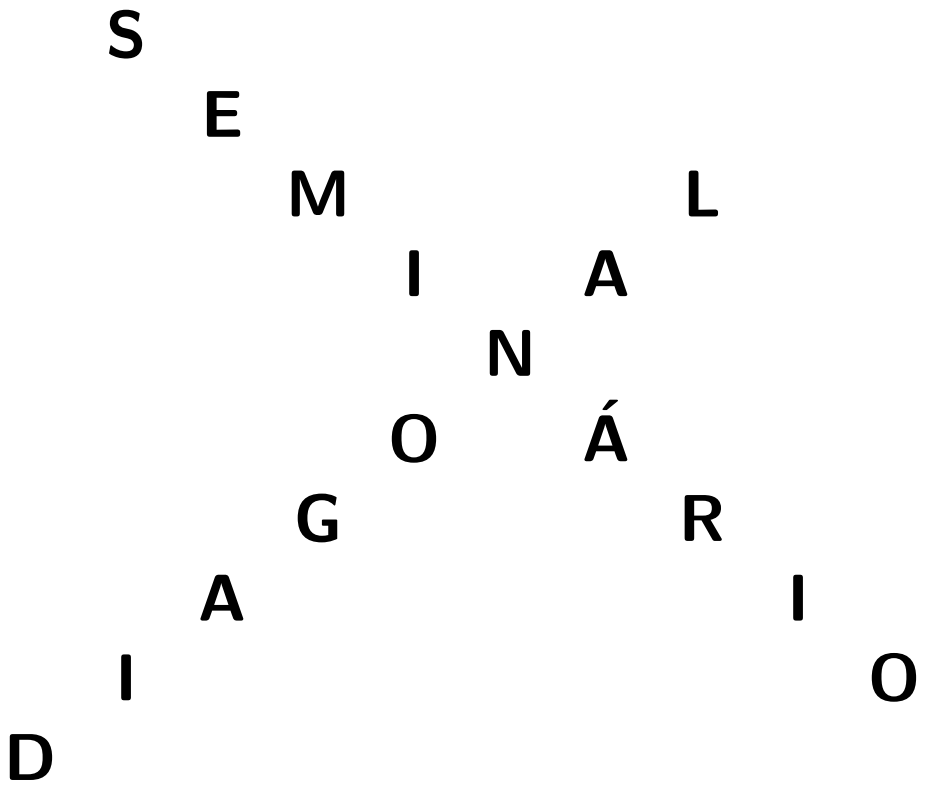


Proceedings IST, II

Outubro 2004

Editores:

Ana Cannas da Silva
Luís Cruz-Filipe
Ricardo Gonçalves
João Pimentel Nunes
Ana Rita Pires
Tiago Reis
Pedro Manuel Resende
Jorge Silva

Título ◇ Seminário Diagonal – Proceedings IST, II ◇ Outubro 2004

Editores ◇ Ana Cannas da Silva[1] ◇ Luís Cruz-Filipe[1]
◇ Ricardo Gonçalves[2] ◇ João Pimentel Nunes[1] ◇ Ana Rita Pires[2]
◇ Tiago Reis[2] ◇ Pedro Manuel Resende[1] ◇ Jorge Silva[1]

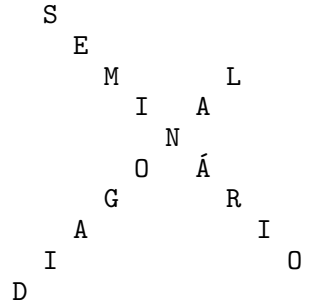
(1) Departamento de Matemática do Instituto Superior Técnico
(2) alunos da Licenciatura em Matemática Aplicada e Computação do IST até 2004

www ◇ `http://www.math.ist.utl.pt/diagonal/`

Documento LaTeX executado em 29 de Outubro de 2004, no Departamento de Matemática do Instituto Superior Técnico.

Prefácio

Com quatro anos de actividade regular, o Seminário Diagonal[1] atingiu a maturidade[2] mantendo-se fiel ao anúncio que o lançou no Instituto Superior Técnico (IST) em Outubro de 2000:




```
-----------------------------------

O que é?
Novo seminário de estudantes.

Para quem?
Todos os interessados em Matemática.

Sobre quê?
Matemática, no sentido lato.

-----------------------------------

Estreia brevemente em todo o país.
```

O Seminário conserva-se um espaço onde os alunos da Licenciatura em Matemática Aplicada e Computação do IST e muitos outros interessados

1 O Seminário Diagonal é uma iniciativa de âmbito nacional, e actualmente realiza-se também na Faculdade de Ciências da Universidade do Porto, na Faculdade de Ciências e Tecnologia da Universidade de Coimbra e na Faculdade de Ciências da Universidade de Lisboa.

2 A passagem do tempo é sublinhada pelo regresso à organização deste seminário do aluno que o inaugurou como orador, tendo concluído um doutoramento entretanto.

em Matemática, podem trocar ideias e experiências matemáticas. Aqui se mostram novos assuntos – sobretudo alguns menos conhecidos – de modo acessível aos alunos dos primeiros anos, ou seja, *diagonalizados*.[3]

Esta colectânea reúne seis *artigos diagonais* correspondentes a seminários entre Outubro de 2001 e Dezembro de 2002. Nem todos os oradores nesse período puderam contribuir com artigos. Juntamos uma lista dos resumos de todos os seminários realizados. A indicação do ano curricular refere-se ao ano lectivo em que o respectivo seminário foi apresentado.

Agradecimentos

Muitas pessoas têm construído o sucesso do Seminário Diagonal. O primeiro agradecimento vai para os oradores e seus colaboradores, com realce para os que contribuiram artigos para este volume, agradando-nos com o facto de serem metade dessas contribuições provenientes de colegas da Universidade do Porto e da Universidade de Coimbra. Vai também para toda a audiência – abrangendo alunos da LMAC, da LEFT, da LCI, da LEC, da LEIC, da LEEC e de outros cursos e professores do Departamento de Matemática do IST – e em especial para os organizadores do Seminário, cuja dedicação o sustenta.

Três instituições possibilitaram materialmente este projecto. A Fundação Calouste Gulbenkian, através do Programa Gulbenkian Novos Talentos em Matemática, patrocinou o trabalho de oito oradores entre Outubro de 2001 e Dezembro de 2002, incluindo os autores de cinco artigos aqui incluídos. O Banco BPI, através do Concurso de Apoio a Actividades Extracurriculares do IST, financiou a publicação da primeira colectânea Proceedings IST 2000-01, financiou a aquisição de livros para a Colecção Diagonal e co-financiou a publicação desta colectânea. O Centro de Análise Matemática, Geometria e Sistemas Dinâmicos do Instituto Superior Técnico viabilizou a publicação desta colectânea assegurando o restante apoio financeiro.

Três pessoas contribuiram com singular generosidade para os alicerces das *inicativas diagonais*. O Raúl Cordovil, a quem se deve a existência da Colecção Diagonal e estímulo de actividades relacionadas. O João Palhoto Matos, graças a quem existe `http://www.math.ist.utl.pt/diagonal/` e graças a quem a produção deste PDF foi para nós tão fácil. O João Boavida, cujo empenho na organização de 2000 a 2002, largamente ultrapassando as usuais responsabilidades, deixou um legado de ambição, incluindo um esmerado suporte gráfico em LaTeX, que só lastimamos não honrar aqui.

Os Editores,
Lisboa, 29 de Outubro de 2004

3 Citando João Boavida: tendo em conta que 'semi-simples' é sinónimo de 'diagonalizável', isto significa que os assuntos devem ser apresentados como 'soma directa' de 'partes simples'.

Conteúdo

Resumos dos Seminários

Redes Neuronais

Ricardo Silva (5º ano da LMAC — Ciência da Computação)
2 de Outubro de 2001

Os computadores actuais conseguem competir com um humano e ter sucesso em algumas áreas específicas (velocidade com que efectuam cálculos fastidiosos, jogos de xadrez, etc...). Contudo existem tarefas para as quais os computadores não parecem talhados. Qualquer criança de 2 ou 3 anos reconhece a cara dos pais, independentemente da distância, ângulo, iluminação, maquilhagem. Esta é uma tarefa complicada para qualquer computador. Na base destas diferencas estão as formas diferentes de funcionar do computador e do cérebro humano. O que podemos esperar se usarmos sistemas computacionais que são baseados no funcionamento do cérebro? Como funcionam esses sistemas?

Matrizes, Operadores Integrais e Difracção

Pedro Serranho (5º ano da LMAC — Análise Numérica)
16 de Outubro de 2001

Considere-se uma função $f : \mathbb{R}^2 \to \mathbb{R}$. Um processo para ver o seu gráfico consiste em tomar pontos $x_1, \ldots, x_N$ igualmente espaçados num intervalo $[a, b]$, e definir a matriz $A_{ij} = f(x_i, x_j)$. No Mathematica bastará então fazer `ListPlot3D[A]` – mas porque não calcular o seu determinante ou os seus valores próprios? Devemos suspeitar que o determinante é quase nulo quando f é contínua?

Este será o ponto de partida para o nosso seminário, onde os operadores integrais serão relacionados com matrizes. Veremos ainda simulações numéricas com equações integrais, relativas a uma aplicação física: a difracção de ondas e a localização de falhas em materiais.

Será que $\cos(m\pi/n)$ pode ser escrito de forma radical?

Ricardo Inglês (2º ano da LMAC)

30 de Outubro de 2001

Porque não escrever as razões trigonométricas, quando os argumentos destas são escritos sobre $\mathbb{Q}$ e em graus, com valores precisos ou, melhor, na forma de raízes? Foi esta pergunta que deu mote ao presente seminário, no qual se pretende retratar a investigação matemática de um estudante que, encontrando resultado atrás de resultado, se deixa embrulhar por estes, e de uma caixa de surpresas retira uma boa quantidade de observações interessantes dentro da área da trigonometria.

O Teorema de Gödel

Ricardo Gonçalves (3º ano da LMAC — Ciência da Computação)

20 de Novembro de 2001

Será que a Matemática não passa de uma mera manipulação simbólica? Será que existe algum sistema formal capaz de produzir todas as afirmações verdadeiras da Aritmética? Felizmente, para todos os amantes da Matemática, a resposta a estas perguntas é não! Iremos ver como Kurt Gödel chegou a este resultado, e qual o seu impacto dentro e fora da Matemática.

[1] Ricardo Gonçalves. O Teorema de Gödel. Neste volume.

O Teorema das Cinco Cores

Maria João Resende (3º ano de Matemática Pura, Universidade do Porto)

18 de Dezembro de 2001

Era uma regra, entre os fabricantes de mapas, que num mapa desenhado numa superfície plana, países adjacentes fossem pintados com cores diferentes; constatava-se que isso era sempre possível usando apenas quatro cores. A demonstração deste facto é muito complexa e exige uma utilização intensiva de computadores. Mas, utilizando apenas mais uma cor, o problema pode ser resolvido com mais facilidade.

[1] Maria João Resende. O Teorema das Cinco Cores. Neste volume.

Partidos, Deputados, Cubóides e Criaturas Afins

João Gouveia (2º ano de Matemática, Universidade de Coimbra)

26 de Março de 2002

Como se poderá representar geometricamente um processo eleitoral? O que terá um dodecaedro rômbico (o que é um dodecaedro rômbico??) a ver com partidos e deputados? E ainda como é que δ-cubos martelados aparecem no meio de tudo isto? Ou, mais genericamente, como de uma inocente curiosidade matemática se pode extrair uma quantidade de factos interessantes, divertidos e perfeitamente inúteis.

[1] João Gouveia e Eduardo Marques de Sá. Notas Sobre o Rateio de Hamilton – Geometria e Combinatória das Células Eleitorais. Neste volume.

O Teorema da Classificação de Superfícies

Ana Rita Pires (2º ano da LMAC)

30 de Abril de 2002

Qualquer superfície compacta é homeomorfa (pode ser deformada) a uma esfera, à soma conexa de toros (uma esfera com buracos), ou à soma conexa de planos projectivos (que não existe nas nossas corriqueiras três dimensões) de que a garrafa de Klein é um exemplo. Com muitos desenhos, 'tesoura' e 'cola', vamos cortar nuns sítios e colar noutros até demonstrar este teorema.

Cadeias e Polícias

Paulo Varandas (4º ano de Matemática Pura, Universidade do Porto)

11 de Junho de 2002

Um polícia foi destacado para controlar vários cruzamentos. Foi-lhe ordenado que, ao fim de um certo tempo num dado cruzamento, passe de forma aleatória para um dos cruzamentos vizinhos. Qual será a probabilidade de encontrar o polícia num dado cruzamento ao fim de algum tempo? As cadeias de Markov ajudarão a resolver o problema...

[1] Paulo Varandas. Cadeias e Polícias. Neste volume.

TRANSFORMAÇÕES DE LORENTZ E DE MÖBIUS

Ida Griffith (2º ano da LMAC)
15 de Outubro de 2002

Dois gémeos são separados no seu 20º aniversário; um fica na Terra, o outro viaja a uma velocidade próxima da velocidade da luz em direcção a um planeta situado a 8 anos-luz e regressa. Que idade terá quando regressar? Terá a mesma idade que o gémeo que ficou na Terra? E que relação tem isto tudo com o Grupo de Möbius?

[1] Ida Griffith. Transformações de Lorentz e de Möbius. Neste volume.

CRIPTOGRAFIA: O ESTADO DA ARTE

Paulo Matos (4º ano da LEIC, IST)
29 de Outubro de 2002

Existem dois tipos de criptografia neste mundo: criptografia que impede a nossa irmã mais nova de ler os nossos ficheiros e criptografia que impede os mais poderosos governos mundiais de os lerem. Neste seminário falaremos sobre o último... Veremos como se processa o envio de uma mensagem utilizando um algoritmo híbrido (RSA+RC4) com assinatura digital (muito utilizado actualmente) e qual a ideia existente por detrás dos dois grandes tipos de criptografia: chave pública e chave privada.

[1] Paulo Matos. Criptografia: O Estado da Arte. Neste volume.

SECÇÕES HIPERPLANAS DO HIPERCUBO

Diogo Veloso (4º ano de Matemática, Universidade de Lisboa)
26 de Novembro de 2002

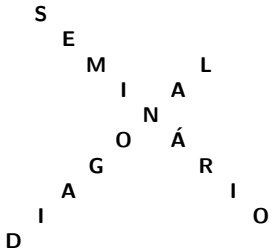
Não é possível ver o hipercubo $[0,1]^4$, nem um hiperplano de $\mathbb{R}^4$. Contudo a sua intersecção é um poliedro convexo, um objecto tridimensional. Como serão, então, essas intersecções? Neste seminário mostrar-se-á como se pode resolver este problema.

O Canário Roxo

Ricardo Inglês (3º ano da LMAC — Análise, Geometria e Álgebra)

17 de Dezembro de 2002

Vamos neste seminário abordar as aparentes contradições lógicas de enigmas com auto-referência e a forma de conseguirmos livrar-nos delas. 'Dizes a verdade ou mentes?', 'Qual a aldeia dos mentirosos?' e 'Será que a testemunha cometeu perjúrio?' são perguntas e enigmas aos quais pretendemos responder e esclarecer por exemplos. Vendo bem, se nos disserem 'Se mentires morres enforcado, se disseres a verdade morres afogado' é bom que saibamos que resposta dar para não morrer. E já que é a nossa vida que está em jogo, porque não usar a matemática?



Teorema de Gödel

Ricardo Gonçalves
4º ano da LMAC — Ciência da Computação
rgon@math.ist.utl.pt

Palavras Chave

sitemas formais, expressabilidade, codificação, auto-referência

Resumo

Será que a Matemática não passa de mera manipulação simbólica? Será que existe um sistema formal capaz de produzir todas as afirmações verdadeiras da Aritmética? Felizmente para todos os amantes da Matemática a resposta a esta pergunta é negativa. Iremos ver como Gödel chegou a este resultado e qual o seu impacto dentro e fora da Matemática.

1 Breve introdução histórica

Desde muito cedo que se pretendia axiomatizar tudo o que existe. Este desejo tornou-se bastante mais intenso em meados do século 20 e em particular na área da Matemática.

Um dos maiores apoiantes da ideia de que a Matemática pode ser totalmente capturada por um sistema formal foi o grande matemático David Hilbert.[1] Hilbert, que era um matemático muito prestigiado na altura, defendia com muita convicção que existia um sistema formal capaz de produzir todas as afirmações verdadeiras da Matemática. No congresso mundial da Matemática, em 1900, Hilbert apresentou um conjunto de problemas matemáticos que ele considerava importantes e que deveriam ser resolvidos no século 20. Entre eles encontrava-se este seu sonho de axiomatizar a Matemática.

Em 1931 um jovem matemático austríaco, Kurt Gödel, mostrou que isso não era possível, isto é, Gödel mostrou que a aritmética não pode ser totalmente capturada por um sistema formal.

1 David Hilbert, matemático alemão que viveu nos séculos XIX e XX.

2 Sistema Formal da Aritmética

Neste capítulo iremos introduzir o sistema formal da aritmética que designaremos por $\mathcal{N}$. Como veremos, este sistema resulta do enriquecimento de sistemas formais mais fracos. Construamos uma sequência de sistemas formais encadeados:

$$\boxed{\text{sistema } L} \rightarrow \boxed{\text{sistema } K_{\mathcal{L}}} \rightarrow \boxed{\text{Sistema } \mathcal{N}}$$

em que o Sistema L representa o sistema formal associado à Lógica Proposicional e o Sistema $K_{\mathcal{L}}$ representa o sistema formal associado à Lógica de 1ª Ordem. Iremos então apresentar a linguagem $\mathcal{L}_N$:

variáveis: x_1, x_2, ...
constantes: a_1 (denota 0)
símbolos de função: f_1^1, f_1^2, f_2^2 (sucessor, adição, produto)
símbolo de predicado: A_1^2 (denota =)
pontuação, conectivos e quantificadores.

Embora não saibamos se, por exemplo, o símbolo de função f_1^2 é, em todos os modelos, sempre interpretado como a soma, utilizaremos a seguinte notação de modo a clarificar um pouco o sistema $\mathcal{N}$:

Notação.

$(t_1 + t_2)$ para $f_1^2(t_1, t_2)$,

$(t_1 \times t_2)$ para $f_2^2(t_1, t_2)$,

t' para f_1^1(t),

onde t, t_1, t_2 são quaisquer termos.
Usaremos também 0 para denotar a_1.

Temos, no entanto, de ter atenção para não confundir o Sistema Formal $\mathcal{N}$, meramente simbólico, com o modelo por nós pretendido que corresponde à nossa intuição de Aritmética. Sabemos que L e $K_{\mathcal{L}}$ são completos. Podemos enunciar então a questão fundamental:

Será $\mathcal{N}$ sistema formal da Aritmética completo ?

A resposta foi dada por Gödel. O Sistema formal da Aritmética não é completo!

Nesta passagem de $K_{\mathcal{L}}$ para $\mathcal{N}$ perdemos a completude, é portanto legítimo colocarmos a seguinte questão: *Será possível determinar o ponto exacto onde se perde a completude?* Esta questão ainda não tem resposta e continua a motivar investigação. Por exemplo, existem aritméticas que ainda são completas e que são chamadas aritméticas bébés. (É obvio que estes sistemas têm um poder expressivo inferior ao de $\mathcal{N}$.)

Vamos ver como Gödel chegou a este resultado. Devido a esta prova ser altamente técnica, iremos omitir as partes mais técnicas para perceber realmente as ideias fundamentais deste teorema.[2]

Iremos por partes. Primeiro falaremos de *expressabilidade*.

3 Expressabilidade

Continuemos o estudo do sistema formal $\mathcal{N}$ da Aritmética e N o seu modelo pretendido. Temos de distinguir, claramente, $\mathcal{N}$ sistema formal de N seu modelo.

Comecemos com o modelo N, cujo domínio é o conjunto dos Naturais que denotaremos por D_N. Observemos que o número 0, as funções sucessor, adição, produto, e a igualdade são representadas de uma forma óbvia por símbolos de $\mathcal{N}$. No entanto, por exemplo, o número 5 não é representado por um único símbolo de $\mathcal{N}$, embora seja a interpretação de uma sequência de símbolos de $\mathcal{N}$. Para tornar mais clara a relação do número n com a sua representação em $\mathcal{N}$, introduziremos a seguinte notação.

Notação.
O símbolo $0^{(n)}$ é a abreviatura de 0 seguido de n apóstrofes. Logo o número $n \in D_N$ é a interpretação em N do termo $0^{(n)}$. Os termos da forma $0^{(n)}$, $n \in \mathbb{N}$ designam-se *termos numerais*.

Apresentamos então a seguinte definição:

Definição (Relação expressável). Uma relação k-ária R em $\mathbb{N}^k$ é expressável em $\mathcal{N}$ se existir uma fórmula $A(x_1, \ldots, x_n)$ com k variáveis livres, tal que para quaisquer $n_1, \ldots, n_k \in D_N$

(i) se $\langle n_1, \ldots, n_k \rangle \in$ R então $\vdash_{\mathcal{N}} A(0^{(n_1)}, \ldots, 0^{(n_k)})$ e

2 Consultar [2] para detalhes técnicos da prova.

(ii) se $\langle n_1, \ldots, n_k \rangle \notin \mathrm{R}$ então $\vdash_{\mathcal{N}} \neg A(0^{(n_1)}, \ldots, 0^{(n_k)})$.

Uma função é um caso particular de relação. Em geral uma relação (k+1)-ária R em D_N é uma função se para qualquer $n_1, \ldots, n_k \in D_N$ existe precisamente um $n_{k+1} \in D_N$ tal que $\langle n_1, \ldots, n_k, n_{k+1} \rangle \in \mathrm{R}$.

A questão da expressabilidade em termos de funções é também muito importante.

DEFINIÇÃO (FUNÇÃO EXPRESSÁVEL). Uma função k-ária em D_N é expressável em $\mathcal{N}$ se é expressável em $\mathcal{N}$ como uma relação por uma fórmula A com k+1 variáveis livres, tal que para quaisquer $n_1, \ldots, n_k \in D_N$

$$\vdash_{\mathcal{N}} (\exists_1 x_{k+1}) A(0^{(n_1)}, \ldots, 0^{(n_k)}, x_{k+1})$$

Por um simples argumento de cardinalidade podemos concluir que nem todas as relações (funções) são expressáveis.

A próxima questão que podemos levantar é se podemos caracterizar o conjunto das funções (relações) em D_N que são expressáveis em $\mathcal{N}$? A resposta é um resultado importante e uma das chaves deste Teorema de Gödel.

TEOREMA.
Uma função (relação) em D_N é expressável em $\mathcal{N}$ se e só se é recursiva.

A prova desta teorema está para lá do âmbito deste artigo. Assumimos que o leitor tem algum conhecimento prévio acerca de funções recursivas.[3]

Juntamos agora uma definição simples.

DEFINIÇÃO. Seja R uma relação k-ária em D_N. A função característica de R, denotada por $\mathcal{C}_R$, é definida por:

$$\mathcal{C}_R(n_1, \ldots, n_k) = \begin{cases} 0 \text{ se } \langle n_1, \ldots, n_k \rangle \in \mathrm{R} \\ 1 \text{ se } \langle n_1, \ldots, n_k \rangle \notin \mathrm{R} \end{cases}$$

DEFINIÇÃO. Uma relação R em D_N é recursiva se a sua função característica for uma função recursiva.

Embora existam mais funções que não sejam recursivas (*o seu cardinal não é contável*) do que funções recursivas, o facto é que é mais complicado encontrar uma função que não seja recursiva, pois praticamente todas as funções facilmente descritíveis são recursivas.

Passemos agora a outro ponto fundamental a caminho do Teorema de Gödel.

3 Consultar [3] para detalhes sobre funções recursivas.

4 Números de Gödel

A palavra-chave aqui é codificação, e codificação utilizando números. A técnica que Gödel utilizou foi codificar, de modo construtivo, a linguagem de 1ª ordem $\mathcal{L}$ (não necessariamente $\mathcal{L}_N$), tal que a cada símbolo, termo, fórmula e sequência de fórmulas de L é atribuído um código numérico. Esta atribuição é feita de tal modo que, dado qualquer código numérico, a expressão correspondente seja facilmente recuperada. Existem várias maneiras de fazer esta codificação, mas não iremos aqui especificar nenhuma. Iremos apenas assumir que temos uma função g, do conjunto de todos os símbolos, sequências de símbolos e sequências finitas de sequências de símbolos, com valores em D_N. Aos valores de g chamaremos números de Gödel.

Uma demonstração ou uma derivação em $K_{\mathcal{L}}$ é uma sequência finita de sequências de símbolos, logo tem um número de Gödel.

O objectivo de Gödel era transformar afirmações acerca de um sistema formal em afirmações acerca de números naturais e depois expressar estas afirmações dentro do próprio sistema formal.

No âmbito de tornar mais clara esta questão, consideremos a afirmação:

$$A \textit{ sequência } A_1, \ldots, A_k, A \textit{ é a prova em } \mathcal{N} \textit{ da fórmula } A.$$

Em termos de números de Gödel isto representa uma relação em D_N, digamos P_f, definida por :

> $\langle m, n\rangle \in P_f$ se e só se m é o número de Gödel de uma sequência de fórmulas de $\mathcal{N}$ que constituem uma demonstração em $\mathcal{N}$ de uma fórmula cujo número de Gödel é n.

É agora que a questão da expressabilidade se torna fundamental. Continuemos o exemplo anterior:

Se a relação P_f fosse expressável em $\mathcal{N}$ existiria uma fórmula $P(x_1, x_2)$ de $\mathcal{L}$ tal que para qualquer $n, m \in D_N$ se teria:

1. se $\langle m, n\rangle \in P_f$ então $\vdash_{\mathcal{N}} P(0^{(m)}, 0^{(n)})$;
2. se $\langle m, n\rangle \notin P_f$ então $\vdash_{\mathcal{N}} \neg P(0^{(m)}, 0^{(n)})$.

O que estamos a fazer é uma tentativa de usar o sistema $\mathcal{N}$ como uma meta sistema por si só. No entanto, sabemos que unicamente as relações recursivas em D_N são expressáveis em $\mathcal{N}$, logo certamente não podemos usar este procedimento para todas as relações em D_N e o uso de $\mathcal{N}$ como seu próprio meta sistema será necessariamente parcial.

O próximo passo na prova do Teorema de Gödel é mostrar que certas relações em D_N, que surgem desta forma, de considerações acerca de fórmulas, teoremas e provas, são recursivas e logo expressáveis em $\mathcal{N}$.

Limitar-nos-emos a listar apenas algumas

Pf $\langle m, n\rangle \in Pf$ se e só se m é o número de Gödel de uma demonstração em $\mathcal{N}$ da fórmula com número de Gödel n.

W $\langle m, n\rangle \in W$ se e só se m é o número de Gödel de uma fórmula $A(x_1)$ onde x_1 ocorre livre e n é o número de Gödel de uma demonstração em $\mathcal{N}$ da fórmula $A(0^{(m)})$

Esta relação W definida acima é a chave para a prova da incompletude. Note-se que esta relação envolve a substituição do termo $0^{(m)}$ (que corresponde ao número m) na fórmula $A(x_1)$ cujo número de Gödel é m.

W é expressável em $\mathcal{N}$, logo existe uma fórmula $\mathcal{W}(x_1, x_2)$ onde apenas x_1 e x_2 ocorrem livres, tal que:

- Se $\langle m, m\rangle \in W$ então $\vdash_{\mathcal{N}} \mathcal{W}(0^{(m)}, 0^{(n)})$
- Se $\langle m, n\rangle \notin W$ então $\vdash_{\mathcal{N}} \neg\mathcal{W}(0^{(m)}, 0^{(n)})$

Considere a fórmula seguinte:

$$(\forall x_2)\neg\mathcal{W}(x_1, x_2)$$

Seja p o número de Gödel desta fórmula e considere-se finalmente a fórmula que se obtém substituindo x_1 por $0^{(p)}$

$$(\forall x_2)\neg\mathcal{W}(0^{(p)}, x_2)$$

e chamemos-lhe $\mathcal{U}$.

