco, para que a Análise Não-Standard seja incluída como disciplina de algumas licenciaturas, nomeadamente em licenciaturas de ensino, de Matemática, Física e Biologia, assim como nas licenciaturas de Engenharia.

Depois do trabalho realizado, podemos concluir que, uma vez que a Análise Não-Standard permite, uma simplificação dos conceitos e das demonstrações da análise clássica, assim como a formulação de novos modelos matemáticos de fenómenos não matemáticos, por que não começar a ensinar em Portugal, tal como já acontece em França, Análise Não-Standard (com as adaptações necessárias) no ensino secundário?

Será esta matemática (a Análise Não-Standard) que acabará com o "medo" da matemática que existe hoje em dia em muitas crianças, jovens e adultos?

Qual a "posição" da Análise Não-Standard no ensino nas próximas décadas? Será esta a matemática do futuro?

Bibliografia

T. M. Apostol, *Cálculo,* vol. 2, Editora Reverté, Ltda., 1993.

J. Campos Ferreira, *Introdução à Análise Matemática,* Fundação Calouste Gulbenkian, 1990.

A. Deledicq & M. Diener, *Leçons de calcul Infinitésimal,* Armand Colin, 1989.

A. Deledicq & V. Gautheron, *Cours élementaire d'Análise Non-Standard,* Nice,1996.

F. R. Dias Agudo, *Análise Real,* vol. I, Escolar Editora, 1994.

F. Diener & M. Diener, *Nonstandard Analysis in Practice,* Springer,1991.

F. Diener & V. Gautheron, *Cours d'Análise Non-Standard,* Nice,1996.

F. Diener & G. Reeb, *Análise Non Standard,* Hermann, 1989.

A. J. F. Oliveira, *Teoria dos Conjuntos,* Escolar Editora, 1982; *Matemática Não-Standard — Uma introdução,* Universidade de Évora, 1997 (Notas do curso de Mestrado de Análise Não-standard).

A. Robert, *Nonstandard Analysis,* J. Wiley & Sons, 1988.

C. Sarrico, *Análise Matemática,* Gradiva, 1999.