

## Aventuras Numéricas no Cálculo do $e$

As calculadoras e os computadores vêm oficialmente afirmando-se como instrumentos didáticos indispensáveis à concretização de objectivos fundamentais de aprendizagem da Matemática, mas como devemos reagir quando os resultados numéricos contrariam a teoria matemática?

A "exploração numérica" do limite da sucessão  $\left(1 + \frac{1}{n}\right)^n$  pode levar a resultados indesejados; por exemplo, a máquina de calcular (ou computador) pode facilmente fazer crer que a sucessão não é monótona. Saiba porquê...

Quando os professores de Matemática são confrontados com as cíclicas mudanças ou ajustamentos dos programas curriculares têm inevitavelmente de ultrapassar muitas dificuldades: técnicas, pedagógicas e científicas.

Ora, as questões de natureza científica, embora intrinsecamente subjacentes a qualquer reforma dos programas escolares, revestem-se ainda de uma importância inquestionável enquanto componentes essenciais do conhecimento do professor, mas nem sempre têm merecido uma reflexão específica e contextualizada nos diversos aspectos das directivas programáticas oficiais.

Num estudo de natureza mais alargada, partimos de um tema chave do programa de Matemática do Ensino Secundário em Portugal: o conceito de limite. Neste artigo reflectimos sobre a definição do número  $e$  como

limite da sucessão de termo geral  $\left(1 + \frac{1}{n}\right)^n$  e exploramos as dificuldades científicas que se colocam a um

professor de Matemática quando confrontado com a preparação deste tópico com recurso às calculadoras gráficas. Preocupamo-nos com a compreensão dos factos, com o "porquê?" dos casos; exploramos as razões dos acontecimentos; questionamos as acções; aprofundamos as análises dos fenómenos e comparamos o conhecimento adquirido desta forma com aquele que temos ao nosso dispor em manuais escolares ou em manuais de utilização das calculadoras gráficas tradicionalmente utilizadas nas escolas.

Percebemos finalmente que a manipulação cega das calculadoras bem como a ênfase prioritária em dificuldades de natureza pedagógica podem, porventura inconscientemente, fazer com que os professores deixem para um segundo plano da sua formação as questões de natureza científica que afinal, hoje como sempre, são o requisito central do ensino da Matemática.

### 1. Introdução

O programa actual do Ensino Secundário prevê a utilização obrigatória de calculadoras gráficas não só como instrumentos de cálculo mas também como um meio incentivador do espírito de pesquisa dos alunos.

Porém, como é sabido, apenas uma quantidade finita de números é exactamente representável na máquina (ou computador); os erros de arredondamento cometidos na representação de números e a forma como estes erros se propagam, dependendo dos cálculos efectuados, não raramente fazem com que os resultados produzidos estejam bem longe dos valores esperados.

Um caso que tem merecido particular atenção é a sucessão de termo geral

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

No 11.º ano, o número  $e$  é definido como o limite desta sucessão. Embora o estudo da convergência desta sucessão ultrapasse o âmbito do actual programa, os conceitos de sucessão monótona e de sucessão limitada, que fazem parte do programa, podem ser usados, como é sugerido em [6], p. 48, para se intuir o teorema das sucessões monótonas.

É sabido que existe  $N$  (que depende das especificidades do sistema de representação da máquina utilizada)

tal que para todo o  $n > N$ , a representação da base  $1 + \frac{1}{n}$  será 1 e, por esta razão, os correspondentes valores de  $u_n$  dados por (1) serão iguais a 1.

Em [6], p. 52, apresentam-se os valores de  $u_n$  obtidos na calculadora TI-83 com  $n = 10^k$  para diferentes valores de  $k$ ; com  $n = 10^{13}$  e  $n = 10^{14}$  obtêm-se os valores 2,760577856 e 1, respectivamente. Como se explicam estes

resultados “estranhos”? De acordo com [6], p. 52, “Quando  $n$  é muito grande, o valor de  $1 + \frac{1}{n}$  deixa de ser

rigorosamente representado pelos 14 dígitos que a TI-83 usa para representar os números. Se  $n$  for maior que  $10^{14}$  a

máquina passa mesmo a obter um valor numérico de  $1 + \frac{1}{n}$  como sendo 1”.

Claramente, esta afirmação está a pressupor que os valores de  $n$  são todos da forma  $10^k$ . Como o sistema de representação da TI-83, bem como o da maioria das calculadoras, é decimal (de base 10), até um certo limite os valores de  $1 + 1/10^k$  são de facto representados sem erro. Mas há que ter consciência de que podem ser obtidos resultados “estranhos” para valores de  $n$  muito menores, que não são da forma  $10^k$ ; por exemplo, na mesma TI-83 obtêm-se

$$u_{10^{10}} = \left(1 + \frac{1}{10^{10}}\right)^{10^{10}} = 2,718281828 \quad (2)$$

$$u_{2^{35}} = \left(1 + \frac{1}{2^{35}}\right)^{2^{35}} = 2,717924089 \quad (3)$$

$$u_{10^{11}} = \left(1 + \frac{1}{10^{11}}\right)^{10^{11}} = 2,718281828. \quad (4)$$

Os valores de  $u_{10^{10}}$  e  $u_{10^{11}}$  concordam com o valor do número  $e$  nos 10 algarismos que aparecem no visor da máquina<sup>1</sup> mas para  $n=2^{35}$  apenas os três primeiros algarismos estão correctos. Ora, tendo em conta que

$$10^{10} < 2^{35} < 10^{11},$$

<sup>1</sup>Na verdade, o valor calculado para  $n=10^{11}$  tem 11 algarismos correctos que não é possível visualizar imediatamente uma vez que os resultados são apresentados no visor da TI-83 com 10 algarismos, embora, como já se disse, esta calculadora use 14 dígitos para representar internamente os números; se à representação do valor calculado para  $(1+1/10^{11})^{10^{11}}$  subtrairmos 2 (o primeiro algarismo correcto), obtemos 0,7182818284, conseguindo mais um algarismo correcto na representação do “ $e$ ”.

seria de esperar que o valor calculado de  $u_{2^{35}}$  fosse pior do que o valor de  $u_{10^{10}}$ . Tal não acontece porque  $1+1/2^{35}$  não é representado exactamente na calculadora, ao contrário do que acontece com  $1+1/10^{10}$  e  $1+1/10^{11}$ . Em suma, o problema não é, de facto,  $2^{35}$  ser “grande demais”.

Antes de analisar com mais detalhe os erros envolvidos no cálculo dos termos da sucessão (1), o que levaremos a cabo nas secções seguintes, fazemos notar que, no âmbito das actividades propostas aos alunos nos manuais escolares<sup>2</sup>, tem o professor de estar preparado para a eventualidade de os alunos, fazendo uso da liberdade de experimentar com a calculadora, usarem valores de  $n$  que não são potências da base 10 e assim obterem resultados numéricos que contradizem a teoria.

## 2. Cálculo dos termos da sucessão $(1+1/n)^n$ no Matlab

Na chamada notação de vírgula flutuante, um número real é representado na forma  $x = \pm m \times b^E$ , em que  $m$  é um número real não negativo, designado por *mantissa*,  $b \geq 2$  é um inteiro positivo (a base do sistema de representação) e o expoente  $E$  é um inteiro. Esta notação permite a representação de números de ordens de grandeza muito diferentes. Por exemplo,  $x = 0,0000000314$  e  $y = 314000$ , podem ser representados exactamente, no sistema decimal ( $b = 10$ ) de vírgula flutuante, com apenas três dígitos na mantissa, na forma  $x = 3,14 \times 10^{-8}$  e  $y = 3,14 \times 10^5$ . Observe-se que numa representação sem expoente a vírgula estaria então fixa (não seria flutuante) e seria necessário dispor de mantissas com 6 dígitos para a parte inteira e oito dígitos para a parte fraccionária, isto é, um total de 14 dígitos, que podem não estar disponíveis, dependendo da capacidade do sistema implementado no computador ou na máquina de calcular. A norma IEEE754<sup>3</sup> é o padrão para a aritmética binária ( $b = 2$ ) de vírgula flutuante (embora existam alguns processadores que não a implementam). No formato duplo da norma IEEE754, um número  $x$  normalizado é representado na forma

$$x = \pm (1, b_1 b_2 \dots b_{52})_2 \times 2^E, \quad (5)$$

onde  $b_i = 0$  ou  $b_i = 1$ , para cada  $i = 1, \dots, 52$ , e o expoente inteiro  $E$  varia entre os limites  $E_{\min} = -1022$  e  $E_{\max} = 1023$ . O maior número que se pode representar neste sistema é

$$\text{realmax} = (1, 11 \dots 1)_2 \times 2^{1023}. \quad (6)$$

A parte não inteira da *mantissa* ocupa 52 bits e é deste número que depende essencialmente a precisão da aritmética. O número 1 é obviamente representável neste sistema com  $b_i = 0$  para cada  $i = 1, \dots, 52$ , isto é,

$$1 = +(1, 0 \dots 0)_2 \times 2^0$$

e o número que lhe sucede é (difere apenas no último bit)

$$1 = +(1, 0 \dots 01)_2 \times 2^0,$$

isto é, trata-se do número

$$1 + 2^{-52}.$$

Se  $x$  é um número tal que

$$1 < x < 1 + 2^{-52},$$

<sup>2</sup>Ver, por exemplo, [2], p. 271, e [3], p. 55.

<sup>3</sup>Para uma excelente introdução veja-se [5].

terá de ser arredondado, isto é, a sua representação será  $\tilde{x} = 1$  ou  $\tilde{x} = 1 + 2^{-52}$ ; se o modo de arredondamento for “para o mais próximo”, tem-se

$$|x - \tilde{x}| \leq 2^{-53}.$$

No caso particular de ser  $x = 1 + 2^{-53}$  (está tão próximo de 1 quanto de  $x = 1 + 2^{-52}$ ), a “regra de desempate” determina que se escolha a representação que tem o último bit da mantissa igual a zero ([5], p. 27) e portanto  $\tilde{x} = 1$ . Por esta razão, no Matlab [4], que implementa a norma IEEE754, tem-se

$$\left(1 + \frac{1}{2^{53}}\right)^{2^{53}} = 1.$$

Se  $n \geq 2^{53}$ , o valor de  $x = 1 + 1/n$  terá a representação  $\tilde{x} = 1$  (continuamos a admitir que estamos a usar o arredondamento “para o mais próximo”) e obtém-se  $u_n = 1$ . Com  $n = 2^{52}$ , o resultado é

$$u_{2^{52}} = \left(1 + \frac{1}{2^{52}}\right)^{2^{52}} = 2,718281828459045$$

que coincide com o número  $e$  nos 16 algarismos representados. Para efeitos de comparação com o valor obtido antes na TI-83, observe-se que para  $n = 2^{35}$  tem-se, no Matlab,

$$u_{2^{35}} = \left(1 + \frac{1}{2^{35}}\right)^{2^{35}} = 2,718281828419489$$

com 11 algarismos correctos.

Neste sistema binário, ao contrário do que sucede com as calculadoras, os termos  $u_n$  com  $n = 10^k$  não conseguem aproximar-se muito do número  $e$ . Na tabela seguinte registam-se os valores calculados e os erros respectivos para os termos  $u_{10^k}$ , com  $k$  inteiro desde 1 até 16.

$k$	$u_{10^k}$	$e - u_{10^k}$
1	2,593742460100002	$1,25 \times 10^{-1}$
2	2,704813829421529	$1,35 \times 10^{-2}$
3	2,716923932235594	$1,36 \times 10^{-3}$
4	2,718145926824926	$1,36 \times 10^{-4}$
5	2,718268237192298	$1,36 \times 10^{-5}$
6	2,718280469095753	$1,36 \times 10^{-6}$
7	2,718281694132082	$1,34 \times 10^{-7}$
8	2,718281798347358	$3,01 \times 10^{-8}$
9	2,718282052011560	$-2,24 \times 10^{-7}$
10	2,718282053234788	$-2,25 \times 10^{-7}$
11	2,718282053357110	$-2,25 \times 10^{-7}$
12	2,718523496037238	$-2,42 \times 10^{-4}$
13	2,716110034086901	$2,17 \times 10^{-3}$
14	2,716110034087023	$2,17 \times 10^{-3}$
15	3,035035206549262	$-3,17 \times 10^{-1}$
16	1,000000000000000	$1,72 \times 10^0$

Como se pode apreciar, com números da forma  $10^k$  não é possível obter a precisão que se consegue com números da forma  $2^k$ . A isto acresce o facto de o erro  $e - u_{10^k}$  variar de sinal, quer dizer, a sucessão dos valores  $u_{10^k}$  calculados não se comporta de facto como uma sucessão monótona.

### 3. Análise dos erros

A característica essencial da aritmética de vírgula flutuante é permitir representar qualquer número  $x$  com o mesmo número de algarismos significativos correctos, independentemente da grandeza de  $x$ . De facto, se  $x$  é o valor exacto e  $\tilde{x}$  é a correspondente aproximação dada por (5), para o erro relativo, definido por

$$\varepsilon_x = \frac{\tilde{x} - x}{x}, \quad (7)$$

tem-se (ver, por exemplo, [5], p. 29)

$$|\varepsilon_x| < 2^{-52}, \quad (8)$$

seja qual for o modo de arredondamento utilizado. No caso do arredondamento para o mais próximo, o limite anterior pode ser dividido por 2, isto é, tem-se

$$|\varepsilon_x| \leq 2^{-53}. \quad (9)$$

Com  $x = 1 + \frac{1}{n}$  tem-se  $u_n = x^n$  e  $\tilde{u}_n = \tilde{x}^n$ , escrevendo (7) na forma  $\tilde{x} = x(1 + \varepsilon_x)$ , para o erro relativo no valor calculado de  $u_n$  obtemos

$$\varepsilon_{u_n} = \frac{\tilde{x}^n - x^n}{x^n} = (1 + \varepsilon_x)^n - 1. \quad (10)$$

Usando a fórmula do binómio de Newton, temos

$$(1 + \varepsilon_x)^n = 1 + n\varepsilon_x + \sum_{k=2}^n \frac{n(n-1)\dots(n-k+1)}{k!} \varepsilon_x^k, \quad (11)$$

logo

$$\varepsilon_{u_n} = n\varepsilon_x + \sum_{k=2}^n \frac{n(n-1)\dots(n-k+1)}{k!} \varepsilon_x^k. \quad (12)$$

Se  $|n\varepsilon_x|$  for bastante menor que a unidade, podemos em (12) desprezar as parcelas que envolvem as potências  $\varepsilon_x^2, \varepsilon_x^3, \dots, \varepsilon_x^n$  e escrever

$$\varepsilon_{u_n} \approx n\varepsilon_x \quad (13)$$

o que mostra que, ainda que seja

$$|\varepsilon_x| \leq 2^{-53} \approx 10^{-16}, \quad (14)$$

o erro relativo em  $u_n$  pode ser grande; com efeito, como a sucessão (1) converge muito lentamente, é necessário usar valores de  $n$  muito grandes para obter boas aproximações do número de Neper  $e$  e, a menos que  $1 + 1/n$  se represente exactamente, o valor de  $(1 + 1/n)^n$  será calculado com um elevado erro relativo. Isto explica os erros apresentados na tabela anterior, para  $n = 10^k$ : as aproximações melhoram à medida que  $k$  cresce desde 1 até 8 e o

número de algarismos correctos é aproximadamente igual a  $k$ ; a partir de  $k=9$  o número de algarismos correctos diminui porque os erros de arredondamento dominam a aproximação produzida. De facto, pode apreciar-se que os erros de arredondamento afectam aproximadamente os últimos  $k$  algarismos decimais apresentados e assim, por exemplo,  $u_{10}^{12}$  exhibe apenas quatro algarismos correctos de um total de 16 representados.

#### 4. Influência do modo de arredondamento

Cumprindo a norma IEEE, o Matlab prevê a possibilidade de se usarem quatro modos distintos de arredondamento, em alternativa, à escolha do utilizador; o sistema implementa o modo de arredondamento para o mais próximo, mas pode ser instruído no sentido de usar algum dos outros modos de arredondamento previstos, um dos quais é o arredondamento "para a direita", isto é, no sentido de  $+\infty$ .

É interessante observar o efeito que tal arredondamento produz no caso do cálculo dos termos  $(1 + 1/n)^n$ , tendo em conta que, de acordo com (8) e (9), os majorantes dos erros não são muito diferentes. Na tabela seguinte registam-se os valores obtidos e os erros respectivos.

$k$	$u_{10}^k$	$e - u_{10}^k$
1	2,593742460100002	$1,25 \times 10^{-1}$
2	2,704813829421529	$1,35 \times 10^{-2}$
3	2,716923932236196	$1,36 \times 10^{-3}$
4	2,718145926830961	$1,36 \times 10^{-4}$
5	2,718268237192298	$1,36 \times 10^{-5}$
6	2,718280469699332	$1,36 \times 10^{-6}$
7	2,718281694132082	$1,34 \times 10^{-7}$
8	2,718281858705339	$-3,02 \times 10^{-8}$
9	2,718282052011560	$-2,24 \times 10^{-7}$
10	2,718282053234788	$-2,25 \times 10^{-7}$
11	2,718282053357110	$-2,25 \times 10^{-7}$
12	2,718523496037238	$-2,42 \times 10^{-4}$
13	2,722147710568192	$-3,87 \times 10^{-3}$
14	2,777094348318434	$-5,89 \times 10^{-2}$
15	3,035035206549262	$-3,17 \times 10^{-1}$
16	9,211438704993531	$-6,49 \times 10^0$

Observe-se, por comparação dos valores aqui apresentados com os da tabela dada anteriormente, que as aproximações obtidas nos dois modos de arredondamento não diferem muito para os valores menores de  $k$  mas afastam-se dramaticamente quando  $k$  cresce. Com efeito, como se explica a seguir, o comportamento das sucessões dos termos calculados depende do modo de arredondamento usado.

No caso do arredondamento para o mais próximo, tem-se, como já se fez notar,  $\tilde{u}_n = 1$  para  $n \geq 2^{53}$ . Isto acontece porque, para  $n \geq 2^{53}$ ,

$$\varepsilon_x = \frac{1 - \left(1 + \frac{1}{n}\right)}{1 + \frac{1}{n}} = -\frac{1}{n+1}. \quad (15)$$

De (10) resulta

$$\varepsilon_{u_n} = \left(1 - \frac{1}{n+1}\right)^n - 1 \quad (16)$$

$$= \frac{1}{\left(1 + \frac{1}{n}\right)^n} - 1 \quad (17)$$

e o erro relativo  $\varepsilon_{u_n}$  tende para  $\frac{1}{e} - 1$ , quando  $n$  cresce. No caso do modo de arredondamento “para a direita”,  $x = 1 + 1/n$  será representado por  $\tilde{x} = 1 + 2^{-52}$ , para  $n \geq 2^{52}$ ; então, tem-se

$$\varepsilon_x = \frac{2^{-52} - \frac{1}{n}}{1 + \frac{1}{n}} \quad (18)$$

Neste caso,  $\varepsilon_x$  não tende para zero mas tende para  $2^{-52}$  quando  $n$  cresce e  $\varepsilon_{u_n}$  tende para  $+\infty$  como se pode concluir de

$$\varepsilon_{u_n} = (1 + \varepsilon_x)^n - 1 = \left(1 + \frac{2^{-52} - \frac{1}{n}}{1 + \frac{1}{n}}\right)^n - 1 = \frac{(1 + 2^{-52})^n}{\left(1 + \frac{1}{n}\right)^n} - 1.$$

Por exemplo, com  $n = 2^{61}$  obtemos, no Matlab,

$$u_{2^{61}} = 2,284413586539627 \times 10^{222}$$

e para  $n = 2^{62}$  dá *Inf*<sup>4</sup>, por ser o valor calculado maior do que *realmax* dado em (6).

### 5. Condicionamento de $(1 + x/n)^n$ e $x^n$ .

O problema da propagação dos erros no cálculo dos termos  $(1 + 1/n)^n$  está na fórmula de cálculo e não na função propriamente dita. Para esclarecer cabalmente este aspecto, vamos usar o importante conceito de número de condição relativo de uma função (de variável real) num ponto.

Se  $f$  é uma função que admite derivadas contínuas de 1.ª e 2.ª ordem, então a fórmula de Taylor com resto de 2.ª ordem escreve-se, com  $\xi_x$  um ponto entre  $x$  e  $\tilde{x}$ ,

$$f(\tilde{x}) = f(x) + (\tilde{x} - x)f'(x) + (\tilde{x} - x)^2 \frac{f''(\xi_x)}{2} \quad (19)$$

Assumindo que podemos desprezar o termo de 2.ª ordem (o que será lícito se o erro  $|x - \tilde{x}|$  for suficientemente pequeno), temos

$$f(\tilde{x}) \approx f(x) + (\tilde{x} - x)f'(x), \quad (20)$$

daqui resulta

$$\frac{f(\tilde{x}) - f(x)}{f(x)} \approx \frac{\tilde{x} - x}{x} \cdot x \frac{f'(x)}{f(x)}, \quad (21)$$

ou seja

$$\varepsilon_{f(x)} \approx \varepsilon_x \cdot k_f(x), \quad (22)$$

onde  $\varepsilon_x$  é o erro relativo na variável independente,

<sup>4</sup>Esta é uma das chamadas “exceções” da aritmética IEEE. O sistema usa uma representação especial *Inf* para qualquer número maior do que *realmax* (ver [5], p. 22).

$$\varepsilon_{f(x)} = \frac{f(\tilde{x}) - f(x)}{f(x)} \quad (23)$$

é o correspondente erro relativo na variável dependente e

$$k_f(x) = x \frac{f'(x)}{f(x)} \quad (24)$$

é o chamado número de condição relativo da função  $f$  no ponto  $x$ . Se  $|k_f(x)| \approx 1$ , então o número de algarismos significativos em  $f(\tilde{x})$  será o mesmo que em  $\tilde{x}$ , mas se  $|k_f(x)| \approx 10^3$ , por exemplo, então  $f(\tilde{x})$  terá menos três algarismos correctos do que  $\tilde{x}$ .

Sendo, para cada  $n$  inteiro positivo, a função definida por

$$f_n(x) = \left(1 + \frac{x}{n}\right)^n \quad (25)$$

tem-se

$$k_{f_n}(x) = \frac{x}{1 + \frac{x}{n}}, \quad (26)$$

o que mostra que para  $n$  suficientemente grande comparado com  $|x|$  é  $k_f(x) \approx x$  e, portanto, a função será bem condicionada para valores de  $|x|$  pequenos. Para cada  $n$ ,  $f_n(1)$  é o valor do termo  $u_n$  e, como acabamos de explicar, a função  $f_n$  é bem condicionada no ponto  $x=1$ .

Como explicar os erros no cálculo de  $u_n$  à luz do condicionamento de funções? Para a função

$$f(x) = x^n, \quad (27)$$

é, para todo  $x$ ,

$$k_f(x) = n \quad (28)$$

e tem-se, de acordo com (22),

$$\varepsilon_{f(x)} = \frac{\tilde{x}^n - x^n}{x^n} \approx n\varepsilon_x, \quad (29)$$

o que mostra que um pequeno erro relativo  $\varepsilon_x$  causa um elevado erro relativo em  $x^n$ . Isto é afinal o que já tinha sido observado antes com  $x = 1 + \frac{1}{n}$  e formalizado em (13).

### 6. Uma fórmula estável para calcular $(1 + 1/n)^n$

Consideremos a função definida por

$$g(x) = \exp(n \log(x)), \quad (30)$$

onde  $\log$  representa o logaritmo natural. Tem-se, para  $x > 0$ ,

$$g(x) = f(x) = x^n$$

e

$$k_g(x) = k_f(x) = n. \quad (31)$$

Portanto, se  $n$  for grande, um pequeno erro relativo em  $x$  causará um elevado erro relativo em  $g(x)$ . O número de condição da função exponencial é

$$k_{\exp}(x) = x$$

e para valores  $x = 1 + \frac{1}{n}$  a função é bem condicionada; quer dizer, o valor de  $g\left(1 + \frac{1}{n}\right)$  será calculado com tantos algarismos correctos quantos os algarismos correctos que se encontrarem para  $n \log\left(1 + \frac{1}{n}\right)$ . Para valores próximos de 1,  $\log(x)$  é mal condicionada, como resulta de

$$k_{\log}(x) = x \frac{1}{\log(x)} = \frac{1}{\log(x)}. \quad (32)$$

O conhecido desenvolvimento em série de potências de  $\theta$

$$\log(1 + \theta) = \theta - \frac{\theta^2}{2} + \frac{\theta^3}{3} - \frac{\theta^4}{4} + \dots \quad (33)$$

pode ser usado para produzir boas aproximações. Com  $\theta = \frac{1}{n}$  dá

$$g\left(1 + \frac{1}{n}\right) = \exp\left[n\left(\frac{1}{n} - \frac{1}{2n^2} + \frac{1}{3n^3} - \frac{1}{4n^4} + \dots\right)\right] = \exp\left(1 - \frac{1}{2n} + \frac{1}{3n^2} - \frac{1}{4n^3} + \dots\right).$$

Por exemplo, tem-se

$$\left(1 + \frac{1}{10^{14}}\right)^{10^{14}} \approx \exp\left(1 - \frac{10^{-14}}{2}\right),$$

com um erro de truncatura que é (por se tratar de uma série alternada convergente) inferior ao valor absoluto do primeiro termo que se despreza, que neste caso vale  $\frac{\theta^3}{3} = 10^{-42}/3$ .

No Matlab obtém-se desta maneira a aproximação **2,718281828459032** que tem 14 algarismos correctos enquanto, como vimos antes, o valor calculado directamente tem apenas três algarismos correctos.

Porque é que isto acontece? Observe-se que o número de condição do polinómio

$$T_2(\theta) = \theta - \frac{\theta^2}{2} \quad (34)$$

é

$$k_{T_2}(\theta) = \frac{\theta \cdot T_2'(\theta)}{T_2(\theta)} = 1 - \frac{1}{2}\theta + O(\theta^2), \quad (35)$$

logo, para  $|\theta|$  pequeno,  $k_{T_2}(\theta) \approx 1$  e pequenos erros relativos no valor de  $\theta$  produzirão erros relativos igualmente pequenos no valor calculado de  $T_2(\theta)$ .

Uma condição indispensável ao sucesso deste procedimento é ter-se o valor do argumento  $x$  na forma  $1 + \theta$ , com  $\theta$  exacto ou com pequeno erro relativo. Para ilustrar isto observemos que, no Matlab, o valor calculado de

$$\theta = (1 + 10^{-14}) - 1$$

é

$$\tilde{\theta} = 9,992007221626409 \times 10^{-15},$$

com erro relativo

$$\varepsilon_{\tilde{\theta}} \approx -7,99 \times 10^{-4}$$

Com aquele valor de  $\tilde{\theta}$  obtém-se

$$\exp(T_2(\tilde{\theta})) = 2,718281828459032$$

que tem apenas três algarismos correctos.

### 7. Cálculo de juros compostos

Muito acertadamente, no nosso entendimento, alguns manuais escolares introduzem a sucessão de termo geral  $(1 + 1/n)^n$  num contexto de modelação matemática: a capitalização contínua dos juros. Em [3], p. 53-56, é obtida uma aproximação de  $e$  com cinco algarismos correctos para  $n = 525600$ , que corresponde a juros capitalizados minuto a minuto. Em [2], p. 270-271 e [1], p. 99-100, são obtidas as aproximações que correspondem à capitalização em cada segundo ( $n = 31536000$ ). Neste contexto, entende-se que em [1] o número  $e$  seja designado por “constante bancária”.

Se um capital  $C_0$  for investido a uma taxa anual de  $x\%$  e a capitalização de juros for feita  $n$  vezes por ano, então ao cabo de um ano o valor do capital acumulado é<sup>5</sup>

$$C_n(x) = C_0 \times \left(1 + \frac{x/100}{n}\right)^n \quad (36)$$

Esta é a conhecida fórmula dos juros compostos. Para um valor de  $x$  fixo, o valor de  $C_n(x)$  tende para  $C_0 e^x$ . Assim sendo, frequências muito altas para a composição dos juros não alteram muito o valor de  $C_n(x)$ . Porém, devido a erros de arredondamento, os valores de  $C_n(x)$  calculados podem afastar-se dramaticamente dos valores correctos. Para uma ilustração detalhada deste problema veja-se [5], p. 82-85. O que se disse anteriormente sobre o cálculo de aproximações do número  $e$  aplica-se imediatamente no caso em que a taxa de juro anual é de 100%. Se um cliente fizer um depósito a prazo num banco e o computador do banco usar aritmética com arredondamentos “para a direita”, então  $C_n(1)$  tenderá, como vimos na secção 4, para  $+\infty$  e o banco irá à falência se permitir, no contrato com o cliente, que  $n$  seja muito grande!

### 8. Aproximações obtidas nas calculadoras

Na secção 3 apresentámos uma análise detalhada dos erros cometidos no cálculo de  $(1 + 1/n)^n$  usando aritmética binária IEEE. Não estamos em condições de fazer uma análise tão detalhada dos erros de arredondamento cometidos nas calculadoras TI-83 e CASIO-CFX-9850 por falta de informação completa relativamente à aritmética implementada.

Como já se fez notar antes, por ser decimal o sistema de representação nas calculadoras, obtém-se melhores aproximações para  $(1 + 1/n)^n$  usando valores de  $n$  que são potências de 10. O exemplo apresentado na secção 1 ilustra claramente este facto. Isto acontece porque  $1 + 1/n = 1 + 10^{-k}$  tem uma representação do tipo

<sup>5</sup>Para uma explicação mais detalhada, ver <http://www.cut-the-knot.org/arithmic/interest.shtml>.

$$(1, d_1 d_2 \dots d_{p-1}) \times 10^0, \tag{37}$$

onde  $p$  denota o número de algarismos decimais disponíveis no sistema de representação da calculadora. Se  $k < p$ , então o número  $1 + 10^{-k}$  terá representação exacta e o valor calculado de  $(1 + 10^{-k})^{10^k}$  terá um erro desprezável. Porém, se  $k$  for tal que a representação de  $1 + 10^{-k}$  não é exacta, então este pequeno erro propagar-se-á de forma “violenta” no cálculo de  $(1 + 10^{-k})^{10^k}$ . Na tabela seguinte apresentam-se as aproximações obtidas nas calculadoras CASIO-CFX-9850 e TI-83 para os valores  $u_{10^k}$  com  $k=1, \dots, 16$ .

$k$	CASIO-CFX-9850	Ti-83
1	2,593742460	=
2	2,704813829	=
3	2,716923932	=
4	2,718145927	=
5	2,718268237	=
6	2,718280469	=
7	2,718281693	=
8	2,718281815	=
9	2,718281827	=
10	2,718281828	=
11	2,718281828	=
12	2,718281828	=
13	2,718281828	2,760577856
14	1	1
15	1	1
16	1	1

Com 10 algarismos, os valores obtidos nas duas máquinas coincidem excepto no caso de ser  $n = 10^{13}$ . A partir de  $k = 14$ , em ambas as máquinas, a representação da base  $1 + 10^{-14}$  será 1 e, por esta razão, os correspondentes valores de  $u_n$  dados são iguais a 1.

Apresentam-se a seguir as aproximações obtidas nas mesmas calculadoras para  $n = 2^k$  com  $k = 10, 15, 20, 25, 30, 35, 40$  e 45.

$n$	CASIO-CFX-9850	Ti-83
$2^{10}$	2,716955729	2,716955729
$2^{15}$	2,718240351	2,718240350
$2^{20}$	2,718280514	2,718280486
$2^{25}$	2,718281570	2,718279746
$2^{30}$	2,718274313	2,718215939
$2^{35}$	2,717924089	2,717924089
$2^{40}$	2,690051841	2,695061956
$2^{45}$	1	1

Como se pode apreciar, com  $n = 2^k$  não se conseguem obter aproximações com mais de sete algarismos correctos. Estes resultados não são surpreendentes já que nenhuma das calculadoras representa exactamente a base  $1 + 1/n$ , sendo  $n$  uma potência de 2. Por outro lado, as calculadoras produzem em muitos casos resultados

distintos porque as representações de  $1 + 1/n$  são diferentes (um dígito mais no caso da CASIO). Admitindo que numa mesma sala de aula ambas as calculadoras poderão estar a ser usadas pelos alunos, o professor terá de estar preparado para, de uma forma simples, explicar estas diferenças que seguramente não deixarão de intrigar os alunos.

Observe-se ainda que, ao contrário do que sucede com  $n = 10^k$ , os valores produzidos com  $n = 2^k$  não estão ordenados de forma crescente; esta situação é especialmente infeliz tendo em conta que se quer fazer passar a mensagem de que a sucessão de termo geral (1) é monótona crescente. Por esta razão, os manuais escolares usam valores  $n = 10^k$ ; se a aritmética das calculadoras fosse binária, os números da forma  $2^k$  seriam preferíveis.

### 9. Conclusões

É surpreendente que a maioria dos relatórios/estudos sobre a formação contínua dos professores de Matemática em Portugal que consultámos não tenha ainda abordado as exigências científicas que se colocam nessa formação. Surpreende-nos também, na literatura consultada e que serve de apoio generalizado aos professores de Matemática em Portugal, a pouca relevância dada a reflexões na área das questões numéricas (nomeadamente da teoria de erros) e que se impõem sempre que se aposta no recurso às novas tecnologias como ferramenta de ensino da Matemática. Surpreende-nos ainda o grau de popularidade que, na classe dos professores dos Ensino Básico e Secundário, tem vindo a atingir a formação em cursos de "Iniciação ao Excel", "sobre PowerPoint", ou ainda "Como usar as calculadoras na sala de aula".

Juntamos a estas "surpresas" a impressão que recolhemos da consulta da documentação ilustrativa (da responsabilidade dos representantes dos fabricantes dos modelos adoptados nas nossas escolas) das "maravilhas" pedagógicas atingíveis com o uso das calculadoras na sala de aula; anotamos também a ausência de quaisquer esclarecimentos específicos que, também no decurso desta investigação, julgámos pertinentes e procurámos obter junto de um desses representantes. Acrescentemos o facto, já referido no início deste artigo, de que as calculadoras são instrumento de uso obrigatório nas aulas de Matemática do Ensino Secundário em Portugal.

Nestas circunstâncias não é difícil perceber que, estando porventura reunidas as condições necessárias à difusão das calculadoras entre os professores de Matemática, não estão igualmente satisfeitas as condições de suficiência na informação e ainda menos na formação científica dadas a estes professores.

As calculadoras gráficas são, no nosso entendimento, uma ferramenta que pode ser extraordinariamente útil no ensino de diferentes tópicos dos programas de Matemática do Ensino Secundário. Porém, para garantir uma boa utilização das máquinas, importa compreender as suas limitações tecnológicas.

A incapacidade de representar exactamente todos os números na máquina faz com que não raramente os resultados produzidos se afastem significativamente dos valores correctos. Neste trabalho explorámos com detalhe os erros que ocorrem no cálculo dos termos da sucessão que tende para o número de Neper.

A atitude que o professor deve cultivar nos seus alunos é a de não aceitar cegamente os resultados dados pela calculadora e discutir sempre tais resultados à luz do que se sabe sobre o problema em estudo. **M**

### Referências

- [1] **Bernardes, A.**, et al. (2004). *Matemática 11*. Vol.3, Edições Contraponto, Porto.
- [2] **Gomes, F. e Viegas, C.**, (2004). *Xeqmat 11.º ano*. Vol.2, Texto Editora, Lisboa.
- [3] **Jorge, A.**, et al. (2004). *Infinito 11.º ano*. Vol.3, Areal Editores, Porto.
- [4] **MATLAB**, *The Language of Technical Computing*, <http://www.mathworks.com>.
- [5] **Overton, M.**, (2001). *Numerical Computing with IEEE Floating Point Arithmetic*, Society of Industrial and Applied Mathematics.
- [6] **Teixeira, P.**, et al. (1999). *Funções: Matemática – 12.º ano de escolaridade*, Departamento do Ensino Secundário, Ministério da Educação.

### Bibliografia

- Consciência, M.**, (2003). "Calculadoras gráficas: algumas limitações". *Gazeta da Matemática*, n.º 145, pp. 34-42.
- Teixeira, P.**, et al. (1998). *Funções: Matemática – 11.º ano de escolaridade*, Departamento do Ensino Secundário, Ministério da Educação.