Dêmos uma interpretação grosseira de $\mathcal{U}$:

> Para qualquer $n \in D_N$, não é o caso que p é o número de Gödel de uma fórmula $\mathcal{A}(x_1)$ em que x_1 ocorre livre e n é o número de Gödel de uma demonstração em $\mathcal{N}$ de $\mathcal{A}(0^p)$

Sabemos que p é o número de Gödel de uma fórmula onde x_1 ocorre livre [$(\forall x_2)\neg\mathcal{W}(x_1, x_2)$] e se esta fórmula for denotada por $\mathcal{A}(x_1)$ então $\mathcal{A}(0^{(p)})$ é a fórmula $\mathcal{U}$. Logo a interpretação que demos de $\mathcal{U}$ é equivalente a:

> Para todo o $n \in D_N$, n não é número de Gödel de uma demonstração em $\mathcal{N}$ da fórmula $\mathcal{U}$.

Note-se que se $\mathcal{N}$ não fosse coerente, então seria trivialmente completo, pois todas as fórmulas seriam teoremas.

Logo o teorema da incompletude de Gödel necessitará da hipótese de que $\mathcal{N}$ seja coerente. De facto, esta prova do teorema de Gödel requer uma hipótese ligeiramente mais forte.

DEFINIÇÃO. Um sistema de 1ª ordem S com a mesma linguagem de $\mathcal{N}$ é ω-coerente se para nenhuma fórmula $\mathcal{A}(x_1)$, onde x_1 ocorre livre, se tem $\neg(\forall x_1)\mathcal{A}(x_1)$ um teorema de S se $\mathcal{A}(0^{(n)})$ é teorema de S para qualquer $n \in D_N$.

Observação 1. O facto de $\mathcal{A}(0^{(n)})$ ser teorema para qualquer $n \in D_N$, não implica que a fórmula $\neg(\forall x_1)\mathcal{A}(x_1)$ seja necessariamente teorema.

Observação 2. Sabemos também que se S é ω- coerente, então S é coerente.

5 Teorema de Gödel

Enunciemos o teorema da Incompletude de Gödel:

TEOREMA (TEOREMA DA INCOMPLETUDE DE GÖDEL). *Se $\mathcal{N}$ é ω-coerente, então nem a fórmula $\mathcal{U}$ nem a sua negação são teoremas em $\mathcal{N}$. Se $\mathcal{N}$ é ω-coerente, então $\mathcal{N}$ não é completo.*

Prova. Vamos dividir esta prova em duas partes. Na primeira parte iremos assumir que $\mathcal{U}$ é teorema de $\mathcal{N}$, na segunda iremos assumir que $\mathcal{U}$ não é teorema de $\mathcal{N}$.

1. **$\mathcal{U}$ é teorema de $\mathcal{N}$** e seja q o número de Gödel de uma prova em $\mathcal{N}$ de $\mathcal{U}$. Como antes, seja p o número de Gödel de $(\forall x_2)\neg\mathcal{W}(x_1, x_2)$. Então $\mathcal{W}(p, q)$ verifica-se. W é expressável em $\mathcal{N}$ por $\mathcal{W}$ então teremos, $\vdash_{\mathcal{N}} \mathcal{W}(0^{(p)}, 0^{(q)})$. Mas $\vdash_{\mathcal{N}} \mathcal{U}$, isto é, $\vdash_{\mathcal{N}} (\forall x_2)\neg\mathcal{W}(0^{(p)}, x_2)$ e então $\vdash_{\mathcal{N}} \neg\mathcal{W}(0^{(p)}, 0^{(q)})$ usando $K5$[4] e MP. Isto contradiz a hipótese de que $\mathcal{N}$ é coerente, logo $\mathcal{U}$ não pode ser teorema de $\mathcal{N}$.

2. **$\mathcal{U}$ não é teorema de $\mathcal{N}$** , isto é, não existe uma prova em $\mathcal{N}$ de $\mathcal{U}$ $((\forall x_2)\neg\mathcal{W}(0^{(p)}, x_2))$. Logo $W(p, q)$ não se verifica para nenhum número q. Então $\vdash_{\mathcal{N}} \neg\mathcal{W}(0^{(p)}, 0^{(q)})$ para qualquer q. Por ω-coerência, temos que $\neg(\forall x_2)\neg\mathcal{W}(0^{(p)}, x_2)$ não é teorema de $\mathcal{N}$, isto é, $(\neg\mathcal{U})$ não é teorema de $\mathcal{N}$.

□

4 Consultar [1] para detalhes sobre o axioma $K5$.

Podemos então enunciar o seguinte corolário:

COROLÁRIO. *$\mathcal{N}$ contém uma fórmula fechada que é verdadeira no modelo N mas que não é teorema de $\mathcal{N}$.*

Prova.
A fórmula $\mathcal{U}$ é fechada e nem $\mathcal{U}$ nem $(\neg\mathcal{U})$ é teorema de $\mathcal{N}$. No entanto, visto que N é uma interpretação, ou $\mathcal{U}$ ou $(\neg\mathcal{U})$ é verdadeiro em N. □

Poderíamos pensar em tentar tornar $\mathcal{N}$ completo adicionando-lhe $\mathcal{U}$ como axioma. Mas o que vai acontecer é o seguinte:

PROPOSIÇÃO. *Seja S qualquer extensão de $\mathcal{N}$ tal que o conjunto de números de Gödel dos axiomas próprios de S é um conjunto recursivo. Então se S é coerente S não é completo.*

6 Observações finais

Observação. Observemos que as hipóteses do Teorema de Gödel incluem a hipótese de que o sistema formal em causa tem um conjunto de axiomas cujos números de Gödel constituem um conjunto recursivo. Esta é uma suposição necessária para a prova, visto que a prova envolvia demonstrações que certas relações seriam recursivas.

Se nós permitirmos que o nosso conjunto de axiomas próprios seja não recursivo, poderemos ter um sistema de 1ª ordem da aritmética que é coerente e completo.

Comentário. É claro que, se considerarmos como axiomas todas as fórmulas que são verdadeiras em N obtemos um sistema formal coerente e completo, mas claramente o conjunto dos números de Gödel dos axiomas não é recursivo.

Uma questão que surge é a seguinte: *o que têm os conjuntos recursivos que faça com que consideremos o Teorema da Incompletude de Gödel significativo?* A resposta assenta em ideias de computabilidade e algoritmos e a sua relação com a ideia de recursividade.[5]

7 Agradecimentos

Gostaria em primeiro lugar de agradecer ao Professor Félix Costa sem o qual, quer a elaboração deste artigo, quer a realização do seminário que lhe deu

5 ver por exemplo Capítulo 7 de [1].

origem não teriam sido possíveis. Queria também agradecer à organização do Seminário Diagonal por toda a disponibilidade e ajuda. Finalmente, um agradecimento muito especial aos meus pais, às minhas irmãs e à minha namorada Lurdes por todo o apoio que me deram. Obrigado a todos.

Referências

[1] A. G. Hamilton, *Logic for Mathematicians*, Cambridge University Press, 1978.

[2] R. Cori, D. Lascar, *Mathematical Logic II: Recursion Theory, Gödel's Theorems, Set Theory, Model Theory*, Oxford University Press, 2000.

[3] N. Cutland, *Computability – An Introduction to Recursive Function Theory*, Cambridge University Press, 1980.

O Teorema das Cinco Cores

Maria João Resende
3º ano de Matemática Pura
Faculdade de Ciências da Universidade do Porto
mjoaoresende@net.sapo.pt

Palavras Chave

mapa, região, fronteira, face, aresta, vértice, identidade de Euler

Resumo

Era uma regra entre os fabricantes de mapas, que num mapa desenhado numa superfície plana, países adjacentes fossem pintados com cores diferentes; constatava-se que isso era sempre possível usando apenas quatro cores. A demonstração conhecida deste facto é muito complexa e exige uma utilização intensiva de computadores. Mas, utilizando apenas mais uma cor, o problema similar pode ser resolvido com mais facilidade.

1 Problema das Quatro Cores

Em 23 de Outubro de 1852, De Morgan, Professor do University College em Londres recebia de um aluno, Frederick Guthrie, o enunciado do Problema das Quatro Cores. O autor da questão era o irmão do aluno, Francis Guthrie, a quem o problema ocorrera quando coloria um mapa de Inglaterra.

Quatro cores bastam para pintar um mapa plano de forma a que dois países vizinhos não partilhem a mesma cor?

Consideramos que países que se tocam apenas num ponto não são vizinhos.

Por exemplo, se num tabuleiro de xadrez cada casa representar uma nação, a coloração habitual mostra-nos que duas cores bastam para este caso como se vê na Figura 1.

No entanto a Figura 2 prova que por vezes as quatro cores são mesmo necessárias, já que cada um dos países tem três vizinhos distintos.

Como resposta ao "Problema das Quatro Cores" surge o seguinte teorema.

TEOREMA. *Dado um mapa plano, dividido em regiões, são necessárias no máximo quatro cores para o colorir, de forma a que regiões vizinhas não partilhem a mesma cor.*

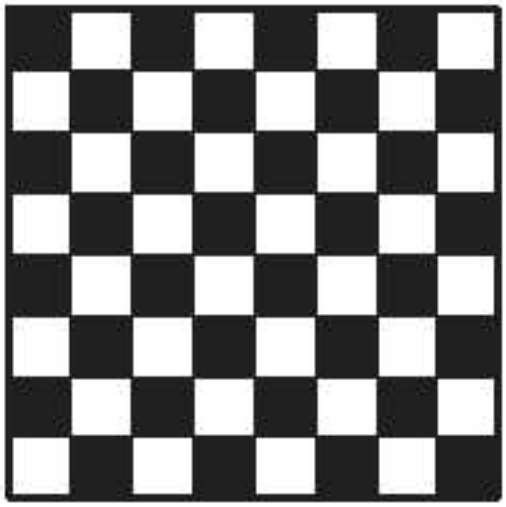


Figura 1: Tabuleiro de Xadrês

Observação. Regiões que só se tocam num ponto não são consideradas vizinhas.

1.1 Nota Histórica

Este teorema foi "demonstrado" com a ajuda de um computador IBM360 em 1976 por Appel (matemático americano) e Haken (matemático alemão). A prova mostrava que, se para cerca de 2000 formas "básicas" de mapas era possível colorir as regiões nas condições do teorema, então para qualquer outro mapa também seria.

Sendo a prova demasiado longa, não pode ser inspeccionada por um ser humano. Trata-se do primeiro célebre resultado, cuja prova não pode ser aferida por outros matemáticos.

Em Abril de 1975, Martin Gardner apresentou um mapa, com 110 regiões, e dizia que para aquele exemplo seriam necessárias cinco cores. Seria assim um contra-exemplo, colocando em dúvida o Teorema das Quatro Cores. Este facto não passou de uma brincadeira no "Dia das Mentiras". Realmente era difícil utilizar apenas quatro cores, mas como se vê na Figura 3 não era impossível.

Em 1994 uma prova simplificada, da autoria dos matemáticos americanos Paul Seymour, Neil Robertson, Daniel Sanders e Robin Thomas foi anunciada. Reduzindo a 633 casos a verificar e utilizando algoritmos mais eficientes. No entanto a colaboração do computador é ainda indispensável. Realmente, este é um problema que qualquer pessoa sem preparação matemática pode

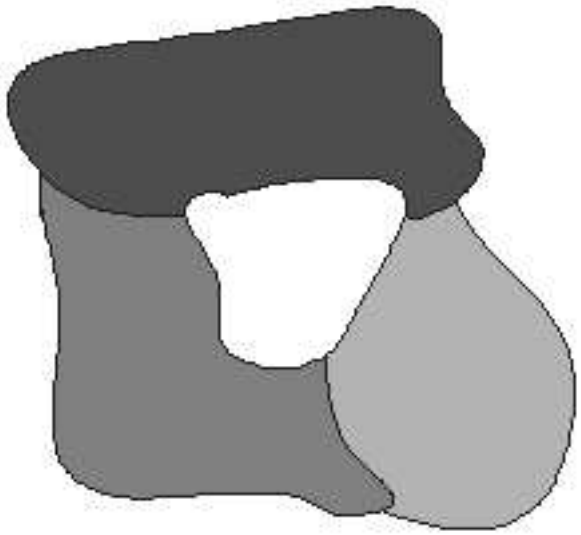


Figura 2: Neste Caso São Necessárias Quatro Cores

entender, e no entanto a sua demonstração é bastante elaborada.

2 O Teorema das Cinco Cores

Se acrescentarmos uma cor obteremos um resultado análogo mas muito mais simples de demonstrar.

Teorema. *Dado um mapa num plano, dividido em regiões, é possível colorir cada uma das regiões de forma a que regiões vizinhas tenham cores diferentes, utilizando apenas cinco cores.*

Num mapa temos representadas regiões que são delimitadas pelas fronteiras. Ao longo deste artigo iremos considerar que as regiões são faces enquanto as fronteiras serão as arestas. À intersecção de três ou mais países chamaremos vértice.

Na prova vamos supor que o mapa representa apenas uma ilha. Se uma ilha e o mar puderem ser pintados com cinco cores, então um mapa formado por várias ilhas poderá também ser pintado, usando as mesmas cores em cada ilha.

Vamos ainda supor que no mapa não existem regiões em forma de anel, isto é, se escolhermos um ponto de uma fronteira é possível aceder a todas as fronteiras percorrendo apenas caminhos sobre as fronteiras (sem "entrar" em nenhuma região). Se uma tal região existir então bastará considerar dois mapas: um formado pela região em forma de anel e pelas regiões que estão

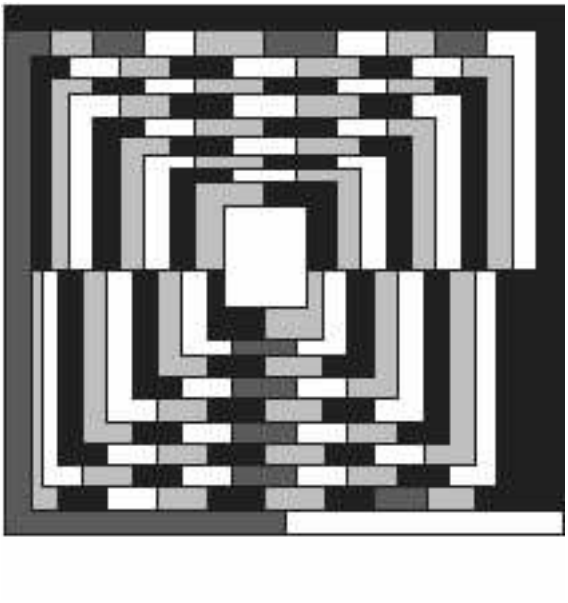


Figura 3: Solução Para o Problema Colocado por Martin Gardner

no "exterior" do anel e o outro será formado pela região em forma de anel e pelas regiões que estão no "interior" do anel. Se estes dois mapas verificarem o teorema, podemos colorir cada um deles com apenas cinco cores e tendo o cuidado para que a região em forma de anel tenha a mesma cor nos dois mapas. Então bastará sobrepôr os dois mapas e estará verificado o teorema.

Para provar o Teorema das Cinco Cores vamos então mostrar a Identidade de Euler.

2.1 Identidade de Euler

LEMA. *Num mapa o número de faces (países) adicionado ao número de vértices é igual à soma de dois com o número de arestas (fronteiras).*

$$v + f = a + 2$$

Demonstração. Sejam:

f = n.º de faces (países);
a = n.º de arestas (fronteiras);
v = n.º de vértices.

Já vimos que no nosso mapa não existem regiões em forma de anel.

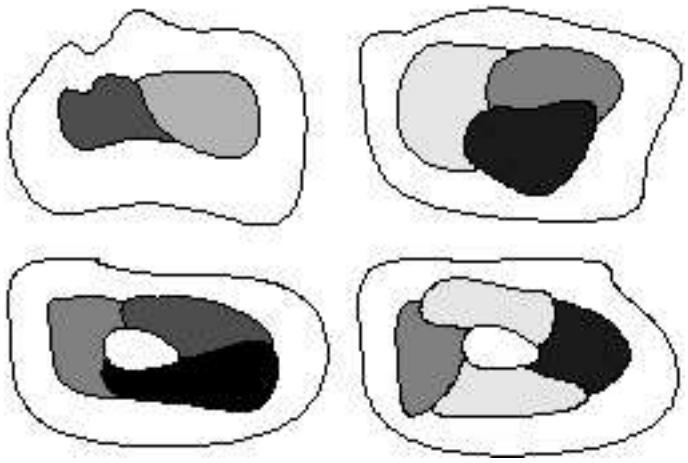


Figura 4: Exemplos de Mapas Formados por uma Ilha

Para a prova da Identidade de Euler vamos deixar um pouco de parte a ideia de mapa e considerar uma ideia semelhante. Pensemos num sistema de campos de cultivo e diques. As arestas, que representavam fronteiras, vamos associar aos diques e os países, representados pelas faces, associamos aos campos de cultivo. Temos assim os diques a delimitarem os campos. A área exterior, que no mapa corresponde ao oceano, é uma zona completamente inundada.

Suponhamos que é necessário inundar todos os campos. Então vamos deitar diques abaixo de forma a conseguirmos o pretendido.

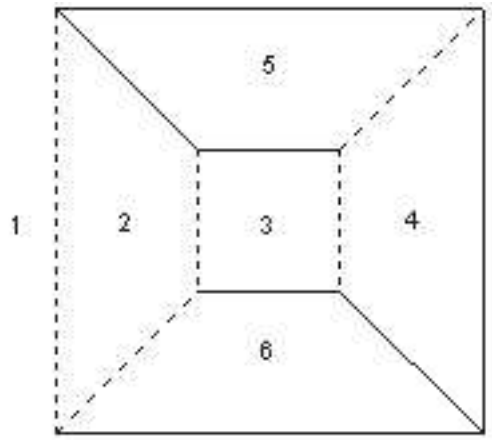


Figura 5: Sistema de campos de cultivo e diques (As arestas a tracejado são diques destruídos e os restantes são diques intactos)

Mas, para isso, não será necessário destruir todos os diques (Figura 5).

Um dique que já tenha água dos dois lados, certamente poderá ficar intacto. Se apenas deitarmos abaixo os diques que tenham água apenas de um lado, então a cada passo, devemos destruir um dique e inundar mais um campo.

A área exterior – que correspondia ao mar – está coberta de água desde o início. Logo existem exactamente $f - 1$ campos para inundar. Depois deste processo sucessivo de destruição dos diques, conseguimos inundar todos os campos e assim deveremos ter exactamente $f - 1$ diques destruídos.

Consideremos agora o sistema de diques que não foi destruído, vamos então estudar duas questões.

I. Uma pessoa pode ir de um vértice a qualquer outro, caminhando sobre os diques, sem molhar os pés (isto é, sem caminhar sobre o chão dos campos)?

   Antes dos diques serem destruídos, isso certamente é possível, pois supomos desde o início que o mapa era formado por apenas uma ilha sem regiões em forma de anel.

   Suponhamos agora que no processo de inundação dos campos, a destruição de um dique AB (Figura 6) separa o sistema em duas ilhas completamente separadas.

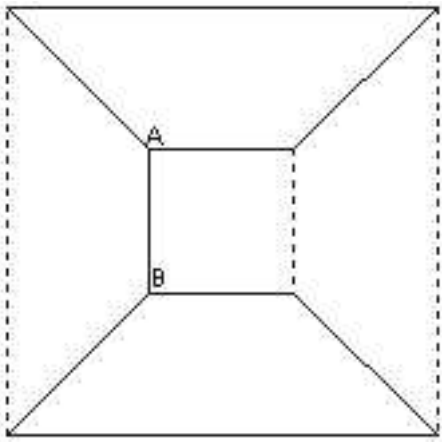


Figura 6: Separação do sistema de campos em dois sistemas disjuntos.

   Se AB fosse destruído, seria impossível ir de A até B ao longo dos diques, pois a água estaria a rodear completamente os dois sistemas. Por isso, deveria haver água nos dois lados do dique AB antes deste ser destruído e portanto, por hipótese, o dique não deveria ter sido destruído.

II. Pode haver dois caminhos diferentes para um mensageiro ir de um vértice P a um vértice Q?

Suponhamos que temos dois caminhos diferentes para ir de P a Q (Figura 7), então esses caminhos, tendo em comum o ponto inicial P e o ponto final Q estariam a cercar alguma área.

Então o "anel" formado pelos diques não destruídos, que fazem parte dos dois caminhos que unem P a Q divide o plano em duas partes, estando então a parte interior separada da exterior, e portanto não está inundada, o que contradiz o facto de todos os campos estarem inundados.

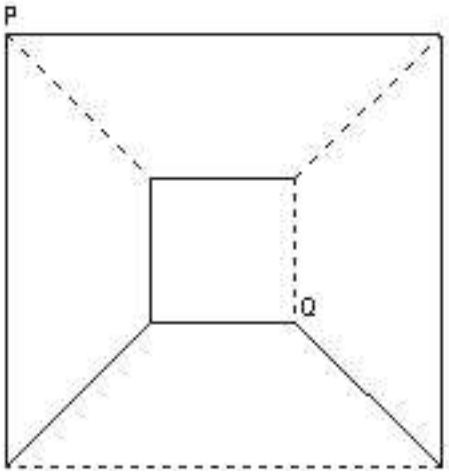


Figura 7: Número de caminhos que unem dois vértices.

Agora vejamos, se mantivermos o ponto inicial P fixo, já sabemos que existe apenas um caminho para ir até cada um dos outros vértices. Nesse caminho, vai haver um último dique para passar até ao vértice final (e esse dique será último para apenas um vértice, embora possamos passar por ele para chegar a outros vértices). Assim temos uma correspondência entre o número de vértices e o número de diques não destruídos. A cada dique não destruído corresponde o último ponto de um caminho que começou em P. Como inicialmente estamos em P, então este é ponto inicial e não será ponto final de nenhum caminho. Logo o número de diques não destruídos é $v - 1$.

Temos então,

$$\text{Número de diques não destruídos} = v - 1$$
$$\text{Número de diques destruídos} = f - 1$$
$$\text{Número total de diques} = a$$

Logo, vem que

$$a = (v - 1) + (f - 1)$$

e feitas as contas...

$$a + 2 = v + f$$

chegamos então à Identidade de Euler.

□

2.2 Demonstração do Teorema das Cinco Cores

Para a demonstração do Teorema das Cinco Cores comecemos por simplificar o problema.

Se tivermos um mapa em que mais de três países se encontram no mesmo vértice (Figura 8), então podemos desenhar um novo mapa, que será exactamente a cópia do original, excepto um pequeno país novo, formado em volta do vértice em questão (Figura 9). Este novo país deverá ser suficientemente pequeno, de forma a não "cobrir" por completo nenhuma fronteira.

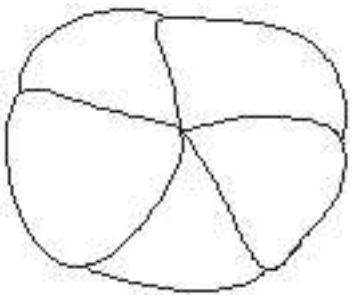


Figura 8: Vértice com mais de três países em comum.

O novo mapa tem mais um país, logo tem mais vértices do que o mapa original, mas apenas três países se encontram em cada um dos novos vértices. Assim, o vértice comum a mais de três países foi eliminado. Podemos fazer o mesmo para cada vértice do nosso mapa que tenha em comum mais de três

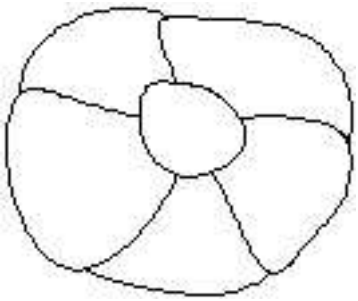


Figura 9: Criação de um novo país.

países. Dessa forma, obtemos um novo mapa com mais vértices, mas todos os vértices têm apenas três países em comum.

Agora, vamos mostrar que se cinco cores são suficientes para colorir o novo mapa, em que cada vértice é comum apenas a três países, então é possível colorir o mapa original em que um vértice podia ter mais de três países em comum.

Suponhamos que o mapa em que os vértices são comuns apenas a três países pode ser pintado nas condições do teorema (Figura 10), então, para pintar o mapa original, basta manter as mesmas cores nos países existentes desde o início e eliminar os novos países formados (Figura 11).

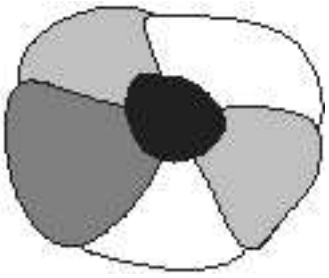


Figura 10

Este processo, não infringe as regras de dois países que se encontram num vértice, mas não ao longo de uma fronteira, poderem ter a mesma cor. Recapitulando, um vértice é um ponto em que se encontram pelo menos três países, mas já vimos que não precisamos considerar os mapas em que se encontram mais de três países em cada vértice. Portanto, temos que considerar apenas os mapas em que se encontram exactamente três países em cada vértice.

Vamos considerar o número de vértices na fronteira de cada país. Seja f_n o número de países com n vértices, com $n \in \mathbb{N}_0$ e f o número total de países.

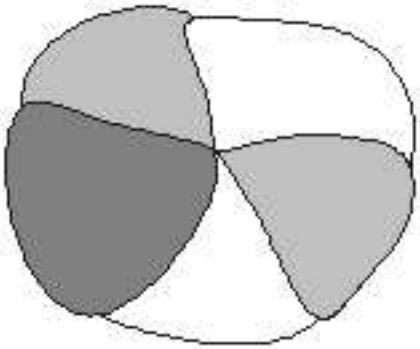


Figura 11: Mapa original já colorido.

Note-se que, se um dado país não tem vértices ou tem apenas um vértice então ele tem apenas um país vizinho e então podemos colorir o país com qualquer cor excepto com a cor do vizinho. Como estes países não causam problemas vamos deixá-los de parte e assumir que não estão presentes no resto da prova.

Temos então,

$$f = f_2 + f_3 + f_4 + \cdots \tag{1}$$

Façamos agora algumas contagens. Vamos começar por contar o números de arestas (fronteiras) existentes no nosso mapa.

Os países com 2 vértices, têm duas fronteiras, logo há $2f_2$ arestas. Os países com 3 vértices, têm três fronteiras, logo há $3f_3$ arestas. E assim sucessivamente. Esta contagem irá finalmente considerar todas as fronteiras do nosso mapa. No entanto cada uma das arestas será contabilizada duas vezes (pelos dois países que a partilham). Assim, temos

$$2a = 2f_2 + 3f_3 + 4f_4 + \cdots \tag{2}$$

Analogamente, podemos contar o número de vértices existentes no mapa. Mas, nesse caso, cada um dos vértices será contabilizado três vezes (pelos três países que o partilham). Portanto, vamos ter

$$3v = 2f_2 + 3f_3 + 4f_4 + \cdots \tag{3}$$

Das igualdades (2) e (3) vemos que

$$2a = 3v \tag{4}$$

Pela Identidade de Euler, vem

$$6v + 6f = 6a + 12$$

De (4) chegamos a

$$6f = 3v + 12$$

Substituindo f e $3v$, usando 1 e 3,

$$6(f_2 + f_3 + f_4 + \cdots) = (2f_2 + 3f_3 + 4f_4 + \cdots) + 12$$

$$4f_2 + 3f_3 + 2f_4 + f_5 = 12 + f_7 + 2f_8 + 3f_9 + \cdots \tag{5}$$

Da igualdade (5), podemos mostrar que, para todo o mapa em que apenas três países se encontram no mesmo vértice, existe pelo menos um país com menos de seis vértices. Pois se tal país não existisse, então não havia países com dois, três, quatro nem com cinco vértices, logo $f_2 = f_3 = f_4 = f_5 = 0$. Logo o membro da esquerda da igualdade (5) seria zero enquanto o da direita é maior ou igual a 12.

Agora já sabemos que no nosso mapa temos pelo menos um país com menos de 6 vértices. Poderá então ter 2, 3, 4 ou 5 vértices. Vejamos então cada um destes casos.

I. Há um país com 2 vértices.

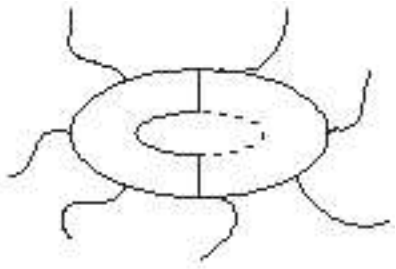


Figura 12: País com dois vizinhos.

Consideremos o país que tem dois vértices. Se tem dois vértices, então duas aresta (que unem os vértices) e portanto tem dois vizinhos. Podemos remover uma das fronteiras (Figura 12).

O novo mapa tem $f-1$ países em vez dos f países iniciais. Suponhamos que o novo mapa pode ser pintado com cinco cores, nas condições do teorema, então poderá ser assim

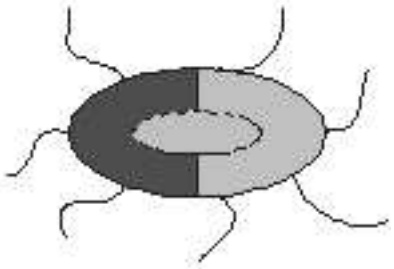


Figura 13: Eliminação de um país.

Logo ao repormos a fronteira podemos pintar o país eliminado com uma das três cores que não se utilizaram no dois países vizinhos.

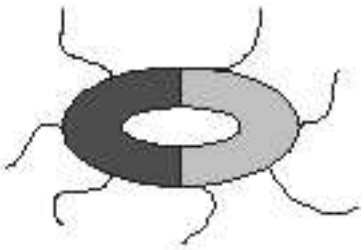


Figura 14: Reposição da fronteira.

Assim, se o mapa com $f-1$ países pode ser pintado como pretendemos, o mapa original com f países também pode.

II. Há um país com 3 vértices.

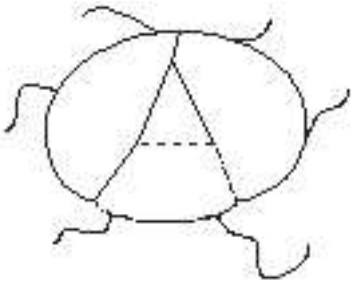


Figura 15: País com três vizinhos

Analogamente ao primeiro caso, retirámos uma das fronteiras do país que tem os três vértices. Supomos que o novo mapa com $f-1$ países pode ser pintado nas condições do teorema (Figura 16). Depois repomos

a fronteira e pintamos o país que tinha sido eliminado com uma das duas cores que não foi utilizada nos seus três vizinhos (Figura 17).

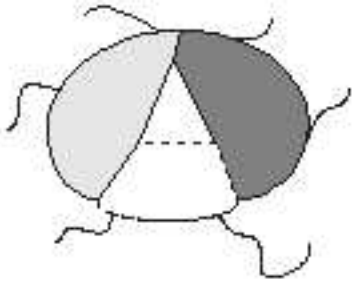


Figura 16

Desta forma, se o mapa com $f-1$ países podia ser pintado com apenas as cinco cores nas condições determinadas, então o mapa original com f países também podia.

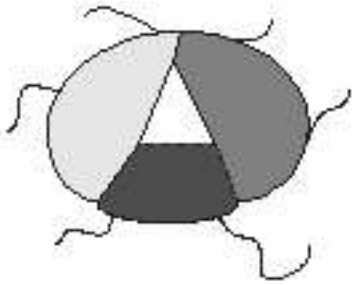


Figura 17

III. Há um país com 4 vértices.

Argumentando da mesma forma, vemos que sobra uma cor para pintar o país a que retirámos a fronteira desde que tenhamos utilizado quatro cores diferentes para os quatro países vizinhos. Mas surge uma dificuldade neste caso. Um dos países vizinhos pode ter duas fronteiras em comum com o mesmo país, como podemos ver na Figura 18.

Se removermos uma das fronteiras entre P e P_2 temos que remover a outra, pois uma fronteira não pode separar um país de si próprio. Não é esta a ideia que temos de fronteira. Assim surge um país com duas fronteiras completamente separadas e em forma de anel. Mas na demonstração da Identidade de Euler, excluímos implicitamente este caso ao supor que o sistema não se podia dividir em dois. No entanto podemos evitar a formação do anel. Se P e P_2 formam um anel então as outras duas fronteiras de P deverão pertencer a outros dois países que

estão separados pelo anel. Logo esses dois países P_1 e P_3 são diferentes e não têm nenhuma fronteira em comum, logo podem ter a mesma cor. Removemos as fronteiras que separam P de P_1 e de P_3 e obtemos um novo mapa com $f-2$ países. Supomos que este mapa com menos de f países pode ser pintado como pretendemos. Assim utilizamos uma cor para P_1 e a mesma para P_3 e outra para P_2. Ao repormos as fronteiras pintamos o país P com uma das três cores que não foram utilizadas nos seus vizinhos. Logo, podemos pintar o mapa original como pretendemos.

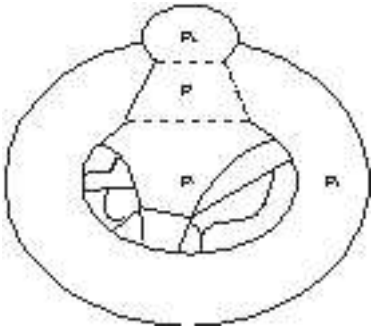


Figura 18

IV. Há um país com 5 vértices.

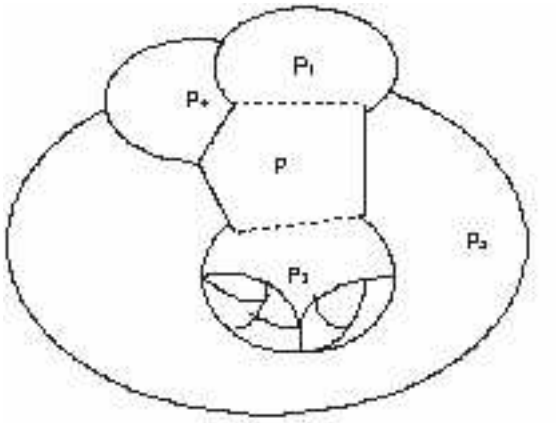


Figura 19

As mesmas dificuldades podem surgir em formas mais complicadas. Um dos países pode ter duas fronteiras em comum com P (Figura 19) ou então poderão formar-se outros anéis (Figura 20).

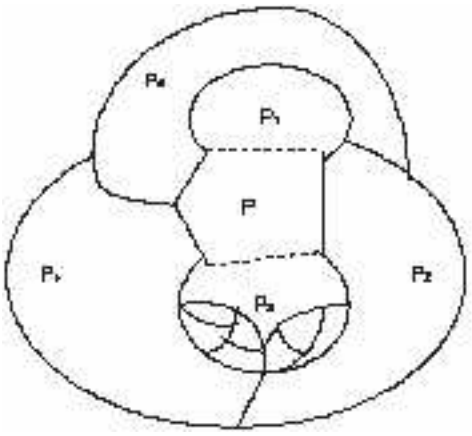


Figura 20

Em ambos os casos, como também no caso em que não há formação de anel, o país P tem dois vizinhos, P_1 e P_3, que não têm nenhuma fronteira em comum entre eles.

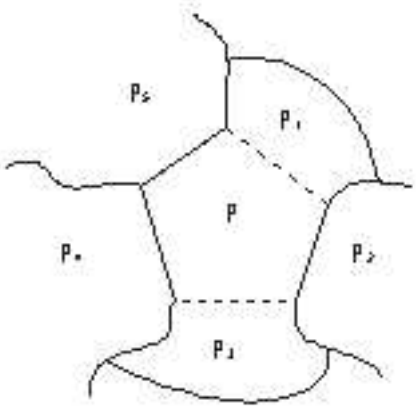


Figura 21

Removemos então as fronteiras de P com P_1 e P_3. O novo mapa que se obtém tem $f - 2$ países. Supomos que o novo mapa pode ser colorido com cinco cores nas condições pretendidas. Então, como P, P_1 e P_3 estão agora a formar apenas um país terão a mesma cor, P_2 terá outra, P_4 outra e P_5 outra. Assim já utilizámos quatro cores. Ao repormos as fronteiras podemos pintar o país P com a cor que sobrou. E assim o mapa com f países pode ser pintado obedecendo às condições do teorema.

Como vimos, todo o mapa insere-se num destes quatro casos. Logo o

processo de redução está completo. Todo o mapa pode ser reduzido a menos países, com a eliminação de uma ou duas fronteiras. E se o mapa reduzido puder ser colorido nas condições do teorema o original também pode. Assim, basta repetir o processo de redução até obtermos um mapa com apenas cinco países. E um mapa com cinco países pode, trivialmente, ser pintado nas condições do teorema. O mesmo é verdadeiro para cada mapa que se obtém ao longo do processo de redução, logo também é válido para o mapa original.

O que conclui a prova do Teorema das Cinco Cores.

3 Coloração de Mapas Noutras Superfícies

Além de mapas planos, podemos considerar mapas que estejam sobre outras superfícies como a superfície esférica, o toro, a tira de Möbius ou até mesmo a garrafa de Klein.

A prova aqui apresentada, do Teorema das Cinco Cores, também é válida para mostrar que numa superfície esférica cinco cores bastam para colorir qualquer mapa. No entanto é verdade que são necessárias apenas quatro cores para colorir qualquer mapa que esteja sobre uma superfície esférica.

A mesma prova já não é válida se pensarmos no caso do toro. Neste tipo de superfície é possível ter sete países, sendo cada um deles vizinho dos restantes seis. Neste caso, a prova apresentada para o Teorema das Cinco Cores no plano não pode ser utilizada. Há dois pontos em que a prova apresentada falha se tivermos o mapa num toro. Primeiro falha na prova da identidade de Euler em que vimos que dois caminhos entre dois pontos dividiam o plano em duas partes. A segunda falha é no estudo do Caso III e do Caso IV, em que se viu que dois países separados por um anel não têm fronteiras em comum entre eles. Num toro estas duas situações não são válidas, logo a prova apresentada para o teorema não serve. No entanto já está provado que o mesmo problema num toro se pode resolver utilizando apenas sete cores.

Se pensarmos num mapa sobre uma tira de Möbius poderemos provar que o número de cores necessárias para colorir qualquer mapa nesta superfície é seis. Sendo também seis o número de cores necessárias para colorir qualquer mapa que esteja sobre uma garrafa de Klein.

Agradecimentos

Em primeiro lugar, gostaria de agradecer à Organização do Seminário Diagonal do IST o convite e a oportunidade de apresentar o seminário e de

elaborar este artigo.

Queria também agradecer à Professora Doutora Gabriela Chaves por me ter ajudado na preparação do seminário, ao Professor Doutor Manuel Arala Chaves por me ajudar a ter disponível este trabalho na internet e ainda ao Professor Doutor Fernando Jorge Moreira a paciência e o tempo dispendidos a ajudar-me na escrita deste artigo.

Referências

[1] `http://www.fc.up.pt/atractor/matviva/geral/t5cores/`

[2] H. Rademacher, O. Toeplitz, *The Enjoyment of Mathematics*, Princeton, 1970.

[3] J.N. Silva, *O teorema das quatro cores (T4C)*, Educação e Matemática – Revista da APM **60**, Nov/Dez 2000.

[4] M. Aigner, G. Ziegler, *Proofs from the Book*, Springer.

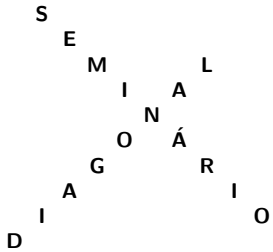
[5] M. Gardner, *Mathematical games*, Scientific American, Abril 1975.

[6] L.C. Kinsey, *Topology of Surfaces*, Springer-Verlag.

[7] `http://trends.dts.cet.pt/users/esmmat/quatrocores.htm`

[8] `http://www.cs.unb.ca/~alopez-o/math-faq/mathtext/node27.html`

[9] `http://www-groups.dcs.st-and.ac.uk/~history/HistTopics/The-four-colour-theorem.html`



Notas Sobre o Rateio de Hamilton

Geometria e Combinatória das Células Eleitorais *

João Gouveia †
1º ano de Matemática
Faculdade de Ciências e Tecnologia da Universidade de Coimbra
jsgouveia@netcabo.pt

E. Marques de Sá ‡
Departamento de Matemática da Universidade de Coimbra
emsa@mat.uc.pt

Palavras Chave

eleições, politopos, vértices, faces, simetrias, cubos martelados

Resumo

O conjunto dos resultados eleitorais que conduzem a uma certa e predeterminada composição de um parlamento chama-se *célula eleitoral*. No caso do processo mais natural de rateio, tais células são poliedros convexos. De cada célula determinamos as faces e o reticulado que elas constituem; determinamos a dimensão e a 'forma' de cada face, as simetrias e outras características geométricas. Para poupar espaço, não apresentamos demonstrações.

1 Introdução

Estas notas surgiram a propósito de um dos problemas do famoso livro *Mathematical Snapshots*, de H. Steinhaus [8]. A questão pode colocar-se assim: num processo eleitoral, três partidos, numerados 1, 2, 3, disputam ℓ lugares num parlamento; face aos resultados eleitorais, denotados por V_1, V_2, V_3, expressos em número de votos por partido — V_i votos no partido i — coloca-se

* Trabalho feito no âmbito do Programa *Novos Talentos em Matemática* da Fundação Calouste Gulbenkian, em 2000-2002.

† Como aluno do 1º ano do DMUC, apresentou este tema, em 2001, no *Seminário Diagonal* do DMUC e no *Encontro Nacional dos Novos Talentos da Matemática*, Luso 2001. O tema foi também apresentado no *Seminário Diagonal* do IST, em 2002.

‡ Professor do DMUC. Apoios da Fundação Calouste Gulbenkian, Fundação para a Ciência e a Tecnologia, e Fundação Luso-Americana para o Desenvolvimento.

o problema dito do *rateio*: quantos dos ℓ lugares devem atribuir-se a cada partido, de modo a que nenhum se sinta prejudicado?

O partido i, face à percentagem de votos que obteve (que, feitas as contas, é $\frac{V_i}{V_1+V_2+V_3}$) espera poder ocupar uma correspondente fracção dos ℓ lugares, ou seja, espera colocar no parlamento x_i candidatos seus, onde

$$x_i := \tfrac{V_i}{V_1+V_2+V_3}\,\ell\,.$$

Mas este número, a que chamamos *resultado eleitoral do partido* i, não é, em geral, inteiro; o partido i reclamará, pelo menos, um número de deputados igual à parte inteira do seu resultado x_i. Não se concebendo, por enquanto, fatiar um deputado como um queijo limiano, o problema do rateio acaba por ser o de recompor as partes fraccionárias dos x_i —chamados *resíduos* dos x_i— em um ou dois deputados inteiros e atribuí-los a um ou dois dos três partidos. De entre os muitos algoritmos para determinar quais os partidos beneficiados, vamos adoptar o mais *natural*, o chamado *método de Hamilton*,[1] que premeia com um deputado extra os partidos com os maiores resíduos. Assim, no final do rateio obtém-se uma composição parlamentar de, digamos, a_1, a_2, a_3 lugares para os partidos concorrentes. Note-se que os a_i são *inteiros* de soma ℓ, igual à dos x_i.

O problema dos *Snapshots* é, então, o seguinte: fixada certa composição parlamentar, (a_1, a_2, a_3), descreva-se a correspondente *célula*, isto é, o conjunto dos resultados eleitorais (x_1, x_2, x_3) que conduzem, pelo rateio de Hamilton, a essa composição do parlamento. Steinhaus dá uma representação 'triangular' plana dos resultados eleitorais, como na figura 1, onde cada (x_1, x_2, x_3) se representa por um ponto R de um triângulo equilátero de altura ℓ. Logo a seguir afirma, em estilo 'instantâneo', serem hexagonais as ditas células que, no seu conjunto, formam o conhecido mosaico plano dos favos de mel. Na figura 2 esboçamos o caso de eleições para um parlamento com 5 deputados; ao resultado que determinou o ponto R corresponde uma composição parlamentar de 2, 1 e 2 deputados para os partidos que tiveram as percentagens eleitorais x_1, x_2 e x_3 da figura 1. Repare-se que o mosaico de hexágonos regulares é truncado pelo domínio triangular, sobrevivendo,

1 Alexander Hamilton [1755(57?)-1804] foi figura de primeiro plano na revolução americana, no processo de independência e na elaboração da Constituição dos EUA. O seu método foi o primeiro adoptado na determinação do número de repesentantes por estado, no Congresso. Em 1880 foi detectado, teoricamente, o 'paradoxo do Alabama': se o total ℓ de lugares no Congresso passasse de 299 para 300, o rateio de Hamilton determinaria a perda de um representante do Alabama! O paradoxo ocorrera já, no concreto, em 1870: Rhode Island passou de dois representantes a um só, como consequência de o parlamento ter sido alargado de um total de 270 para 280 representantes. Estes e outros paradoxos levaram ao abandono do método em 1901, nos EUA. Veja-se [1, 2] para mais pormenores.

na periferia do triângulo, metades e sextavos de hexágonos. Por exemplo, o sextavo localizado no topo do triângulo corresponde à situação em que o partido 1 ocupa todos os lugares do parlamento.

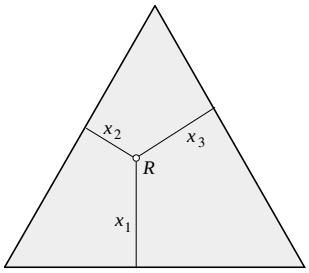


figura 1

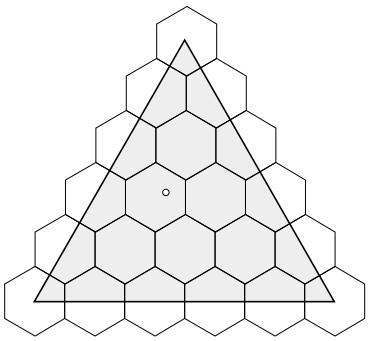


figura 2

O ponto R cairá sobre a fronteira de uma célula hexagonal sempre que ocorram certos tipos de empates nos resíduos dos x_i's. Por exemplo, quando a soma dos resíduos for 1 e os dois partidos de maiores resíduos tenham resíduos iguais, ou quando a soma dos resíduos for 2 e os dois partidos de menores resíduos tenham resíduos iguais; os vértices dos hexágonos correspondem aos casos em que os três partidos têm todos o mesmo resíduo positivo. Em situações dessas ocorrem bloqueios no preenchimento das vagas parlamentares que, no real concreto, só podem resolver-se com legislação adequada, por vezes pouco clara.[2] As eleições que mais frequentemente ocorrem são as que conduzem a resultados eleitorais pertencentes ao interior de uma célula, "longe"das fronteiras; são, no rateio, muito menos animadas que as outras, pois determinam distribuições parlamentares inequívocas ditadas pelas coordenadas do centro da única célula atingida.

Mais adiante [8, p.210], os *Snapshots* discutem com brevidade o caso de 4 partidos, modelando o problema em coordenadas 'tetraédricas': num tetraedro de referência, regular, de altura ℓ, escolhe-se uma face para cada partido; o resultado eleitoral (x_1, x_2, x_3, x_4) identifica-se com o ponto R que dista x_1 da face do partido 1, x_2 da face do partido 2, etc. Os *Snapshots* sugerem depois, *à la minute*, uma identificação algo imprecisa das células eleitorais: *"a tiling composed of regular tetrahedra and regular octahedra"*. De facto, como veremos mais adiante, as células são dodecaedros rômbicos pavimentando o espaço tridimensional de um modo simples de imaginar: pense na pavimentação do espaço com cubos sólidos todos iguais empacotados do modo óbvio; pinte-os de branco e negro 'alternadamente', de modo que cada branco fique em contacto facial com seis vizinhos negros e cada negro com seis vi-

2 Recorde-se o recente caso do Estado da Florida, para selecção dos 'grandes eleitores' nas presidenciais americanas...

zinhos brancos, como na figura 3; decomponha cada cubo branco em seis pirâmides de base quadrada com vértices no centro do cubo. Se a cada cubo negro colarmos as seis pirâmides brancas que o cercam, uma em cada face, o sólido obtido é um dodecaedro rômbico, como ilustra a figura 4.

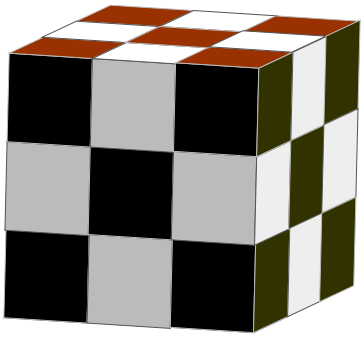


figura 3

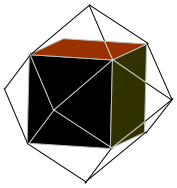


figura 4

Depois destas operações, a pavimentação cúbica transforma-se na pavimentação dodecaédrica que referimos, em que os dodecaedros são as células eleitorais. Nas figuras 5 e 6 mostramos as posições relativas do tetraedro de referência e das células eleitorais, no caso de um parlamento com 5 lugares:

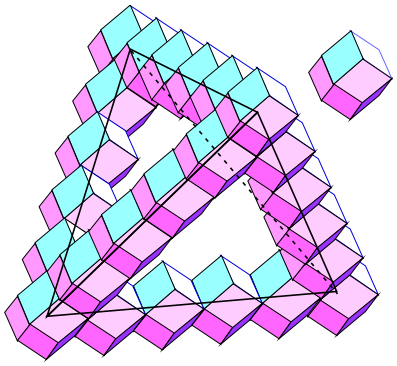


figura 5

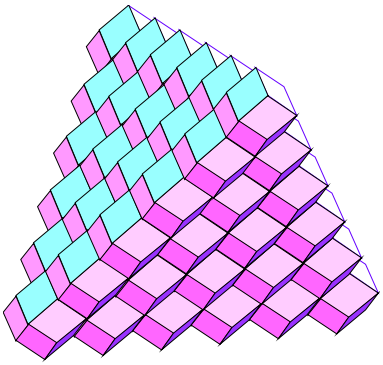


figura 6

Na figura da esquerda, o tetraedro de referência está sendo coberto por 'pequenos' dodecaedros rômbicos justapostos (note-se que cada vértice do tetraedro é centro de um dodecaedro); a figura 6 mostra o tetraedro completamente coberto por células eleitorais. Se seccionarmos este conglomerado de dodecaedros pelo plano de qualquer uma das faces do tetraedro de referência, obtemos o triângulo com favo de mel da figura 2, onde cada hexágono é uma secção plana de um dodecaedro.

Compreendido o caso de 4 partidos, a questão que naturalmente se coloca

é descrever as células eleitorais no caso geral, com p partidos em confronto. Acontece, sem grande surpresa, que as células são politopos convexos de dimensão $p-1$, todos iguais entre si, dotados de notáveis propriedades de simetria. A descrição de tais criaturas consistiu, para nós, em identificar as suas faces, rotulá-las, contá-las, determinar as dimensões e os grupos de isometrias de cada uma.[3] A forma de cada face própria depende apenas da sua dimensão δ: é um *δ-cubo martelado*, politopo que resulta de um δ-cubo por compressão ao longo de uma diagonal. O dodecaedro rômbico ilustra bem este facto: as facetas são rombos todos iguais e as arestas são do mesmo comprimento.

A descrição das células por meio de desigualdades lineares (Teorema 1) coloca-nos em terreno conhecido; em particular, mostra tratar-se de células de Voronoi particulares, frequentemente tratadas na literatura (veja-se, e.g., [3] e [4]). No entanto, não encontrámos traços das descrições da sua estrutura extremal e do seu tipo combinatório, assuntos que aqui tratamos exaustivamente, sem demonstrações, para que o texto pese o menos possível.

Dos dois autores, um orientou e outro foi orientado, mas os resultados foram obtidos por ambos, no meio de grande gozo e discussão, por nós partilhados no decorrer do primeiro ano do milénio.

2 Resultados Eleitorais

Temos então p partidos que disputam ℓ lugares de um parlamento. O total V dos votos apurados, abstenções eliminadas, decompõe-se em p parcelas, $V = V_1 + \cdots + V_p$, em que V_i é o número de votos no partido i, $i = 1, \ldots, p$. Por definição, o *resultado do partido i* é o real $x_i := \ell\, V_i/V$; o vector $x = (x_1, \ldots, x_p)$ é o *resultado das eleições.* Cada x_i pode escrever-se na forma

$$x_i = n_i + r_i\,,$$

onde n_i é a parte inteira de x_i e r_i é a sua parte fraccionária, $0 \leqslant r_i < 1$, a que chamaremos *resíduo de x_i*. A soma dos r_i's é um inteiro, que sempre denotaremos por s, tal que $0 \leqslant s < p$.

O rateio faz-se de acordo com o método de Hamilton, dos maiores resíduos, conforme explicámos atrás: cada n_i é o número de lugares imediatamente atribuídos ao partido i; nestas condições, s é o número sobrante de lugares ainda não atribuídos aos partidos; para distribuir esses lugares sobrantes,

3 Os conceitos básicos sobre convexos, politopos, faces, etc, podem ver-se nos livros [4]-[6]. O prefixo 'δ', como em *'δ-politopo'*, *'δ-cubo'*, etc, designa a dimensão do objecto em causa. As *facetas* de um δ-politopo são as suas faces de dimensão $\delta - 1$. Veja-se mais na secção 4.

ordenam-se os partidos por ordem não crescente dos resíduos e atribui-se mais um lugar a cada um dos s melhores partidos nessa ordenação. Eventuais empates nos resíduos podem implicar o bloqueio deste processo de distribuição dos s lugares sobrantes; nesses casos diremos que o resultado x é *indeterminado*; no caso oposto, o resultado diz-se *determinado*. Claro que x é determinado se, e somente se, o s-ésimo maior resíduo é estritamente superior ao $(s+1)$-ésimo maior resíduo.

Cada resultado x é um elemento de $\mathbb{R}^p$, de coordenadas racionais de soma ℓ. Mais ainda: fixada a lista de eleitores, há apenas um número finito de resultados eleitorais possíveis. Eliminamos este incómodo aceitando como resultados eleitorais exequíveis *todos* os p-uplos reais, de soma ℓ, mesmo os que têm coordenadas irracionais. Tudo vai passar-se num hiperplano de dimensão $p-1$, nomeadamente o dos x's de soma ℓ, que denotamos por Σ_ℓ.

As Coordenadas 'Simpliciais'
Fazemos aqui uma breve pausa para generalizar as coordenadas triangulares e tetraédricas dos *Snapshots*. E para concluir que isso não vai ser preciso! A generalização natural dos conceitos de triângulo equilátero e de tetraedro regular é o de *símplice regular*. Trata-se, na dimensão $p-1$ (a que nos interessa) de um politopo convexo que tem por vértices p pontos distintos e equidistantes uns dos outros. É muito aborrecido determinar p pontos nessas condições em $\mathbb{R}^{p-1}$, mas em $\mathbb{R}^p$ a coisa é trivial: para cada λ positivo, o conjunto Δ_λ constituído pelos $y \in \mathbb{R}^p$ não negativos de soma λ é um símplice regular de vértices $(\lambda, 0, \ldots, 0)$, $(0, \lambda, \ldots, 0)$, $\ldots$, $(0, 0, \ldots, \lambda)$. Note-se que o conjunto dos resultados eleitorais admissíveis constitui o símplice Δ_ℓ.

Para cada k, o conjunto Φ_k dos $y \in \Delta_\lambda$ com coordenada y_k nula é uma faceta de Δ_λ. Vamos tomar $\lambda := \ell\sqrt{1-1/p}$. O nosso símplice tem, então, altura ℓ — e é naturalmente ele que nos vai servir como *símplice de referência*. Tal como na figura 1, cada resultado eleitoral admissível, $x \in \Delta_\ell$, representa-se pelo ponto R de Δ_λ que está à distância x_k da faceta Φ_k, para $k = 1, \ldots, p$. Feitas as contas, algo cansativas, R é o vector $\sqrt{1-1/p}\cdot x$.

Quer dizer, a transformação $x \mapsto R$ é uma homotetia, contractiva, de Δ_ℓ sobre Δ_λ! Assim, tudo o que vai passar-se daqui para a frente poderá processar-se analiticamente em Δ_ℓ ou em Δ_λ, com os resultados eleitorais x ou com os seus representantes R. Os tratamentos são equivalentes e conduzem às mesmas conclusões. Preferimos trabalhar com os resultados x, em Δ_ℓ, para fugir ao incómodo factor $\sqrt{1-1/p}$.

TEOREMA 1. *Dado um resultado x, de soma ℓ, o método de Hamilton determina a distribuição $a = (a_1, \ldots, a_p)$ dos ℓ lugares no parlamento, se e só se, para todos os $i \neq j$ vale*

$$(a_i - x_i) + (x_j - a_j) < 1\,.$$

Estas p^2-p inequações e a identidade $x_1+\cdots+x_p=\ell$ determinam um convexo, aberto na topologia de Σ_ℓ. Preferimos trabalhar com o fecho deste conjunto, a que chamamos *célula-a* e denotamos por $\mathfrak{C}_a$. Claro que $x \in \mathfrak{C}_a$ sse x tem soma ℓ e, para $i \neq j$,

$$(a_i - x_i) + (x_j - a_j) \leqslant 1\,.$$

Esta descrição da célula $\mathfrak{C}_a$ permite levantar a restrição de os a_i's e x_i's serem não negativos, reduzindo tudo a mera manipulação formal. Podemos eliminar do problema partidos e deputados e trabalhar com parlamentos virtuais onde certos partidos participam com um número negativo de representantes. Efectivamente, vamos supor que o nosso parlamento tem zero lugares! Isto corresponde a fazer uma translação do hiperplano Σ_ℓ de modo a que se transforme no subespaço Σ_0 que lhe é paralelo. A situação é multiplamente delirante: há partidos com cotas positivas de presença e outros com cotas negativas que, no cômputo geral, constituem uma assembleia sem deputados, o que acaba por não fazer grande diferença para um matemático.

Em $\mathbb{R}^p$, as células são conjuntos fronteiros, sem pontos interiores. Mais interessante é a topologia de Σ_0, relativamente à qual cada célula tem uma fronteira e um interior não vazios, a que se chama fronteira e o interior *relativos* (ver [6] sobre a topologia relativa de um convexo).

A célula $\mathfrak{C}_0$ irá desempenhar um papel central no processo por estar, como facilmente se constata, centrada na origem e, por isso, ter uma representação matemática mais simples do que as outras: trata-se, de facto, do poliedro convexo de Σ_0 descrito pelas p^2-p desigualdades

$$x_j \leqslant 1 + x_i\,. \tag{1}$$

Outra vantagem de $\mathfrak{C}_0$ é poder caracterizar-se de modo simples através de resíduos, do seguinte modo:

Teorema 2. *Um p-uplo x, de soma nula, pertence a $\mathfrak{C}_0$ se e só se é possível reordenar as suas coordenadas de modo a obter um vector da forma*

$$(r_1, r_2, \ldots, r_p) - (\underbrace{1, \ldots, 1}_{s}, 0, \ldots, 0)\,, \tag{2}$$

onde s denota o número de coordenadas negativas de x e os r_i's são reais tais que $1 > r_1 \geqslant \cdots \geqslant r_p \geqslant 0$. A representação (2) é unicamente determinada por x.

Um ponto x de $\mathfrak{C}_0$ pertence ao interior relativo de $\mathfrak{C}_0$ sse $r_s > r_{s+1}$;

Um ponto x de $\mathfrak{C}_0$ pertence à fronteira relativa de $\mathfrak{C}_0$ sse $r_s = r_{s+1}$.

Facilmente se demonstra que a intersecção de duas células distintas tem interior relativo vazio, que Σ_0 é a união de todas as suas células e que toda a célula se obtém por translação de $\mathfrak{C}_0$, mais precisamente

$$\mathfrak{C}_a = a + \mathfrak{C}_0 \,.$$

Estamos, pois, perante um mosaico de Σ_0, obtido por justaposição de ladrilhos todos iguais a $\mathfrak{C}_0$, dispostos de forma regular, centrados nos pontos inteiros de Σ_0. A partir de (1) é fácil mostrar que, para cada resultado x (determinado) o processo de rateio de Hamilton selecciona, de entre os pontos inteiros a de soma ℓ, o mais próximo de x na métrica euclidiana (veja-se [7, p.248]). Dito de outro modo, em nomenclatura consagrada, as células de Σ_0 constituem o mosaico das *células de Voronoi* determinado pela 'rede' dos pontos inteiros de Σ_0 (e.g., [3, 4]).

A figura 7 ilustra, na dimensão 2, uma propriedade muito curiosa das células de Voronoi determinadas pelo rateio de Hamilton. Rebatido o plano Σ_0 sobre o plano do papel, os seus pontos inteiros são os nodos de uma 'rede triangular'. Imaginemos pequenos círculos de borracha, todos iguais, centrados nos nodos da rede (a). Façamos aumentar os raios desses círculos, mantendo-os iguais entre si, até se tocarem (b) e passarem a deformar-se (c), em competição pelo espaço disponível. No limite obtemos o mosaico (d) do favo de mel.

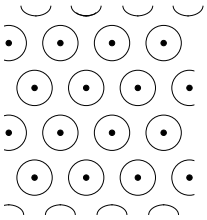


figura 7a

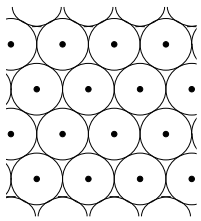


figura 7b

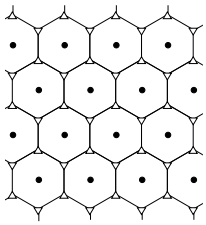


figura 7c

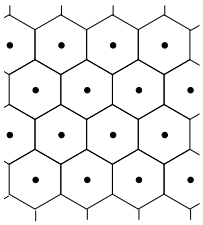


figura 7d

O mesmo se pode imaginar no caso de 4 partidos: colocam-se balões de borracha, todos iguais, centrados nos centros dos cubos negros da figura 3 (que se supõe estendida a todo o espaço); sopram-se os balões, mantendo-os iguais, até que se toquem e a seguir se deformem, atingindo, no limite, a forma de dodecaedros rômbicos.

3 Os Vértices de $\mathfrak{C}_0$

No caso $p = 3$ não é difícil determinar os vértices do hexágono $\mathfrak{C}_0$. No favo de mel, os vértices são aqueles resultados eleitorais indeterminados em que o

equilíbrio dos resíduos é total: uma pequena perturbação poderá fazer cair o resultado para uma de *três* células vizinhas, e qualquer dos três partidos poderá ver eleito um ou nenhum deputado extra. Trata-se, pois, dos resultados em que os partidos obtêm resíduos todos iguais. Em $\mathfrak{C}_0$, há 6 trios (x_1, x_2, x_3) nessas condições, que são as 6 linhas das matrizes:

$$\begin{bmatrix} -\frac{2}{3} & \frac{1}{3} & \frac{1}{3} \\ \frac{1}{3} & -\frac{2}{3} & \frac{1}{3} \\ \frac{1}{3} & \frac{1}{3} & -\frac{2}{3} \end{bmatrix}, \begin{bmatrix} \frac{2}{3} & -\frac{1}{3} & -\frac{1}{3} \\ -\frac{1}{3} & \frac{2}{3} & -\frac{1}{3} \\ -\frac{1}{3} & -\frac{1}{3} & \frac{2}{3} \end{bmatrix}.$$

Não é fácil estender este tipo de argumentação a $\mathbb{R}^p$, mas colhe-se a conjectura de que *os vértices de* $\mathfrak{C}_0$ *são os seus pontos de resíduos todos iguais*, que se revelou correcta. Vamos mostrar porquê. Por definição, $v \in \mathfrak{C}_0$ um *vértice* de $\mathfrak{C}_0$ se, para qualquer vector 'perturbador' $\varepsilon \neq 0$, uma das perturbações $v+\varepsilon$ ou $v-\varepsilon$ não pertence a $\mathfrak{C}_0$. Se v tem dois resíduos distintos, consegue-se construir um ε tal que: $v+\varepsilon$ e $v-\varepsilon$ ambos pertencem a $\mathfrak{C}_0$, o que prova a conjectura feita.

Por exemplo, no caso $p = 4$, $\mathfrak{C}_0$ tem 14 vértices, que são as linhas das matrizes

$$\begin{bmatrix} -\frac{3}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \\ \frac{1}{4} & -\frac{3}{4} & \frac{1}{4} & \frac{1}{4} \\ \frac{1}{4} & \frac{1}{4} & -\frac{3}{4} & \frac{1}{4} \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & -\frac{3}{4} \end{bmatrix}, \begin{bmatrix} -\frac{1}{2} & -\frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} & -\frac{1}{2} & \frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} & \frac{1}{2} & -\frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} & \frac{1}{2} & -\frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} \end{bmatrix}, \begin{bmatrix} \frac{3}{4} & -\frac{1}{4} & -\frac{1}{4} & -\frac{1}{4} \\ -\frac{1}{4} & \frac{3}{4} & -\frac{1}{4} & -\frac{1}{4} \\ -\frac{1}{4} & -\frac{1}{4} & \frac{3}{4} & -\frac{1}{4} \\ -\frac{1}{4} & -\frac{1}{4} & -\frac{1}{4} & \frac{3}{4} \end{bmatrix}.$$

Observadas com cuidado as regularidades, podemos dizer que $\mathfrak{C}_0$ tem por vértices os p-uplos

$$\left(\tfrac{s}{p}, \tfrac{s}{p}, \ldots, \tfrac{s}{p}\right) - (\underbrace{1, \ldots, 1}_{s}, 0, \ldots, 0), \qquad \text{para } 0 < s < p,$$

e todos os que destes se obtêm permutando coordenadas. Dito de um modo mais elegante: cada subconjunto S de $\{1, \ldots, p\}$, próprio e não vazio, determina um vértice, denotado $\mathcal{V}_S$ ou $\mathcal{V}(S)$, de coordenada i dada por

$$\left[\,\mathcal{V}_S\,\right]_i := \begin{cases} s/p - 1, & \text{se } i \in S \\ s/p, & \text{se } i \notin S\,, \end{cases}$$

onde s denota o cardinal de S; e todo o vértice de $\mathfrak{C}_0$ é deste tipo. Há, pois, um total de $2^p - 2$ vértices. Uma propriedade muito interessante dos vértices é a seguinte, onde S e T são subconjuntos arbitrários de $\{1, \ldots, p\}$:

$$\mathcal{V}_S + \mathcal{V}_T = \mathcal{V}_{S \cup T} + \mathcal{V}_{S \cap T} \tag{3}$$

onde se convenciona $\mathcal{V}_S = 0$ no caso de [o complementar de] S ser vazio.

4 As Faces de $\mathfrak{C}_0$

Sempre que nos seja dado um convexo C, imediatamente ocorre o problema da determinação das suas faces. Dizemos que um subconjunto F de C é uma *face* de C, se F é convexo e se todo o segmento aberto $]a, b[$ contido em C que intersecta F está totalmente contido em F.

Quando C é um politopo, as suas faces são, também, politopos. Por exemplo, as faces de um politopo de $\mathbb{R}^3$ são os vértices, as arestas, as 'faces' propriamente ditas, de forma poligonal, e ainda o conjunto vazio e o próprio politopo.

Em geral, as faces das faces de C são faces de C. Em particular, cada face tem os seus vértices, que são vértices de C. O problema da identificação das faces ficará resolvido se determinarmos quais são os conjuntos de vértices de C que são vértices de uma mesma face. Por exemplo, as faces de dimensão 1 são as 'arestas' de C, colocando-se, então, a seguinte questão subsidiária: quais os pares de vértices de C que estão ligados por uma aresta de C?

Se C é um poliedro delimitado por uma família finita de hiperplanos (quer dizer, C é a intersecção de semi-espaços delimitados por esses hiperplanos) as suas faces próprias são os conjuntos que resultam da intersecção de C com alguns dos hiperplanos dessa família. Tendo o nosso politopo $\mathfrak{C}_0$ como hiperplanos limítrofes Σ_0 e os hiperplanos $\mathcal{H}_{ij}$, de equação $x_j = 1 + x_i$, para $i \neq j$, as suas faces são os conjuntos da forma

$$\mathfrak{F} := \mathcal{H}_{i_1 j_1} \cap \cdots \cap \mathcal{H}_{i_t j_t} \cap \mathfrak{C}_0 \,. \tag{4}$$

Esta face é própria sse $t \geqslant 1$. Dada $\mathfrak{F}$, definamos os conjuntos de índices

$$I := \{i_1, \ldots, i_t\} \quad \text{e} \quad J := \{j_1, \ldots, j_t\} \,.$$

Podemos provar que $\mathcal{V}_S$ é vértice de $\mathfrak{F}$ sse $S \supseteq I$ e $S \cap J = \varnothing$. Isto implica que a face em causa apenas depende dos conjuntos I e J, pelo que a denotaremos por $\mathfrak{F}_{IJ}$, ou $\mathfrak{F}(I, J)$.

TEOREMA 3. *A face $\mathfrak{F}_{IJ}$ é própria e não vazia se e só se I e J são disjuntos e não vazios; ela tem dimensão $\delta := p - |I| - |J|$; e tem 2^δ vértices.*

$\mathfrak{F}_{I'J'} \subseteq \mathfrak{F}_{IJ}$ se e só se: $I' \supseteq I$ e $J' \supseteq J$. $(I, J) \mapsto \mathfrak{F}_{IJ}$ é uma bijecção do conjunto dos pares (I, J), com I, J não vazios e disjuntos, sobre o das faces próprias e não vazias de $\mathfrak{C}_0$.

A principal dificuldade da demonstração está no cálculo da dimensão de $\mathfrak{F}_{IJ}$, por envolver representações (4) com t mínimo, ou seja, com hiperplanos linearmente independentes. A segunda parte do teorema, que descreve o reticulado das faces de $\mathfrak{C}_0$, depende desse cálculo dimensional. O teorema

permite uma contagem simples do número de faces próprias e não vazias de $\mathfrak{C}_0$, que é $3^p - 2^{p+1} + 1$, e do número de faces próprias de dimensão δ que é:

$$f_\delta = \binom{p}{\delta}\left[2^{p-\delta} - 2\right]. \tag{5}$$

De entre as consequências do teorema, destacamos a determinação das arestas e do *grau* de cada vértice, que é o número de arestas nele incidentes.

COROLÁRIO 4. *Dois vértices, $\mathcal{V}_S$ e $\mathcal{V}_T$, são os extremos de uma mesma aresta, se e só se um dos conjuntos, S ou T, resulta do outro por adjunção de um índice. Há $2p$ vértices de grau $p-1$, nomeadamente $\mathcal{V}_{\{1\}}, \ldots, \mathcal{V}_{\{p\}}$ e os seus opostos; todos os outros têm grau p.*

5 A Forma das Faces Próprias

Nesta secção, a notação X^c designa o complementar de X relativamente a $\{1, \ldots, p\}$. Vamos apresentar características geométricas das faces próprias de $\mathfrak{C}_0$, que conduzirão, em particular, à determinação da sua forma. Dito em poucas palavras, tudo se dirige à demonstração do seguinte resultado:

> *As faces de certa dimensão δ são geometricamente iguais entre si. Cada uma delas é um δ-cubo 'martelado', mais precisamente, resulta de um δ-cubo de aresta* 1 *por uma compressão afim ao longo de uma diagonal interna, com factor de compressão $\sqrt{1-\delta/p}$.*

5.1 As Faces Próprias são Paraleletopos

Fixemos uma face própria $\mathfrak{F}_{IJ}$. Para cada K disjunto de $I \cup J$, define-se $K^* := (I \cup J \cup K)^c$; os pares $\{\mathcal{V}(I \cup K)\,,\, \mathcal{V}(I \cup K^*)\}$ constituem uma partição dos vértices de $\mathfrak{F}_{IJ}$. Pode provar-se que os segmentos

$$\left[\mathcal{V}(I \cup K)\,,\, \mathcal{V}(I \cup K^*)\right] \tag{6}$$

têm o mesmo ponto médio (independente de K), que é $c_{IJ} := [\mathcal{V}_I - \mathcal{V}_J]/2$. Isto prova que a face $\mathfrak{F}_{IJ}$ tem c_{IJ} por centro de simetria. Por essa razão, diremos que (6) é *diagonal interna* da face $\mathfrak{F}_{IJ}$; cada diagonal passa pelo centro c_{IJ} e une vértices opostos da face. O comprimento do segmento (6) é $\sqrt{\delta - (\delta - 2|K|)^2/p}$. Tomando K vazio, obtemos $\sqrt{\delta(1-\delta/p)}$ como mínimo comprimento das diagonais de uma face de dimensão δ e a menor diagonal interna é a que une os vértices $\mathcal{V}_I$ e $\mathcal{V}_{J^c}$.

As facetas de $\mathfrak{F}_{IJ}$ são as faces de uma das seguintes formas: $\mathfrak{F}(I \cup \{t\}, J)$ e $\mathfrak{F}(I, J \cup \{t\})$, onde t percorre o complementar de $I \cup J$. Para cada t,

estas duas facetas são mutuamente opostas a respeito do centro de simetria da face $\mathfrak{F}_{IJ}$, cada uma delas é translação da outra, e a face $\mathfrak{F}_{IJ}$ é um tronco de prisma que tem as ditas facetas por bases e $\mathcal{V}_{\{t\}}$ por vector gerador. Isto pode obter-se por uso sistemático das fórmulas (3). Em símbolos:

$$\mathfrak{F}(I, J \cup \{t\}) = \mathcal{V}_{\{t\}} + \mathfrak{F}(I \cup \{t\}, J)\,. \tag{7}$$

$$\mathfrak{F}_{IJ} = \mathfrak{F}(I \cup \{t\}, J) + [0, \mathcal{V}_{\{t\}}] \tag{8}$$

Recorde-se que um *δ-paraleletopo* é um convexo que resulta da soma conjuntista de δ segmentos, $[0, v_1], \ldots, [0, v_\delta]$, com extremos linearmente independentes, seguida de eventual translação. Acrescentando às fórmulas (7)-(8) um argumento indutivo, podemos concluir que $\mathfrak{F}_{IJ}$ *é um δ-paraleletopo.* Mais explicitamente, para qualquer K contido no complementar de $I \cup J$, $\mathfrak{F}_{IJ}$ é, após translação, a soma das arestas de $\mathfrak{F}_{IJ}$ emergentes do vértice $\mathcal{V}(I \cup K)$. Em particular, para K vazio, temos:

$$\mathfrak{F}_{IJ} = \mathcal{V}_I + \sum_{t \in (I \cup J)^c} [0, \mathcal{V}_{\{t\}}]. \tag{9}$$

5.2 Alongamento de uma Face

Fixada uma face $\mathfrak{F}_{IJ}$ vamos considerar o *alongamento-θ segundo a sua diagonal menor* $[\mathcal{V}_I, \mathcal{V}_{J^c}]$. Trata-se, por definição, da aplicação linear que fixa todos os vectores ortogonais a essa diagonal e que transforma o vector $D := \mathcal{V}_{J^c} - \mathcal{V}_I$ em θD, onde θ é uma constante $\geqslant 1$. Dado um $x \in \mathbb{R}^p$, a imagem x' pelo dito alongamento é dada por:

$$x' = x + (\theta - 1)\left(\frac{x \cdot D}{\|D\|^2}\right) D\,.$$

Vamos escolher para θ o valor $1/\sqrt{1 - \delta/p}$. Com cálculos algo complicados podemos provar o teorema seguinte que estabelece a forma da face em estudo:

TEOREMA 5. *O alongamento-θ transforma os* $\mathcal{V}_{\{t\}}$ *que ocorrem em* (9) *num sistema ortonormado. Portanto,* $\mathfrak{F}_{IJ}$ *transforma-se num δ-cubo de lado* 1.

Note-se que cada face $\mathfrak{F}_{IJ}$ pode alongar-se de muitos modos para se transformar num cubo unitário. Escolhe-se uma qualquer face $\mathfrak{F}_{RS}$ que contenha $\mathfrak{F}_{IJ}$ e alonga-se $\mathbb{R}^p$ ao longo da diagonal menor de $\mathfrak{F}_{RS}$, com parâmetro $\theta = 1/\sqrt{1 - \epsilon/p}$, onde ϵ é a dimensão de $\mathfrak{F}_{RS}$. $\mathfrak{F}_{RS}$ transforma-se num ϵ-cubo unitário, pelo que todas as suas faces, incluída $\mathfrak{F}_{IJ}$, se transformam em cubos unitários de dimensões correspondentes.

6 Os Automorfismos de $\mathfrak{C}_0$

Um *automorfismo* de $\mathfrak{C}_0$ é um operador linear de Σ_0 que transforma $\mathfrak{C}_0$ em si mesmo. O conjunto de todos os automorfismos de $\mathfrak{C}_0$ constitui um grupo que denotamos por $\operatorname{Aut}\mathfrak{C}_0$.

Por exemplo, como o sistema de desigualdades (1) é invariante por permutação das coordenadas $x_1, \ldots, x_p$, cada permutação determina e pode identificar-se com um automorfismo de $\mathfrak{C}_0$; por razão análoga, a simetrização, $x \mapsto -x$, é um automorfismo de $\mathfrak{C}_0$. Pode provar-se que

TEOREMA 6. $\operatorname{Aut}\mathfrak{C}_0$ *é constituído pelas aplicações de permutação e pelas suas simétricas.*

Vamos apresentar um esqueleto de demonstração. O Corolário 4 diz que $\pm\mathcal{V}_{\{1\}}, \ldots, \pm\mathcal{V}_{\{p\}}$ são os vértices de grau $p-1$ de $\mathfrak{C}_0$. Designemos por $\mathbb{H}$ o invólucro convexo destes vértices. De entre eles, os p que vão com sinal '+' são os vértices de um símplice $\mathbb{S}$, regular e centrado na origem; os outros são os vértices do oposto $-\mathbb{S}$. Como um automorfismo de $\mathfrak{C}_0$ preserva graus, todo o automorfismo de $\mathfrak{C}_0$ é automorfismo $\mathbb{H}$. Prova-se, depois, que um automorfismo de $\mathbb{H}$ transforma $\mathbb{S}$ em $\mathbb{S}$, ou $\mathbb{S}$ em $-\mathbb{S}$. O teorema resulta de os automorfismos de um símplice regular se poderem identificar com as permutações dos seus vértices.

7 Ilustração e Exemplos

Os resultados das secções anteriores sobre as dimensões, o número e formato das faces, os graus dos vértices, etc, permitem retirar a conclusão, há muito intuida, de que $\mathfrak{C}_0$ é um hexágono regular para $p = 3$, e um dodecaedro rômbico para $p = 4$. Quando $p = 5$, $\mathfrak{C}_0$ é um politopo de dimensão 4; de acordo com as fórmulas (5), tem 20 facetas que são cubos 'martelados', 60 faces bidimensionais que são quadrados 'martelados', 60 arestas e 30 vértices. Não sabemos se tal criatura e as que se lhe seguem nas dimensões maiores têm nome de baptismo.

Vamos ver o que nos diz o Teorema 5 sobre a forma e medidas das faces *próprias* de $\mathfrak{C}_0$. Note-se, para já, que um δ-cubo de aresta 1 tem diagonais todas de comprimento $\sqrt{\delta}$. Fazendo-se a compressão segundo uma diagonal do δ-cubo, com factor $\sqrt{1-\delta/p}$, uma boa maneira de imaginar as faces do politopo resultante é ter em conta as diagonais explicitadas em (6), cujos comprimentos são, conforme dissemos,

$$\sqrt{\delta - (\delta - 2k)^2/p}\,, \quad \text{para } k = 0, 1, \ldots, \delta\,. \tag{10}$$

Daqui resulta, em particular, que as arestas de $\mathfrak{C}_0$ têm todas o mesmo comprimento: $\sqrt{1-1/p}$. Cada face de dimensão 2 resulta de um quadrado de lado 1 comprimido segundo uma diagonal, isto é, um rombo com diagonal maior $\sqrt{2}$ (a diagonal inalterada do quadrado original) e diagonal menor $\sqrt{2}\sqrt{1-2/p}$. Cada face de dimensão 3 (para $p > 4$) é um cubo 'martelado' com diagonais de dois tamanhos: a diagonal menor mede $\sqrt{3-9/p}$ e as outras três diagonais medem $\sqrt{3-1/p}$. De um modo geral, os $\delta+1$ valores possíveis de k produzem, em (10), $\lfloor \delta/2 \rfloor + 1$ comprimentos possíveis para as diagonais. Na figura 8 representamos em perspectiva as faces tridimensionais para os valores de p: 5, 6 e 20; para cada um destes valores mostramos apenas um exemplar de um total de 20, 120 e 149419800 cubos 'martelados', respectivamente. Na segunda linha figuram as faces bidimensionais para os valores de p indicados. Para $p = 4$, a face tridimensional de $\mathfrak{C}_0$ é o dodecaedro rômbico com facetas iguais ao rombo representado.

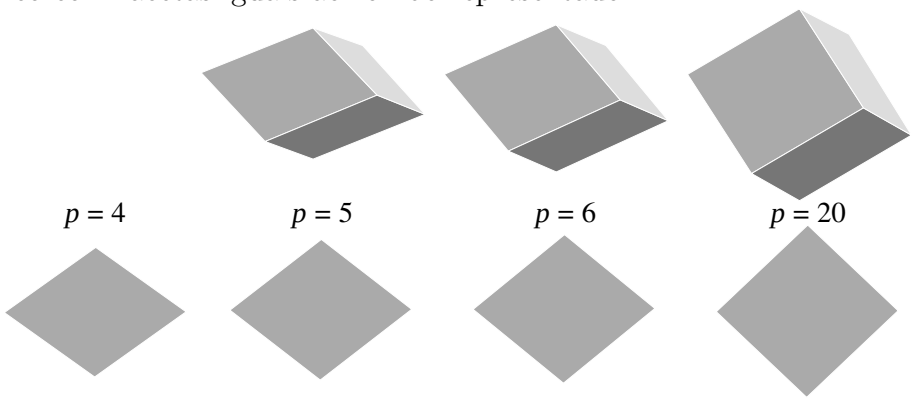


figura 8

As compressões foram feitas segundo diagonais 'verticais'. Para $p = 20$, o factor de compressão do cubo é de cerca de 92%, o que dá um sólido muito próximo do cubo original, cujas facetas são todas iguais ao rombo da linha de baixo, com diagonal menor cerca de 95% da diagonal maior. De facto, é óbvio que a sequência de faces de dimensão δ, que se obtém fazendo sucessivamente $p = \delta+1, \delta+2, \delta+3, \ldots$, tem por 'limite' um δ-cubo de aresta 1.

Vamos agora mostrar, para dimensões pequenas, o modo como os politopos $\mathbb{S}$ e $\mathbb{H}$, definidos na secção 6, se posicionam dentro de $\mathfrak{C}_0$.

No caso $p = 3$, $\mathfrak{C}_0$ é um hexágono regular assente no plano $x_1+x_2+x_3 = 0$; $\mathbb{S}$ é um triângulo equilátero e $\mathbb{H}$ é, por isso, o próprio hexágono, conforme mostra a figura:

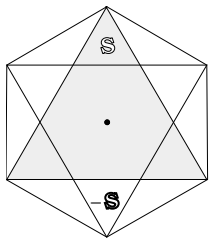


figura 9

De acordo com o Teorema 6, cada automorfismo do hexágono é determinado por uma permutação dos vértices do triângulo $\mathbb{S}$, seguida, ou não, de uma simetrização relativamente à origem. Por exemplo, uma rotação de 60^{o} (em torno da origem) é determinada por uma permutação cíclica dos vértices de $\mathbb{S}$ — ou seja, por uma rotação de 120^{o} — seguida de simetrização.

No caso $p = 4$, $\mathbb{S}$ e $-\mathbb{S}$ são tetraedros regulares e $\mathbb{H}$ é um cubo como mostra a figura 10. Os vértices do cubo são os vértices de grau 3 do dodecaedro $\mathfrak{C}_0$. Os outros 6 vértices do dodecaedro, todos de grau 4, determinam um octaedro regular, o dual do cubo! Isto faz pensar em algo mais, como a determinação de relações de dualidade existentes entre 'subpolitopos' notáveis de $\mathfrak{C}_0$ em dimensões elevadas. A figura 11 mostra a situação, ilustrando, de algum modo, o facto de os automorfismos do dodecaedro serem exactamente os mesmos que os do cubo $\mathbb{H}$.

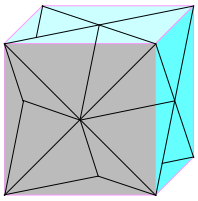


figura 10

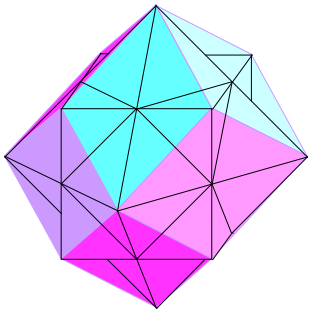


figura 11

8 Células Vizinhas

Nesta secção tratamos as relações de $\mathfrak{C}_0$ com as células suas vizinhas: quantas e quais as células que tocam $\mathfrak{C}_0$?, de que modo se tocam?, etc.

Claro que nem todas as células intersectam $\mathfrak{C}_0$. De facto, a célula-a intersecta $\mathfrak{C}_0$ se e só se a é de soma nula e tem coordenadas 0, 1 e -1.

Toda a célula intersecta $\mathfrak{C}_0$ segundo uma face comum às duas células,

mais precisamente: $\mathfrak{C}_a \cap \mathfrak{C}_0$ é a face $\mathfrak{F}_{PN}$ de $\mathfrak{C}_0$, onde

$$P := \{i : a_i = 1\} \quad \text{e} \quad N := \{j : a_j = -1\}\,.$$

A dimensão desta face de contacto é o número de coordenadas nulas de a.

Também não é difícil mostrar que uma face $\mathfrak{F}_{IJ}$ de $\mathfrak{C}_0$ é também face de uma célula vizinha se e só se $|I| = |J|$. Disto resulta que a intersecção de duas células tem dimensão da mesma paridade que p. Assim, nem todas as faces de $\mathfrak{C}_0$ resultam da intersecção de $\mathfrak{C}_0$ com uma célula vizinha. Por exemplo, no caso $p = 3$ (de células hexagonais), duas células, quando se intersectam, fazem-no segundo uma aresta, mas não segundo um vértice. No caso $p = 4$ (de células dodecaédricas) duas células, quando se intersectam, fazem-no segundo um rombo ou um vértice — nunca segundo uma aresta.

A partir destes resultados simples, torna-se fácil fazer algumas contagens. Por exemplo, a face $\mathfrak{F}_{IJ}$ é comum a $\binom{|I|+|J|}{|I|}$ células $\mathfrak{C}_a$ distintas ($\mathfrak{C}_0$ incluída). E o número de células $\mathfrak{C}_a$ distintas que intersectam $\mathfrak{C}_0$ segundo uma face de dimensão $\delta = p - 2k$ é

$$\binom{p}{k}\binom{p-k}{k}.$$

Referências

[1] A. Bogomolny, *The Constitution and Paradoxes*, January 2002, `http://www.maa.org/editorial/knot/Democracy.html`

[2] L. Bowen, *Introduction to Contemporary Mathematics*, University of Alabama, 1998-2001, `http://www.ctl.ua.edu/math103/`

[3] J. Conway, N. Sloane, *Sphere Packings, Lattices and Groups*, Springer Verlag, 1988.

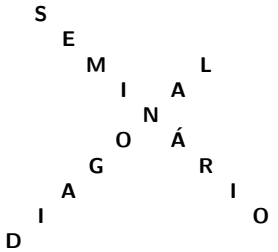
[4] P. Gruber, J. Wills (Eds.), *Handbook of Convex Geometry*, Vols A & B, North-Holland, 1993.

[5] B. Grünbaum, *Convex Polytopes*, Interscience Publ., John Wiley & Sons, 1967.

[6] R.T. Rockafellar, *Convex Analysis*, Princeton University Press, 1970.

[7] D. Saari, *Basic Geometry of Voting*, Springer, 1995.

[8] H. Steinhaus, *Mathematical Snapshots*, G.E. Stechert, New York 1939;
para facilitar a tarefa do leitor interessado, as citações que fazemos neste artigo referem-se a uma edição recente, nomeadamente, a Third American Edition, Revised and Enlarged, Dover Publ., 1999.



Cadeias de Markov e Polícias

Paulo Varandas
4º Ano de Matemática
Faculdade de Ciências da Universidade do Porto
pcvarand@hotmail.com

Palavras Chave

matriz de transição, vector posição, probabilidades, matriz regular, cadeia de Markov, vector próprio, valor próprio

Resumo

Um polícia foi destacado para controlar vários cruzamentos. Foi-lhe ordenado que, ao fim de um certo tempo num dado cruzamento, passe de forma aleatória para um dos cruzamentos vizinhos. Caso seja possível, quer-se determinar a probabilidade de encontrar o polícia num dado cruzamento ao fim de algum tempo. Para tal, usaremos cadeias de Markov.

Introdução

Como é usual em certos problemas da vida real em que se possui um sistema físico com vários objectos e estados, é possível determinar as condições em que há transições desses objectos de um estado para um outro, sem todavia ser possível determinar o comportamento assimptótico desse mesmo sistema, ou seja, prever o que acontecerá ao fim de "muito tempo"[1].

Neste artigo pretende-se ilustrar, com a ajuda de um exemplo bastante prático e no qual é possível determinar as referidas condições de transição entre estados, que sob certas condições é de facto possível determinar o que acontece no sistema.

Há que notar que nem sempre é possível determinar com tanta exactidão como nos exemplos que vão ser apresentados o que acontecerá no sistema físico, sendo explanados em [3] resultados análogos ao apresentado, mas sob condições mais fracas.

1 Por "muito tempo" deve-se entender o limite quando o tempo tende para o infinito.

1 O Problema

Um polícia foi destacado para controlar oito cruzamentos, como é ilustrado pela figura (Figura 1).

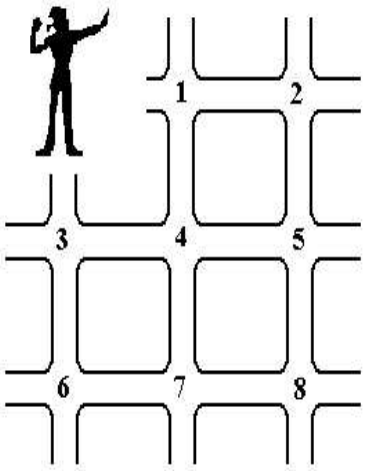


Figura 1

Foi-lhe ordenado que, ao fim de um certo tempo num dado cruzamento, passe de forma aleatória para um dos cruzamentos vizinhos. Aqui considera-se que cada cruzamento é vizinho de si próprio.

A questão que se coloca é a de saber se esta estratégia da distribuição do tempo do polícia pelos vários cruzamentos é a mais adequada, isto é, se o polícia passará igual tempo nos vários cruzamentos.

Outras questões que se podem levantar são acerca da configuração dos cruzamentos em causa, e de que forma (caso esta distribuição do tempo do polícia não seja a mais adequada) se poderá optimizar o serviço do polícia.

1.1 As Ferramentas Matemáticas

Com o objectivo de resolver este problema, introduzam-se os conceitos que vão permitir passar este problema da vida real para um problema de matemática.

Definição. Suponha-se que se tem um sistema físico que, a cada momento, possui um número finito de possibilidades (ou estados). Se a probabilidade de um certo estado ocorrer num instante futuro puder ser determinada sabendo apenas o presente estado, este processo diz-se uma *cadeia de Markov*.

A uma cadeia de Markov que possua n estados, designados por $1, 2, 3, \ldots, n$, pode-se associar uma matriz $n \times n$ que se diz *matriz de transição da cadeia de Markov* e é dada por:

$$\begin{pmatrix} p_{11} & p_{12} & \cdots & p_{1n} \\ p_{21} & p_{22} & \cdots & p_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ p_{n1} & p_{n2} & \cdots & p_{nn} \end{pmatrix} \tag{1}$$

onde p_{ij} é a probabilidade de transição do estado j para o estado i.

Há que notar que, se P é uma matriz de transição como em (1) de uma cadeia de Markov com n estados, então:

$$p_{1j} + p_{2j} + \cdots + p_{nj} = 1$$

pelo que P também se diz uma *matriz de probabilidade*. Será ainda útil a introdução dos seguintes conceitos.

Definição. Um vector coluna x cuja i-ésima componente, x_i, é a probabilidade de se estar no i-ésimo estado do sistema físico, diz-se *vector posição*, ou vector de probabilidade.

Definição. Uma matriz de transição P como em (1) diz-se *regular* se existir alguma potência de P que não tenha entradas nulas.

1.2 Outros Exemplos

Com o objectivo de facilitar a introdução destes conceitos e permitir uma familiaridade com os mesmos, introduzem-se dois exemplos mais simples.

Exemplo 1. Uma escola secundária faz, no final de cada ano lectivo, um peditório para as crianças mais carenciadas. Verificando os registos de vários anos, conclui-se que:

(a) 80% dos alunos que contribuem para o peditório contribuem no ano seguinte;

(b) 30% dos alunos que não contribuem um ano, contribuem no ano seguinte.

Esta situação pode ser traduzida por uma cadeia de Markov com dois estados:

1. aluno contribuiu no ano

2. aluno não contribuiu no ano

à qual se associa a matriz de transição:

$$\begin{pmatrix} .8 & .3 \\ .2 & .7 \end{pmatrix}.$$

Exemplo 2. Alguns proprietários de apartamentos contrataram uma empresa de administração de propriedades com reputação de melhorar as condições das habitações sob o seu controlo. Usando por classificações *má*, *média* e *excelente* para as condições dos apartamentos arrendados, sabe-se de fonte fiável que 10% dos apartamentos que apresentam condições más continuam em condições más, 50% são melhorados para condições médias e os restantes 40% são renovados para condições excelentes. De todos os apartamentos que iniciaram o ano em condições médias, 70% destes deterioram-se no fim do ano, 20% mantêm-se em condições médias e 10% melhoram a sua condição para excelentes. De todos os apartamentos que começam o ano em condições excelentes, 60% deterioram-se para más condições, 20% descem para condições médias e 20% mantêm a sua excelente condição.

A situação de um apartamento pode ser traduzida por uma cadeia de Markov com três estados, correspondentes ao estado do apartamento, e que possui como matriz de transição:

$$\begin{pmatrix} 0.1 & 0.7 & 0.6 \\ 0.5 & 0.2 & 0.2 \\ 0.4 & 0.1 & 0.2 \end{pmatrix}.$$

2 Analisando o Problema...

Deste modo, como a escolha do cruzamento seguinte é completamente aleatória, pode-se interpretar o nosso problema do polícia sinaleiro como sendo uma cadeia de Markov com 8 estados, numerados de 1 a 8, cada um deles correspondendo à presença do polícia no respectivo cruzamento.

Assim sendo, facilmente[2] se obtém para o nosso problema a matriz de

2 Basta usar a lei de Laplace, que nos diz que a probabilidade de um certo acontecimento é dada pelo quociente entre o número de casos favoráveis e o número de casos possíveis.

transição:

$$\begin{pmatrix} \frac{1}{3} & \frac{1}{3} & 0 & \frac{1}{5} & 0 & 0 & 0 & 0 \\ \frac{1}{3} & \frac{1}{3} & 0 & 0 & \frac{1}{4} & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{3} & \frac{1}{5} & 0 & \frac{1}{3} & 0 & 0 \\ \frac{1}{3} & 0 & \frac{1}{3} & \frac{1}{5} & \frac{1}{4} & 0 & \frac{1}{4} & 0 \\ 0 & \frac{1}{3} & 0 & \frac{1}{5} & \frac{1}{4} & 0 & 0 & \frac{1}{3} \\ 0 & 0 & \frac{1}{3} & 0 & 0 & \frac{1}{3} & \frac{1}{4} & 0 \\ 0 & 0 & 0 & \frac{1}{5} & 0 & \frac{1}{3} & \frac{1}{4} & \frac{1}{3} \\ 0 & 0 & 0 & 0 & \frac{1}{4} & 0 & \frac{1}{4} & \frac{1}{3} \end{pmatrix}. \tag{2}$$

Gostaríamos de ser capazes de poder determinar onde é que o polícia se encontrará ao fim de muito tempo, i.e., prever o comportamento assimptótico da órbita de um certo vector posição $x^{(0)}$, denotada por $x^{(1)}, x^{(2)}, \ldots, x^{(n)}, \ldots$.

O seguinte lema permite-nos estabelecer uma relação entre vectores de posição consecutivos, o que é um primeiro passo para uma análise assimptótica deste sistema.

LEMA 1. *Se P é uma matriz de transição de uma cadeia de Markov e $x^{(n)}$ é o vector posição na n-ésima observação, então $x^{(n+1)} = Px^{(n)}$.*

Demonstração. Esta prova assenta essencialmente em conceitos de probabilidade. Dados dois acontecimentos A e B, tais que $p(B) \neq 0$, então:

$$p(A \mid B) = \frac{p(A \cap B)}{p(B)}.$$

Dado i tal que $1 \leq i \leq k$, considerem-se os acontecimentos:
A_i: Estar no estado i na $(n+1)$-ésima observação.
B_i: Estar no estado i na n-ésima observação.
Tem-se então que:

$$x^{(n)} = \begin{pmatrix} p(B_1) \\ p(B_2) \\ \vdots \\ p(B_k) \end{pmatrix}, x^{(n+1)} = \begin{pmatrix} p(A_1) \\ p(A_2) \\ \vdots \\ p(A_k) \end{pmatrix} \quad \text{e} \quad p_{ij} = p(A_i \mid B_j).$$

Obtém-se então o pretendido pela lei da probabilidade total,

$$p(A_i) = \sum_{j=1}^{k} p(A_i \mid B_j) p(B_j).$$

□

	0	1	2	3	4	5	10	20	22
$x_1^{(n)}$	0	.000	.133	.116	.130	.123	.113	.108	.107
$x_2^{(n)}$	0	.250	.146	.163	.140	.138	.115	.108	.107
$x_3^{(n)}$	0	.000	.050	.039	.067	.073	.100	.107	.107
$x_4^{(n)}$	0	.250	.113	.187	.162	.178	.178	.179	.179
$x_5^{(n)}$	1	.250	.279	.190	.190	.168	.149	.143	.143
$x_6^{(n)}$	0	.000	.000	.050	.056	.074	.099	.107	.107
$x_7^{(n)}$	0	.000	.133	.104	.131	.125	.138	.143	.143
$x_8^{(n)}$	0	.250	.146	.152	.124	.121	.108	.107	.107

Tabela 1:

Deste modo, o vector de posição $x^{(n)}$ fica completamente determinado por $x^{(0)}$ e pela matriz de transição P, uma vez que:

$$\begin{aligned} x^{(1)} &= Px^{(0)} \\ x^{(2)} &= Px^{(1)} = P^2x^{(0)} \\ &\dots \\ x^{(n)} &= Px^{(n-1)} = P^nx^{(0)}. \end{aligned}$$

Por exemplo, assumindo que o polícia começa no cruzamento 5, os seus destinos prováveis são dados pelos vectores posição $x^{(n)} = (x_1^{(n)}, x_2^{(n)}, \dots, x_8^{(n)})$ apresentados na Tabela 1 com arredondamentos de três casas decimais.

2.1 Voltando aos Exemplos Anteriores...

Note-se que a matriz de transição do exemplo 1 é uma matriz regular e que, se um aluno está inicialmente no estado 2 então os seus vectores posição

(arredondados com três casas decimais) serão dados por:

$$x^{(0)} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$
$$x^{(1)} = Px^{(0)} = \begin{pmatrix} .8 & .3 \\ .2 & .7 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} .3 \\ .7 \end{pmatrix}$$
$$x^{(2)} = Px^{(1)} = \begin{pmatrix} .8 & .3 \\ .2 & .7 \end{pmatrix} \begin{pmatrix} .3 \\ .7 \end{pmatrix} = \begin{pmatrix} .45 \\ .55 \end{pmatrix}$$
$$x^{(3)} = Px^{(2)} = \begin{pmatrix} .8 & .3 \\ .2 & .7 \end{pmatrix} \begin{pmatrix} .45 \\ .55 \end{pmatrix} = \begin{pmatrix} .525 \\ .475 \end{pmatrix}$$
$$\dots$$
$$x^{(8)} = \begin{pmatrix} .598 \\ .402 \end{pmatrix}, x^{(9)} = \begin{pmatrix} .599 \\ .401 \end{pmatrix}, x^{(10)} = \begin{pmatrix} .595 \\ .405 \end{pmatrix}, x^{(n)} = \begin{pmatrix} .600 \\ .400 \end{pmatrix}, \forall n \geq 11.$$

No exemplo 2 põe-se a questão do que poderão os arrendatários esperar dos seus apartamentos a longo prazo, sabendo que no início do arrendamento eles estão classificados como médios. De uma forma heurística, à semelhança do que acontece no exemplo 1, parece que os vectores posição tendem para um vector de posição fixo pois

$$x^{(0)} = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \qquad \text{e} \qquad x^{(n)} = \begin{pmatrix} .421 \\ .326 \\ .251 \end{pmatrix}, \forall n \geq 14.$$

Logo, isto leva-nos a supor que a maior parte dos apartamentos ficarão degradados.

3 A Solução

Gostar-se-ia de encontrar condições sobre os vectores posição $x^{(n)}$ que nos permitam, de alguma forma, saber mais sobre o comportamentos destes. O lema seguinte permite-nos estabelecer uma relação entre os valores mínimo e máximo dum vector posição com o vector posição anterior e a matriz de transição.

Lema 2. *Seja P uma matriz de transição que não tem qualquer entrada nula, e ϵ a menor entrada em P. Se x é um vector de probabilidade em que $M_0 = \max_{1\leq i\leq n} x_i$, $m_0 = \min_{1\leq i\leq n} x_i$, $M_1 = \max_{1\leq i\leq n}(x^T P)_i$ e $m_1 = \min_{1\leq i\leq n}(x^T P)_i$ então:*

1. *$M_1 \leq M_0$ e $m_1 \geq m_0$*
2. *$M_1 - m_1 \leq (1-2\epsilon)(M_0 - m_0)$.*

Demonstração. A primeira afirmação é trivial, pelo que se vai provar somente a segunda. Seja y o vector obtido de x substituindo todas as coordenadas, com excepção de uma de valor m_0, por M_0.

Então, para todo $i \in \{1, \ldots, n\}$, $x_i \leq y_i$ e cada componente de $y^T P$ é da forma

$$a\, m_0 + (1-a)\, M_0 = M_0 - a(M_0 - m_0), \text{ com } \quad a \geq \epsilon.$$

Logo, como $x_i \leq y_i$, tem-se que

$$M_1 \leq M_0 - \epsilon(M_0 - m_0).$$

Usando um raciocínio completamente análogo ao anterior obtém-se que

$$-m_1 \leq -m_0 - \epsilon(M_0 - m_0).$$

Assim sendo, decorre destas duas últimas expressões que $M_1 - m_1 \leq (1 - 2\epsilon)(M_0 - m_0)$, como se pretendia. □

TEOREMA 3. *Se P é uma matriz de transição regular, então*

$$Q = \lim_{n\to\infty} P^n = \begin{pmatrix} q_1 & q_1 & \cdots & q_1 \\ q_2 & q_2 & \cdots & q_2 \\ \vdots & \vdots & \ddots & \vdots \\ q_k & q_k & \cdots & q_k \end{pmatrix}$$

onde q_i são números positivos tais que $q_1 + q_2 + \cdots + q_k = 1$.

Demonstração. Há dois casos a considerar:

i) P não tem entradas nulas

Seja ϵ a menor entrada de P, $\{e_i\}_{1\leq i\leq n}$ a base canónica de $\mathbb{R}^n$ e $1 \leq j \leq n$ qualquer.

Sejam M_k e m_k as coordenadas máxima e mínima de $e_j^T P^k$, respectivamente, ou seja, as entradas máxima e mínima da j-ésima linha da matriz P.

Então, pelo lema anterior, obtém-se uma sucessão (M_k) tal que:

$$M_1 \geq M_2 \geq M_3 \geq \ldots$$
$$M_k - m_k \leq (1-2\epsilon)(M_{k-1} - m_{k-1}), \text{ para } k \geq 1.$$

Assim sendo, têm-se as duas sucessões (m_k) e (M_k) relacionadas por

$$M_k - m_k \leq (1-2\epsilon)^k,$$

pelo que convergem para um mesmo valor $0 < q_j < 1$. Como $e_j^T P^k$ é a j-ésima linha de P^k, tem-se que P^k converge para uma matriz Q cujas linhas têm entradas constantes.

O facto de a soma das linhas ser 1 deve-se ao facto de esta propriedade passar ao limite.

ii) P tem alguma entrada nula

Toma-se $N \in \mathbb{N}$ tal que P^N não tem entradas nulas e $\eta > 0$ a menor entrada de P^N.

Analogamente ao que foi feito, obtém-se que:

$$M_{kN} - m_{kN} \leq (1 - 2\eta)^k.$$

O resto da prova decorre como no caso anterior.

□

PROPOSIÇÃO 4. *Se P é uma matriz de transição regular e x é um qualquer vector de probabilidade, ou seja $\sum_{i=1}^n x_i = 1$, então*

$$q = \lim_{n\to\infty} P^n x = \begin{pmatrix} q_1 \\ q_2 \\ \vdots \\ q_k \end{pmatrix}$$

onde q é um vector de probabilidade fixo (independente de x), com todas as entradas positivas.

Demonstração. Este resultado é essencialmente um corolário do teorema anterior. Assim sendo, considerem-se Q e q como no teorema anterior, i.e.,

$$Q = \begin{pmatrix} q_1 & q_1 & \cdots & q_1 \\ q_2 & q_2 & \cdots & q_2 \\ \vdots & \vdots & \ddots & \vdots \\ q_k & q_k & \cdots & q_k \end{pmatrix} \quad \text{e} \quad q = \begin{pmatrix} q_1 \\ q_2 \\ \vdots \\ q_k \end{pmatrix}.$$

Então, para qualquer vector de probabilidade x tem-se que:

$$Qx = \begin{pmatrix} q_1 & q_1 & \cdots & q_1 \\ q_2 & q_2 & \cdots & q_2 \\ \vdots & \vdots & \ddots & \vdots \\ q_k & q_k & \cdots & q_k \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_k \end{pmatrix} = \begin{pmatrix} q_1x_1 + q_1x_2 + \cdots + q_1x_k \\ q_2x_1 + q_2x_2 + \cdots + q_2x_k \\ \vdots \\ q_kx_1 + q_kx_2 + \cdots + q_kx_k \end{pmatrix}$$

$$= (x_1 + x_2 + \cdots + x_k) \begin{pmatrix} q_1 \\ q_2 \\ \vdots \\ q_k \end{pmatrix} = 1 \cdot q = q.$$

Logo, pelo teorema anterior conclui-se o pretendido. □

Tem-se ainda uma forma prática de calcular q que é dada pelo seguinte teorema.

TEOREMA 5. *O vector de probabilidade q de uma matriz de transição regular é o único vector de probabilidade que satisfaz a equação $Pq = q$.*

Demonstração. Decorre facilmente da proposição anterior que, de facto, q é vector próprio de P. A unicidade de q decorre do facto de que, se r é outro vector de probabilidade tal que $Pr = r$ então $P^n r = r$, para todo o n e, pela proposição anterior, esta sucessão converge para q. Como o espaço $\mathbb{R}^n$ é separado decorre que $r = q$. □

Tente-se agora aplicar esta teoria ao estudo do exemplo do polícia. Em primeiro lugar convém realçar que a matriz de transição em (2) é regular[3], pelo que a análise assimptótica do sistema fica completamente determinada. Deste modo, para calcular o vector posição, que dará a proporção de tempo gasto em cada cruzamento a longo prazo, bastará calcular o único vector próprio da matriz de transição P associado ao valor próprio 1. Assim sendo, de $(I - P)q = 0$, tira-se que:

$$q = \begin{pmatrix} \frac{3}{28} \\ \frac{3}{28} \\ \frac{3}{28} \\ \frac{5}{28} \\ \frac{4}{28} \\ \frac{3}{28} \\ \frac{4}{28} \\ \frac{3}{28} \end{pmatrix} = \begin{pmatrix} .1071\ldots \\ .1071\ldots \\ .1071\ldots \\ .1785\ldots \\ .1428\ldots \\ .1071\ldots \\ .1428\ldots \\ .1071\ldots \end{pmatrix}. \tag{3}$$

3 Para verificar este facto basta calcular P^4, onde P denota a matriz em (2).

Logo, a estratégia da distribuição do tempo do polícia de forma aleatória não permite cumprir o objectivo de gastar igual tempo em cada cruzamento.

4 Questões Póstumas

Na secção anterior concluiu-se que a estratégia adoptada na distribuição do polícia pelos vários cruzamentos não terá sido a mais adequada. A solução "óptima" seria a correspondente ao facto do vector próprio q como em (3) possuir todas as componentes iguais, ou seja, de que cada componente do vector de probabilidade obtido fosse de valor $\frac{1}{8}$.

Para que haja uma distribuição uniforme bastará somente que a soma de cada linha da matriz de transição P seja 1 (porquê?). De facto, se tal acontecer, como esta propriedade também é obtida em P^n, é fácil concluir[4] que cada linha de Q tem soma 1. O pretendido decorre então de imediato.

Mas de que forma é que esta "receita" se traduziria para o comandante da polícia na distribuição do serviço do seu agente? Na verdade, uma qualquer matriz de transição P cuja soma de cada linha é 1 seria como que um guia para o polícia, indicando-lhe diferentes probabilidades de transitar entre os cruzamentos, e permitiria a distribuição uniforme do seu tempo de serviço. Há então uma infinidade de soluções "óptimas" para este nosso problema.

Uma outra curiosidade reside ainda na configuração dos cruzamentos apresentada como exemplo. De facto, existem outras situações em que a estratégia adoptada inicialmente até poderia ser correcta, mas que correspondem a casos mais regulares. Isto passa-se, por exemplo, quando cada cruzamento tem uma única ligação por estrada com todos os outros cruzamentos, como ilustrado pela Figura 2 seguinte onde os cruzamentos são assinalados por números, sendo que as outras intersecções das figuras não correspondem a cruzamentos mas a passagens aéreas.

4 Mediante um simples argumento de continuidade.

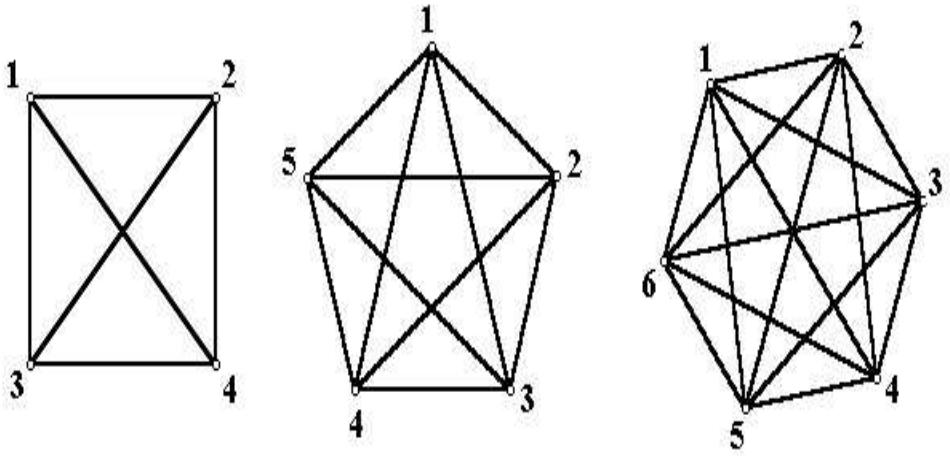


Figura 2

Agradecimentos

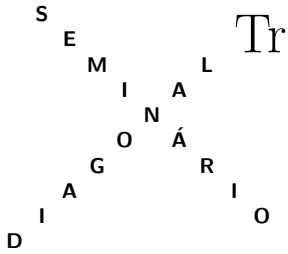
Gostava de deixar os meus agradecimentos ao Professor José Ferreira Alves pelo desafio e incentivo para o seminário diagonal na FCUP, e aos Professores Jorge Rocha e Maria João Costa por todo o apoio e motivação. Não posso deixar de agradecer o convite do Instituto Superior Técnico para o Seminário Diagonal e de salientar a forma como fui bem recebido por toda a organização, realçando todavia o João Boavida e a Professora Ana Cannas da Silva pela sua atenção inesgotável. Por fim, gostaria de agradecer à Patrícia Gonçalves pela ajuda na revisão deste texto. A todos estes — e a quem por um mero lapso me posso ter esquecido! — o meu "Muito Obrigado!".

Referências

[1] H. Anton, C. Rorres, *Elementary Linear Algebra – Applications Version*, 7ª edição, John Wiley & Sons Inc., 1994.

[2] R. Bronson, *Matrizes*, McGraw-Hill, 1993.

[3] J. Kemeny, J. Snell, *Finite Markov Chains*, Springer-Verlag, 1976.



Transformações de Lorentz e de Möbius

Ida Griffith
2º ano da LMAC
L51279@isabelle.math.ist.utl.pt

Palavras Chave

transformações de Lorentz, transformações de Möbius, projecção estereográfica, esferas celestes

Resumo

Dois gémeos são separados no seu 20º aniversário; um fica na Terra, o outro viaja a uma velocidade próxima da velocidade da luz em direcção a um planeta situado a 8 anos-luz e regressa. Que idade terá quando regressar? Terá a mesma idade que o gémeo que ficou na Terra? E que relação tem isto tudo com o grupo de Möbius?

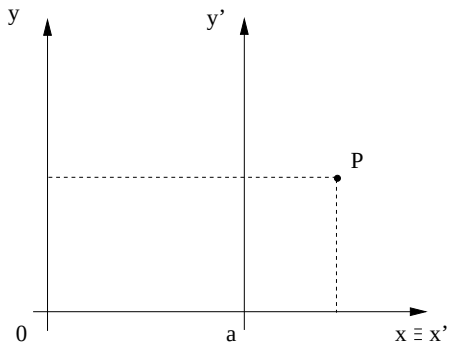
1 Transformações de Galileu

Como todos aprendemos no Secundário, facilmente se relaciona a posição de dois observadores inerciais. Considere-se um observador O, e seja O' um observador na posição $(a, 0, 0)$. Considere-se um ponto P. Se O atribui a este ponto as coordenadas (x, y, z), O' atribui (x', y', z'), onde

$$\begin{cases} x = x' + a \\ y = y' \\ z = z' \end{cases} \Leftrightarrow \begin{cases} x' = x - a \\ y' = y \\ z' = z \end{cases}$$

As coordenadas de um ponto (acontecimento) dependem do referencial (observador) que se considera.

Suponhamos que O' se está a mover segundo o eixo dos xx com uma velocidade v. Se O' parte da origem, a abcissa do referencial O' (vista por O) é dada por $a = vt$. Assim, O' vê um acontecimento P com coordenadas (t', x', y', z'), enquanto O vê o mesmo acontecimento com coordenadas



(t, x, y, z) tais que

$$\begin{cases} t = t' \\ x = x' + vt \\ y = y' \\ z = z' \end{cases} \Leftrightarrow \begin{cases} t' = t \\ x' = x - vt \\ y' = y \\ z' = z \end{cases} \Leftrightarrow$$

$$\Leftrightarrow \begin{bmatrix} t' \\ x' \\ y' \\ z' \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ -v & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} t \\ x \\ y \\ z \end{bmatrix}$$

Isto na Mecânica Clássica, e pensou-se ser assim durante séculos. No fim do século XIX, descobriu-se que quando a velocidade aumenta muito, os cálculos, aplicando as transformações de Galileu, começavam a apresentar um erro cada vez mais significativo.

Sejam:

u' - velocidade de um objecto em relação a O'

v - velocidade de O' em relação a O

Qual a velocidade u do objecto em relação a O? De

$$\begin{cases} x = ut \\ x' = u't' \end{cases}$$

obtém-se a **lei da adição das velocidades**:

$$u = \frac{x}{t} = \frac{x' + vt'}{t'} = u' + v$$

Isto não faz sentido quando se consideram velocidades muito elevadas. Por exemplo, se $u' = c$ (velocidade da luz) e $v > 0$,

$$u = c + v > c$$

Mas é um facto experimental que a velocidade da luz é a mesma em todos os referenciais. Não se podem ter velocidades superiores à velocidade da luz.

2 Transformações de Lorentz

Introduzimos o factor de correcção $\gamma = \gamma(\vec{v}) = \gamma(\|\vec{v}\|)$:

$$x = vt + \frac{x'}{\gamma} \tag{1}$$

Determinação de γ:

Pelo **Princípio da Relatividade**, se v é a velocidade de O' em relação a O, então $(-v)$ é a velocidade de O em relação a O'. Por analogia com (1)

$$x' = -vt' + \frac{x}{\gamma} \tag{2}$$

Por outro lado, (1) é equivalente a

$$x' = \gamma(x - vt) \tag{3}$$

De (2) e (3):

$$\begin{aligned} \gamma(x - vt) &= -vt' + \frac{x}{\gamma} \Leftrightarrow \\ \Leftrightarrow vt' &= \frac{x}{\gamma} - \gamma x + \gamma vt \Leftrightarrow \\ \Leftrightarrow t' &= \frac{1}{v}\left(\left(\frac{1}{\gamma} - \gamma\right)x + \gamma vt\right) \Leftrightarrow \\ \Leftrightarrow &\begin{cases} x' = \gamma(x - vt) \\ t' = \frac{1-\gamma^2}{\gamma v}x + \gamma t \end{cases} \end{aligned} \tag{4}$$

Se em $t = x = 0$ for emitido um sinal luminoso (velocidade= c)

$$\begin{cases} x = ct \\ x' = ct' \end{cases}$$

(todos os observadores vêem a luz à mesma velocidade, direcção e sentido), e portanto

$$\begin{cases} ct' = \gamma(ct - vt) \\ t' = \frac{1-\gamma^2}{\gamma v}ct + \gamma t \end{cases} \Leftrightarrow$$

$$\Leftrightarrow \begin{cases} t' = \frac{\gamma}{c}(c - v)t \\ t' = \left(\frac{1-\gamma^2}{\gamma v}c + \gamma\right)t \end{cases} \Leftrightarrow$$

$$\Leftrightarrow \frac{\gamma}{c}(c - v) = \frac{1 - \gamma^2}{\gamma v}c + \gamma \Leftrightarrow$$

$$\Leftrightarrow \gamma^2 \frac{v}{c^2}c\left(1 - \frac{v}{c}\right) = 1 - \gamma^2 + \gamma^2\frac{v}{c} \Leftrightarrow$$

$$\Leftrightarrow \gamma^2\left(\frac{v}{c} - \frac{v^2}{c^2} + 1 - \frac{v}{c}\right) = 1 \Leftrightarrow$$

$$\Leftrightarrow \gamma = \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}} \tag{5}$$

De (4) e (5),

$$t' = \frac{1 - \frac{1}{1-\frac{v^2}{c^2}}}{v\sqrt{1 - \frac{v^2}{c^2}}}x + \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}}t =$$

$$= \frac{\frac{-\frac{v^2}{c^2}}{1-\frac{v^2}{c^2}}}{v\sqrt{1 - \frac{v^2}{c^2}}}x + \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}}t =$$

$$= \frac{-\frac{v}{c^2}}{\sqrt{1 - \frac{v^2}{c^2}}}x + \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}}t$$

e portanto as transformações de Lorentz são dadas por

$$\begin{cases} x' = \frac{x - vt}{\sqrt{1-\frac{v^2}{c^2}}} \\ t' = \frac{t - \frac{v}{c^2}x}{\sqrt{1-\frac{v^2}{c^2}}} \end{cases}$$

Portanto **diferentes observadores fazem diferentes medições de espaço e de tempo.**

Quando $v \ll c$, a Tranformações de Lorentz são bem aproximadas por

$$\begin{cases} x' = x - vt \\ t' = t \end{cases}$$

(ou seja, as transformações de Galileu estão correctas para velocidades baixas relativamente à velocidade da luz).

Pelo Princípio da Relatividade as transformações inversas são

$$\begin{cases} x = \frac{x'+vt'}{\sqrt{1-\frac{v^2}{c^2}}} \\ t = \frac{t'+\frac{v}{c^2}x'}{\sqrt{1-\frac{v^2}{c^2}}} \end{cases}$$

Sejam:

u' - velocidade de um objecto em relação a O'

v - velocidade de O' em relação a O

Qual a velocidade u do objecto em relação a O? De

$$\begin{cases} x = ut \\ x' = u't' \end{cases}$$

obtém-se a **regra relativista para a adição de velocidades**:

$$u = \frac{x}{t} = \frac{x' + vt'}{t' + \frac{vx'}{c^2}} = \frac{u' + v}{1 + \frac{vu'}{c^2}}$$

Assim, se $u' = c$

$$u = \frac{c + v}{1 + \frac{v}{c}} = \frac{c(c + v)}{c + v} = c!$$

Para simplificar as fórmulas, vamos usar unidades nas quais $c = 1$ (por exemplo, medindo o tempo em anos e as distâncias em anos-luz).

Definição. Espaço-tempo de Minkowski

O espaço-tempo de Minkowski é $\mathbb{R}^4$ com coordenadas t, x, y, z (a primeira coordenada é de tempo e as restantes são de espaço).

Definição. Produto interno de Minkowski

Dados dois vectores de $\mathbb{R}^4$,

$$\vec{u} = (t_1, x_1, y_1, z_1)$$
$$\vec{v} = (t_2, x_2, y_2, z_2),$$

define-se

$$< \vec{u}, \vec{v} >= t_1t_2 - x_1x_2 - y_1y_2 - z_1z_2.$$

(Na realidade é um pseudo-produto interno, pois não satisfaz positividade).

Pode mostrar-se que o produto interno de Minkowski é invariante sob as transformações de Lorentz, i.e., tem sempre o mesmo valor independentemente do referencial em que é calculado.

Norma de um vector: dado

$$\vec{x} = (x_0, x_1, x_2, x_3)$$

define-se

$$\|\vec{x}\|^2 = x_0^2 - x_1^2 - x_2^2 - x_3^2.$$

Existem 3 tipos de vectores:

$$\begin{cases} \text{tipo tempo:} & \|\vec{x}\|^2 > 0 \\ \text{tipo luz ou nulo:} & \|\vec{x}\|^2 = 0 \\ \text{tipo espaço:} & \|\vec{x}\|^2 < 0 \end{cases}$$

No caso de um vector do *tipo tempo*, $\sqrt{\|\vec{x}\|^2}$ corresponde ao tempo medido por um observador que vê o acontecimento em posição constante; no caso de ser um vector do *tipo espaço*, $\sqrt{-\|\vec{x}\|^2}$ corresponde à distância medida por um observador que vê o acontecimento no mesmo instante.

Exemplo.

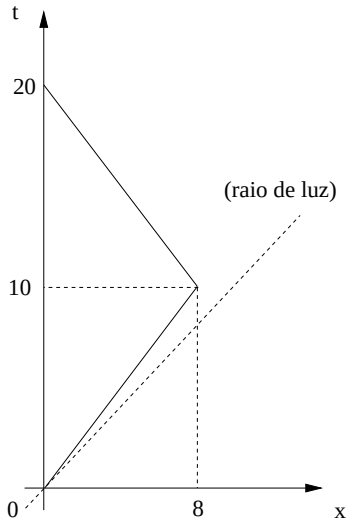
Paradoxo dos Gémeos: *Dois gémeos, Alice e Bernardo, separam-se no seu 20º aniversário: enquanto Alice fica na Terra (que constitui muito aproximadamente um referencial inercial), Bernardo parte a 80% da velocidade da luz na direcção de um planeta situado a 8 anos-luz da Terra, que alcança 10 anos mais tarde (medidos no referencial da Terra). Após uma curta estadia, Bernardo regressar à Terra, novamente a 80% da velocidade da luz. Consequentemente, Alice tem 40 anos quando revê o irmão. Quantos anos terá Bernardo?*

Seja t_1 o tempo da primeira parte do percurso e t_2 o tempo relativo à segunda parte.

$$t_1 = \sqrt{< (10, 8), (10, 8) >} = \sqrt{36} = 6$$

$$t_2 = \sqrt{< (10, -8), (10, -8)} = \sqrt{36} = 6$$

Para o Bernardo passaram 12 anos, logo terá 32 anos quando regressar à Terra. Portanto Alice e Bernardo têm 8 anos de diferença quando se reencontram!



Define-se o grupo de Lorentz como o conjunto de todas as transformações lineares $A : \mathbb{R}^4 \to \mathbb{R}^4$ que preservam o produto interno de Minkowski. Pretende-se relacionar o grupo de Lorentz com as transformações de Möbius; para isso é necessário clarificar alguns conceitos.

3 Projecção Estereográfica

Definimos $S^2 = \{(x, y, z) \in \mathbb{R} : x^2 + y^2 + z^2 = 1\}$, $\alpha = \{(x, y, z) \in \mathbb{R} : z = 0\}$ e fazemos a identificação $\mathbb{R}^2 \approx \mathbb{C}$.

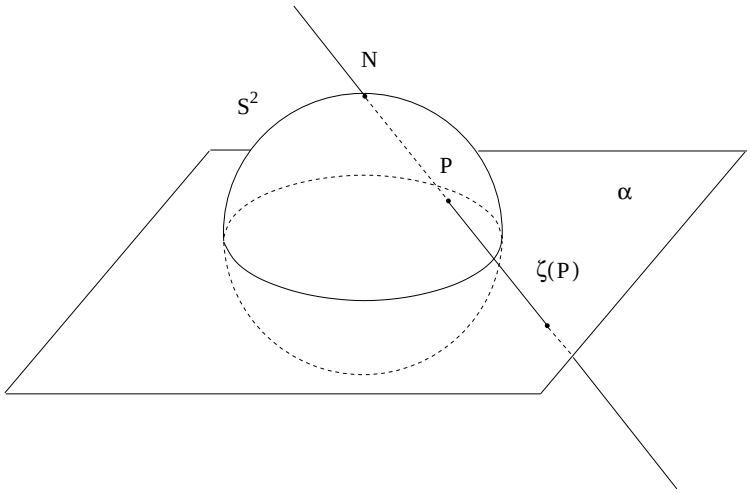
Definição. Projecção Estereográfica

A projecção estereográfica $\zeta : S^2 \setminus N \to C$ é uma aplicação que para cada ponto $P = (x, y, z)$ associa a intersecção ζ da recta que une $N = (0, 0, 1)$ e P com α. Portanto

$$\zeta(x, y, z) = \frac{x + iy}{1 - z},$$

ou em coordenadas esféricas ($r = 1$),

$$\zeta(\theta, \varphi) = \frac{\sin\theta}{1 - \cos\theta} e^{i\varphi}$$



PROPOSIÇÃO.

A projecção estereográfica de uma circunferência $\gamma \subset S^2$ é uma recta (se $N \in \gamma$) ou uma circunferência (caso contrário).

Esta proposição, apesar de não ter uma demonstração simples, é relativamente fácil de verificar intuitivamente. Definimos ainda

$$\zeta(N) = \zeta(0,0,1) \equiv \infty$$

Assim, pode-se pensar na esfera S^2 como o plano complexo reunido com o ponto no infinito:

$$S^2 \approx \mathbb{C} \cup \{\infty\}$$

4 Transformações de Möbius

DEFINIÇÃO. Transformações de Möbius

Uma *transformação de Möbius* é uma função $f : \mathbb{C} \cup \{\infty\} \to \mathbb{C} \cup \{\infty\}$ da forma

$$f(\zeta) = \frac{a\zeta + b}{c\zeta + d}$$

com $a, b, c, d \in \mathbb{C}, ad - bc \neq 0$.

O conjunto $GL(2, \mathbb{C}) = \{A \in \mathcal{M}_2(\mathbb{C}) : \det A \neq 0\}$ é um grupo.

Considere-se a função

$$H : GL(2,\mathbb{C}) \to \mathcal{M}$$
$$H\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \frac{a\zeta + b}{c\zeta + d}$$

($\mathcal{M}$ designa o grupo das transformações de Möbius com a operação de composição)

Como $H(A \cdot B) = H(A) \circ H(B)$, H é um homomorfismo de grupos. O núcleo de H, $\mathcal{N}_{\mathcal{H}}$, pode ser determinado resolvendo

$$H\begin{pmatrix} a & c \\ c & d \end{pmatrix} = \zeta \Leftrightarrow \frac{a\zeta + b}{c\zeta + d} = \zeta \Leftrightarrow b = c = 0 \wedge a = d,$$

ou seja,

$$\mathcal{N}_H = \left\{ \begin{bmatrix} a & 0 \\ 0 & a \end{bmatrix} : a \in \mathbb{C} \setminus \{0\} \right\} = \{aI : a \in \mathbb{C} \setminus \{0\}\}$$

Assim,

$$\mathcal{M} \approx \frac{GL(2,\mathbb{C})}{\mathcal{N}_H}.$$

O conjunto $SL(2,\mathbb{C}) = \{A \in GL(2,\mathbb{C}) : \det A = 1\}$ é um subgrupo de $GL(2,\mathbb{C})$. Se considerarmos a restrição de H a $SL(2,\mathbb{C})$,

$$H|_{SL(2,\mathbb{C})} : SL(2,\mathbb{C}) \to \mathcal{M},$$

vemos que o núcleo deste homomorfismo é determinado por

$$H\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \zeta \Leftrightarrow \frac{a\zeta + b}{c\zeta + d} = \zeta \Leftrightarrow b = c = 0 \wedge a = d$$

e ainda por

$$\det\begin{pmatrix} a & b \\ c & d \end{pmatrix} = 1 \Leftrightarrow ad - bc = 1 \Leftrightarrow a = d = 1 \vee a = d = -1$$

Portanto $\mathcal{N}_{H|_{SL(2,\mathbb{C})}} = \{\pm I\}$

Assim,

$$\mathcal{M} \approx \frac{SL(2,\mathbb{C})}{\{\pm I\}}.$$

Seja $\{\vec{e}_0, \vec{e}_1, \vec{e}_2, \vec{e}_3\}$ uma base ortonormada do espaço-tempo de Minkowski e $\vec{v} = t\vec{e}_0 + x\vec{e}_1 + y\vec{e}_2 + z\vec{e}_3$ um vector.

Pode-se identificar $\vec{v}$ com uma matriz hermitiana V usando a função $I : \mathbb{R}^4 \to \mathbb{H}_2 \equiv \{V \in \mathcal{M}_2(\mathbb{C}) : V^* = V\}$ definida por

$$I(\vec{v}) = V = \frac{1}{\sqrt{2}} \begin{pmatrix} t+z & x+iy \\ x-iy & t-z \end{pmatrix}$$

Uma vez que

$$\det V = \frac{1}{2}(t^2 - x^2 - y^2 - z^2) = \frac{1}{2} < \vec{v}, \vec{v} >,$$

podemos considerar a função $F : SL(2,\mathbb{C}) \to O(3,1)$ dada por

$$F(g)V = gVg^*,$$

(note-se que $\det(F(g)V) = \det V$). Esta aplicação é um homomorfismo de grupos, já que $F(gh)V = F(g) \circ F(h)V$, $\forall_{V \in \mathbb{H}_2}$ $\forall_{g,h \in SL(2,\mathbb{C})}$.

É possível mostrar que $\mathcal{N}_F = \{\pm I\}$, e que F é sobrejectiva. Logo,

$$\frac{SL(2,\mathbb{C})}{\{\pm I\}} \approx O(3,1),$$

e portanto

$$O(3,1) \approx \mathcal{M}$$

O grupo das transformações de Lorentz é isomorfo ao grupo das transformações de Möbius.

Este isomorfismo tem a seguinte interpretação: dois observadores distintos estão relacionados por uma transformação de Lorentz. Por outro lado, cada observador situa a um objecto cuja luz lhe chega de uma determinada direcção no ponto correspondente de S^2 (a sua "esfera celeste"). Acontece que diferentes observadores situam o mesmo objecto em pontos diferentes das suas esferas celestes. A função $f : S^2 \to S^2$ que relaciona as duas esferas celestes é exactamente a transformação de Möbius associada à transformação de Lorentz que relaciona os dois observadores.

Usando as propriedades da projecção estereográfica e das transformações de Möbius (vistas como transformações $f : \mathbb{C} \to \mathbb{C}$), é possível mostrar que estas transformações (vistas como transformações $f : S^2 \to S^2$) levam círculos para círculos. Portanto se um observador vê um objecto com um contorno circular, todos os outros observadores verão o mesmo objecto com um contorno circular. Em particular, isto será verdade se o objecto for esférico.

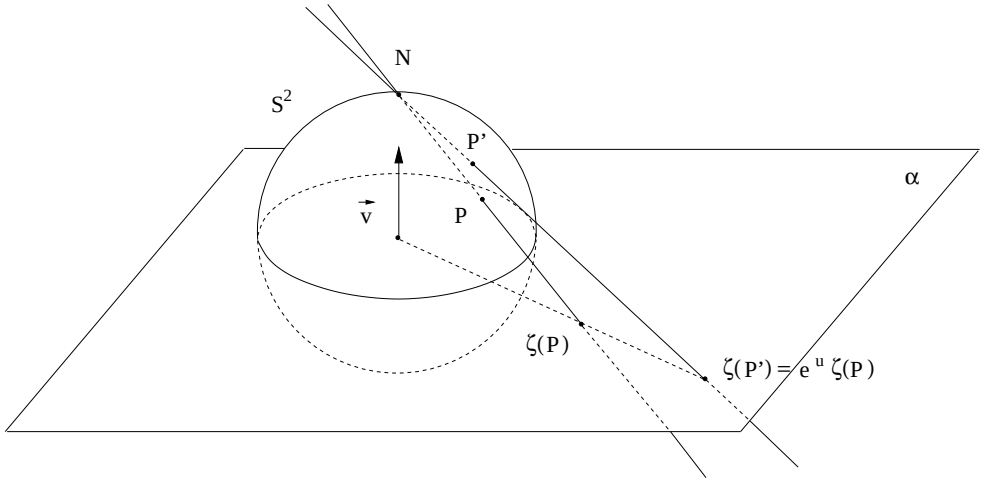
Exemplo.

Velocidade de O' em relação a O: $\vec{v} = v\vec{e}_z$.

Seja $u = \arctan v$. A transformação de Möbius que corresponde a este transformação de Lorentz é

$$f(\zeta) = e^u \zeta$$

Assim, um objecto no ponto P da esfera celeste do observador O será visto pelo observador O' num ponto P' mais próximo de N.



Agradecimentos

Queria agradecer ao Prof. José Natário toda a ajuda e tempo que me dispensou. Gostaria também de agradecer à Fundação Calouste Gulbenkian pela oportunidade que me foi dada de aprofundar os meus conhecimentos.

Referências

[1] G. Ávila, *Variáveis Complexas e Aplicações*, LTC Editora (2000).

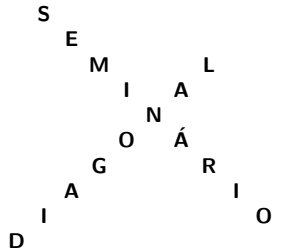
[2] H. Bondi, *Relativity and Common Sense*, Dover (1962).

[3] W. Oliva, *Geometric Mechanics*, Springer (2002).

[4] R. Penrose, *The Apparent Shape of a Relativistically Moving Sphere*, Proc. Cambridge Philos. Soc. 55 (1959) 137-139.

[5] J.M. Resina Rodrigues, *Introdução à Teoria da Relatividade Restrita*, IST Press (1998).

[6] E.F. Taylor, J.A. Wheeler, *Spacetime Physics*, Freeman (1992).
`http://math.ucr.edu/home/baez/relativity.html`



Criptografia — O Estado da Arte

Paulo Jorge de Oliveira Cantante de Matos
4º ano da LEIC - Inteligência Artificial
pocm@mega.ist.utl.pt

Palavras Chave

teoria dos números, criptografia, RSA, Knapsack, logaritmo discreto

Resumo

Existem dois tipos de criptografia neste mundo: criptografia que impede a nossa irmã mais nova de ler os nossos ficheiros e criptografia que impede os mais poderosos governos mundiais de os lerem. Este artigo é sobre o último...

"Two can keep a secret if one is dead"
— Unknown

1 Introdução

Desde sempre, os seres humanos necessitaram de privacidade e quiseram dela usufruir, procurando que aquilo que fazem em determinado momento seja totalmente confidencial, se assim entenderem, e que apenas quem eles desejem tenha acesso a essa informação. De facto, podemos pensar em vários exemplos que vão desde a jovem que não quer que os seus pais leiam o seu diário até grandes potências mundias que não desejam que os seus segredos de estado deixem de ser secretos. Esta situação seria bastante simplificada se todos respeitassem a privacidade dos outros e se todos nós tivessemos a certeza absoluta disso. No entanto, sabemos que isso não acontece. Tal como os pais, cheios de curiosidade, gostam de saber o que a sua filha faz, também os estados, em situações de conflito, desejam saber o que os outros estados fazem ou planeiam. Surge, então, a necessidade de tornar secreta certa informação, tendo sempre em conta que alguém a vai querer descobrir e fazer todos os esforços para o conseguir.

Este artigo levar-nos-á a uma viagem pelo mundo da criptografia dentro de uma perspectiva matemática, tentando explicar alguns dos métodos hoje utilizados. No entanto, não iremos fazer uma cobertura pormenorizada dos algoritmos e métodos, nem de todos os tipos de criptografia utilizados

hoje em dia. Imagine que a criptografia é uma enorme casa na qual cada divisão tratará de um determinado assunto específico. Este artigo levá-lo-á pelo corredor, espreitando apenas para algumas divisões. Assim sendo, e terminada esta introdução, estará pronto para entrar...

2 Conceitos Elementares da Teoria dos Números

Para procedermos de maneira formal à análise dos algoritmos de criptografia convém termos algumas noções básicas de teoria dos números.

2.1 Divisibilidade e Congruências

Iniciamos a nossa introdução à teoria dos números, como é habitual na literatura sobre o assunto, com as noções de divisibilidade e congruência[1]. Comecemos, então, por perceber o que é a divisibilidade, para depois mergulharmos no assunto:

DEFINIÇÃO 1. Dados dois inteiros a e b, $a|b$ (lê-se a divide b ou b é divisível por a) sse existir um inteiro k tal que $b = ak$.

Vejamos, então, algumas propriedades da divisibilidade:

- Todos os números maiores que 1 têm pelo menos 2 divisores: o 1 e o próprio número.
- Se $a|b$ e c é inteiro então $a|bc$.
- Se $a|b$ e $b|c$, então $a|c$.
- Se $a|b$ e $a|c$, então $a|(b \pm c)$.

DEFINIÇÃO 2. Chama-se divisor trivial de um número inteiro b àquele que é igual a b ou a 1. Caso contrário diz-se não trivial.

DEFINIÇÃO 3. Um número inteiro maior que 1 diz-se primo se apenas tiver divisores triviais. Se um número não for primo, diz-se composto.

DEFINIÇÃO 4. Dados dois inteiros a e b (não simultaneamente iguais a zero), o mdc (maior divisor comum) de a e b, denotado $mdc(a, b)$, é o maior inteiro d que divide a e b:

$$mdc(a, b) = \max\{k : k|a \wedge k|b\}.$$

DEFINIÇÃO 5. Dois números inteiros a e b dizem-se primos entre si (e designa-se $a \perp b$) quando $mdc(a, b) = 1$, ou seja, quando não têm divisores comuns maiores que 1.

1 Claro que todos nós sabemos dividir mas será que percebemos realmente o que isso significa?

Felizmente, existe um algoritmo simples para calcular o *mdc* de dois inteiros. Esse algoritmo é denominado Algoritmo de Euclides e apresenta-se, formalmente, já de seguida: dados dois números inteiros positivos a e b, é possível descobrir o seu *mdc* da seguinte forma:

1. Dividir a por b e fazer r igual ao resto.
2. Se $r = 0$, o algoritmo termina; b é a resposta.
3. Fazer $a \leftarrow b$, $b \leftarrow r$ e voltar ao passo 1.

Informalmente, a ideia é dividir a por b e obter um quociente q_1 e um resto r_1, depois dividir b por r_1 e obter um quociente q_2 e um resto r_2 e continuar a divisão até r_n dividir r_{n-1}. Verifica-se facilmente que este algoritmo termina sempre num número finito de passos.

Exemplo: Vamos, então, calcular $mdc(73, 51)$:

$$\begin{aligned} 73 &= 1 \times 51 + 22 \\ 51 &= 2 \times 22 + 7 \\ 22 &= 3 \times 7 + 1 \end{aligned}$$

É de notar que 1 divide 7, pelo que o algoritmo termina e o resultado é 1. Assim sendo, temos que $mdc(73, 51) = 1$ e podemos concluir que 53 e 71 são primos entre si. Verifica-se também que $1 = 7 \times 73 - 10 \times 51$ e isto não é por acaso, como mostra a seguinte proposição.

Proposição 6. *Seja $d = mdc(a, b)$, onde $a > b$. Então existem dois inteiros u e v, tais que $d = ua + bv$. Por outras palavras, o mdc de dois números pode ser expresso como uma combinação linear desses números com coeficientes inteiros.*

Teorema 7 (Teorema Fundamental da Aritmética). *Qualquer número inteiro $n > 1$ pode ser univocamente escrito como o produto de um número finito de primos:*

$$n = p_1 \dots p_m = \prod_{k=1}^{m} p_k, \qquad p_1 \leq \dots \leq p_m.$$

Demonstração. Este teorema é tão óbvio que muitos de nós poderemos questionar por que precisará de ser demonstrado. Como poderiam existir dois conjuntos distintos de primos, cujo produto fosse igual? Realmente, isso não pode acontecer, mas não é simplesmente pela "definição de número primo". Vejamos então a demonstração:

Supondo que $n = 2$ o teorema verifica-se trivialmente, dado que n é primo, logo a sua factorização é ele próprio e, consequentemente, única. Consideremos agora que $n > 2$ e que todos os números menores que n têm uma factorização única (hipótese de indução). Suponhamos que temos duas factorizações:

$$n = p_1 \cdots p_m = q_1 \cdots q_m, \qquad p_1 \leq \ldots \leq p_m \quad e \quad q_1 \leq \ldots \leq q_m,$$

onde os p's e q's são todos primos. Se $p_1 \neq q_1$ então podemos assumir que $p_1 < q_1$ e, consequentemente, menor que todos os q's. Como p_1 e q_1 são primos, o seu mdc é 1 e então o algoritmo de Euclides dá-nos inteiros a e b tais que $ap_1 + bq_1 = 1$. Logo:

$$ap_1q_2 \cdots q_k + bq_1q_2 \cdots q_k = q_2 \cdots q_k.$$

Como p_1 divide ambos os termos da esquerda, pois $q_1q_2 \cdots q_k = n$, então p_1 divide o membro da direita, $q_2 \cdots q_k$ e $q_2 \cdots q_k/p_1$ é um inteiro. Assim, $q_2 \cdots q_k$ tem uma factorização em números primos na qual p_1 aparece, contradizendo a hipótese de indução segundo a qual $q_2 \cdots q_k < n$ tem uma factorização única. Esta contradição mostra que p_1 é mesmo igual a q_1. Podemos assim dividir ambas as factorizações por p_1, obtendo $p_2 \cdots p_m = q_2 \cdots q_k < n$. O mesmo se demonstra para os outros factores e, por indução, a nossa prova está completa. □

Daqui, podemos extrair outras propriedades da divisibilidade:

- Se p é um número primo que divide ab, então $p|a$ ou $p|b$.
- Se $m|a$ e $n|a$ e se $mdc(m, n) = 1$, então $mn|a$.

DEFINIÇÃO 8. Seja n um inteiro positivo. A função Phi de Euler $\varphi(n)$ é definida como sendo o número de inteiros b não negativos tais que $b < n$ e $b \perp n$, ou seja,

$$\varphi(n) \stackrel{def}{=} \#\{b \in \mathbb{N}_0 : mdc(b, n) = 1\}.$$

É fácil ver que $\varphi(1) = 1$ e que $\varphi(p) = p - 1$ para qualquer primo p. Além disso, para qualquer potência de um número primo temos:

$$\varphi(p^\alpha) = p^\alpha - p^{\alpha-1} = p^\alpha \left(1 - \frac{1}{p}\right).$$

Com efeito, é de notar que os números de 0 a $p^\alpha - 1$ que têm um factor comum com p^α diferente de 1, são precisamente aqueles que são divisíveis por p, e que existem $p^{\alpha-1}$ destes números.

Após esta série de definições, e resultados sobre a divisibilidade de números inteiros, vamos ver um assunto extremamente importante para podermos perceber os sistemas de encriptação apresentados mais à frente. Vejamos, então, o que são e como funcionam as congruências.

Definição 9. Dados 3 inteiros a, b e m dizemos que "a é congruente com b módulo m" e escrevemos $a \equiv b \pmod{m}$ se a diferença $a - b$ for divisível por m, ou dito de outra forma, se o resto da divisão de a por m for igual ao resto da divisão de b por m.

Esta definição será uma das mais necessárias, pelo que é bom perceber o melhor possível este conceito. Vejamos algumas propriedades importantes das congruências:

i) $a \equiv a \pmod{m}$

ii) $a \equiv b \pmod{m} \iff b \equiv a \pmod{m}$

iii) Se $a \equiv b \pmod{m}$ e $b \equiv c \pmod{m}$, então $a \equiv c \pmod{m}$.

iv) Se $a \equiv b \pmod{m}$ e $c \equiv d \pmod{m}$ então $a \pm c \equiv b \pm d \pmod{m}$ e $ac \equiv bd \pmod{m}$.

v) Se $a \equiv b \pmod{m}$, então $a \equiv b \pmod{d}$ para qualquer d tal que $d|m$.

vi) Se $a \equiv b \pmod{m}$, $a \equiv b \pmod{n}$ e $m \perp n$ então $a \equiv b \pmod{mn}$.

vii) Se $ac \equiv bc \pmod{m}$ e se m e c são primos entre si então $a \equiv b \pmod{m}$.

Para m fixo, as três primeiras propriedades significam que a congruência mod m é uma relação de equivalência e designa-se o conjunto das classes de equivalência correspondentes por $\mathbb{Z}/m\mathbb{Z}$. Cada uma destas classes tem um e apenas um representante entre 0 e $m-1$, ou seja, qualquer inteiro é congruente módulo m com um e apenas um inteiro entre 0 e $m-1$. Note que, pela quarta propriedade acima enunciada, $\mathbb{Z}/m\mathbb{Z}$ é um anel comutativo.

Teorema 10 (Teorema do Resto Chinês). *Considere-se o sistema de congruências:*

$$\begin{aligned} x &\equiv a_1 \pmod{m_1}, \\ x &\equiv a_2 \pmod{m_2}, \\ &\cdots \\ x &\equiv a_r \pmod{m_r}, \end{aligned}$$

onde $mdc(m_i, m_j) = 1$ *para todo o* $i \neq j$. *Então existe uma solução para todas as congruências e quaisquer duas soluções são congruentes módulo* $M = m_1 m_2 \cdots m_r$.

Demonstração. Em primeiro lugar, demonstraremos a unicidade módulo M (a última frase). Sejam x' e x'' duas soluções do sistema e considere-se $x = x' - x''$. Então, x é congruente com 0 módulo m_i para $i = 1, \ldots, r$, pelo que também o é módulo M (pela sexta propriedade das congruências atrás enunciada). Vejamos, pois, como construir uma solução.

Seja $M_i = M/m_i$ o produto de todos os módulos excepto o m_i. É claro que $m_i \perp M_i$, pelo que existe um inteiro N_i (que pode ser encontrado pelo algoritmo de Euclides) tal que $M_i N_i \equiv 1 \mod m_i$. Considere-se então $x = \sum_i a_i M_i N_i$. Para cada i podemos ver que todos os termos da soma, à excepção do i-ésimo, são divisíveis por m_i, pelo que $x \equiv a_i \mod m_i$ como se pretendia demonstrar. □

EXEMPLO: Vamos resolver o seguinte sistema de congruências, usando o Teorema do Resto Chinês:

$$\begin{aligned} x &\equiv 2 \mod 3 \\ x &\equiv 3 \mod 5 \\ x &\equiv 4 \mod 11 \\ x &\equiv 5 \mod 16. \end{aligned}$$

Obtemos facilmente que $M = 3 \cdot 5 \cdot 11 \cdot 16 = 2640$ e:

$$\begin{aligned} M_1 &= M/m_1 = 880 \\ M_2 &= M/m_2 = 528 \\ M_3 &= M/m_3 = 240 \\ M_4 &= M/m_4 = 165. \end{aligned}$$

Vamos agora calcular todos os N_i tais que $M_i N_i \equiv 1 \pmod{m_i}$, usando para isso o Algoritmo de Euclides. Obtemos assim $N_1 = 1, N_2 = 2, N_3 = 5, N_4 = -3$.

Calculamos então $x = \sum_i a_i M_i N_i = 7253$ que, como se pode facilmente verificar, é uma solução do nosso sistema.

COROLÁRIO 11. *A função Phi de Euler é "multiplicativa", isto é, se $m \perp n$, $\varphi(mn) = \varphi(m)\varphi(n)$.*

Demonstração. Queremos calcular o número de inteiros entre 0 e $mn - 1$ que não têm factores comuns com mn. Para cada j em $]0, mn - 1[$, consideremos j_1 e j_2 os seus menores restos não negativos, módulo m e módulo n, respectivamente:

$$0 \leq j_1 < m, \quad 0 \leq j_2 < n, \quad j \equiv j_1 \mod m \text{ e } j \equiv j_2 \mod n$$

Pelo Teorema do Resto Chinês, conclui-se que para cada par j_1, j_2 existe um e apenas um j entre 0 e $mn-1$ tal que $j \equiv j_1 \mod m$ e $j \equiv j_2 \mod n$. Note-se que j não tem factores comuns com mn sse não tem factores comuns com m (o que é equivalente a dizer que $j_1 \perp m$), e não tem factores comuns com n (o que é equivalente a dizer que $j_2 \perp n$). Logo, os j's que queremos determinar estão numa correspondência de 1-para-1 com os pares (j_1, j_2) para os quais $0 \leq j_1 < m, 0 \leq j_2 < n$ e $j_1 \perp m, j_2 \perp n$. Como o número de j_1's nestas condições é $\varphi(m)$ e o número de j_2's é $\varphi(n)$, o número de pares é necessariamente $\varphi(m)\varphi(n)$. □

Dado que todo o n pode ser escrito como um produto finito de potências de números primos (os quais não têm factores comuns) e, dado que conhecemos a fórmula $\varphi(p^\alpha) = p^\alpha(1 - \frac{1}{p})$, podemos concluir que para $n = p_1^{\alpha_1} p_2^{\alpha_2} \cdots p_r^{\alpha_r}$,

$$\varphi(n) = p_1^{\alpha_1}(1 - \frac{1}{p_1})p_2^{\alpha_2}(1 - \frac{1}{p_2}) \cdots p_r^{\alpha_r}(1 - \frac{1}{p_r}) = n \prod_{p|n} \left(1 - \frac{1}{p}\right).$$

Teorema 12 (Pequeno Teorema de Fermat). *Se p é um primo e se n é um inteiro qualquer, então $n^p \equiv n \pmod p$. Em particular, se p não divide n, temos que $n^{p-1} \equiv 1 \pmod p$.*

Demonstração. Vamos começar por demonstrar o caso em que $p \perp n$. Sabemos que os $p-1$ números, $n \pmod p$, $2n \pmod p$, $\ldots$, $(p-1)n \pmod p$ são os números 1, 2,$\ldots$, $p-1$, não necessariamente por esta ordem. Então, se os multiplicarmos a todos, temos:

$$n \cdot (2n) \cdots ((p-1)n) \equiv (p-1)!,$$

onde a congruência é módulo p. Isto significa que:

$$(p-1)!n^{p-1} \equiv (p-1)! \pmod p,$$

e então $n^{p-1} \equiv 1 \pmod p$, dado que $(p-1)!$ não é divisível por p (ver propriedade vii) das congruências). Além disso, $n^p \equiv n \pmod p$ (pela propriedade iv)). Se $p|n$ então claramente $n^p \equiv 0 \equiv n \pmod p$. □

Vamos, de seguida, enunciar uma proposição que generaliza a proposição I3.5 de [4].

Proposição 13. *Sejam p e q dois primos distintos e $n = pq$. Sejam d e e dois inteiros positivos tais que $de \equiv 1 \mod \varphi(n)$. Então $a^{de} \equiv a \mod n$ para qualquer inteiro a.*

Demonstração. Note-se que $de \equiv 1 \mod \varphi(n) \Leftrightarrow de \equiv 1 \mod (p-1)(q-1)$, pelo que, pela propriedade v) das congruências,

$$\begin{cases} de \equiv 1 \mod (p-1) \\ de \equiv 1 \mod (q-1). \end{cases}$$

Vamos dividir a demonstração nos 3 casos possiveis:

1. $mdc(a,n) = 1$;
2. $mdc(a,n) = p$ ou $mdc(a,n) = q$;
3. $mdc(a,n) = n = pq$.

CASO 1: Se $mdc(a,n) = 1$, então $mdc(a,p) = 1$ (p não divide a) e $mdc(a,q) = 1$ (q não divide a). Pelo Pequeno Teorema de Fermat, $a^{p-1} \equiv 1 \mod p$ e então, como também $de - 1 \equiv 0 \mod p-1$, existem inteiros k, t tais que $a^{de-1} = a^{t(p-1)} = (1+kp)^t \equiv 1 \mod p$.

Analogamente, $a^{de-1} \equiv 1 \mod q$, pelo que

$$\begin{cases} a^{de-1} \equiv 1 \mod p \\ a^{de-1} \equiv 1 \mod q \end{cases} \text{ ou seja } \begin{cases} a^{de} \equiv a \mod p \\ a^{de} \equiv a \mod q. \end{cases}$$

CASO 2: Se $mdc(a,n) = p$, então existe $m \in \mathbb{Z}$ tal que $a = mp$ e q não divide a. Assim,

$$a^{de} \equiv 0 \equiv a \mod p.$$

Além disso, como vimos no caso 1, temos que $a^{de} \equiv a \mod q$, pois q não divide a. Então,

$$\begin{cases} a^{de} \equiv a \mod p \\ a^{de} \equiv a \mod q. \end{cases}$$

Se $mdc(a,n) = q$ o resultado é similar.

CASO 3: Se $mdc(a,n) = pq$, então existe $m \in \mathbb{Z}$ tal que $a = mpq$, pelo que podemos facilmente concluir que

$$\begin{cases} a^{de} \equiv 0 \equiv a \mod p \\ a^{de} \equiv 0 \equiv a \mod q. \end{cases}$$

Nos três casos obtivemos

$$\begin{cases} a^{de} \equiv a \mod p \\ a^{de} \equiv a \mod q \end{cases}$$

o que, pela propriedade vi) das congruências atrás enunciada, implica que $a^{de} \equiv a \mod n$. □

2.2 Corpos Finitos

Nesta secção vamos ver uma série de definições sobre corpos finitos. Não iremos dizer muito sobre eles, mas sim o suficiente para ficarmos com uma boa ideia do que são e percebermos o que iremos fazer mais à frente.

Definição 14. Um corpo é um conjunto $\mathcal{F}$ com operações de multiplicação e adição que satisfazem as regras usuais de associatividade, comutatividade, distributividade e existência de um elemento neutro aditivo 0, e de um multiplicativo 1. Além disso, todos os elementos têm um inverso para a adição, e todos os elementos de $\mathcal{F}\backslash\{0\}$ têm um inverso para a multiplicação.

Definição 15. Dado um corpo $\mathcal{F}$, chama-se característica de $\mathcal{F}$ ao menor inteiro positivo n (se existir) tal que $n \times 1 = 0$. Se não existir tal inteiro, diz-se que $\mathcal{F}$ tem característica zero.

Exemplo: $\mathbb{R}$, $\mathbb{Q}$ e $\mathbb{C}$ são exemplos de corpos de característica zero. $\mathbb{Z}/p\mathbb{Z}$ com p primo é um corpo (finito) com característica $p \neq 0$, denominado corpo primo.

Prova-se que, se $\mathcal{F}$ tem característica zero, então contém uma cópia do corpo dos números racionais. Se a característica de $\mathcal{F}$ não for zero, então ela é, necessariamente, um número primo, pois $\mathcal{F}$ não tem divisores de zero (se $r \times s \times 1 = 0$ então $r \times 1 = 0$ ou $s \times 1 = 0$). Neste caso, $\mathcal{F}$ contém uma cópia de $\mathbb{Z}/p\mathbb{Z}$.

Em $\mathcal{F}_q$, os $q-1$ elementos não nulos formam, por definição de corpo, um grupo abeliano com respeito à multiplicação: o produto de dois elementos diferentes de 0 é diferente de 0, a multiplicação é associativa, existe um elemento identidade 1 e qualquer elemento não nulo tem inverso.

Vamos a partir de agora designar o conjunto não vazio de elementos não nulos de $\mathcal{F}_q$ por $\mathcal{F}_q^*$. Dado um corpo $\mathcal{F}_q$, define-se a ordem de um elemento não nulo como sendo a menor potência positiva desse elemento que é igual a 1. Temos então a seguinte proposição:

Proposição 16. *A ordem de qualquer elemento $a \in \mathcal{F}_q^*$ divide $(q-1)$.*

Demonstração. Note-se que existe sempre uma potência finita de a igual a 1. Com efeito, como as potências de a no conjunto finito $\mathcal{F}_q^*$ não podem ser todas distintas, temos $a^i = a^j$ para algum $j > i$ e então $a^{j-i} = 1$. Seja d a menor potência de a igual a 1. Considere-se o conjunto $S = \{1, a, a^2, \ldots, a^{d-1}\}$ de todas as potências de a e, dado $b \in \mathcal{F}_q^*$, considere-se o conjunto bS de todos os elementos da forma ba^j (por exemplo, $1S = S$). É fácil ver que quaisquer dois destes conjuntos ou são iguais ou são disjuntos. Com efeito, se algum b_1a^i estiver em b_1S e também em b_2S então $b_1a^i = b_2a^j$ e, então, qualquer

elemento $b_1a^{i'}$ de b_1S está em b_2S (pois $b_1a^{i'} = b_1a^ia^{i'-i} = b_2a^{j+i'-i}$). Como cada um destes conjuntos contém exactamente d elementos e a união de todos os conjuntos do tipo bS é igual a $\mathcal{F}_q^*$, temos que $\mathcal{F}_q^*$ é uma união disjunta de conjuntos com d elementos e então $d|(q-1)$. □

PROPOSIÇÃO 17. *O inverso de um elemento a em $\mathbb{Z}/p\mathbb{Z}$ com p primo é a^{p-2}.*

Demonstração. É fácil ver que assim é, recorrendo ao Pequeno Teorema de Fermat enunciado e demonstrado anteriormente. Em geral, o inverso de um número a é tal que $a^{-1}a = 1$. Em $\mathbb{Z}/p\mathbb{Z}$ podemos escrever isto da forma $a^{-1}a \equiv 1 \mod p$. Então, dado que pelo pequeno teorema de Fermat $a^{p-1} \equiv 1 \pmod p$ (p primo com a), temos $a^{p-2}a \equiv 1 \mod p$. Conclui-se assim que o inverso de a em $\mathbb{Z}/p\mathbb{Z}$ é a^{p-2}. □

DEFINIÇÃO 18. Chama-se gerador de um corpo finito $\mathcal{F}_q$, a um elemento de ordem $(q-1)$ ou, equivalentemente, a um elemento cujas potências percorrem todos os elementos de $\mathcal{F}_q^*$.

Todos os corpos finitos têm um gerador. Se g é um gerador de $\mathcal{F}_q^*$, então g^j é também um gerador sse $j \perp (q-1)$. Em particular, existe um total de $\varphi(q-1)$ geradores diferentes de $\mathcal{F}_q^*$.

3 Criptografia Simples

Nesta secção falaremos sobre criptografia simples, na perspectiva de que, não só é muito simples percebê-la, como também utilizá-la; ou seja, é simples[2] descodificar as mensagens codificadas com este tipo de criptografia. O que aqui vamos ver são algumas formas de criptografia de chave privada.

A criptografia de chave privada é muito antiga. Quando Júlio César enviava mensagens aos seus generais (ele não confiava nos seus generais), substituía todos os A's por D's, B's por E's, e assim sucessivamente. Só alguém que soubesse o sistema de 'translação por 3'[3], poderia decifrar a mensagem.

É então assim que começamos...

3.1 Utilização de Funções Afins

Ao longo desta secção vamos dar início à criptografia usando funções afins (o método mais simples de todos).

2 Pelo menos mais simples do que na da próxima secção.

3 Shift by 3.

Chamamos sistema criptográfico a uma configuração do tipo:

$$\mathcal{T} \xrightarrow{f} \mathcal{C} \xrightarrow{f^{-1}} \mathcal{T},$$

onde $\mathcal{T}$ é o nosso texto original, a partir do qual obtemos o código $\mathcal{C}$ quando aplicamos uma função f. Após aplicar f^{-1} ao nosso código, recuperamos o nosso texto original. É fácil ver que f deverá ser injectiva para que a sua inversa possa ser aplicada sobre $\mathcal{C}$ sem que exista qualquer tipo de escolha.

Podemos então associar a cada letra do alfabeto inglês (com 26 letras) um número de 0 a 25. À letra A corresponde o número 0, a B o número 1, etc. Neste caso, diz-se que as nossas unidades de mensagem são mono-alfabéticas e a função usada poderia ser, por exemplo, a função identidade.

No entanto, podemos ter também unidades bi-alfabéticas sobre o alfabeto anterior com o espaço entre palavras. Neste caso, usando a mesma associação de uma letra a um número, e considerando o espaço como o último caracter do alfabeto, podemos usar a função:

$$f(x, y) = 27x + y \quad \in \{0, 1, \ldots, 728\}.$$

Por exemplo, codificamos "NO" com o número: $27 \times 13 + 14 = 365$. Generalizando, podemos usar unidades de mensagem de qualquer tamanho, bastando para isso ter a nossa função injectiva. Por exemplo, para unidades tri-alfabéticas teriamos:

$$f(x, y, z) = 729x + 27y + z \quad \in \{0, 1, \ldots, 19682\}.$$

Verifica-se facilmente que, em geral, para um alfabeto de k caracteres e uma mensagem m-alfabética,

$$f(x_1, x_2, \ldots, x_m) = k^{m-1}x_1 + k^{m-2}x_2 + \cdots + k^0 x_m = \sum_{i=1}^{m} k^{m-i}x_i.$$

Podemos, no entanto, verificar que começamos a trabalhar com números muito grandes e cálculos muito pesados, não sendo por isso conveniente para encriptar e decriptar mensagens. Podemos então recorrer ao conjunto de inteiros $\mathbb{Z}/N\mathbb{Z}$ e utilizar as operações de adição e multiplicação módulo N.

Como introdução a esta ideia, vejamos o sistema de César, que usa uma translação alfabética de 3 letras e cuja função é $\mathcal{C} = f(\mathcal{T}) = \mathcal{T} + 3 \mod 26$, onde 3 é o deslocamento e 26 é o tamanho do alfabeto. Para com $\mathcal{C}$ voltar a encontrar $\mathcal{T}$ basta achar a nossa função inversa, tarefa que é bastante simples: $\mathcal{T} = f^{-1}(\mathcal{C}) = \mathcal{C} - 3 \mod 26$.

Em geral, podemos definir formalmente um sistema deste tipo da forma seguinte:

DEFINIÇÃO 19. Um sistema criptográfico de translação é um sistema dado por:

$$f(\mathcal{T}) = \mathcal{T} + k \mod m$$
$$f^{-1}(\mathcal{C}) = \mathcal{C} - k \mod m,$$

onde $\mathcal{T}$ é o nosso texto original (separado em unidades de mensagem), $\mathcal{C}$ é o código respectivo à unidade de mensagem que foi encriptada e k é a chave de translação ($0 \leq k \leq m$, onde m é o tamanho do nosso alfabeto).

É possível verificar a existência de casos particulares para os quais $f = f^{-1}$. Esses casos ocorrem quando $k/2 = m$. Vamos ver um exemplo, que utiliza uma aplicação denominada `rot13`[4]. Esta usa o seguinte sistema criptográfico:

$$f(\mathcal{T}) = \mathcal{T} + 13 \mod 26$$
$$f^{-1}(\mathcal{C}) = \mathcal{C} + 13 \mod 26.$$

EXEMPLO: Usando a aplicação `rot13` mencionada anteriormente e a palavra "LISP" verificamos rapidamente que:

$$LISP \xrightarrow{rot13} YVFC$$
$$YVFC \xrightarrow{rot13} LISP$$

Utilizando a mesma função, podemos encriptar e decriptar a nossa mensagem. Vejamos os detalhes desta operação:

Encriptação							
L	→	11	$11 + 13 \equiv 24 \pmod{26}$	24	→	Y
I	→	8	$8 + 13 \equiv 21 \pmod{26}$	21	→	V
S	→	18	$18 + 13 \equiv 5 \pmod{26}$	5	→	F
P	→	15	$15 + 13 \equiv 2 \pmod{26}$	2	→	C
Decriptação							
Y	→	24	$24 + 13 \equiv 11 \pmod{26}$	11	→	L
V	→	21	$21 + 13 \equiv 8 \pmod{26}$	8	→	I
F	→	5	$5 + 13 \equiv 18 \pmod{26}$	18	→	S
C	→	2	$2 + 13 \equiv 15 \pmod{26}$	15	→	P

Podemos também utilizar unidades de mensagem bi-alfabéticas. Se o nosso texto original possuir um número ímpar de letras, acrescentamos ao texto um símbolo que não cause confusão após a decriptação (por exemplo, um espaço, "X", ou "Q").

4 Disponível na maioria dos sistemas UNIX.

Cada mensagem bi-alfabética terá, neste caso, um equivalente númerico dado por $\mathcal{T} = xN + y$ (onde N é o tamanho do nosso alfabeto). Podemos então considerar a seguinte transformação no conjunto de inteiros $\mathbb{Z}/N^2\mathbb{Z}$,

$$\mathcal{C} \equiv a\mathcal{T} + b \pmod{N^2}$$
$$\mathcal{T} \equiv a^{-1}\mathcal{C} - a^{-1}b \pmod{N^2},$$

onde $a \perp N$ (de modo a que a tenha inverso no anel $\mathbb{Z}/N^2\mathbb{Z}$). Apesar dos sistemas criptográficos afins com mensagens bi-alfabéticas (ou seja, $\bmod N^2$) serem melhores, também têm alguns problemas. A segunda letra de cada mensagem bi-alfabética de $\mathcal{C}$ depende apenas do valor $\bmod N$ de $\mathcal{C} \equiv a\mathcal{T} + b \pmod{N^2}$ que, por sua vez, depende só de $\mathcal{T} \pmod{N}$, isto é, apenas da segunda letra da mensagem bi-alfabética do texto. Logo poderíamos obter muita informação (a e $b \pmod{N}$) a partir de uma análise de frequências em letras pares do nosso $\mathcal{C}$. O mesmo acontece para transformações afins $\bmod N^k$ em blocos de k letras.

Exemplo: Vamos, mais uma vez, encriptar a mensagem "LISP", usando agora funções afins no alfabeto de 26 letras ($N = 26$) com unidades de mensagem bi-alfabéticas. Note que, como só queremos encriptar uma palavra, não é necessário considerar o espaço como último caracter do alfabeto. Dividindo "LISP" em unidades bi-alfabéticas, obtemos "LI" e "SP". Assumindo, por exemplo, $a = 5$ (que verifica a condição $a \perp N$) e $b = 10$, temos:

Encriptação		
$\{\text{L, I}\} \to \{11, 8\}$	$N^2 = 676$ $11 \times 26 + 8 = 294 = \mathcal{T}_1$ $5 \times 294 + 10 \equiv 128 \pmod{N^2}$ $128 = x'N + y'$	$\{x', y'\} \to \{4, 24\}$
$\{4, 24\} \to \{E, Y\}$	$\{\text{E, Y}\} = \mathcal{C}_1$	
$\{\text{S, P}\} \to \{18, 15\}$	$18 \times 26 + 15 = 483 = \mathcal{T}_2$ $5 \times 483 + 10 \equiv 397 \pmod{N^2}$ $397 = x'N + y'$	$\{x', y'\} \to \{15, 7\}$
$\{15, 7\} \to \{P, H\}$	$\{\text{P, H}\} = \mathcal{C}_2$ $\mathcal{C} = \text{EYPH}$	
Decriptação		
$\{\text{E, Y}\} \to \{4, 24\}$	$a^{-1} \equiv 541 \pmod{N^2}$ $4 \times 26 + 24 = 128 = \mathcal{C}_1$ $541 \times 128 + 674 \equiv 294 \pmod{N^2}$ $294 = xN + y$	$\{x, y\} \to \{11, 8\}$
$\{11, 8\} \to \{L, I\}$	$\{\text{L, I}\} = \mathcal{T}_1$	

$\{\text{P, H}\} \to \{15, 7\}$	$15 \times 26 + 7 = 397 = \mathcal{C}_2$	
	$541 \times 397 + 674 \equiv 483 \pmod{N^2}$	
	$483 = xN + y$	$\{x, y\} \to \{18, 15\}$
$\{18, 15\} \to \{S, P\}$	$\{\text{S, P}\} = \mathcal{T}_2$	
	$\mathcal{T} = \text{LISP}$	

3.2 Utilização de Matrizes

Quando utilizamos um alfabeto de N símbolos e mensagens bi-alfabéticas, podemos representá-los como vectores em vez de os representar como elementos de $\mathbb{Z}/N^2\mathbb{Z}$ como foi feito anteriormente. Vamos explorar e aprofundar um pouco esta possibilidade.

Começamos por representar cada mensagem bi-alfabética como um ponto numa tabela $N \times N$. Obtemos assim um "plano-xy", mas em cada eixo, em vez de termos uma cópia da recta real temos uma cópia de $\mathbb{Z}/N\mathbb{Z}$. Assim, a nossa tabela será designada por $(\mathbb{Z}/N\mathbb{Z})^2$. Uma transformação de encriptação será então uma função injectiva de $(\mathbb{Z}/N\mathbb{Z})^2$ em si próprio.

Tal como nos exemplos atrás, existem dois tipos simples de transformações:

1. Lineares: $\mathcal{C} \equiv A\mathcal{T}$ onde A é uma matriz invertível em $(\mathbb{Z}/N\mathbb{Z})^2$.
2. Afins: $\mathcal{C} \equiv A\mathcal{T} + b$ onde A é uma matriz invertível em $(\mathbb{Z}/N\mathbb{Z})^2$.

Vejamos a seguinte proposição de fácil demonstração, que nos deixará mais à vontade para discutir o que se segue:

PROPOSIÇÃO 20. *Seja:*

$$\mathbf{A} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in M_2(\mathbb{Z}/N\mathbb{Z}) \text{ e seja } D = ad - bc.$$

As seguintes afirmações são equivalentes:

1. $mdc(D, N) = 1$;
2. *matriz A tem inversa no anel $M_2(\mathbb{Z}/N\mathbb{Z})$;*
3. *se* $\begin{pmatrix} x \\ y \end{pmatrix} \neq \begin{pmatrix} 0 \\ 0 \end{pmatrix}$, *então* $A \begin{pmatrix} x \\ y \end{pmatrix} \neq \begin{pmatrix} 0 \\ 0 \end{pmatrix}$;
4. *a matriz A dá-nos uma correspondência biunívoca de $(\mathbb{Z}/N\mathbb{Z})^2$ em si próprio.*

Com base na proposição anterior podemos obter transformações lineares de encriptação dos nossos vectores bi-alfabéticos utilizando matrizes $A \in M_2(\mathbb{Z}/N\mathbb{Z})$ com determinante D, tal que $D \perp N$. Cada unidade de mensagem $\mathcal{T} = \binom{x}{y}$ é transformada numa unidade codificada $\mathcal{C} = \binom{x'}{y'}$ da forma seguinte:

$$\mathcal{C} = A\mathcal{T}, \qquad \Leftrightarrow \quad \begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}.$$

Para decifrar a mensagem basta aplicar a matriz inversa:

$$\mathcal{T} = A^{-1}AT = A^{-1}C, \qquad \Leftrightarrow \quad \begin{pmatrix} x \\ y \end{pmatrix} = D^{-1} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix} \begin{pmatrix} x' \\ y' \end{pmatrix}.$$

É de notar que, para encriptar um texto com k mensagens bi-alfabéticas $\mathcal{T} = \mathcal{T}_1\mathcal{T}_2\mathcal{T}_3\cdots\mathcal{T}_k$, podemos escrever os k vectores como colunas de uma matriz $2 \times k$, $\mathcal{T}$, e depois multiplicá-la à esquerda pela matriz $A_{2\times 2}$ obtendo a matriz $\mathcal{C}_{2\times k} = A\mathcal{T}$ de vectores encriptados.

Exemplo: Vamos utilizar de novo a nossa já conhecida mensagem "LISP" e encriptá-la usando matrizes. Escolhamos, por exemplo, a matriz

$$A = \begin{pmatrix} 2 & 3 \\ 7 & 8 \end{pmatrix}$$

cujo determinante é -5 (e $-5 \perp 26$).

Encriptação

$$\begin{aligned} \text{"LISP"} &= \begin{pmatrix} 11 & 18 \\ 8 & 15 \end{pmatrix} \\ \mathcal{C} &\equiv A\mathcal{T} \equiv \begin{pmatrix} 20 & 3 \\ 11 & 12 \end{pmatrix} \pmod{26} \\ \mathcal{C} &= \text{"ULDM"} \end{aligned}$$

Decriptação

$$\begin{aligned} \text{"ULDM"} &= \begin{pmatrix} 20 & 3 \\ 11 & 12 \end{pmatrix} \\ \mathcal{T} &\equiv A^{-1}\mathcal{C} \\ &\equiv \begin{pmatrix} 14 & 11 \\ 17 & 10 \end{pmatrix} \begin{pmatrix} 20 & 3 \\ 11 & 12 \end{pmatrix} \\ &\equiv \begin{pmatrix} 11 & 18 \\ 8 & 15 \end{pmatrix} \pmod{26} \\ \mathcal{T} &= \text{"LISP"} \end{aligned}$$

Podemos também utilizar transformações afins para encriptar mensagens. Dada uma mensagem bi-alfabética $\mathcal{T} = \binom{x}{y}$, basta multiplicá-la à esquerda por uma matriz $A_{2\times 2} \in M_2(\mathbb{Z}/N\mathbb{Z})$ invertível e adicionar um vector constante $B = \binom{e}{f}$:

$$\begin{array}{rcl} \mathcal{C} & \equiv & A\mathcal{T} + B \qquad \Leftrightarrow \\ \begin{pmatrix} x' \\ y' \end{pmatrix} & \equiv & \begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} e \\ f \end{pmatrix} \equiv \begin{pmatrix} ax+by+e \\ cx+dy+f \end{pmatrix}. \end{array}$$

A transformação inversa obtém-se subtraindo B e multiplicando A^{-1} à esquerda,

$$\mathcal{T} = A^{-1}(\mathcal{C} - B).$$

EXEMPLO: Vamos utilizar uma transformação afim com matrizes para encriptar e decriptar a nossa mensagem "LISP", escolhendo os parâmetros para o nosso sistema:

$$A = \begin{pmatrix} 2 & 3 \\ 7 & 8 \end{pmatrix} \quad \text{e} \quad B = \begin{pmatrix} 5 \\ 14 \end{pmatrix}$$

Encriptação

$$\begin{array}{rcl} \text{"LISP"} & = & \begin{pmatrix} 11 & 18 \\ 8 & 15 \end{pmatrix} \\ \mathcal{C}_1 & \equiv & \begin{pmatrix} 2 & 3 \\ 7 & 8 \end{pmatrix}\begin{pmatrix} 11 \\ 8 \end{pmatrix} + \begin{pmatrix} 5 \\ 14 \end{pmatrix} \equiv \begin{pmatrix} 25 \\ 25 \end{pmatrix} \\ \mathcal{C}_2 & \equiv & \begin{pmatrix} 2 & 3 \\ 7 & 8 \end{pmatrix}\begin{pmatrix} 18 \\ 15 \end{pmatrix} + \begin{pmatrix} 5 \\ 14 \end{pmatrix} \equiv \begin{pmatrix} 8 \\ 0 \end{pmatrix} \\ \mathcal{C} & = & \text{"ZZIA"} \end{array}$$

Decriptação

$$\begin{array}{rcl} \text{"ZZIA"} & = & \begin{pmatrix} 25 & 8 \\ 25 & 0 \end{pmatrix} \\ \mathcal{T}_1 & \equiv & \begin{pmatrix} 14 & 11 \\ 17 & 10 \end{pmatrix}\begin{pmatrix} 25 \\ 25 \end{pmatrix} - \begin{pmatrix} 14 & 11 \\ 17 & 10 \end{pmatrix}\begin{pmatrix} 5 \\ 14 \end{pmatrix} \equiv \begin{pmatrix} 11 \\ 8 \end{pmatrix} \\ \mathcal{T}_2 & \equiv & \begin{pmatrix} 14 & 11 \\ 17 & 10 \end{pmatrix}\begin{pmatrix} 8 \\ 0 \end{pmatrix} - \begin{pmatrix} 14 & 11 \\ 17 & 10 \end{pmatrix}\begin{pmatrix} 5 \\ 14 \end{pmatrix} \equiv \begin{pmatrix} 18 \\ 15 \end{pmatrix} \\ \mathcal{T} & = & \text{"LISP"} \end{array}$$

Resumindo os exemplos desta secção temos:

Resultados	**Método**
"LISP" $\longrightarrow$ "YVFC"	Translação mono-alfabética
"LISP" $\longrightarrow$ "EYPH"	Afim mono-alfabética
"LISP" $\longrightarrow$ "ULDM"	Linear matricial bi-alfabética
"LISP" $\longrightarrow$ "ZZIA"	Afim matricial bi-alfabética

4 Criptografia no Presente

Neste capítulo veremos alguns algoritmos interessantes, que são hoje utilizados nos mais variados sítios. Estes algoritmos são os chamados algoritmos de encriptação com chave pública. Poder-se-á encontrar mais informação em [4] para uma abordagem mais matemática, ou em [6] para uma abordagem mais algorítmica.

4.1 RSA

Este é um dos mais antigos algoritmos de chave pública mas é, no entanto, um dos mais populares. A sua designação vem dos nomes dos seus inventores Rivest, Shamir e Adleman. A popularidade deste algoritmo vem da grande diferença entre a facilidade em arranjar um número primo grande e a extrema dificuldade de factorizar o produto de dois números primos grandes.

Criação das Chaves

A criação de chaves é muito simples e todos os utilizadores têm que criar o seu par de chaves. Cada utilizador escolhe 2 primos p e q (com cerca de 100 dígitos decimais cada) e faz $n = pq$. Sabendo a factorização de n é fácil calcular $\varphi(n) = (p-1)(q-1) = n + 1 - p - q$. Depois o utilizador escolhe aleatoriamente um número e, entre 1 e $\varphi(n)$, tal que $e \perp \varphi(n)$ e, em seguida, calcula $d = e^{-1} \mod \varphi(n)$. Temos assim as chaves calculadas. A chave de encriptação corresponde ao tuplo (n, e) (e é tornada pública) e a chave de decriptação corresponde ao tuplo (n, d) (e é mantida secreta).

Exemplo: Como não nos interessa, neste momento, preocuparmo-nos com a possibilidade de quebrar o algoritmo, vamos escolher números pequenos para ser mais fácil compreender o processo.

$$\begin{aligned} p &= 5 \text{ e } q = 11 \\ n &= 5 \times 11 = 55 \\ \varphi(n) &= 55 + 1 - 5 - 11 = 40. \end{aligned}$$

Escolhendo, por exemplo, $e = 3$ (em que $e \perp \varphi(n)$) temos

$$d \equiv e^{-1} \equiv 27 \pmod{40}.$$

A chave de encriptação é então $(n, e) = (55, 3)$ e a chave de decriptação é $(n, d) = (55, 27)$.

Procedimento de Encriptação e Decriptação

Dados os tuplos de encriptação e decriptação, (n, e) e (n, d), podemos agora analisar como se fazem as respectivas operações.

A operação de encriptação é dada pela transformação de $\mathbb{Z}/n\mathbb{Z}$ em si próprio, dada por:

$$f(\mathcal{T}) \equiv \mathcal{C} \equiv \mathcal{T}^e \pmod{n}$$

e a transformação de decriptação é dada por:

$$f^{-1}(\mathcal{C}) \equiv \mathcal{T} \equiv \mathcal{C}^d \pmod{n}.$$

Cada uma destas transformações é a inversa da outra. Com efeito, devido à nossa escolha de d ($de \equiv 1 \pmod{\varphi(n)}$) temos pela proposição 13 que

$$k \xrightarrow{f} k^e \xrightarrow{f^{-1}} (k^e)^d = k^{ed} \equiv k \pmod{n}.$$

EXEMPLO: Vamos ver agora como encriptar e decriptar a já nossa conhecida mensagem "LISP" usando RSA.

Encriptação							
L	→	11	$11^3 \equiv 11 \pmod{55}$	11	→	L
I	→	8	$8^3 \equiv 17 \pmod{55}$	17	→	R
S	→	18	$18^3 \equiv 2 \pmod{55}$	2	→	C
P	→	15	$15^3 \equiv 20 \pmod{55}$	20	→	U
Decriptação							
L	→	11	$11^{27} \equiv 11 \pmod{55}$	11	→	L
R	→	17	$17^{27} \equiv 8 \pmod{55}$	8	→	I
C	→	2	$2^{27} \equiv 18 \pmod{55}$	18	→	S
U	→	20	$20^{27} \equiv 15 \pmod{55}$	15	→	P

É de notar que, como $n > k$ (onde k é o tamanho do nosso alfabeto), após a encriptação podemos obter mensagens não representáveis no alfabeto. No entanto, este inconveniente é resolvido na prática enviando as mensagens em blocos.

Assinatura Digital

Imaginemos agora que queremos enviar uma mensagem com l bits acompanhada da nossa assinatura (muito menor) formada por k bits com $k \ll l$. Para isso vamos utilizar uma função de dispersão.

Definição 21. Uma função $\mathcal{H}(x) : \{0,1\}^l \to \{0,1\}^k$ diz-se de dispersão se for fácil calcular $\mathcal{H}(x)$ para qualquer x, mas:

- for difícil conseguir encontrar dois valores distintos x e x' tais que $\mathcal{H}(x) = \mathcal{H}(x')$ ("resistência à colisão") e,
- dado y (uma imagem de $\mathcal{H}$), for difícil encontrar x tal que $\mathcal{H}(x) = y$ ("resistência da imagem").

Vejamos então como tudo funciona. Suponhamos que o Carlos tenta enviar à Ana uma mensagem $\mathcal{T}$ com l símbolos e que ambos estão a usar a mesma função de dispersão (a qual não é necessariamente secreta). Depois do Carlos enviar a mensagem $\mathcal{T}$ à Ana, ele envia-lhe também o valor de dispersão $\mathcal{H}(\mathcal{T})$. A Ana gostaria de ter a certeza de que foi realmente o Carlos a enviar-lhe a mensagem $\mathcal{T}$ e que a Catarina (pessoa não autorizada que pode interceptar a mensagem) não a alterou antes da Ana a receber. Suponhamos que ela tem a certeza que, pelo menos, o valor de dispersão vem do Carlos. Nesse caso, tudo o que ela tem a fazer é aplicar a função de dispersão à mensagem recebida. Se o resultado obtido estiver em concordância com o valor de dispersão recebido ela sabe que a Catarina não modificou a mensagem $\mathcal{T}$, pois seria difícil produzir uma versão distorcida $\mathcal{T}'$ tal que $\mathcal{H}(\mathcal{T}) = \mathcal{H}(\mathcal{T}')$. O problema que falta resolver, e que discutiremos em seguida, é o de como a Ana pode ter a certeza que o valor de dispersão vem realmente do Carlos. Este problema pode ser resolvido usando RSA. Por conveniência, escolhemos k, de modo a que uma sequência de k bits seja suficientemente pequena para preencher uma unidade de mensagem. O Carlos, depois de calcular o valor de dispersão $\mathcal{H} = \mathcal{H}(\mathcal{T})$ para a sua mensagem, não o envia simplesmente com a mensagem (encriptada ou não, como ele preferir). Ele eleva $\mathcal{H}$ à potência da sua chave de decriptação. Então, o que o Carlos envia à Ana não é $\mathcal{H}$, mas sim $\mathcal{H}' = \mathcal{H}^{d_{\text{Carlos}}} \pmod{n_{\text{Carlos}}}$. Depois de a Ana receber a mensagem (e de a decriptar, caso o Carlos a tenha encriptado), ela pega na última unidade de mensagem (que se parece com lixo) e eleva-a ao expoente de encriptação e_{Carlos}, de modo a recuperar $\mathcal{H}$. Depois, calcula o valor de dispersão da mensagem recebida e verifica se o valor calculado coincide com o recebido. Há que notar que a Ana sabe que *apenas* o Carlos saberia o expoente que é invertido, elevando-o à potência e_{Carlos}. Logo, ela sabe que foi realmente o Carlos que lhe enviou $\mathcal{H}$ e sabe também que a mensagem não foi modificada.

Note-se que esta assinatura tem a propriedade da não-repudiação, ou seja, o Carlos não pode, depois de enviar a mensagem, negar que a enviou.

Conclusão

Se a Catarina for de novo a pessoa não autorizada que tenta interceptar a mensagem encriptada $\mathcal{C}$, o que é que (tendo $\mathcal{C}$ e a chave pública (n, e)) iria impedi-la de aceder a $\mathcal{T}$? O problema da Catarina seria que sem saber os factores p e q de n não existe, aparentemente, nenhuma maneira de encontrar o expoente d que inverta a função de encriptação. Além disso, também parece não haver maneira de inverter a encriptação sem o expoente de decriptação. Usamos aqui as palavras "aparentemente" e "parece" dado que estas afirmações não foram provadas. Assim, *aparentemente* quebrar um cripto-sistema RSA é tão difícil como factorizar n. Para terminar, vejamos uma tabela de resumo do algoritmo:

	Chave Pública
$\mathcal{T}$	mensagem
n	produto de dois primos, p e q (p e q têm que permanecer secretos)
e	primo relativamente a $(p-1)(q-1)$
	Chave Privada
d	$\equiv e^{-1} \pmod{((p-1)(q-1))}$
	Encriptação
$\mathcal{C}$	$\equiv \mathcal{T}^e \pmod{n}$
	Decriptação
$\mathcal{T}$	$\equiv \mathcal{C}^d \pmod{n}$

4.2 Logaritmo Discreto

Vejamos agora um outro tipo de algoritmo que se baseia no chamado *Problema do Logaritmo Discreto.*

O algoritmo de RSA baseia-se (tal como já foi mencionado) no facto de que é fácil multiplicar dois números primos muito grandes, mas não o contrário. Existem outros problemas fundamentais em teoria dos números que se baseiam nesta propriedade da uni-direccionalidade da função (funções denominadas usualmente *funções de sentido único*). Um dos problemas mais importantes é o das potências em corpos finitos.

Quando trabalhamos sobre o corpo dos reais, a exponenciação não é significativamente mais fácil que a operação inversa. No entanto, num corpo finito (por exemplo $\mathbb{Z}/n\mathbb{Z}$) é simples calcular b^x mas, dado um elemento y que sabe-

mos ser da forma b^x (supondo que a base é fixa), como poderemos encontrar uma potência de b igual a y, isto é, como poderemos calcular $x = \log_b y$ nesse corpo? Esta questão é conhecida como o *problema do logaritmo discreto*. A palavra *discreto* distingue o logaritmo num corpo finito do logaritmo clássico (que é contínuo).

Definição 22. Se $\mathcal{G}$ é um corpo finito, $b \in \mathcal{G}$ e $y \in \mathcal{G}$ é uma potência de b, então o *logaritmo discreto* com base b de y é um inteiro x tal que $b^x = y$.

O Sistema de Troca de Chaves Diffie–Hellman

Como os sistemas de chave pública são relativamente mais lentos e pesados computacionalmente que os sistemas de chave privada, eles podem ser conjugados. Costuma dizer-se que: "A melhor maneira de quebrar um sistema de chave privada é roubando a chave". E se nós usássemos um sistema de chave pública para transmitir a nossa chave privada e depois usássemos o sistema de chave privada correspondente a essa chave para transmitir a mensagem? A primeira proposta detalhada deste tipo, usando logaritmos discretos, deve-se a W. Diffie e M. E. Hellman.

Vamos supor que a nossa chave privada é um número (ou uma colecção deles) aleatório muito grande. O método de Diffie–Hellman vai gerar um número aleatório num corpo finito $\mathcal{F}_q$ muito grande. Suponhamos que q é público (todos sabem em que corpo finito estará a nossa chave) e que $g \in \mathcal{F}_q$ também é público. Idealmente, g deverá ser um gerador de $\mathcal{F}_q$, mas isso não é absolutamente necessário. O método que descreveremos para criar a chave privada irá apenas gerar chaves que são potências de g (por isso convém que g seja um gerador).

Imaginemos então que a Catarina e o Nuno querem criar uma chave (ou seja, um elemento aleatório de $\mathcal{F}_q^*$) que utilizarão para encriptar e decriptar as mensagens que vão trocar entre si. A Catarina escolhe um inteiro aleatório a entre 1 e $q-1$, que mantém secreto, e calcula $g^a \in \mathcal{F}_q^*$ que torna público. O Nuno procede da mesma forma, escolhendo um inteiro aleatório b e tornando público g^b. A chave privada será g^{ab}.

Exemplo: Vamos criar uma chave com a qual a Catarina e o Nuno possam comunicar. Como, neste momento, o que nos interessa é o método, e este é independente dos números escolhidos, usaremos números pequenos para simplificar o exemplo mas, na prática, os números utilizados são muito maiores.

	Catarina	Nuno	Público
Corpo			$q = 19$
Gerador			$g = 2$
Número	$a = 7$		$g^a \equiv 14 \pmod{19}$
Número		$b = 13$	$g^b \equiv 3 \pmod{19}$
Chave	$(g^b)^a \equiv 2 \pmod{19}$	$(g^a)^b \equiv 2 \pmod{19}$	

É de notar que ambos ficam a conhecer a chave, mas que mais ninguém a consegue calcular apenas com a informação pública. O g escolhido, neste caso, é gerador de $\mathcal{F}_{19}$. No entanto, e tal como já foi mencionado, isso não é estritamente necessário. Após esta troca de chaves, o Nuno e a Catarina podem comunicar usando um sistema criptográfico de chave privada sem que ninguém, que não conheça a chave, consiga interceptar e quebrar o sistema. Isto é *quase* assegurado pela Suposição de Diffie-Hellman.

SUPOSIÇÃO DE DIFFIE-HELLMAN: Calcular g^{ab} sabendo apenas g^a e g^b é computacionalmente impossível.

A suposição de Diffie-Hellman é, *a priori*, pelo menos tão forte como a suposição que os logaritmos discretos não são facilmente calculados no corpo utilizado. Ou seja, se os logaritmos discretos podem ser calculados, então a suposição de Diffie-Hellman falha. Algumas pessoas conjecturam que a implicação contrária também é verdadeira, mas isso é ainda uma questão em aberto.

Cripto-Sistema de Massey-Omura

Vamos supor que todos os utilizadores do sistema conhecem o corpo $\mathcal{F}_q$ (fixo) sobre o qual se está a enviar mensagens. Cada utilizador escolhe então $e \perp (q-1)$ e calcula $d \equiv e^{-1} \pmod{q-1}$. Imaginemos agora que a Catarina deseja enviar uma mensagem $\mathcal{T}$ ao Nuno. Em primeiro lugar, ela deverá calcular $\mathcal{T}^{e_{Cat}} \pmod{q}$ e enviá-la. O Nuno recebe então uma sequência de caracteres incompreensível. No entanto, eleva $\mathcal{T}^{e_{Cat}}$ a e_{Nuno} e envia o resultado de novo para a Catarina. Quando esta recebe $\mathcal{T}^{e_{Cat}e_{Nuno}}$ eleva a mensagem a d_{Cat} e devolve o resultado ao Nuno, o qual já pode achar $\mathcal{T}$ porque conhece d_{Nuno}. A tabela seguinte tentará explicar este procedimento.

Catarina	$\xrightarrow{\mathcal{T}}$	Nuno
	q conhecido	
$e_{Cat} \perp (q-1)$		$e_{Nuno} \perp (q-1)$
$d_{Cat} \equiv e_{Cat}^{-1} \pmod{q-1}$		$d_{Nuno} \equiv e_{Nuno}^{-1} \pmod{q-1}$
$\mathcal{T}$		
$\mathcal{T}^{e_{Cat}}$	$\longrightarrow$	$\mathcal{T}^{e_{Cat}}$
$\mathcal{T}^{e_{Cat}e_{Nuno}}$	$\longleftarrow$	$(\mathcal{T}^{e_{Cat}})^{e_{Nuno}}$
$(\mathcal{T}^{e_{Cat}e_{Nuno}})^{d_{Cat}}$	$\longrightarrow$	$\mathcal{T}^{e_{Nuno}}$
		$(\mathcal{T}^{e_{Nuno}})^{d_{Nuno}} = \mathcal{T}$

Este sistema utiliza conceitos bastante simples e é, ele próprio, bastante simples. No entanto, é necessário ter um cuidado especial. É necessário ter um bom esquema de assinatura. Caso contrário, qualquer pessoa que não seja suposto ver a mensagem pode interceptá-la e passar-se por receptor utilizando o seu valor de e. É também importante que, após um utilizador (por exemplo o Nuno) ter decriptado várias mensagens $\mathcal{T}$, e sabendo vários pares $(\mathcal{T}, \mathcal{T}^{e_{Cat}})$, não consiga usar essa informação para determinar e_{Cat}. