Gazeta de Matemática

Nº 0157

Publicação quadrimestral
da Sociedade Portuguesa de Matemática
Ano LV | Abril 2009
4,20€

A Demanda do Centro de Portugal

18 | Apanhados na Rede: Pitágoras: Factos e Lendas
[António Machiavelo]

32 | Livros Contados: Entrevista com Ron Aharoni
[Jorge Nuno Silva]

Editorial

por **Jorge Buescu**
[Universidade de Lisboa]

É um lugar-comum bastante frequente – pelo menos para nós, amantes da Matemática – dizermos que "a Matemática está em todo o lado". Mas, lugar-comum ou não, a verdade dos factos é que ela está mesmo por todo o lado; basta fazermos as perguntas apropriadas para a sua presença ser inevitável.

Uma dessas perguntas é "Onde fica o centro de Portugal?". É uma pergunta frequente, que até tem uma resposta frequente. Quem não viu já imagens de um político, em plena campanha eleitoral, a fazer discursos inflamados num marco geodésico próximo de Vila de Rei, "o centro de Portugal" e portanto seu símbolo? João Filipe Queiró, da Universidade de Coimbra, trata este assunto de forma brilhante, mostra como a pergunta é essencialmente matemática e fornece a resposta, que depende de forma não-trivial de ferramentas analíticas e numéricas. E não, o centro de Portugal não é em Vila de Rei.

Como mostram os problemas do nosso Canto Délfico, a Matemática está até presente em situações tão simples como um jogo de futebol! De facto, esta edição do "Canto" começou, aparentemente, quando a meio de um jogo da Académica de Coimbra alguém perguntou o que é o esférico (ignora-se se a pergunta foi feita aquando da marcação de um canto).

A secção *O que é*, desta vez a cargo de Paulo Mateus, do IST, trata o problema P=NP – que o matemático Stephen Smale descreveu como "o maior presente da Teoria da Computação à Matemática". É um dos Problemas do Milénio do Clay Institute, e uma eventual resposta negativa teria consequências desastrosas sobre a segurança das comunicações, junto da qual o colapso do sistema financeiro no Outono passado seria provavelmente um fenómeno menor. A Matemática está *mesmo* por todo o lado.

Esta edição da Gazeta é ainda marcada pelo regresso do *Inquérito* e por uma entrevista a Ron Aharoni, um matemático profissional que decidiu dedicar-se ao ensino básico em Israel. Aharoni esteve em Portugal em Novembro passado, na Fundação Gulbenkian, onde deu conta das suas experiências na Conferência *Ensino da Matemática: Questões e Soluções*, e onde também lançou o seu livro *Aritmética para Pais*, publicado pela Gradiva.

A Matemática está *mesmo* por todo o lado. Os seus amantes podem orgulhar-se disso.

No âmbito de uma colaboração acordada entre a Gazeta e o Atractor, este é um espaço da responsabilidade do Atractor, relacionado com conteúdos interactivos do seu site (www.atractor.pt). Quaisquer reacções ou sugestões serão bem-vindas para atractor@atractor.pt.

A Matemática do Disfarce

Que este processo de fazer arte cause estranheza, não admira; o que admira é que haja cousa alguma que não cause estranheza. Fernando Pessoa

As anamorfoses são representações distorcidas de imagens num plano ou noutra superfície que, quando observadas de um ponto de vista não convencional, ou reflectidas por um espelho curvo, aparecem com proporções correctas e identificáveis. Embora de efeito mágico, são resultado de fórmulas precisas, obtidas com recurso à perspectiva, de figuração em superfícies planas de objectos tridimensionais. São contudo deformações não lineares e correspondem a uma utilização insólita da geometria projectiva, uma vez que, ao contrário desta – que teve, na sua origem, a intenção de representar mais fielmente a realidade, devolvendo-lhe a forma e a profundidade que o carácter plano de uma pintura lhe retira –, as anamorfoses pretenderam mesmo, em tempos idos, difundir com disfarce mensagens sigilosas ou cenas de carácter suspeito.

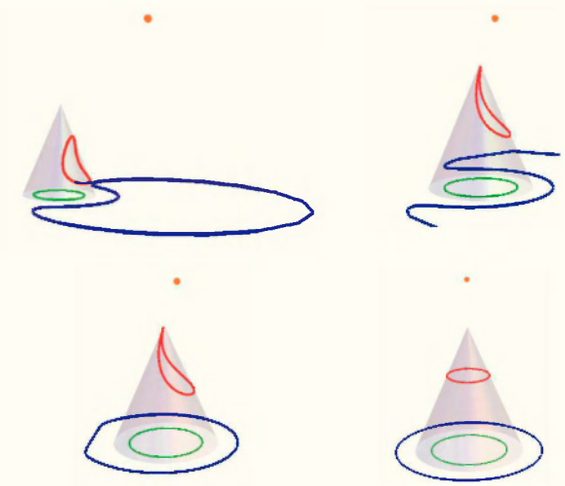
Numa representação em perspectiva, as rectas são transformadas em rectas, mas as linhas paralelas podem ser levadas em linhas convergentes num ponto de fuga, o que significa que as distâncias não são preservadas. O objecto é desenhado, tendo sido escolhido previamente um ponto de vista, através de uma «janela» com uma grelha rectangular; e, portanto, é necessário que este ponto de vista seja respeitado pelo observador. Um tipo de anamorfose com muito sucesso no século XV, e primeiro sistematizada por Leonardo da Vinci, é o que apresenta de um ponto de vista convencional uma imagem sem sentido e exige do observador a busca de outro ponto de vista, em geral em posição extrema, a partir do qual o truque se desvenda e a imagem é identificável. Uma tal anamorfose pode ser produzida com uma direcção de visão não perpendicular à janela (que pode até nem ser plana); a um observador que coloque a janela na vertical e olhe a pintura de frente, a imagem aparece grotesca ou mesmo irreconhecível.

Uma versão moderna das técnicas para produzir anamorfoses é o *Quarto de Ames*[1], onde figuras com representação correcta são

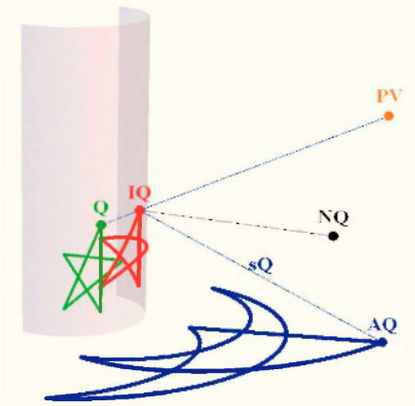
Caderno_1
1ª a-feira, : de Maio de 2009 22:54:00

colocadas num espaço com contornos manipulados (mas que parecem normais ao observador), de modo a que elas apareçam com proporções distorcidas.

As anamorfoses correspondentes às figuras que ilustram este texto têm outra génese. É escolhido um ponto de vista e a figura tridimensional é projectada a partir dele numa superfície curva e espelhada (em geral um cilindro ou um cone), que reflecte esta projecção num plano segundo a regra da igualdade dos ângulos de incidência e de reflexão. Deste modo, a distorção que resulta da primeira projecção, como descrito anteriormente, é acentuada pelo desvio provocado pelo espelho curvo.



Este procedimento pode ser invertido, e é esse o uso mais corrente das anamorfoses em arte: dada uma anamorfose (desenho plano com imagem deformada), podemos procurar a superfície espelhada e o ângulo de visão apropriados para deslindar o mistério e a reconhecer.



A figura ao lado mostra como construir a anamorfose cilíndrica de um objecto. Para vermos em reflexão a figura verde a partir do ponto de vista *PV*, teremos, para cada ponto *Q*, de determinar a intersecção *IQ* da semi-recta *PV Q* com o cilindro e depois determinar a normal *IQ NQ* ao cilindro no ponto *IQ* e a semi-recta *sQ* de origem *IQ* no plano *PV IQ NQ*, tal que ang (*PV*,*IQ*,*NQ*) = ang (*NQ*,*IQ*,*sQ*). A intersecção *AQ* desta semi-recta *sQ* com o plano horizontal fixado é a anamorfose de *Q*: é o (único) ponto desse plano horizontal cuja reflexão no cilindro é vista a partir de *PV* exactamente na direcção de *Q*. A transformação que leva *Q* em *AQ* é invertível. Um processo análogo pode ser utilizado com outras superfícies reflectoras, por exemplo um cone.

O Atractor desenvolveu um módulo interactivo para construir anamorfoses cilíndricas e observar a execução do desenho em tempo real. Um programa especialmente elaborado para o efeito permite que, usando um rato e um editor de imagem incorporado, o utilizador faça um desenho num ecrã vertical. Ao mesmo tempo, esse desenho é *traduzido* pelo programa numa anamorfose que vai sendo desenhada num outro ecrã (horizontal), onde está apoiado um cilindro reflector de eixo vertical. Nesse cilindro reflector é *recuperada* a imagem original, por reflexão da anamorfose em construção. Estava previsto que este módulo acompanhasse a exposição *Experimentar Matemática*, que esteve recentemente no nosso país. Acabou, no entanto, por só ser exibido em conjunto com ela no CMUP (Porto); esteve depois no Fórum Ciência Viva (Lisboa, 2008). Em ambos os casos, despertou grande interesse nos visitantes.M

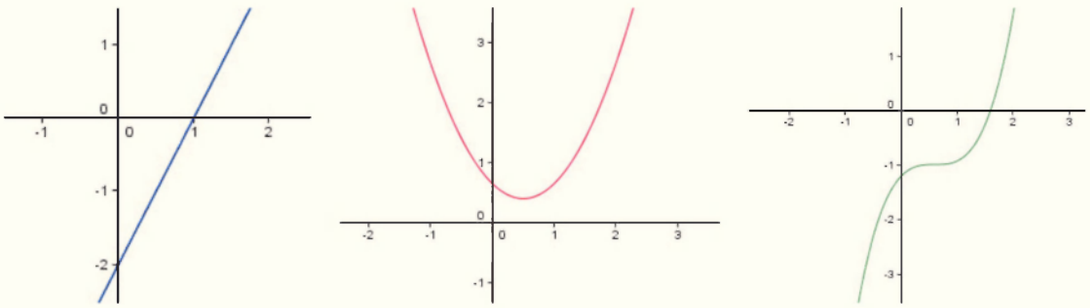
Caderno_1
ter a-feira, de Maio de 2009 21:54:04

por Jorge Nuno Silva
[Universidade de Lisboa]

Polinómios

Todos pensamos conhecer bem esta classe de funções. Contudo, viajar por estes terrenos pode trazer surpresas, e não só aos mais distraídos…

Os polinómios são estudados bem cedo no currículo escolar. Correspondem-lhes funções com representações gráficas simpáticas e, aparentemente, poucos segredos.



Gráficos de polinómios de graus 1, 2 e 3.

As expressões gerais dos polinómios representados são

$$y=ax+b;\ y=ax^2+bx+c;\ y=ax^3+bx^2+cx+d$$

onde a, b, c e d são números.

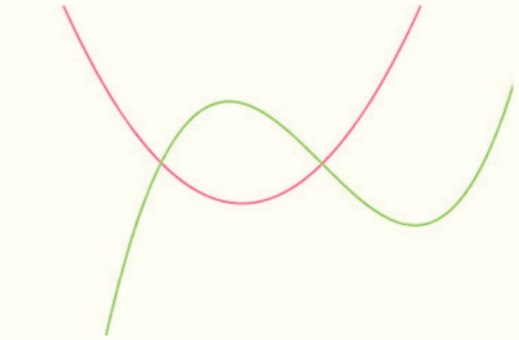
Que uma função linear fica definida pelo seu valor em duas concretizações da variável é fácil de admitir, já que temos dois parâmetros, a e b, na equação geral, que podemos afinar. Por exemplo, se quisermos a função linear P tal que $P(0)=-2$ e $P(1)=0$, basta-nos resolver o sistema

$$-2=a\times 0+b\ ;\ 0=a\times 1+b,$$

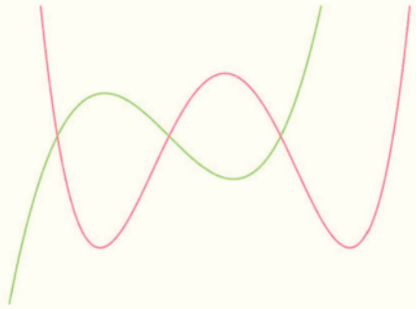
cuja solução é $P(x)=2x-2$, que está representado na figura mais à esquerda.

Mas será que o conhecimento de dois valores de uma função polinomial de grau superior nos permite determinar o respectivo polinómio? "Claro que não!", dirão os leitores. Por exemplo, tanto x^2-1 como x^3-x se anulam em 1 e em -1.

Caderno_1
ter a-feira, de Maio de 2009 21:54:17



Gráficos de dois polinómios que partilham dois valores



Gráficos de dois polinómios que partilham três valores

Contudo… um dia a Ara e o Breu tiveram o seguinte diálogo.

Ara: Pensa num polinómio de coeficientes inteiros não negativos, de qualquer grau, mas não me digas nada sobre ele.

Breu: Hmmm, está bem. Já está.

Ara: Que valor toma para x=1?

Breu: 14.

Ara: Que valor toma para x=17?

Breu: Caramba, tenho de fazer contas… dá 1 236 696 894.

Ara: Pois bem, o teu polinómio é $3x^7+4x^5+5x^2+2$.

Breu: Está certo. Mas foi à sorte que adivinhaste! Dois valores não são suficientes para determinar um polinómio…

Ara: Acertarei sempre! Queres apostar?

Acontece que a Ara tem razão. Os leitores compreendem o seu *modus operandi*?

Sobre os problemas do número anterior:

1 – As moedas em fila podem ser numeradas de 1 a 100. A Ara pode calcular a soma das moedas de numeração par e das de ordem ímpar. É fácil escolher um desses conjuntos para si. Se o número for ímpar (como 101) e as moedas das extremidades originais tiverem valores pequenos, o Breu pode ganhar, pela mesma razão.

2 – Há somente um número par na lista. Sendo o quadrado de ordem par, as linhas em que o 2 ocorrer terão soma ímpar, ao contrário das outras.

3 – O número de casas negras é sempre ímpar.

Caderno_1
ter a-feira, de Maio de 2009 21:54:38

por João Filipe Queiró
[Departamento de Matemática, Universidade de Coimbra]

A Demanda do Centro de Portugal

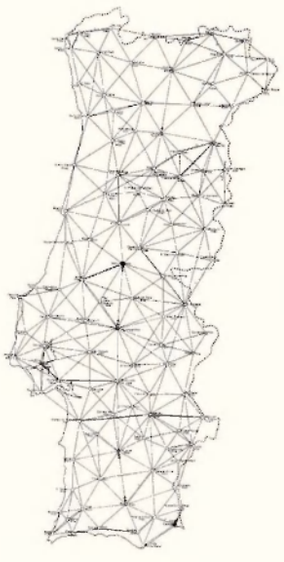
Onde fica o centro de Portugal? O que é o centro de Portugal? Uma incursão pelas aplicações da Matemática à Geografia e à Demografia.

Quando visitamos o Cabo da Roca, podemos adquirir um certificado atestando que estivemos no ponto mais ocidental da massa continental euro-asiática. A autarquia local, de forma imaginativa, consegue assim gerar algumas receitas, já que o serviço é popular entre os turistas.

Há muitos anos, quando estava na fila para obter a minha certidão de ocidentalidade, ocorreu-me a questão de procurar outros pontos que tivessem interesse exclusivamente por motivos geográficos. E lembrei-me de um candidato óbvio: o centro de Portugal.[1]

Surge logo a questão do que se deve entender por "centro" de uma figura irregular, sem simetrias. O centro de um círculo, de um rectângulo, de um losango, todos sabemos o que é. Mas como saber o que é o centro de uma figura como o mapa de Portugal? Já voltarei a esse assunto.

Nestas coisas não se inventa nada, e o tipo de problema que nos interessa já foi abordado noutros países. Para o caso de Portugal, encontrei uma publicação do Instituto Geográfico e Cadastral de 1983, onde vi o seguinte mapa, com a legenda "Rede geodésica primordial":



O nó central desta rede, visível a meio do mapa, está no Pico da Melriça, no concelho de Vila de Rei, distrito de Castelo Branco. Trata-se de uma elevação com 593 metros de altitude, no topo da qual se ergue um imponente marco geodésico.

O Pico da Melriça é uma atracção turística, promovida nos materiais publicitários da Região de Turismo dos Templários como o "Centro Geodésico de Portugal". O seu valor simbólico é repetidamente utilizado em vários tipos de eventos, incluindo iniciativas partidárias em campanhas eleitorais.

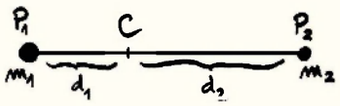
Estará o nosso problema resolvido? Será o Pico da Melriça o centro de Portugal? Foi este o problema que estudei e o que venho aqui descrever são as minhas conclusões.

O ponto de partida tem de ser a definição de centro de uma figura plana[2]. O conceito que nos vai interessar é o de centro de gravidade, ou centro de massa, ou baricentro, ou centróide. Informalmente, trata-se do ponto em que, para efeitos de equilíbrio, podemos supor concentrada toda a massa da figura. A única hipótese que postulamos é a da "associatividade" do centro de gravidade: se dividirmos uma figura em várias partes, então o centro da figura total é o centro do conjunto dos centros das partes, supondo colocada em cada um desses pontos a massa da respectiva parte.

[1] Ao longo do artigo usarei a palavra Portugal para significar apenas o território continental, isto é, sem incluir as ilhas portuguesas no Atlântico. A inclusão das ilhas, algumas das quais são grandes, traria resultados muito diferentes.
[2] Poderíamos estudar a questão para figuras no espaço, não necessariamente planas, mas não o faremos neste artigo.

Caderno_1
ter a-feira, de Maio de 2009 21:54:13

Começamos pela situação mais simples de todas: a figura constituída apenas por dois pontos P_1 e P_2, estando em cada um colocada uma massa. Chamemos m_1 e m_2, respectivamente, a essas massas.



Pelo chamado "princípio do equilíbrio de Arquimedes", o centro de gravidade deste sistema de dois pontos é o ponto C do segmento $[P_1,P_2]$ que satisfaz o seguinte: sendo d_1 e d_2 as distâncias de C a P_1 e P_2, respectivamente, deverá ter-se

$$m_1 d_1 = m_2 d_2.$$

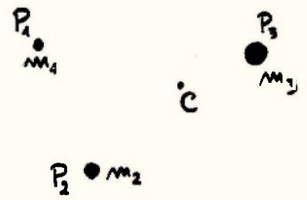
Se considerarmos um eixo que contenha o segmento $[P_1,P_2]$ e nesse eixo introduzirmos coordenadas, de forma que a de P_1 seja x_1, a de P_2 seja x_2 e a de C seja x_C, a condição fica

$$m_1(x_C - x_1) = m_2(x_2 - x_C).$$

Resolvendo esta equação, obtemos

$$x_C = \frac{m_1 x_1 + m_2 x_2}{m_1 + m_2}.$$

Consideremos agora três pontos, P_1, P_2 e P_3, com massas m_1, m_2 e m_3, respectivamente. Fixemos um sistema de eixos no plano dos pontos e suponhamos que as coordenadas de P_1, P_2 e P_3 são, respectivamente, (x_1,y_1), (x_2,y_2) e (x_3,y_3). Para achar o centro de gravidade, C, deste sistema de três pontos com massas, podemos usar a associatividade acima referida: determinamos o centro do sistema formado por P_1 e P_2, colocamos aí a massa m_1+m_2, e depois determinamos o centro do novo sistema formado pelo ponto assim obtido e por P_3.



Facilmente se vê que as coordenadas de C são dadas por

$$x_C = \frac{m_1 x_1 + m_2 x_2 + m_3 x_3}{m_1 + m_2 + m_3} \quad \text{e} \quad y_C = \frac{m_1 y_1 + m_2 y_2 + m_3 y_3}{m_1 + m_2 + m_3}.$$

Se em vez de três tivermos k pontos, $P_1, P_2, \ldots, P_k$, com coordenadas (x_1,y_1), (x_2,y_2), ..., (x_k,y_k) e massas $m_1, m_2, \ldots, m_k$, um raciocínio análogo conduz-nos às seguintes fórmulas para as coordenadas do centro de gravidade desse sistema de pontos:

$$x_C = \frac{m_1 x_1 + m_2 x_2 + \cdots + m_k x_k}{m_1 + m_2 + \cdots + m_k} \quad \text{e} \quad y_C = \frac{m_1 y_1 + m_2 y_2 + \cdots + m_k y_k}{m_1 + m_2 + \cdots + m_k}.$$

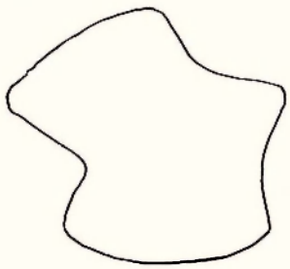
Se as massas $m_1, m_2, \ldots, m_k$ forem todas iguais, facilmente se vê que as fórmulas se simplificam para

$$x_C = \frac{x_1 + x_2 + \cdots + x_k}{k} \quad \text{e} \quad y_C = \frac{y_1 + y_2 + \cdots + y_k}{k},$$

Caderno_1
ter a-feira, 5 de Maio de 2009 22:54:16

Ou seja: no caso de as massas serem todas iguais, a localização do centro de gravidade não depende das massas mas apenas das posições dos pontos. E, nesse caso as coordenadas do centro são simplesmente as médias aritméticas das coordenadas dos pontos.

Passemos agora a figuras planas de outro tipo. Como achar o centro de gravidade de uma figura como a seguinte?

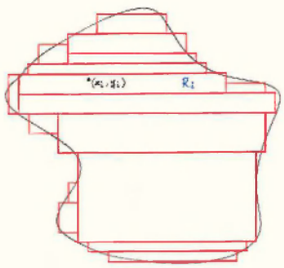


Pensamos nesta figura como uma placa plana R feita de um material cuja densidade pode variar de ponto para ponto. Uma ideia para determinar o centro é aproximar a figura dada usando rectângulos.

Fixando um sistema de eixos no plano da figura, a densidade é uma função real de duas variáveis reais $\delta(x,y)$. Aproximamos então a figura R pela reunião de um número finito de rectângulos R_i e em cada um deles escolhemos um ponto qualquer (x_i,y_i).

Claramente, uma aproximação para a massa do rectângulo R_i é dada por

$$\delta(x_i,y_i) \cdot \text{área de } R_i.$$



Colocando, para cada i, esta massa aproximada no ponto (x_i,y_i), obtemos um número finito de pontos com massas. O centro deste sistema de pontos é uma aproximação para o centro de gravidade, C, da figura R. Recordando as fórmulas vistas anteriormente, as coordenadas de C podem portanto ser aproximadas da seguinte forma:

$$x_C \approx \frac{\sum_i x_i \cdot \delta(x_i,y_i) \cdot \text{área de } R_i}{\sum_i \delta(x_i,y_i) \cdot \text{área de } R_i} \quad \text{e} \quad y_C \approx \frac{\sum_i y_i \cdot \delta(x_i,y_i) \cdot \text{área de } R_i}{\sum_i \delta(x_i,y_i) \cdot \text{área de } R_i}.$$

Se agora tomarmos rectângulos cada vez mais pequenos, vamos obtendo aproximações cada vez melhores. Fazendo o tamanho dos rectângulos tender para zero (uma forma de controlar isso é, por exemplo, olhando para a maior das diagonais dos rectângulos), obtemos, no limite, expressões exactas para as coordenadas do centro de gravidade da figura R:

$$x_C = \frac{\iint_R x\,\delta(x,y)\,dx\,dy}{\iint_R \delta(x,y)\,dx\,dy} \quad \text{e} \quad y_C = \frac{\iint_R y\,\delta(x,y)\,dx\,dy}{\iint_R \delta(x,y)\,dx\,dy}.$$

Note-se que o denominador de ambas as expressões, $\iint_R \delta(x,y)\,dx\,dy$, é a massa total da figura.

No caso de a densidade, $\delta(x,y)$, ser constante – isto é, de a placa ser homogénea – estas expressões simplificam-se para:

$$x_C = \frac{\iint_R x\,dx\,dy}{\iint_R dx\,dy} \quad \text{e} \quad y_C = \frac{\iint_R y\,dx\,dy}{\iint_R dx\,dy}.$$

Ou seja: no caso homogéneo, que é o que nos interessa, a localização do centro de gravidade não depende da densidade, mas apenas da forma da figura. O centro, neste caso, é portanto um ponto com uma caracterização apenas geométrica. Podemos notar que o denominador de ambas as expressões, $\iint_R dx\,dy$, é simplesmente a área da figura.

Caderno_1
ter a-feira, 5 de Maio de 2009 22:54:19

Para calcular estes integrais duplos recorremos a um teorema bem conhecido de Análise, que permite reduzir esse cálculo a outro envolvendo apenas a fronteira da figura R. Trata-se do teorema de Green, que afirma que, sendo f e g duas funções reais definidas numa região plana R com fronteira $fr(R)$, se tem[3]

$$\iint_R \left(\frac{\partial g}{\partial x} - \frac{\partial f}{\partial y}\right) dx\,dy = \int_{fr(R)} f\,dx + \int_{fr(R)} g\,dy.$$

(Os integrais que figuram no segundo membro são integrais curvilíneos.)

O teorema de Green pode ser imediatamente usado (e de mais que uma maneira, conforme as escolhas de f e g) para calcular os integrais duplos que nos interessam. Assim, temos, por exemplo, as seguintes expressões:

$$\text{área de } R = \iint_R 1\,dx\,dy = \int_{fr(R)} x\,dy = -\int_{fr(R)} y\,dx$$

$$\iint_R x\,dx\,dy = \frac{1}{2}\int_{fr(R)} x^2\,dy = -\int_{fr(R)} xy\,dx$$

$$\iint_R y\,dx\,dy = \int_{fr(R)} xy\,dy = -\frac{1}{2}\int_{fr(R)} y^2\,dx$$

Para cada um dos integrais duplos, temos aqui duas expressões com integrais curvilíneos. No cálculo aproximado que vamos fazer, usaremos, para cada um desses três integrais, a média aritmética dos dois integrais curvilíneos correspondentes, com a intenção de diminuir os erros cometidos na aproximação.

Usando a definição de integral curvilíneo como limite de somas, obtemos as seguintes aproximações para as coordenadas do centro da figura R:

$$x_C \approx \frac{\sum_{i=1}^{n}\left[\frac{1}{2}x_i^2(y_i - y_{i-1}) - x_i y_i (x_i - x_{i-1})\right]}{\sum_{i=1}^{n}\left[x_i(y_i - y_{i-1}) - y_i(x_i - x_{i-1})\right]}$$

$$y_C \approx \frac{\sum_{i=1}^{n}\left[x_i y_i (y_i - y_{i-1}) - \frac{1}{2}x_i^2(x_i - x_{i-1})\right]}{\sum_{i=1}^{n}\left[x_i(y_i - y_{i-1}) - y_i(x_i - x_{i-1})\right]}$$

onde $(x_1,y_1), (x_2,y_2), \ldots, (x_n,y_n)$ são pontos da fronteira de R.

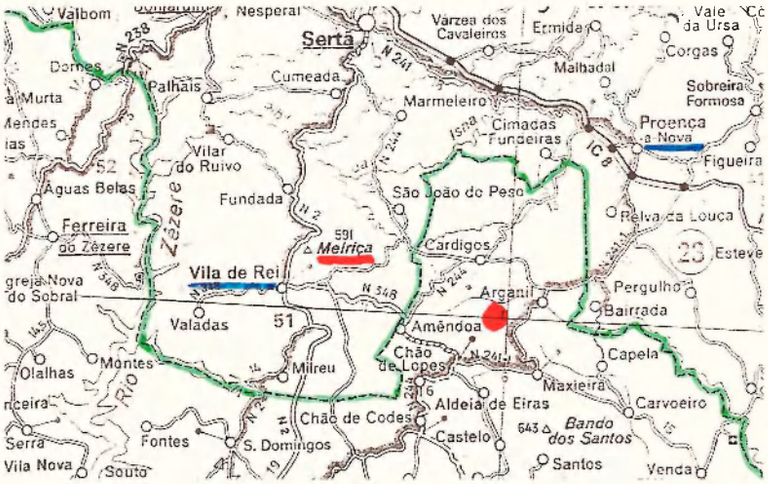
Antes de aplicar as fórmulas vistas à determinação do centro de Portugal, observemos que, como o país é muito pequeno relativamente ao perímetro da Terra, não há erro sensível em desprezar a esfericidade do globo terrestre e considerar o território plano.

Para realizar o cálculo usei o mapa de Portugal que encontrei na já referida publicação do Instituto Geográfico e Cadastral. Fotocopiei-o em papel milimétrico, de modo a dispor de um sistema de coordenadas. Registei as coordenadas de 762 pontos da fronteira da figura (isto é, da reunião das fronteiras terrestre e marítima do território continental de Portugal). Em seguida, utilizei as fórmulas anteriores com os pontos encontrados. O número elevado de pontos torna a aproximação razoável.

[3] Omitimos as condições exigidas a f, g e R, que supomos satisfeitas sem mais pormenores.

Caderno_1
ter a-feira, 5 de Maio de 2009 22:54:11

O resultado a que cheguei foi um ponto cerca de 11 km a leste e 3 km a sul do Pico da Melriça, a que podemos chamar centro geométrico ou geográfico de Portugal. Encontra-se entre as povoações de Arganil e Amêndoa, no concelho de Mação, distrito de Santarém.



Não sei se o interesse geográfico deste ponto se pode aproveitar para algum fim turístico. O exemplo de outros países sugere que não. A pequena cidade de Lebanon, Kansas, nos Estados Unidos, não conseguiu tirar grande partido do facto de se situar no centro geográfico do país.

*

Depois de realizado o trabalho que descrevi, ocorreu-me outra noção possível de centro de Portugal, aquilo a que se pode chamar o centro de gravidade demográfico do país. A definição rigorosa seria o centro de gravidade do sistema de pontos constituído por todos os habitantes do território, cada um na sua posição (e todos com a mesma massa, o que torna a questão puramente geométrica). Como calcular tal ponto em cada momento seria obviamente impraticável, simplifiquei o problema considerando os 278 concelhos do continente e, para cada um deles, supondo a respectiva população concentrada na sede do concelho.

Note-se que, ao contrário do centro geográfico, o centro demográfico de Portugal não é fixo, porque a distribuição da população no território vai mudando com o tempo.

As fórmulas que utilizei foram as vistas no início para um conjunto finito de pontos com massas. Os pontos são as sedes dos concelhos, cujas coordenadas geográficas foi necessário registar. As massas são as populações dos concelhos, apuradas pelos recenseamentos oficiais.

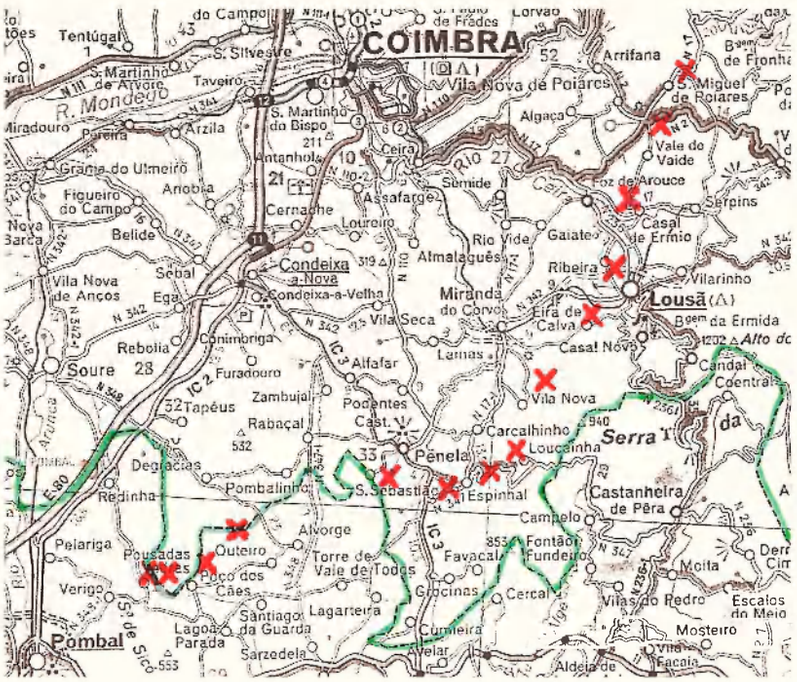
Fiz o cálculo do centro demográfico para os catorze recenseamentos realizados em Portugal desde o século XIX[4], nos anos de 1864, 1878, 1890, 1900, 1911, 1920, 1930, 1940, 1950, 1960, 1970, 1981, 1991 e 2001. O número de concelhos desde 1864 manteve-se bastante estável, o que facilitou as coisas. Os novos concelhos desde então são apenas o do Entroncamento (que aparece pela primeira vez no

[4]Agradeço a Luísa Saraiva, da Delegação de Coimbra do Instituto Nacional de Estatística, o empréstimo dos volumes com os dados populacionais, e à minha filha Luísa a grande ajuda na introdução desta enorme quantidade de dados numa folha de cálculo.

Caderno_1
ter a-feira, de Maio de 2009 21:54:23

recenseamento de 1930), o da Amadora e o de Vendas Novas (1970), e os de Odivelas, da Trofa e de Vizela (2001).

O centro demográfico de Portugal encontrava-se, em 1864, perto de S. Miguel de Poiares, cerca de 17 km a leste de Coimbra. Depois moveu-se lentamente para sul e um pouco para oeste, até 1930, altura em que o movimento passou a ser para oeste. O salto mais pronunciado deu-se entre 1960 e 1970, o que corresponde sem dúvida à forte emigração que se verificou nessa década a partir dos meios rurais, nomeadamente do interior. Em 2001, o centro demográfico encontrava-se cerca de 30 km a Su-sudoeste de Coimbra, na fronteira entre os concelhos de Soure e Pombal.



À primeira vista, esta localização do centro demográfico de Portugal pode surpreender-nos. Contudo, ela explica-se pelo facto de a população de Portugal continental estar muito concentrada na faixa litoral entre Setúbal e Braga. É curioso que o centro demográfico de Portugal em 2001 esteja próximo do ponto médio do segmento que une essas duas cidades.

Os centros geográfico e demográfico de um país podem ter interesse político. Há países em que a capital é uma cidade escolhida algo artificialmente para ocupar um lugar geograficamente central: é o caso de Espanha. Noutros países a capital está muito longe do centro geográfico, parecendo óbvio que cresceu naturalmente onde a população estava concentrada (embora as causas e os efeitos por vezes se possam misturar): o Egipto é um bom exemplo.

Voltando ao centro demográfico de Portugal, será interessante ver onde se situará em 2011, ano do próximo recenseamento da população. É de admitir que não se mova muito em relação à posição de 2001: é difícil que se desloque ainda mais para oeste e, por outro lado, a região do Grande Porto é muito populosa, pelo que não deve haver grande deslocação do centro demográfico para sul.

Caderno_1
ter a-feira, 5 de Maio de 2009 22:54:: 1

por Amílcar Branquinho e Alexander Kovacec
[Departamento de Matemática, Universidade de Coimbra]

O Esférico...

Caro leitor,
Este canto começou quando alguém perguntou a meio de uma partida de futebol da Académica de Coimbra o que era o esférico.

Realmente, é o objecto fundamental do jogo!

O objectivo de cada uma das equipas é introduzi-lo o maior número de vezes possível na baliza contrária, ainda que para tal seja preciso muito treino, trabalho de equipa, companheirismo, boas instalações, uma boa táctica, um bom treinador... Pelo menos isto é o que ouvimos com insistência nos *media*.

Na altura a questão parecia despropositada, pois o jogo estava a decorrer, mas vendo bem o esférico é o foco de todos os olhares e desejos e sem ele este desporto, como tantos outros, não existiria.

No entanto, apesar de ser o objecto em redor do qual gravita toda a actividade no terreno de jogo, não lhe prestamos a atenção que merece. Vamos hoje observá-lo, mas de um ponto de vista que pode parecer estranho: vamos vê-lo com *olhos matemáticos*.

Sim, caro leitor, não se surpreenda. Nesse esférico que passou tantas vezes pelas suas mãos, que tantas alegrias e tristezas lhe proporcionou, há mais surpresas matemáticas do que possa imaginar.

Quando está bem cheio, parece uma esfera perfeita, o corpo ideal para os filósofos gregos, uma criação dos deuses. Mas será realmente uma esfera?

Olhe a figura 1 com atenção. Observe as peças que a compõem. São polígonos regulares, pois têm todos os lados iguais.

Figura 1

Efectivamente, são pentágonos e hexágonos unidos entre si. Se tiver pouco ar, pode ficar perfeitamente equilibrada sobre cada uma das suas faces...

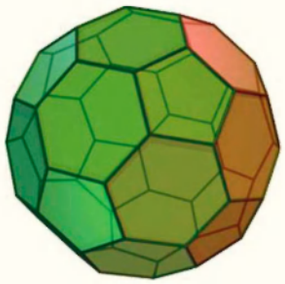


Figura 2

Sendo assim, deixa de ser uma esfera! Agora é um poliedro. Um poliedro que tem nome próprio, ainda que um tanto estranho: *icosaedro truncado* (cf. figura 2).

Figura 3

Vejamos como desenhar a bola de futebol (ver figura 3): comece por um pentágono regular e adicione linhas que partam de cada uma das pontas do pentágono, que servirão para formar 5 hexágonos em torno do pentágono. Na confluência de cada par

Caderno_1
ter a-feira, de Maio de 2009 21:54:55

de hexágonos, complete (com três linhas) um pentágono, ficando com seis pentágonos. Neste momento só temos de circunscrever uma circunferência e dar o toque final colorindo a preto os pentágonos.

Mas voltemos à estrutura do esférico, para contar o número de pentágonos e hexágonos que o compõem. Parece fácil? Já terminou? Bem... não se terá enganado? Os pentágonos não oferecem grande dificuldade; efectivamente, são 12.

Agora vamos contar os hexágonos... Talvez seja melhor pensar um pouco... Como vimos, cada pentágono está rodeado por cincos hexágonos; logo, deveríamos ter doze vezes cinco, isto é, sessenta hexágonos. Mas cada um deles está unido a três pentágonos diferentes. Nesse caso, temos sessenta a dividir por três, isto é, 20. Um total de 32 faces. Bom, não foi assim tão complicado.

Agora que começou a contar pode determinar quantas arestas (costuras) tem a bola? Um conselho: não tente contá-las uma a uma...

Se há 20 hexágonos e cada um tem 6 arestas, obtemos 120 arestas, mais 12 pentágonos por 5 arestas cada um, isto é 60. No total temos 180 arestas. Mas cuidado: cada aresta é comum a 2 polígonos. Assim, contámos cada aresta 2 vezes, logo, temos 90 arestas. Quem diria?

Agora que a curiosidade se apoderou de nós, determine quantos *vértices* (lugar geométrico onde as arestas confluem) tem a bola?

Para tal apliquemos um **teorema de Euler para poliedros**: *Considere-se um poliedro com s faces, l arestas e p vértices, então, s+p=l+2.*

Para o demonstrar comece por demonstrar o **teorema de Euler no plano**: *Dado um mapa qualquer, considere-se um número s de países, o número l de fronteiras e o número p dos vértices. Então, s+p=l+2.*

Para tal use indução sobre *l*. Coloque o poliedro no interior de uma esfera, de raio suficientemente grande, e projecte desde o seu centro (que está sem perda de generalidade no interior do poliedro) sobre a superfície esférica todos os pontos do poliedro. Ficamos com um mapa na superfície esférica. A figura assim desenhada pode projectar-se desde um ponto qualquer da esfera não pertencente ao mapa sobre o plano tangente à superfície esférica no ponto diametralmente oposto (projecção estereográfica, cf. figura 4). Aplicando o teorema de Euler no plano ao mapa resultante terminamos a demonstração.

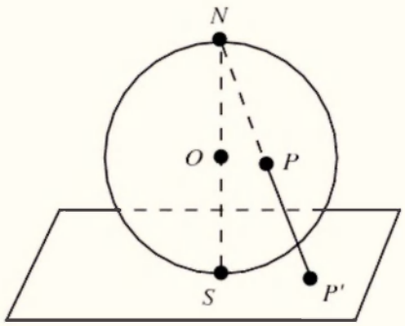


Figura 4

Vemos assim que há 60 vértices, pois 32+60=90+2.

Decididamente, não devemos preocupar os nossos Figos e Ronaldos com estas questões.

Voltemos ao nosso esférico, isto é, ao icosaedro truncado.

Ainda que à primeira vista não pareça, este poliedro obtém-se cortando os 12 vértices de um icosaedro - um dos 5 poliedros regulares descobertos por Platão há mais de 2500 anos, formado por 20 triângulos equiláteros, de onde vem o seu nome. Os 12 pentágonos correspondem aos 12 cortes nos vértices do icosaedro e os 20 hexágonos constituem o resto das faces do icosaedro truncado.

A partir de agora queremos acompanhar o leitor na determinação do volume de 2 sólidos, a pirâmide e a esfera, e também na construção de um icosaedro.

Problema 1: Demonstre por indução sobre *n* que $\sum_{k=1}^{n} k^2 = \frac{n(n+1)(2n+1)}{6}$.

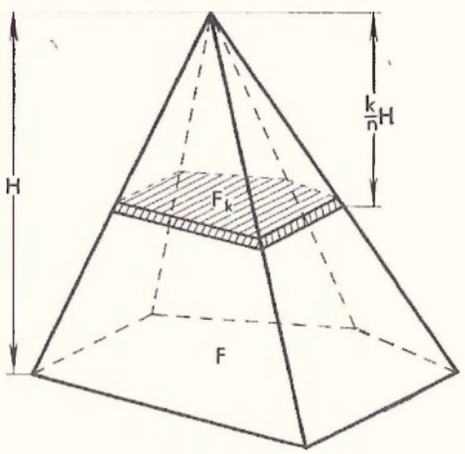


Figura 5

Caderno_1
ter a-feira, de Maio de 2009 21:54:39

Problema 2: Considere-se uma pirâmide com altura H e área da base F (cf. figura 5).

a) Dividindo a altura em n partes iguais e traçando por elas planos paralelos ao plano da base, obtemos n prismas, V_k, com altura $\frac{H}{n}$ (na verdade não temos prismas, mas sim pirâmides truncadas), cuja união funciona como aproximação da pirâmide. Determine o volume dos prismas, V_k, em termos de n, k, H e F.

b) Atendendo a que o volume da pirâmide é dado aproximadamente pela soma dos volumes dos prismas, V_k, escreva essa aproximação em termos de n, F e H.

c) Tomando o limite, quando n tende para infinito, mostre que o volume da pirâmide e igual a $\frac{FH}{3}$.

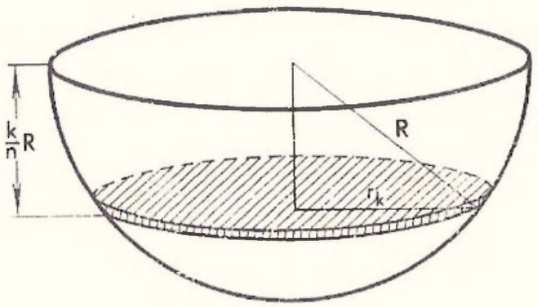


Figura 6

Problema 3: Considere-se a semi-esfera na figura 6, de raio R, dividida em n planos com espessura $\frac{R}{n}$, isto é, considere-se aproximada por n cilindros de altura $\frac{R}{n}$. Sendo o raio de cada um destes cilindros r_k, e utilizando a figura 4, mostre por um processo análogo ao do problema 2 que o volume da esfera é $\frac{4}{3}\pi R^3$.

Como aperitivo deixamos-lhe a construção do icosaedro a partir da sua planificação na figura 7. Na sua construção, use materiais transparentes, considerando os triângulos equiláteros que o compõem de 6 centímetros de lado.

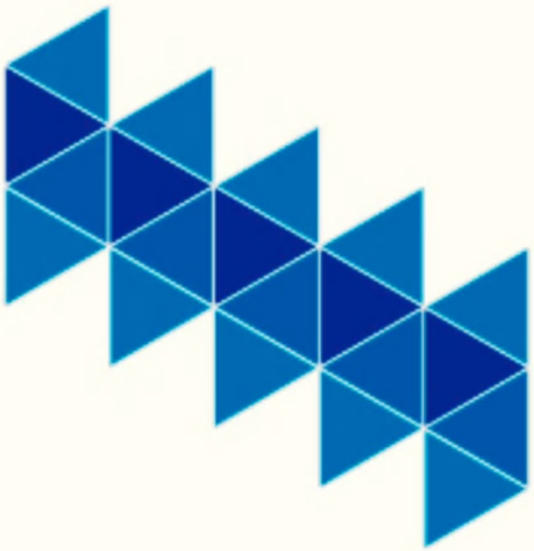


Figura 7

Problema 4: Ao terminar a construção do icosaedro, verá 12 pirâmides pentagonais. Cortando o icosaedro próximo dos vértices por planos paralelos às bases das pirâmides obtemos um icosaedro truncado. Determine o local onde o corte deve ser feito de modo a produzir hexágonos regulares.

Com a resolução destes quatro problemas, bem como com a dos problemas propostos no canto anterior, lançámos as bases para a resolução do problema maior, não trivial, que é comparar os volumes do icosaedro truncado com o da esfera que o circunscreve. Esperamos abordá-lo num próximo canto.

Envie as soluções para:

Projecto Delfos
Departamento de Matemática da FCTUC
Apartado 3008
EC Universidade
3001-454 COIMBRA

Caderno_1
ter a-feira, de Maio de 2009 2:54:2

por Fabio Chalub
[Universidade Nova de Lisboa]

Novas Informações da Teoria da Informação

A imagem abaixo pode dar-nos a impressão de estarmos a entrar na praia do Meco. No entanto o significado desejado é exactamente o oposto: num escritório para tirar cartas de condução, próximo de uma das praias mais famosas do mundo, a Direcção-Geral de Viação local tenta impor aos utentes um código de vestimenta que dê ao lugar a respeitabilidade que considera devida. Que tem a matemática com isto?

O que fazer?

Tudo o que se refere a comunicação é antigo como o ser humano. Talvez seja a habilidade de trocar informações a característica que mais nos diferencia do resto do reino animal. No entanto, os estudos matemáticos sobre a comunicação só começaram a ser sistematizados há exactos 60 anos, quando Claude Shannon e Waren Weaver publicaram o livro *"A Teoria Matemática da Comunicação"* [2]. Na verdade, o livro é uma reedição alargada de um artigo do ano anterior cujo título era parecido: *"Uma Teoria Matemática da Comunicação"*. A mudança da primeira palavra e a rápida transformação de um artigo científico em livro, ainda hoje reeditado, mostra a importância deste trabalho.

Shanon e Weaver definem três "problemas da comunicação": i) *o problema técnico*: quão exactamente estão a ser transmitidos nos símbolos? ii) *o problema semântico*: com que precisão os símbolos transmitem a mensagem desejada? iii) *o problema da efectividade*: será que o sentido transmitido fará o recipiente da mensagem comportar-se da forma adequada? Na figura 1 os problemas ii) e iii) estão claramente postos e o leitor pode divertir-se a pensar nos significados alternativos e nas suas consequências.

Os símbolos transmitidos podem ser um conjunto linear discreto, como as letras que compõem esta página, ou até mesmo uma função de três variáveis, sendo uma o tempo e duas que definem o plano da televisão. Vamos pensar no primeiro exemplo, por ser ele mais simples. Suponha que a transmissão dos caracteres é feita através de um canal que troca todas as letras A por uma outra qualquer. Sendo esta a letra mais comum em português, é natural que o efeito para a transmissão da informação seja maior do que se o mesmo acontecesse no X. Desta forma, as letras têm um conteúdo informativo distinto.

Como medir este conteúdo informativo? Shannon recorreu aos trabalhos de Ludwig Boltzman em física estatística no século 19, para definir a *entropia da informação*, ou, como a conhecemos actualmente, a *entropia de Shannon*. Para Boltzman, a entropia mede a desordem de um sistema; para Shannon, o que está a ser medido pelo inverso da entropia é a informação. Uma sucessão de caracteres alfabéticos completamente desordenada é aquela da qual não podemos obter nenhum significado. Para transmitir alguma informação é necessário colocar as letras certas nos lugares certos.

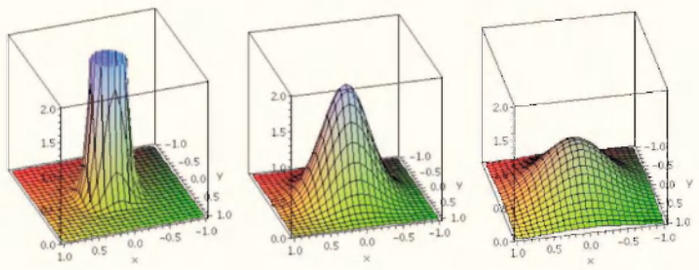
Na verdade, não exactamente, pois muitas vezes há redundância no texto (só em algumas linguagens artificiais não existe redundância). Desta forma é possível comprimir informação sem a perder. Torna-se mais económico guardá-la no disco rígido do computador ou enviar um ficheiro por *e-mail*. É o que fazem os programas de compressão de dados, como o *zip*. O limite teórico de compressão sem perda é dado pelo *teorema da codificação de fonte*, do próprio Shanon.

A expressão para a entropia é obtida da seguinte forma: considere n possíveis símbolos, ou expressões, ou seja o que for que tenha conteúdo informativo, cada um destes com probabilidade $p(1)$, $p(2)$, até $p(n)$ de ser o escolhido. Como estamos a falar de probabilidade, evidentemente temos de ter $p(1)+p(2)+\ldots+p(n)=1$. A entropia é dada por

Caderno_1
ter a-feira, de Maio de 2009 21:54:48

$H=-p(1)\log p(1)-p(2)\log p(2)-...-p(n)\log p(n)$.

Após esta definição o trabalho de Shannon passa a ter cariz essencialmente matemático. Antes de formular a sua conjectura recentemente demonstrada e já generalizada, vamos analisar um exemplo simples, conhecido como o *modelo dos passos de bêbado*. Depois de alguns copos a mais, uma pessoa põe-se a andar, mas dá passos totalmente ao acaso, sem uma direcção privilegiada. Em média, não sai do lugar. Esta é a maneira informal de dizer que a distribuição de probabilidade de sua posição no espaço tem média zero. No entanto, se considerarmos a média do quadrado da distância, esta tem de ser positiva. Mais exactamente, a raiz quadrada da média do quadrado da distância (média determinada a partir de um grande número de bêbados a andarem ao acaso) é proporcional à raiz quadrada do número de passos dados por cada pessoa (supomos todos os passos do mesmo tamanho). Como o número de passos é directamente proporcional ao tempo, concluímos que a dispersão do *movimento Browniano* é proporcional à raiz quadrada do tempo. Este nome foi dado originalmente ao movimento aleatório de partículas de pólen dispersas na água, descrito pelo botânico britânico Robert Brown, explicado por Albert Einstein e que inspirou o matemático Mandelbrot a criar os fractais. Antes mesmo de Einstein, Luis Bachelier usou ideias parecidas no estudo da bolsa de valores.



Distribuição de probabilidade para o movimento Browniano em função do tempo: t=0.01 (esquerda), t=0.05 (centro), t=0.1 (direita). A medida que o tempo passa a dispersão aumenta, mas o ponto médio continua o mesmo: a origem. O eixo vertical é o mesmo nas três figuras (o que omite a parte mais alta do gráfico na figura da esquerda).

Finalmente chegamos à conjectura de Shannon: considere N eventos ("os bêbados"), cada um descrito por uma função X_1, X_2, ..., X_N, e considere a soma destas N variáveis e divida pela raiz quadrada de N. Chamamos ao resultado Z_N. O teorema central do limite garante que Z_N se aproxima de uma distribuição gaussiana quando N cresce. Além disto, e esta é a conjectura de Shannon, a entropia da variável aleatória Z_N é uma função não crescente de N.

Se $p(x)$ descreve a distribuição de probabilidade de um certo evento (no caso acima, o evento descrito pela função Z_N), então a sua entropia é dada pelo integral de $-p(x)\log p(x)$ em todos os valores possíveis de x, generalizando a fórmula descrita anteriormente.

No final da década de 50 foi finalmente demonstrado que a conjectura de Shannon era verdadeira quando comparávamos a entropia de 2 eventos com a de 1 evento. Foram necessários mais 45 anos para que finalmente, em 2004, 4 cientistas em 4 países diferentes a conseguissem demonstrar em toda a generalidade [1]. Mais 3 anos e foi finalmente possível caracterizar os casos em que a entropia é uma função decrescente (e não apenas não crescente) [3].

A demonstração em [1] usa um pouco de tudo: cálculo variacional, cálculo vectorial, análise funcional, semi-grupos, probabilidades e muitas, mas muitas, desigualdades entre funções. No entanto, para quem esperava a necessidade de desenvolver técnicas novas e profundas na solução de um problema com mais de 50 anos, a demonstração pode ser decepcionante. A verdadeira arte foi trabalhar conceitos já consolidados há décadas de forma criativa.

A extensão feita em [3] já é mais sofisticada. Foi necessário usar uma versão não-comutativa da teoria das probabilidades, desenvolvida por Voiculescu no início da década de 90. Foi a assim chamada teoria de probabilidades livres que permitiu não apenas a caracterização acima descrita, mas também a extensão do resultado anteriormente demonstrado para o caso de várias variáveis aleatórias em paralelo (como se muitos bêbados partissem de muitos pontos da cidade em simultâneo). M

Referências

[1] **Arstein, S., Ball, K., Barthe, F., Naor, A.** (2004). "Solution of Shannon's problem on the monotonicity of the entropy" *J. Am. Math. Soc* 17:975-982.

[2] **Shannon, C. E., e Weaver, W.** (1949). *The Mathematical Theory of Communication*. University of Illinois Press

[3] **Shlyakhtenko, D.** (com um apêndice por **H. Schultz**) "Shannon's monotonicity problem for free and classical entropy" (2007). *Proc. Nat. Acad. Sci.* 104(39) 15254-15258.

Caderno_2
quarta-feira, 6 de Maio de 2009 12:54:23

por António Machiavelo
[Departamento de Matemática Pura da Faculdade de Ciências da Universidade do Porto]

Pitágoras: Factos e Lendas

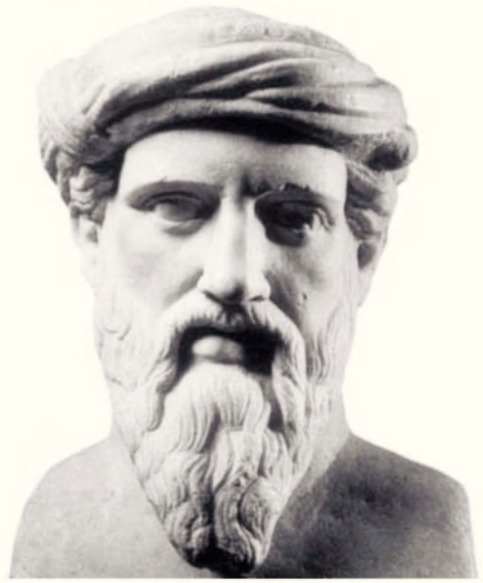
O chamado "teorema de Pitágoras" era já conhecido muito antes de Pitágoras ter nascido. Porque é então atribuído a Pitágoras? E como terá sido descoberto?

A imagem que todos temos de Pitágoras é a de um dos primeiros grandes matemáticos, imediatamente posterior a Tales de Mileto, que descobriu que num triângulo rectângulo a soma do quadrado dos catetos é igual ao quadrado da hipotenusa. É pois muito provável que o leitor receba com algum choque a notícia de que grande parte do que se sabe sobre Pitágoras foi inventado, em larga medida deliberadamente!

Como se pode ler no artigo sobre Pitágoras da *Stanford Encyclopedia of Philosophy*[1], já referido no último número desta rubrica, as indicações mais recuadas mostram que a fama de Pitágoras, em vida e mesmo 150 anos mais tarde, no tempo de Platão e Aristóteles, nada tinha a ver com a matemática. De facto, Pitágoras era nessa altura conhecido como entendido em questões ligadas à alma e aos rituais religiosos, como taumaturgo[2] e fundador de um modo de vida rigoroso e disciplinado[3].

Os relatos mais detalhados e extensos da vida e do pensamento de Pitágoras, dos quais provém a imagem que dele hoje temos, foram escritos cerca de 800 anos após este ter vivido, em obras cujo objectivo não era relatar objectiva e rigorosamente factos históricos. Como observa Jean-François Mattei[4]:

Todos estes textos tradicionais estão compostos segundo um género literário bem conhecido na época helénica que idealiza, com fins edificantes, o retrato moral do Sábio ou das confrarias religiosas...



Pitágoras

[1]Disponível em http://plato.stanford.edu/entries/pythagoras.
[2]As fontes "históricas" através das quais "conhecemos" Pitágoras contêm relatos de milagres e profecias que pretendem evidenciar os poderes sobre-humanos deste "homem-deus".
[3]Com regras estritas, sendo uma das mais famosas a interdição de comer favas. Sobre este tema, ver o curioso artigo "Favism – from the "avoid fava beans" of Pythagoras to the present", na revista da *Sociedade Helénica de Hematologia Haema* (2004, vol. 7, pp. 17 - 21), disponível online.
[4]Em *Pythagore et les Pythagoriciens*, Presses Universitaires de France, colecção "Que sais-je?", n.° 2732, 1993, p. 8.
[5]Secção 2.2 do documento da *Stanford Encyclopedia of Philosophy*.
[6]Citado por Jean-François Mattei na obra acima referida (p. 7) e por Walter Burkert em *Lore and Science in Ancient Pythagorism*, Harvard University Press, 1972 (o original alemão é de 1962), p. 2-3.
[7]Escrita de 1844 a 1852 e continuamente ampliada e melhorada até à sua última edição, em 1902. Disponível em http://books.google.com.

Caderno_2
quarta-feira, 6 de Maio de 2009 12:54:27

Pior ainda[5]:

Estes relatos (...) basearam-se em fontes mais antigas, que se perderam. Algumas dessas fontes mais antigas estavam grandemente contaminadas pela visão neopitagórica de Pitágoras como fonte de toda a verdadeira filosofia, cujas ideias Platão, Aristóteles e todos os filósofos posteriores haviam plagiado.

Ou seja, os discípulos, mais ou menos fanáticos, distorceram relatos e inventaram histórias, chegando mesmo a forjar documentos, com o objectivo de evidenciar a superioridade de Pitágoras relativamente a filósofos posteriores.

Já o filósofo e teólogo alemão Eduard Zeller[6] (1814- -1908), autor da obra erudita *Die Philosophie der Griechen*[7], tinha notado que quanto mais nos afastamos do período em que Pitágoras viveu, mais detalhes têm os relatos sobre o pitagorismo e o seu fundador, uma expansão que é fruto de preconceitos dogmáticos e interesses ideológicos com fins propagandísticos.

Em resultado destas circunstâncias, não é nada fácil perceber se Pitágoras teve ou não algo a ver com a matemática. Para se apreciar bem a dificuldade de separar o trigo do joio nesta matéria, nada melhor do que ler a secção 5 do referido documento da *Stanford Encyclopedia of Philosophy*, assim como um artigo de recensão crítica a dois livros recentes sobre Pitágoras, da autoria de Myles F. Burnyeat[8], intitulado "Other Lives"[9], no *London Review of Books*. Melhor ainda seria ler, se possível, as obras citadas nesses documentos.

Por outro lado, é sobejamente sabido que o teorema "de Pitágoras" era de alguma forma conhecido no antigo Egipto, na Babilónia[10], na Índia e na China centenas de anos, ou mesmo milénios, antes de Pitágoras ter nascido[11]. É possível que tenha sido descoberto mais de uma vez, de modo independente, embora haja quem defenda uma origem comum, de que não há qualquer registo, somente indícios não conclusivos. Mas como terá sido descoberto? Nesta questão os registos históricos são totalmente omissos. Há apenas especulações, umas quantas invenções e muitos disparates, que por vezes se lêem em manuais cujos autores deviam ter um pouco mais de cuidado.

Uma dessas especulações, um pouco infeliz e algo pobre do ponto de vista epistemológico, deve-se a Moritz Cantor (1829-1920), que a expõe na sua monumental obra *Vorlesungen über Geschichte der Mathematik*[12], considerada o trabalho fundador da história da matemática como disciplina rigorosa, e sem qualquer dúvida uma obra absolutamente notável. Cantor conjectura (vol. 1, pp. 105-106) que os Egípcios poderiam ter conhecido o triângulo rectângulo de lados 3, 4 e 5, que usariam para medir ângulos rectos com cordas em que uma certa unidade de comprimento seria marcada por nós igualmente espaçados, o que seria útil nas mais diversas construções. Não há no entanto qualquer testemunho histórico que sustente a hipótese de Cantor, que foi fortemente criticada por Bartel van der Waerden no prefácio de *Science Awakening*. Observe-se que fica ainda por explicar como terão os Egípcios descoberto essa particularidade do triângulo de lados 3, 4 e 5. Lamentavelmente, e por a obra de Cantor ser muito influente (com todo o mérito, repita-se!) e algumas pessoas serem ávidas de certezas mesmo onde elas não são possíveis, esta conjectura transformou-se em

[8]Professor aposentado de Filosofia Antiga da Universidade de Cambridge.
[9]Disponível em http://www.lrb.co.uk/v29/n04/burn02.html.
[10]Vários autores gregos relatam que Tales esteve no Egipto, de onde trouxe a ciência da geometria para a Grécia, e que Pitágoras viajou pelo Egipto e pela Babilónia, onde colheu ensinamentos. Verdadeiros ou não, esses relatos mostram, todavia, a alta consideração que os Gregos da antiguidade tinham por essas duas civilizações.
[11]Ver o livro de Bartel van der Waerden *Science Awakening*, originalmente publicado em holandês em 1950 e traduzido em inglês em 1954, e *Geometry and Algebra in Ancient Civilizations* (Springer, 1983) e o artigo de Abraham Seidenberg "The Origin of Mathematics", *Archive for History of Exact Sciences* 18 (1978) 301-342.
[12]Lições sobre História da Matemática, disponível *online* em http://www.archive.org (pesquisando "Moritz Cantor", por exemplo).
[13]No livro *Pitágoras Africano: um Estudo em Cultura e Educação Matemática*, Instituto Superior Pedagógico, Maputo/Beira, Moçambique, 1992, Gerdes tinha já formulado hipóteses alternativas à de Cantor em *Cultura e o Despertar do Pensamento Geométrico*, Instituto Superior Pedagógico, 1991, que é uma versão condensada da sua tese de doutoramento (de 1985).

Caderno_2
quarta-feira, 6 de Maio de 2009 12:54:33

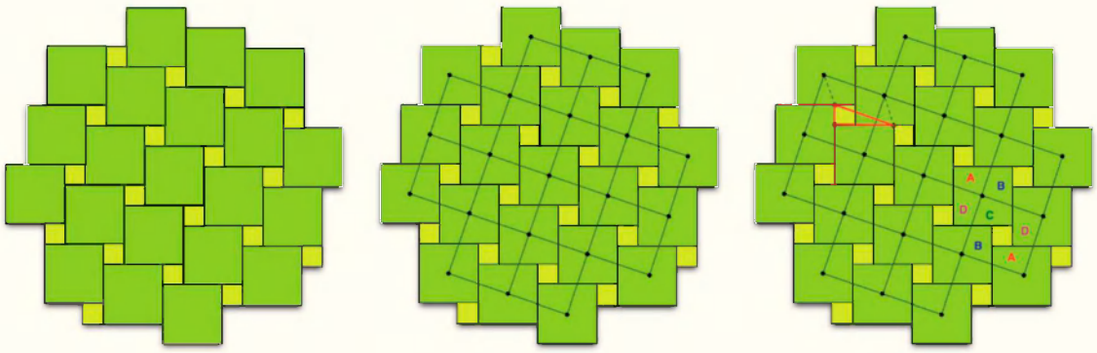


Figura 1: Pavimentação contendo uma demonstração do teorema de Pitágoras

facto, repetido até à exaustão noutras obras e inúmeros manuais como algo de certo e sabido.

Mais recentemente, Paulus Gerdes, da Universidade Pedagógica de Maputo, Moçambique, formulou[13] uma hipótese alternativa, a meu ver muito mais plausível e que deixa menos por explicar. Depois de observar que as espirais constituem um dos elementos mais importantes da decoração egípcia, Gerdes conjectura que, para poderem desenhar certos padrões relativamente complexos de espirais, os artesãos usariam grelhas compostas por dois quadrados de tamanhos distintos para auxiliar na tarefa. Essas grelhas formam pavimentações, como ilustrado na imagem da esquerda da figura 1. Unindo os centros das cópias dos quadrados maiores, obtém-se uma figura (imagem central da fig. 1) que contém simultaneamente o enunciado do teorema de Pitágoras e a sua demonstração! De facto, como sugerido na imagem da direita da figura 1, os quadrados originais são os catetos de um triângulo rectângulo cuja hipotenusa é precisamente o lado dos quadrados obtidos unindo os referidos centros (ver canto superior esquerdo).

Por outro lado, cada um destes quadrados contém um dos quadrados da pavimentação inicial, um dos menores, e quatro fragmentos de um dos maiores (ver canto inferior direito da fig. 1). Como observa Gerdes, os artesões egípcios terão feito construções análogas durante milénios, sendo muito provável que durante todo esse tempo alguns tenham observado que elas continham algo de interessante: um quadrado contendo dois quadrados de uma forma curiosa...

Ao longo do seu livro *O Pitágoras Africano*, assim como em *Geometry from Africa: Mathematical and Educational Explorations*[14], Gerdes dá muitos outros exemplos de situações em que o trabalho de um artífice conduz, mais ou menos naturalmente, ao enunciado e a uma demonstração do teorema que certamente não é de Pitágoras!

Ainda a propósito de pavimentações, Roger B. Nelsen, num curto mas delicioso artigo intitulado "Paintings, plane tilings, and proofs"[15], mostra como usar algumas pavimentações simples para deduzir vários resultados geométricos, obtendo demonstrações destes. Espero que as pequenas pérolas aí contidas ajudem a sarar o trauma que o leitor pode ter sofrido em consequência das revelações acima feitas.

Mas por que razão, poderá agora perguntar-se, se perpetuam as lendas e histórias falsas sobre Pitágoras? Não sei exactamente qual a resposta, mas deve ter algo a ver com ser mais fácil transmitir lendas simplistas do que verdades complicadas e com o culto da personalidade, do líder, na base da importância excessiva que os relatos históricos usuais atribuem aos indivíduos que são glorificados, e em que se presta mais atenção ao «quem» do que ao «como». Não deixa de ser irónico que se tenha acabado por atribuir a Pitágoras, que criou uma comunidade em que os bens e os conhecimentos eram propriedade colectiva, toda a produção intelectual dessa comunidade, dos seus descendentes e sabe-se lá de quem mais! M

[14]Publicado pela *The Mathematical Association of America*, 1999.
[15]Publicado na revista *Math Horizons* (November 2003), pp. 5 - 8, disponível em http://www.lclark.edu/~mathsci/nelsen.html.

Caderno_2
quarta-feira, 6 de Maio de 2009 12:54:33

por **Pedro Quaresma** [Departamento de Matemática da Universidade de Coimbra]
Augusto Pinho [Núcleo de Estágio Pedagógico Lic. Matemática, F.C.T.U.C.]

Criptoanálise

Há mais de 2000 anos que várias organizações tentam manter secretas as suas comunicações. Lado a lado com estes esforços estão os esforços dos "quebradores de códigos", que tentam achar padrões em mensagens aparentemente sem sentido quebrando o código. Isto é a criptologia: por um lado a criptografia, por outro a criptoanálise.

No presente artigo vamos abordar algumas das técnicas disponíveis ao criptoanalista, nomeadamente o estudo da frequência relativa das letras ou grupos de letras numa dada língua natural e os algoritmos de factorização de números primos.

1. Introdução.

A capacidade de comunicar está na base da definição de qualquer sociedade humana, a necessidade de esconder a informação que um dado grupo troca entre si dos outros grupos vem logo a seguir, concomitante com esta vem a necessidade de descobrir os segredos que os outros pretendem guardar. Os sistemas criptográficos surgiram então como forma de garantir a confidencialidade da informação, a criptoanálise surgiu como forma de "quebrar" essa mesma confidencialidade.

Um sistema criptográfico é um conjunto de técnicas que permitem tornar incompreensível uma dada mensagem, de modo que só o verdadeiro destinatário da mesma a consiga decifrar, obtendo dessa forma o texto original [4, 6, 7].

A criptoanálise, pelo seu lado, desenvolve as técnicas capazes de "quebrar" as diferentes cifras de forma a conseguir recuperar o texto original, mesmo que parcialmente, a partir do texto cifrado.

É importante notar que a criptoanálise não é o mesmo que a decifração de uma mensagem cifrada. Neste último caso está-se perante um processo "normal" de, dado um texto cifrado, obter o texto original, isto é, o decifrador pertence ao mesmo grupo que o cifrador e tem a informação necessária para, por um processo simétrico ao da cifração, obter a mensagem original a partir da mensagem cifrada. No caso da criptoanálise está-se perante um elemento de um grupo contrário que ou não tem nenhum conhecimento sobre o processo de cifração ou tem um conhecimento parcial do mesmo.

No presente artigo vamos abordar algumas das técnicas disponíveis ao criptoanalista, nomeadamente o estudo da frequência relativa das letras ou grupos de letras numa dada língua natural e os algoritmos de factorização de números primos. A cada método de criptografia estão associados um ou mais métodos de criptoanálise. No caso presente trata-se dos sistemas de substituição monoalfabéticos e o método RSA [1, 4, 6, 7].

100
95
75
25
5
0

2. Criptoanálise

Como já foi dito, o criptoanalista tenta obter o texto original a partir do texto cifrado, não tendo para tal o conhecimento do processo exacto usado para a cifração, em particular não tendo conhecimento da chave secreta necessária para a decifração.

Quais são então os objectos e as técnicas do criptoanalista? Para já o(s) texto(s) cifrado(s), eventualmente o conhecimento do tipo de cifração usada, e também o conhecimento, mesmo que parcelar, do(s) texto(s) original(is). As histórias de espionagem estão cheias de casos em que uma dada informação é "entregue" ao inimigo só para provocar uma mensagem cifrada cujo texto original se poderia adivinhar, ou então de golpes de sorte que resultam na obtenção de informações importantes sobre o processo de cifração usada pelo inimigo.

Estes são alguns dos objectos que o criptoanalista tem ao seu dispor; vejamos agora algumas das técnicas de que ele se pode socorrer. No que se segue vamos assumir que temos o conhecimento do tipo de cifração usado e de um, ou mais, textos cifrados.

2.1 Sistemas de substituição monoalfabéticos

Os sistemas de substituição monoalfabéticos são caracterizados por serem sistemas em que cada letra do texto original é substituída por uma outra letra no texto cifrado, sendo que ambos os textos usam o mesmo alfabeto de base. Um exemplo de um sistema desse tipo é a designada *Cifra de Júlio César*. Neste sistema, cada letra é substituída por uma outra três posições mais à direita no alfabeto romano. Este é um sistema de substituição monoalfabético aditivo [6, 7] muito simples de quebrar. A designação "aditivo" advém do modo como é feita a substituição, neste caso preciso, pela adição (deslocação para a direita) de três posições à posição original da letra no alfabeto. Considerando que o alfabeto tem 26 caracteres distintos, existem somente 25 cifras distintas. Como tal um "ataque directo" por tentativa e erro é possível, permitindo "quebrar" a cifra rapidamente.

Nem todos os sistemas de substituição monoalfabéticos são assim tão fáceis de quebrar. No entanto, a sua própria essência vai permitir um outro tipo de "ataque", sendo este aplicável a todos os sistemas deste tipo que possamos inventar.

Como foi dito atrás, cada letra do alfabeto é substituída por uma outra. Quer isto dizer que, num dado texto, todos os "a" são substituídos por, por exemplo, "t", isto é, pode dizer-se que todos os "a" continuam no texto cifrado, só que mascarados de "t". Será que isso é importante? Será que desse facto podemos extrair alguma informação que nos permita decifrar o texto cifrado?

A resposta a estas duas questões é afirmativa em ambos os casos. Para uma dada língua natural pode verificar-se que cada letra tem uma frequência própria, por exemplo para o Português temos (não se têm em conta os caracteres acentuados) [6,7]:

A	12,71%	B	0,81%	C	4,16%	D	5,52%	E	11,99%
F	1,34%	G	1,32%	H	0,74%	I	7,18%	J	0,21%
K	0,00%	L	3,23%	M	4,48%	N	5,24%	O	11,32%
P	3,07%	Q	1,41%	R	6,47%	S	7,99%	T	5,31%
U	3,44%	V	1,36%	W	0,02%	X	0,28%	Y	0,02%
Z	0,37%								

É possível, dados estes valores, agrupar as letras em grupos bem definidos, das mais frequentes para as menos frequentes:

1.° A, E, O
2.° S, R, I
3.° N, D, M, U, T, C
4.° L, P, V, G, H, Q, B, F
5.° Z, J, X, K, W, Y

Caderno_2
quarta-feira, 6 de Maio de 2009 12:54:50

É ainda possível obter as frequências relativas de alguns *digramas* (pares de letras) e *trigramas* (ternos), assim como as letras de início de palavras, as de fim de palavras e as palavras mais frequentes.

Toda esta informação permite "quebrar" uma cifra deste tipo tendo somente o conhecimento do texto cifrado. Obviamente, quanto maior o texto mais fácil se torna a análise. Por outro lado, este estudo não elimina por completo um processo de tentativa e erro; o que faz é reduzir o número de tentativas a efectuar antes de se ser bem sucedido.

Tentemos então "quebrar" a cifra usada para criar o texto D FKDYH WHP GH VHU PDQWLGD VHFUHWD, sabendo de antemão que foi usado um sistema de cifração monoalfabético aditivo.

Fazendo o cálculo das frequências relativas para o texto cifrado, temos:

D	17,9%	F	7,1%	G	7,1%	H	21,4%
K	3,6%	L	3,6%	P	7,1%	Q	3,6%
U	7,1%	V	7,1%	W	10,7%	Y	3,6%

Temos então que o "D", com 17,9%, e o "H", com 21,4%, fazem com certeza parte do grupo de letras de alta frequência ("A", "E" e "O"). Dado que estamos a analisar uma frase pequena, o estudo de frequência só nos fornece uma aproximação, um ponto de partida, para começarmos o processo de tentativa e erro.

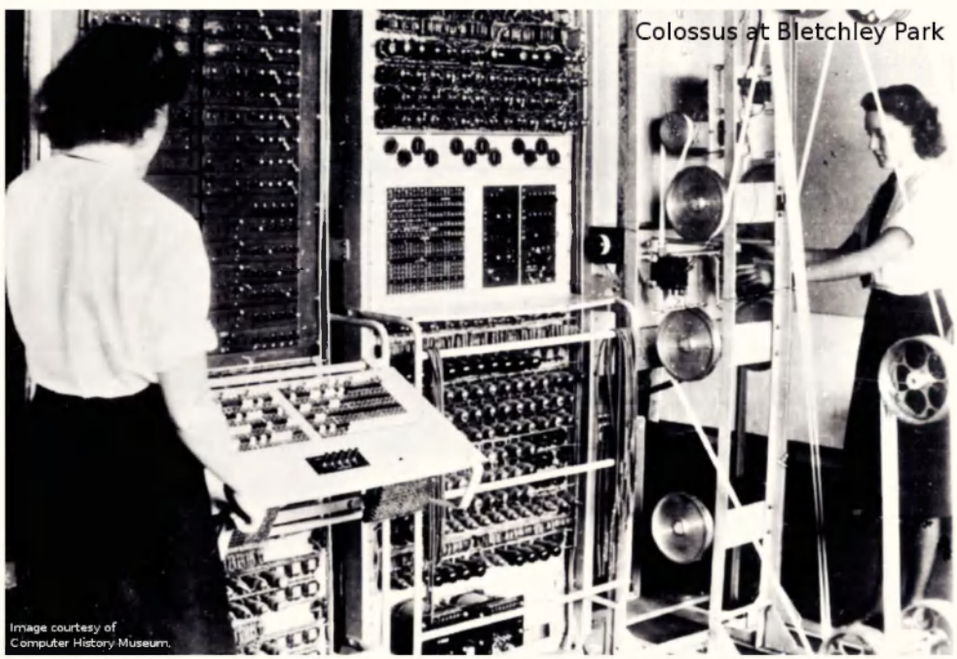
Voltando ao nosso exemplo, temos "D" e "H" como candidatos para "A", tentemos primeiro a chave 3 ("A" transformado em "D"):

A	B	C	D	E	F	G	H	I	J	K	L	M
D	E	F	G	H	I	J	K	L	M	N	O	P
N	O	P	Q	R	S	T	U	V	W	X	Y	Z
Q	R	S	T	U	V	W	X	Y	Z	A	B	C

obtém-se então: A CHAVE TEM DE SER MANTIDA SECRETA.

Ou seja, conseguimos quebrar a cifra!

Com o advento dos computadores, em que todo este tratamento pode ser executado a velocidades estonteantes, facilmente se conclui que este tipo de método deixou de ser seguro.



Caderno_2
quarta-feira, 6 de Maio de 2009 12:54:53

Saltando um pouco na evolução dos sistemas criptográficos, vamos agora analisar os sistemas que baseiam o processo de cifração em problemas computacionais difíceis, isto é, problemas que só são resolvidos computacionalmente à custa de enormes custos em tempo gasto e/ou espaço. Dito doutra forma, problemas cuja resolução é computacionalmente muito difícil, seja em termos de tempo computacional despendido, seja em termos do volume de memória computacional necessária para a sua resolução, ou ambos.

2.2 Problemas computacionais difíceis

A ideia por trás de alguns dos sistemas de criptografia da actualidade é explorar problemas que se crê serem computacionalmente difíceis. Um problema é dito computacionalmente difícil quando, por um lado, a sua solução computacional exige muitos recursos computacionais, seja no tempo necessário para o resolver, seja na quantidade de memória (RAM) exigida para a sua resolução (ou ambos). Por outro lado, num problema deste tipo, a um pequeno aumento na dimensão do problema vai corresponder um grande aumento nos recursos computacionais necessários para a sua resolução. Temos então que, se se basear um sistema criptográfico num problema deste tipo, de tal forma que a quebra da cifra implique a resolução do problema, a quebra da cifra será sempre computacionalmente muito difícil; a um pequeno aumento da dimensão do problema (por exemplo, a escolha de uma chave de cifra melhor) vai corresponder um grande aumento da dificuldade da quebra da cifra.

Entre outros problemas deste tipo temos [6]:

- RSA - dado um $n \in N$, $n=pq$, com p e q primos, factorizar n nos seus factores primos.
- El Gamal - problema do logaritmo discreto. Dado um grupo cíclico finito G, um gerador $g \in G$ e um elemento $x \in G$, determinar o único $i \in \mathbb{Z}_G$ tal que $x = g^i$;
- Knapsack - dada uma mochila com um dado volume, e dados n objectos com volumes diferentes, descobrir um subconjunto de elementos que encha de forma exacta a mochila.

Estes métodos têm em comum o facto de ser possível construir um sistema criptográfico em que a cifração é feita de forma eficiente, mas em que a "quebra" da cifra se crê ser computacionalmente muito difícil.

De seguida vamos apresentar algumas das técnicas de criptoanálise disponíveis para o sistema RSA. Em [6] o leitor poderá encontrar todos os detalhes referentes a este e aos restantes sistemas acima referidos.

2.3 O sistema RSA

O sistema RSA é assimétrico, isto é, possui uma chave pública (e,n) e uma chave privada (d,n), em que [4, 6]:

- $n=pq$, com p e q números primos;
- $1<e<\varphi(n)$, em que $\varphi(n)$, a função de Euler, é igual ao número de co-primos de n;
- $d\,e \equiv 1 \pmod{\varphi(n)}$.

Sabendo a chave pública, isto é, (e,n), basta factorizar n no produto de números primos p e q, para depois facilmente se obter d, ou seja, a chave secreta (d,n). O problema é que a factorização de um número nos seus factores primos é um problema computacionalmente difícil [3, 5, 6].

Vejamos de seguida alguns dos métodos de criptoanálise para o sistema RSA. Iremos acabar esta secção com um estudo comparativo entre eles.

2.3.1 Método das divisões

Este é o *método da força bruta*, ou seja, vai-se tentar a divisão sucessiva por todos os números primos até $\lfloor \sqrt{n} \rfloor$, ou até que a solução seja encontrada.

Por exemplo, em *Octave*[1] ter-se-ia[2]:

```
## Método da "força bruta"
## -> n, o inteiro a factorizar
## <- resp, o primeiro factor primo
function resp=Divisoes(n)
## cálculo do limite
limite=floor(sqrt(n));
## vector com todo os primos até ao limite
```

[1]Octave, www.octave.org, é um sistema de programação numérica de distribuição gratuita, compatível com o *MatLab*.
[2]Os algoritmos estão disponíveis em http://www.mat.uc.pt/~pedro/cientificos/Cripto

Caderno_2
quarta-feira, 6 de Maio de 2009 12:54:17

```
  c=crivoEratostenes(limite);
  ## número de primos obtidos  nprimos=columns(c);
  i=0;
  ## ciclo de teste, desde o primeiro até ao penúltimo primo
  do
   i=i+1;
   if(mod(n,c(i))==0) # (mod - resto da divisão inteira)
      resp=c(i);
   endif
  until (i>=nprimos || mod(n,c(i))==0)
  ## teste para o último primo
  if (i>=nprimos && mod(n,c(nprimos))==0)
    resp=c(nprimos);
  endif
endfunction
```

Um dos pontos fracos desta aproximação é a necessidade de gerar os números primos até um determinado *m*. Infelizmente, a despeito do esforço das melhores mentes matemáticas ao longo dos séculos, não é actualmente conhecida nenhuma fórmula para a geração de números primos. É então necessário recorrer a um método que construa a lista de todos os números primos até um dado limite, por exemplo o *Crivo de Eratóstenes* [3]. Tentando descrever sumariamente o algoritmo, tem-se: começa-se por gerar um vector com todos os inteiros de 2 até *n*; 2 é primo; aplica-se de seguida o "crivo 2" à lista retirando desse modo o 2 e todos os seus múltiplos; 3, o primeiro dos inteiros que sobraram, é primo, então podemos agora aplicar o "crivo 3" à lista. É fácil de ver que, se repetirmos o processo até ao fim da lista original, os elementos que não forem eliminados pelas sucessivas aplicações dos crivos constituem a lista dos números primos de 2 até *n*.

A necessidade de criar a lista de todos os inteiros de 2 até *n* e as inúmeras vezes (*n*-1 para ser preciso) que é necessário percorrer essa lista para aplicar os sucessivos crivos leva a que a utilização do *Crivo de Eratóstenes* acrescente um peso muito significativo ao algoritmo, tanto temporal como espacialmente.

Uma alternativa é a utilização de uma fórmula que gere todos os números primos sucessivos, além de outros não primos também. A utilização de uma tal fórmula torna o ciclo de tentativas de divisão menos eficiente, no entanto os ganhos, tanto espacial como temporalmente, obtidos pela não utilização do Crivo de Eratóstenes compensam essas perdas. Podemos, por exemplo, utilizar a fórmula $p=6k\pm1$ [5].

Em *Octave* ter-se-ia:

```
## Utilização da fórmula 6k+1 e 6k-1 para determinar números primos.
## -> n
## <- resp, o primeiro factor primo
function resp=Divisoes1(n)
  limite=floor(sqrt(n));
  ## testa para 2 e para 3 (mod - resto da divisão inteira)
  if (mod(n,2)==0) resp=2; endif
  if (mod(n,3)==0) resp=3; endif
  ## de 6*1-1=5, 6*1-1+2=6*1+1 até ao limite
  p=5;
  while (p<=limite && resp==0)
    ## testa 6k-1 e 6k-1+2=6k+1
    if (mod(n,p)==0) resp=p; endif
    if (mod(n,p+2)==0) resp=p+2; endif
    p=p+6;
  endwhile
endfunction
```

Depois de apresentarmos mais alguns métodos alternativos iremos comparar a eficiência temporal de todos eles, tentando dessa forma responder à questão de se saber se é viável, com estes programas, quebrar a cifra RSA.

Caderno_2
quarta-feira, 6 de Maio de 2009 12:51:00

2.3.2 Método de Euclides

Este método ganha o seu nome da utilização do algoritmo de Euclides para o cálculo do máximo divisor comum de dois inteiros [2]. Este algoritmo é muito eficiente e pode ajudar-nos a obter um dos factores primos de n. Basta multiplicar todos os números primos entre 2 e $\lfloor\sqrt{n}\rfloor$, calcular de seguida o $m.d.c$ entre esse produto e n, de forma a obter o factor primo desejado.

Esta aproximação apresenta dois problemas: primeiro, a necessidade de gerar a lista de todos os primos até um dado limite – como já dissemos, essa tarefa é pesada tanto temporal como espacialmente; o outro problema advém da representação computacional do número que se obtém da multiplicação dos números primos – rapidamente o número obtido pela multiplicação ultrapassa a capacidade de representação da maioria das linguagens de programação existentes.

Para obstar a este último problema pode dividir-se a multiplicação em várias multiplicações parcelares.

Para tal procede-se do seguinte modo:

- Começa-se por definir os conjuntos auxiliares:
  $R = \{r_1, \ldots, r_m\}$, representando r_i um limite inferior ($r_1 = 1, r_i < r_{i+1}$);
  $S = \{s_1, \ldots, s_m\}$, representando s_i um limite superior ($s_i < s_{i+1}, s_{m-1} < \lfloor\sqrt{n}\rfloor < s_m$);
- Para cada par r_i e s_i, multiplicam-se todos os números primos entre estes dois limites, $P_i = \prod_{r_i \le p_j \le s_i} p_j$;
- Para cada um dos P_i calcula-se o $mdc(P_i, n) = a_i$;
- Se $a_i \neq 1$, então a_i é o factor primo de n que se pretende obter.

Vejamos um exemplo: seja $n = 1457$, $\lfloor\sqrt{1457}\rfloor = 38$, e façamos as somas parcelares de 10 em 10.

$R = \{1, 11, 21, 31\}$ $S = \{10, 20, 30, 40\}$

$$P_1 = \prod_{1 \le p_j \le 10} p_j = 2 \times 3 \times 5 \times 7 = 210 \qquad mdc(210,1457) = 1$$

$$P_2 = \prod_{11 \le p_j \le 20} p_j = 11 \times 13 \times 17 \times 19 = 46189 \qquad mdc(46189,1457) = 1$$

$$P_3 = \prod_{21 < p_j < 30} p_j = 23 \times 29 = 667 \qquad mdc(667,1457) = 1$$

$$P_4 = \prod_{31 \le p_j \le 40} p_j = 31 \times 37 = 1147 \qquad mdc(1147,1457) = 31$$

Temos então 31 como um dos factores primos de 1457.

Em *Octave*:

```
## Método de Euclides
## -> n
## <- resp, o primeiro factor primo
function resp=Euclides(n)
  ## cálculo do limite
  limite=floor(sqrt(n));
  ## vector com todo os primos até ao limite
  c=crivoEratostenes(limite);
  ## número de primos obtidos
  nprimos=columns(c);
  ## inicialização das variáveis que controlam a construção dos
  ## conjuntos auxiliares (de 10 em 10)
  passo=10; k=1; i=1; si=passo;
  ## inicializa o vector dos produtos só com 1s (elemento neutro
  ## da multiplicação)
  prods=ones(1,limite);
  ## cálculo dos productórios (guardados no vector "prods")
```


Caderno_2
quarta-feira, 6 de Maio de 2009 12:5[illegible]2

```
  while (i<=nprimos)
    if (c(i)>si)
      si=si+passo;
      k=k+1;
    endif
    prods(k)=prods(k)*c(i);
    i=i+1;
  endwhile  m=1;i=1;
  ## cálculo do mdc entre os productórios e n
  while (i<=k && m==1)
     m=mdc(prods(i),n);
    if (m!=1)
      resp=m;
    endif
    i=i+1;
  endwhile
endfunction
```

Novamente põe-se a questão de gerar os números primos até um determinado limite. Além desta limitação os produtos de números primos vão rapidamente levantar problemas de representação numérica, isto é, mesmo que a representação dos números primos por si só não levante problemas a sua multiplicação, mesmo que em grupos reduzidos, vai rapidamente levantar problemas de representação na gama finita disponível nos tipos de dados das linguagens de programação mais usuais.

2.3.3 Método de Fermat

O método de Fermat consiste em encontrar dois inteiros a e b que permitam representar o número natural n como a diferença de dois quadrados:

$$n = a^2 - b^2 \Leftrightarrow n = (a - b)(a + b)$$

Teorema 1 *Qualquer inteiro n ímpar maior que 1 pode ser escrito como a diferença de dois quadrados.*

Demonstração:

Seja $n=pq$, com $p>q$ (no caso de n ser primo considera-se $n=n\times1$)

Por hipótese n é ímpar; então p e q também o são, logo: $\frac{p+q}{2}$ e $\frac{p-q}{2}$ são inteiros, mas então temos:

$$\begin{aligned}\left(\frac{p+q}{2}\right)^2 - \left(\frac{p-q}{2}\right)^2 &= \frac{p^2+2pq+q^2}{4} - \frac{p^2-2pq+q^2}{4} \\ &= \frac{p^2+2pq+q^2 - p^2+2pq - q^2}{4} \\ &= \frac{4pq}{4} \\ &= pq \\ &= n\end{aligned}$$

Para encontrar a e b tais que $n=a^2-b^2$ procede-se do seguinte modo:

- Dado um inteiro n ímpar começamos por tomar

$$a = \lfloor \sqrt{n} \rfloor + 1;$$

- Se $b = \sqrt{a^2 - n}$ é um inteiro, obtém-se o pretendido;
- Caso contrário, incrementamos a de uma unidade até que b seja um inteiro.

Por exemplo: seja n=2027651281; $a = \lfloor \sqrt{n} \rfloor + 1 = 45030$ é o primeiro valor a testar.

Caderno_2
quarta-feira, 6 de Maio de 2009 12:51:04

	a	b		a	b
1.°	45030	$\sqrt{45030^2 - 2027651281} = 222{,}75$	7.°	45036	$\sqrt{45036^2 - 2027651281} = 768{,}12$
2.°	45031	$\sqrt{45031^2 - 2027651281} = 373{,}73$	8.°	45037	$\sqrt{45037^2 - 2027651281} = 824{,}67$
3.°	45032	$\sqrt{45032^2 - 2027651281} = 479{,}31$	9.°	45038	$\sqrt{45038^2 - 2027651281} = 877{,}58$
4.°	45033	$\sqrt{45033^2 - 2027651281} = 565{,}51$	10.°	45049	$\sqrt{45039^2 - 2027651281} = 927{,}49$
5.°	45034	$\sqrt{45034^2 - 2027651281} = 640{,}21$	11.°	45040	$\sqrt{45040^2 - 2027651281} = 974{,}84$
6.°	45035	$\sqrt{45035^2 - 2027651281} = 707{,}06$	12.°	45041	$\sqrt{45041^2 - 2027651281} = 1020$

Concluindo:

$$n = 45041^2 - 1020^2\,,\ n = 46061 \times 44021.$$

No caso do algoritmo de Fermat prova-se que, quanto maior for a diferença entre p e q, maior é o número de tentativas necessárias para obter um primeiro valor inteiro para a raiz [5]. Ou seja, o método de Fermat leva-nos a escolher, para a nossa cifra, p e q primos distantes entre si.

Em *Octave*:

```
## Método de Fermat
## -> n
## <- resp=[a,b]
function resp=Fermat(n)
  ## começa-se por tomar a = (maior inteiro contido em raiz
  ## quadrada de n) + 1
  a=floor(sqrt(n))+1;
  ## se a é um quadrado então a resposta é [a,a]
  if (issqr(m))
    resp=[a,a];
  else # senão
    do
      b=a^2-n;
      ## se b=raiz quadrada(a^2-n) é um inteiro então a resposta
      ## é [a,b]
      if (issqr(b))
        resp=[a,sqrt(b)];
      endif
      ## senão incrementa-se a de uma unidade e recomeça-se
      a=a+1;
    until (issqr(b)) # até que b seja um inteiro
  endif
endfunction
```

2.3.4 Estudo comparativo dos vários métodos

Será que já temos as ferramentas necessárias para quebrar a cifra RSA? Teoricamente a resposta é positiva, existem métodos construtivos de factorização de números primos que podem ser usados para obter o que se pretende. A questão é então de índole prática, isto é, será que já temos as ferramentas computacionais necessárias para quebrar, em tempo útil, a cifra RSA?

Estando o estudo teórico da complexidade dos algoritmos apresentados fora do âmbito do presente texto, vamos somente apresentar os resultados de um estudo comparativo dos vários métodos baseado num conjunto de testes de execução.

Todos os valores respeitam a testes efectuados sob as mesmas condições computacionais: sistema GNU/Linux 2.6.8; Intel Pentium 4 a 3,0GHz; 2GiB RAM; Octave 2.1.69. Os tempos referem-se ao tempo gasto pelo processador tal como nos é dado pelo sistema operativo.

Caderno_2
quarta-feira, 6 de Maio de 2009 12:5! :: 7

n	Factores		Divisões	Divisões1	Euclides	Fermat
1457	47	31	0,009s	0,002s	0,015s	0,001s
13199	197	67	0,023s	0,005s	0,030s	0,002s
281161	3559	79	0,077s	0,006s	0,092s	0,218s
701123	3559	197	0,129s	0,016s	0,169s	0,179s
23420707	41017	571	0,699s	0,047s	0,839s	2,728s
488754769	110503	4423	3,343s	0,361s	4,477s	6,093s
2027651281	46061	41017	8,705s	3,611s	19,575s	0,004s
103955963689	47188363	2203	49,050s	0,179s	51,891s	3926,6s
128228613281	58206361	2203	55,016s	0,180s	58,180s	17175,0s
210528952589	95564663	2203	72,176s	0,182s	75,888s	7953,8s
2746662891777043	47188363	58206361	—	3861,6s	—	50,333s
4509540007616669	47188363	95564663	—	3857,9s	—	712,73s

Olhando para o quadro apresentado podemos concluir que, dada uma escolha apropriada dos factores primos, a cifra RSA está segura. Vejamos mais em pormenor:

- os métodos *Divisões* e *Euclides* não só vêem os seus tempos de execução subir de forma significativa com o crescimento da grandeza relativa dos factores primos, como deixam de ser capazes de resolver o problema a partir de valores relativamente baixos de $n=p\times q$. A necessidade de gerar os números primos até $\lfloor\sqrt{n}\rfloor$ e, no caso do método de *Euclides*, a multiplicação de números de grande dimensão, estão na origem destas limitações.
- o tempo gasto pelo método *Divisões1* está directamente relacionado com o valor do primeiro factor primo; para valores elevados deste é um método pouco eficiente.
- o método de *Fermat* vê os seus tempos de execução crescerem de modo muito significativo com o crescimento da dimensão dos factores primos. É importante notar no entanto que há excepções a esse comportamento genérico, olhando com atenção verifica-se que, quando os factores primos estão próximos entre si, o método de *Fermat* é muito eficiente.

Concluímos então que para assegurar a segurança da cifra RSA os factores primos têm de ser distantes um do outro e o n tem de ter uma dimensão acima dos 20 dígitos.

Na verdade a dimensão de n tem de ser bastante mais elevada. Devido a algoritmos mais eficientes do que os aqui apresentados, a cifra RSA utilizando valores de n com 129, 155, e mesmo 576 dígitos já foi quebrada. A cifra RSA actualmente usa valores de n com 1024 ou mais dígitos [4, 6].

3. A quebra da cifra RSA

Num anterior artigo da *Gazeta de Matemática* colocámos ao leitor um pequeno desafio [4]. Será que já estamos em condições de o resolver?

O texto cifrado é:

```
359394 185904 0 231105 382481 474195 382481 10935 75745 382481 185904 0 201637
382481 302441 522545 270765 382481 185904 0 185904 382481 265174 79985 0 365807
292080 66056 261188 75745 382481 371293 60839 185904 185904 265174 185904 0
90175 75745 75745 382481 185904 270765 522545 10935 66056 474195,
```

sabe-se que foi usado o método RSA, com uma cifração letra a letra (caracteres ASCII entre ' ' e ¯ que correspondem a, ' '=0, '!'=1,…), e que a chave pública usada foi (5,561971).

Qual será então o texto original?

Temos agora as ferramentas necessárias para fazer a criptoanálise deste texto. Na verdade, a dimensão de $n=p\times q$=561971 indica-nos que qualquer um dos métodos aqui apresentados será capaz de obter os factores primos. Assim é:

```
octave:1> Fermat(561971)
tempoCPU = 0.001s
```


Caderno_2
quarta-feira, 6 de Maio de 2009 12:55:10

```
ans = 750   23

octave:2> Divisoes(561971)
tempoCPU = 0.060s
ans = 727

octave:3> Euclides(561971)
tempoCPU = 0.217s
ans = 727
```

Os dois factores primos são então 727 e 773. Dada a sua proximidade, o algoritmo de Fermat foi o mais eficaz, não necessitando mais do que 0,001s para obter a factorização.

Utilizando os algoritmos do método RSA temos como chave pública (5,561971) e como chave privada (224189,561971). Posto isto podemos obter o texto original por aplicação directa do método RSA.

As Palavras Magicas sao: Quebra-Ossos Irascivel

Esta frase é a tradução para o Português (descontando a falta de acentos) da frase "The magic words are squeamish ossifrage", a qual constituía o desafio original dos autores do sistema RSA. A cifra original, de chave pública (9007,*n*), em que *n* é um dado número inteiro com 129 dígitos, foi quebrada utilizando o método do *Crivo Quadrático*, método esse que foi inventado tendo já como objectivo a criptoanálise do sistema RSA [1, 6].

4. Conclusões

A história da criptografia e da criptoanálise é então uma história de passos sucessivos, uma autêntica guerra travada no interior de gabinetes. Desde os métodos simples (mecânicos) da época pré-computadores aos métodos mais complexos que exploram a capacidade dos computadores para "baralhar" a informação, até aos métodos que exploram os limites teóricos dos computadores, com a utilização dos problemas computacionais difíceis, estamos perante um processo constante em que a um dado avanço na criptografia se segue um avanço na criptoanálise e vice-versa. Talvez um próximo passo seja dado nas fronteiras da matéria explorando a criptografia quântica, com a certeza de que a criptoanálise quântica se seguirá de imediato a esse passo, num processo infindável, até ao momento que deixe de haver necessidade de comunicar de forma secreta [6]. M

Referências

[1] **D. Atkins, M. Graff, A. Lenstra e P. Leyland** (1994). "The magic words are squeamish ossifrage". *ASIACRYPT* 1994, pp. 263–277.

[2] **Donald E. Knuth** (1973). *The Art of Computer Programming. Fundamental Algorithms*. Vol. 1, 2.ª edição, Addisson-Wesley Publishing Company, Reading, EUA.

[3] **Donald E. Knuth** (1981). *The Art of Computer Programming. Seminumerical algorithms*. Vol. 2, 2.ª edição, Addisson-Wesley Publishing Company, Reading, EUA.

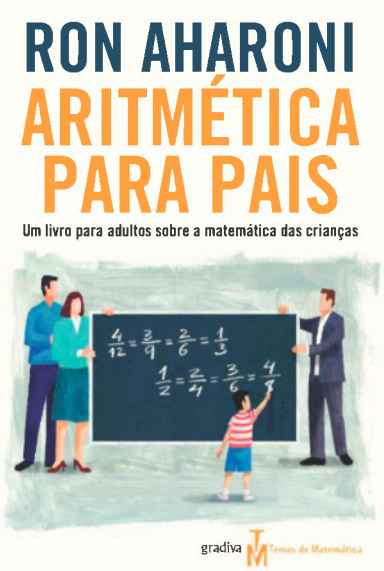
[4] **Pedro Quaresma e Elsa Lopes** (2008). "Criptografia". *Gazeta de Matemática*, n.º 154, pp. 7-11.

[5] **Hans Riesel** (1994). *Prime Numbers and Computer Methods for Factorization*. Vol. 126, 2.ª edição. Progress in Mathematics, Birkhäuser.

[6] **Richard Spillman** (2005). *Classical and Contemporary Cryptology*, Prentice Hall.

[7] **Viktoria Tkotz** (2005). *CRIPTOGRAFIA - Segredos Embalados para Viagem*. NOVATEC Editora, São Paulo.

Caderno_2
quarta-feira, 6 de Maio de 2009 12:5!:13



matemática elementar, diz, "não é complexa, mas é profunda". Os conceitos associados às operações elementares estão longe de ser ligeiros. Como costuma dizer um outro especialista nestas matérias, "Soma, subtracção, multiplicação e divisão têm muito assunto!" O recurso a estratégias diversificadas (esquemas, estórias, desenhos, etc.) é enfaticamente defendido por Aharoni. As operações têm diversos significados e estes devem ser explorados e ensinados.

Se bem que a matemática elementar seja muito respeitável, é legítimo supor que a formação apropriada de professores seja capaz de colocar no terreno profissionais competentes e sabedores. Acontece que, tanto em Israel como em Portugal, essa está longe de ser a realidade. Tanto lá como cá, o que se ensina aos futuros professores não responde às suas necessidades. Porquê?

Na opinião de Ron Aharoni, os professores universitários ensinam o que sabem e não o que deviam ensinar. Por comodismo ou deficiência na percepção das necessidades dos futuros docentes, a universidade elabora currículos que impressionam pela qualidade científica, mas são muitas vezes virtuais e apoucam quem os deve vencer para obter a licença de magistério.

A prática instituída no ensino superior fomenta a bulimia académica que todos conhecemos, em que os alunos deglutem apressadamente conteúdos que regurgitam prontamente após a passagem nas disciplinas. As matérias são muitas vezes irrelevantes para o seu desempenho futuro, pelo que para obter o "canudo" são congeminadas técnicas espúrias, tácita ou explicitamente aceites pelas instituições. Para usar a terminologia gilista: não há inscrição.

Ouve-se muito a expressão "não baixar o nível" entre os professores universitários. A intenção é louvável, mas o que usualmente acontece é que se afere esse "nível" nos próprios docentes e não na aprendizagem dos alunos. Os cursos ficam bem no papel, mas estão desgraçadamente divorciados da realidade. Espanta até que não seja motivo de reflexão o facto, consensual, de os alunos chegarem ao ensino superior mal preparados, e, no entanto, ter sido esse mesmo ensino superior que formou muitos dos que os prepararam…

Claro que há outros problemas no ensino, como os baixos salários dos professores, que não ajudam a atrair e fixar os mais competentes.

Elon Lages Lima, numa obra da mesma chancela, escrevia: "O bom professor é aquele que vibra com a matéria que ensina, conhece muito bem o assunto e tem o desejo autêntico de transmitir esse conhecimento." Ora nem sempre o professor conhece bem a matéria que ministra, mesmo quando esta tem o nome de Matemática Elementar. É que, como Aharoni insiste, há muito a aprender nesta área. Embora não se trate de matemática sofisticada, não a devemos subalternizar, porque, para além de ser fundacional, tem muita substância. Como pode então o professor vibrar com a matéria que ensina? Não pode, muitas vezes porque não a domina e também porque no seu percurso académico encontrou mais razões para se sentir deslocado do que entusiasmado. Resta a vontade de partilhar o conhecimento, que está muitas vezes presente e permite, por voluntarismo, um desempenho mais digno do que aquele para que o professor foi preparado.

Termino com um apelo partilhado com Ron Aharoni: ensinem-se aos futuros professores, e bem, as matérias que eles vão ensinar!

Caderno_3
quarta-feira, 6 de Maio de 2009 1:5::13

Inquérito

por Graciano de Oliveira
[Universidade de Coimbra]

O Magalhães

Nos últimos tempos, tem-se falado muito no novo computador escolar, o Magalhães. O tema é controverso, e os seus defensores e detractores defendem com afinco as suas posições. Mas qual o impacto real que está a ter na vida de estudantes, professores e pais?

A vida dos professores tem sido muito agitada e a causa (avaliação e respectivo modelo) dessa agitação já foi objecto do *Inquérito*. Assim, a pergunta de hoje versa um tema que já foi quente e neste momento esfriou um pouco, o que permite que o abordemos com uma atitude mais tranquila. Trata-se do computador Magalhães e da sua importância na escola. A importância e o peso que os recursos tecnológicos devem ter no ensino e no estudo são questões muito controversas. No entanto, "…nenhum destes recursos dispensa o professor", afirma Helena Damião (http://dererummundi.blogspot.com). Há quem pense que o estudo é sobretudo meditação, seja o estudo da matemática seja o estudo da poesia, não exigindo por isso grandes instrumentos. E há quem, pelo contrário, pense que um computador ou um quadro interactivo são decisivos. Obviamente isso tem a ver com a definição de escola. O que é e que objectivos tem? As crianças frequentam-na para adquirir conhecimentos ou para adquirir competências? E adquirem conhecimentos ou competências para quê? Para ganharem dinheiro ou para outras coisas? Para serem empreendedores ou empregados muito certinhos? Ou para nada? Deverá a escola ser um depósito de crianças (onde estão em segurança enquanto os pais trabalham)? Que devem fazer enquanto lá permanecem? Devem distrair-se e

desfrutar... de quê? Devem aprender a brincar ou aprender brincando? Ou devem aprender a aprender para mais tarde (fora da escola?) aprenderem? São perguntas que poderiam gerar, e têm gerado, estudos profundos e longas discussões filosóficas. Mas não é esse, nem podia ser, o objectivo do *Inquérito*. Por isto solicitámos, como é hábito, respostas muito curtas (na esperança de que gerem meditações longas) às nossas perguntas de hoje:

O que pensa sobre a utilidade do Magalhães para a escola?

Já o utilizou?

Caderno_3
quarta-feira, 6 de Maio de 2009 1:55:17

Diana Rodelo, Faro, mãe de um aluno do ensino básico.

O meu filho frequenta o terceiro ano do 1.º ciclo e tem um Magalhães. A escola dele foi uma das primeiras a entregar um Magalhães a cada aluno que o tivesse solicitado, logo no início do primeiro período. Isto até prometia…

Durante o primeiro período, o meu filho utilizou o Magalhães poucas vezes. Em casa ele recorre preferencialmente aos livros, quando os tem, para fazer os trabalhos que requerem algum tipo de pesquisa extra. Com 8 anos é complicado filtrar toda a informação que lhe está acessível sem a ajuda dos pais. Na escola, o Magalhães era utilizado uma vez por semana nas horas de apoio ao estudo. Ultimamente a professora foi dispensada de dar este apoio por estar a fazer uma formação, pelo que o Magalhães não está a ser usado.

Na minha opinião, a utilização de um computador, concebido especificamente para crianças, como ferramenta de apoio ao ensino pode ser bastante benéfica. É claro que o computador não substitui o papel e o lápis. Para aprender é preciso ler e escrever muito. É preciso sublinhar informações importantes, anotar observações, efectuar cálculos auxiliares, desenhar esquemas, etc. É preciso errar, apagar e voltar a ler e a escrever. Em papel. Com lápis. Depois desta batalha com o papel e o lápis, o computador poderá ajudar na melhor apreensão de alguns conceitos. Do pouco que conheço, o Magalhães tem uma enciclopédia, um processador de texto, acesso à Internet e jogos didácticos. Com estas funcionalidades, o computador permitirá aprofundar alguns temas abordados nas aulas, transcrever e organizar informações pesquisadas, escrever trabalhos em computador, desenvolver capacidades de cálculo, etc. Certamente haveria muitas outras aplicações. No entanto, parece-me que os professores não dispõem de muito tempo para adaptar o plano das aulas de modo a explorar bem as potencialidades do Magalhães durante este ano lectivo. Talvez no próximo.

Filipe Sarmento, Coordenador de TIC do Agrupamento Vertical de Escolas de Vila Pouca de Aguiar Sul

No âmbito do Plano Tecnológico surge o Magalhães, um computador portátil com características particulares.

• Equipado com um processador Intel Celeron M 900 MHz, 1Gb de memória RAM, 30Gb de disco rígido, ecrã de 9 polegadas, Wi Fi: 802.11 b/g para ligações sem fios, 1 porta RJ-45 para ligações à rede com fios, *webcam*, colunas de som, microfone incorporado, 2 portas USB, leitor e gravador de cartões de memória SD *(Secure Digital)*, bateria de 3 células, resistente ao choque e à água e com um peso de 1,4 kg, este computador surge especialmente dirigido aos alunos do 1.º ciclo.

• No Magalhães encontramos *software* específico para crianças, que permite uma utilização controlada e dinamizadora de aprendizagens.

• No ambiente de trabalho podemos encontrar o "Magic Desktop", ambiente especialmente concebido para crianças, lúdico e atractivo.

• O seu software permite que os alunos aprendam conteúdos leccionados no 1.º ciclo e os ponham em prática com jogos lúdicos, tendo logo o respectivo *feedback*, e em função do grau de aprendizagem atingido o aluno obtém minutos de navegação na Internet como recompensa.

• Também o *software* de controlo parental instalado no Magalhães permite definir horas de utilização da Internet fora da escola.

• O *software* Mythware permite ao professor interagir com todos os Magalhães existentes na sala de aula, podendo, inclusive, a partir do seu computador, ver o ambiente de trabalho de cada um dos Magalhães e limitar ou impedir a sua utilização. Desta forma a gestão da aula, em termos de utilização do computador, é facilitada.

• Todos os recursos são poucos quando o objectivo é promover aprendizagens nos homens e mulheres de amanhã. Inevitavelmente, o Magalhães é um recurso com enormes potencialidades. O desafio será

Caderno_3
quarta-feira, 6 de Maio de 2009 1:56:21

estarmos à altura da sua utilização, pois, como professor, estou convicto de que todos sairemos a ganhar com a sua utilização.

Maria da Graça Martins, professora de Matemática do ensino secundário (ESEN Viseu)

Tenho de começar por dizer o que penso da «distribuição» do Magalhães aos meninos do 1.º ciclo: um embuste e uma grande bandeira de propaganda.

O facto de cada aluno ter em casa um Magalhães não me parece ser uma medida prioritária, que contribua para melhorar o ensino e a aprendizagem neste ciclo.

Estamos a falar de meninos com idades entre os 5 e os 10 anos, que depois das aulas não precisam de consultar a Internet para a pesquisa ou a elaboração de trabalhos. E, se no 4.º ano isso acontecer, esse trabalho terá de (deverá?) ser feito na escola, na presença e com a ajuda do professor.

Os meninos com esta idade, depois das aulas, devem brincar e fazer jogos colectivos, e não ficar no quarto a «jogar» com o computador. Isto não significa que na escola do 1.º ciclo não deva haver computadores com *software* interessante para, orientados e ajudados por professores bem preparados, quer na área das TIC quer e especialmente na área de Matemática, os alunos trabalharem.

Há programas de geometria para desenhar figuras geométricas planas (é uma boa maneira de entender as propriedades que caracterizam um quadrado ou um losango…); há *software* apropriado para fazer transformações de figuras geométricas ou composições a partir de um padrão dado; há sítios educativos na Internet. Isto pode ser (é?) um trabalho importante e interessante, mas só com o professor e na sala de aula.

Relativamente à aquisição por parte de cada menino de um Magalhães, só poderei afirmar que é verdade que colocará (quando estiverem disponíveis…) o computador e a Internet ao alcance de TODOS, mas estou convencida que tal medida não terá, de per si, grande utilidade para o ensino.

Para finalizar: nunca utilizei o Magalhães.

David Manuel Antunes da Graça, Escola Secundária com 3.º ciclo Padre António Martins de Oliveira, Lagoa

Simone Lopes Azevedo, Escola Secundária Poeta António Aleixo, Portimão

Pensamos que não há nada que não tenha aspectos positivos e negativos. É o que acontece com o Magalhães, ou qualquer outro computador portátil utilizado por crianças.

Assim, parece-nos que as novas tecnologias podem ser benéficas no desenvolvimento das crianças. Poderão ser uma grande ajuda no que diz respeito à ortografia, ao desenvolvimento do raciocínio lógico e à autonomia da criança, e ainda em funções mais específicas para as quais deverão ser usados programas desconhecidos da maior parte da população. Mas todo o manuseamento deveria ser supervisionado por um adulto com objectivos determinados. É neste aspecto que a autonomia poderá ser um inconveniente, pois muitos pais não têm o cuidado de acompanhar os filhos no uso do computador e outros não terão os conhecimentos específicos para esse acompanhamento, o que leva as crianças a usarem o Magalhães para brincar e não para aprender.

Apenas devemos acrescentar que isto é uma opinião, até porque não temos o Magalhães nem o utilizámos, e não conhecemos portanto as suas potencialidades.

Caderno_3
quarta-feira, 6 de Maio de 2009 1:56:24

por **Rui Ralha e M. Elfrida Ralha** [Departamento de Matemática da Universidade do Minho]
Paula Gomes [Escola EB 2,3/S de Vieira de Araújo - Cantelães]

Aventuras Numéricas no Cálculo do *e*

As calculadoras e os computadores vêm oficialmente afirmando-se como instrumentos didácticos 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.

Caderno_3
quarta-feira, 6 de Maio de 2009 1:53:28

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

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

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 obtém-se

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

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

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

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[1] mas para $n=2^{35}$ apenas os três primeiros algarismos estão correctos. Ora, tendo em conta que

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

[1]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*".

Caderno_3
quarta-feira, 6 de Maio de 2009 1:5[illegible]:31

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[2], 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*, $b > 2$ 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[3] 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

$$x = \pm\left(1, b_1 b_2 \ldots b_{52}\right)_2 \times 2^E, \qquad (5)$$

onde $b_i = 0$ ou $b_i = 1$, para cada $i = 1, \ldots, 52$, 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

$$realmax = \left(1,11\ldots1\right)_2 \times 2^{1023}. \qquad (6)$$

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 $i = 1, \ldots, 52$, isto é,

$$1 = +(1,0\ldots0)_2 \times 2^0$$

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

$$1 = +(1,0\ldots01)_2 \times 2^0,$$

isto é, trata-se do número

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

Se x é um número tal que

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

[2] Ver, por exemplo, [2], p. 271, e [3], p. 55.
[3] Para uma excelente introdução veja-se [5].

Caderno_3
quarta-feira, 6 de Maio de 2009 1:55:34

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 $n > 2^{53}$, 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

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

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,

$$u_{2^{35}} = \left(1+\frac{1}{2^{35}}\right)^{2^{35}} = \mathbf{2{,}718281828}419489$$

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	$1{,}72 \times 10^{0}$

Caderno_3
quarta-feira, 6 de Maio de 2009 1:56:36

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

$$\varepsilon_x = \frac{\tilde{x} - x}{x}, \tag{7}$$

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

$$|\varepsilon_x| < 2^{-52}, \tag{8}$$

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

$$|\varepsilon_x| < 2^{-53}. \tag{9}$$

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

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

Usando a fórmula do binómio de Newton, temos

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

logo

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

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

$$\varepsilon_{u_n} \approx n\varepsilon_x \tag{13}$$

o que mostra que, ainda que seja

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

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

Caderno_3
quarta-feira, 6 de Maio de 2009 1:55:38

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, $u_{10^{12}}$ exibe 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 $n > 2^{53}$. Isto acontece porque, para $n > 2^{53}$,

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

De (10) resulta

$$\varepsilon_{\tilde{u}_n} = \left(1-\frac{1}{n+1}\right)^n - 1 \qquad (16)$$

Caderno_3
quarta-feira, 6 de Maio de 2009 1:5::40

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

e o erro relativo ε_{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 $\dot{x} = 1+2^{-52}$, para $n > 2^{52}$; então, tem-se

$$\varepsilon_{\dot{x}} = \frac{2^{-52} - \frac{1}{n}}{1+\frac{1}{n}} \,. \tag{18}$$

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

$$\varepsilon_{u_{\dot{x}}} = (1+\varepsilon_{\dot{x}})^n - 1 = \left(1+\frac{2^{-52} - \frac{1}{n}}{1+\frac{1}{n}}\right)^n - 1 = \frac{\left(1+2^{-52}\right)^n}{\left(1+\frac{1}{n}\right)^n} - 1.$$

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

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

e para $n = 2^{62}$ dá *Inf*[4], 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 ξ_x um ponto entre x e $\tilde{x}$,

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

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

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

daqui resulta

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

ou seja

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

onde ε_x é o erro relativo na variável independente,

[4]Esta é uma das chamadas "excepções" da aritmética IEEE. O sistema usa uma representação especial *Inf* para qualquer número maior do que *realmax* (ver [5], p. 22).

Caderno_3
quarta-feira, 6 de Maio de 2009 1:56:41

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

é o correspondente erro relativo na variável dependente e

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

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

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

tem-se

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

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

$$f(x) = x^n, \tag{27}$$

é, para todo o x,

$$k_f(x) = n \tag{28}$$

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

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

o que mostra que um pequeno erro relativo ε_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

$$g(x) = \exp(n \log(x)), \tag{30}$$

Caderno_3
quarta-feira, 6 de Maio de 2009 1:56:43

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

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

e

$$k_g(x)= k_f(x)= n. \tag{31}$$

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

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

O conhecido desenvolvimento em série de potências de θ

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

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

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

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

$$T_2(\theta)=1-\frac{\theta^2}{2} \tag{34}$$

é

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

Caderno_3
quarta-feira, 6 de Maio de 2009 1:56:44

logo, para $|\theta|$ pequeno, $k_{T_2}(\theta) \approx 1$ e pequenos erros relativos no valor de θ 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 θ exacto ou com pequeno erro relativo. Para ilustrar isto observemos que, no Matlab, o valor calculado de

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

é

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

com erro relativo

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

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

$$\exp\left(T_2\left(\ddot{\theta}\right)\right) = 2{,}718281828459032$$

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[5]

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

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, obtêm-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

[5]Para uma explicação mais detalhada, ver http://www.cut-the-knot.org/arithmetic/interest.shtml.

Caderno_3
quarta-feira, 6 de Maio de 2009 1:55:47

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

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 $k = 1, \ldots, 16$.

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 $n = 10^{15}$. A partir de $k = 14$, em ambas as máquinas, a representação da base $1 + 10^{-15}$ 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

Caderno_3
quarta-feira, 6 de Maio de 2009 1:55:49

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.° ano*. Vol.2, Texto Editora, Lisboa.

[3] **Jorge, A.**, et al. (2004). *Infinito 11.° 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.° ano de escolaridade*, Departamento do Ensino Secundário, Ministério da Educação.

Bibliografia

Consciência, M., (2003). "Calculadoras gráficas: algumas limitações". *Gazeta da Matemática*, n.° 145, pp. 34-42.

Teixeira, P., et al. (1998). *Funções: Matemática – 11.° ano de escolaridade*, Departamento do Ensino Secundário, Ministério da Educação.

Caderno_3
quarta-feira, 6 de Maio de 2009 1:5[illegible]:54

por **Paulo Mateus**
[Instituto Superior Técnico]

O problema P≠NP?

Um milhão de dólares e o nome na eternidade para quem resolver um dos mais importantes problemas matemáticos.

O problema P≠NP é o mais recente dos problemas do milénio listados pelo Clay Mathematics Institute e é, sem sombra de dúvida, o mais importante e influente na área da teoria da computação. A resposta a este problema terá impacto muito significativo em áreas como a criptografia, a optimização, a verificação de modelos, a bioinformática etc., podendo ainda ter consequências importantes na sociedade (por exemplo, o fim do comércio electrónico). A questão foi posta rigorosamente pela primeira vez por Stephen Cook em 1971 [2].

Para entender com algum detalhe o problema P≠NP é essencial compreender o conceito de função computável, bem como o de função computável em tempo polinomial. Neste texto vai-se evitar a abordagem seminal de Turing [3] e das suas máquinas, que à primeira vista é um pouco pesada, considerando-se em alternativa uma arquitectura moderna para o computador[1]. A memória de um computador ξ é uma sucessão de registos $\xi = R_0, R_1, \ldots, R_n, \ldots$ onde em cada registo se pode guardar um natural arbitrariamente grande. O valor guardado no registo R_0 na configuração ξ é denotado por R_0^{ξ}. Dado que o tamanho de um natural vai ser importante, denotamos por $|R_i|$ o número de *bits* necessários para escrever R_i em base 2, sendo este valor igual a $\lfloor \log_2(R_i)+1 \rfloor$. A escolha da base é irrelevante, mas a base 2 imita a implementação do computador actual. Para inicializar a memória do computador assume-se que os primeiros k registos $(R_0 \ldots R_{k-1})$ tomam os valores $x_0 \ldots x_{k-1}$ e os restantes registos estão a zero. Esta configuração de memória é denotada por $\xi(x_0 \ldots x_{k-1})$.

Para usufruir da memória de um computador é necessário poder alterá-la, o que se faz por intermédio de programas. Apesar de ser costume introduzir uma panóplia variada de programas atómicos, neste texto vamos considerar apenas quatro:

- $R_i := n$ atribui a R_i o natural n;
- $R_i := R_j + R_k$ atribui a R_i a soma de R_j com R_k;
- $R_i := |R_j|$ atribui a R_i o número de *bits* de R_j;
- $R_{iR_k} := R_{jR_i}$ modifica o R_k-ésimo *bit* menos significativo de R_i, dando-lhe o valor do R_i-ésimo *bit* menos significativo de R_j.

Os programas atómicos permitem alterar directamente a memória do computador, enquanto as primitivas de composição de programas controlam o programa a ser executado a seguir. Basta considerar três maneiras de compor programas:

- $(M_1; M_2)$ (composição sequencial) indica que se deve executar o programa M_1 e de seguida M_2;
- (`if` $(R_i \leq R_j)$ `then` M) (composição alternativa) indica que se deve executar M quando o valor guardado no registo R_i for menor ou igual ao valor guardado em R_j e não executar nada caso contrário;
- (`while` $(R_i \leq R_j)$ `do` M) (composição iterativa) indica que o computador deve executar M enquanto o valor guardado no registo R_i for menor ou igual ao valor guardado em R_j.

O menor conjunto contendo todos os programas atómicos e fechado para a composição sequencial,

[1]Curiosamente, a arquitectura moderna dos computadores baseia-se num modelo que von Neumann propôs para o cérebro humano.

Caderno_4
quarta-feira, 6 de Maio de 2009 13:29:45

alternativa e iterativa é o conjunto de todos os programas $\mathcal{M}$. Um programa $M \in \mathcal{M}$ induz uma função (parcial) de naturais para naturais. Assuma que a memória inicial do computador é $\xi(x)$. A execução de um programa M a partir desta configuração da memória, se terminar, deixará um valor $f_M(x)$ no registo R_0. Diz-se então que o programa M computa a função parcial $f_M:\mathbb{N} \nrightarrow \mathbb{N}$. A função é parcial, pois está indefinida para os naturais que não fazem o programa terminar. Por exemplo, o programa

$$(R_2:=3,(\texttt{while}(R_1 \leq R_2)\ do\ R_2:=3))$$

induz a função parcial f tal que $f(x)=x$ para $x>3$ e não está definida para $x<3$. Uma função para a qual existe um programa que a induz diz-se *computável*. A generalização para funções com aridade arbitrária é simples, bastando considerar a configuração inicial de memória $\xi(x_0,..., x_{k-1})$ para uma função de aridade k. Após a execução do programa, o resultado deste (caso exista) é o valor que se encontra no registo R_0.

Observe que o conjunto das funções (parciais) de $\mathbb{N}^k$ para $\mathbb{N}$ tem a cardinalidade dos reais, mas o conjunto de todos os programas, isto é, $\mathcal{M}$, tem a cardinalidade dos naturais. Assim, a esmagadora maioria das funções de $\mathbb{N}$ para $\mathbb{N}$ não é computável, e portanto desafia-se o leitor a encontrar uma função que não seja computável.

Um conjunto $A \subseteq \mathbb{N}^k$ diz-se *recursivo* se a sua função característica é computável, isto é, se a função $f:\mathbb{N}^k \rightarrow \{0,1\}$, tal que $f(x)=1$ se $x \in A$ é computável.

Para encontrar um conjunto não recursivo usa-se um argumento diagonal à la Cantor. Começa-se por considerar uma enumeração dos programas, isto é, uma bijecção $\gamma:\mathbb{N} \rightarrow \mathcal{M}$. Dada esta bijecção, é fácil mostrar que o seguinte conjunto

$$H=\{ x \in \mathbb{N}: f_{\gamma(x)} \text{está definida no valor } x \}$$

não é recursivo. Suponha que H é recursivo. Então há um programa M_H que induz a função característica de H e, mais, é possível construir o programa seguinte:

$$M'=M_H;(R_1:=1;\ while\ (R_1<R_0)\ do\ R_2:=1).$$

Observe que M' não termina se o programa M_H tiver colocado 1 no registo R_0, e termina caso tenha colocado 0 nesse mesmo registo. Com um pouco de esforço verifica-se que $\gamma^{-1}(M') \in H$ se $\gamma^{-1}(M') \notin H$, logo o programa M' não pode existir, e consequentemente a função característica de H não é computável, ou seja, H não é recursivo.

No que se segue, consideram-se apenas funções totais e computáveis. Como seria de esperar, algumas funções são induzidas por programas que utilizam mais tempo e espaço (número de registos e *bits* destes) do que outras. A área da complexidade computacional estuda precisamente os recursos necessários (em tempo e espaço) para programar uma função total e computável. Este texto vai cingir-se ao tempo, pois é o essencial para entender o problema P≠NP. Dada uma configuração de memória do computador ξ, é possível associar a cada programa M o seu tempo de execução $T(\xi,M)$ da seguinte forma:

- $T(\xi,R_i:=n)=|n|$;
- $T(\xi,R_i:=R_j+R_k)=|R_j^\xi|+|R_k^\xi|$;
- $T(\xi,R_i:=|R_j|)=\|R_j^\xi\|$ (o número de *bits* necessário para representar o número de *bits* de R_j^ξ);
- $T(\xi,R_{iR_k}:=R_{jR_l})=R_k^\xi+R_l^\xi$;
- $T(\xi,M_1;M_2)=T(\xi,M_1)+T(\xi',M_2)$ onde ξ' é o estado da memória que se obtém após executar M_1 sobre ξ;
- $T(\xi,\texttt{if}\ (R_1 \leq R_j)\ \texttt{then}\ M)=$ $\min(|R_i^\xi|,|R_j^\xi|)+\chi(R_i^\xi<R_j^\xi)\times T(\xi,M)$ onde $\chi(R_i^\xi<R_j^\xi)$ toma o valor 1 se $R_i^\xi < R_j^\xi$ e 0 caso contrário;
- $T(\texttt{while}(R_i \leq R_j)\ \texttt{do}\ M)=\min(|R_i^\xi|,|R_j^\xi|)+$ $+\sum_{v=0}^{k}(T(\xi_v,M)+\min(|R_i^v|,|R_j^v|)))$ onde k é o número de vezes que o ciclo `while` é executado e ξ_v é o estado da memória após a v-ésima execução de M sobre ξ_{v-1} com $\xi_0=\xi$.

Um programa M com entradas diz-se de *tempo polinomial* se existe um polinómio p e natural k tal que, para todo o $n>k$,
$\max_{\{x_1,...x_k:\sum_{i=0}^{k}|x_i| \leq n\}} T(\xi(x_1,...x_k),M) < p(n)$.

Se existe um programa de tempo polinomial que induz a função característica de um conjunto $A \subset \mathbb{N}^k$, diz-se que o problema da pertença a A é de tempo polinomial. A classe P contém todos os conjuntos cujo problema da pertença é de tempo polinomial. Por exemplo, o subconjunto dos pares está em P (tente encontrar um programa que prova tal facto). Em 2004 [1] foi demonstrado que o conjuntos dos números primos está em P.

Para definir a classe NP vai ser necessário relaxar a definição de função computável em tempo polinomial. Um programa M com entradas $x_1,... x_k$, diz-se de *tempo polinomial para os primeiros t<k*

100
95
75
25
5
0
Caderno_4
quarta-feira, 6 de Maio de 2009 13:25:50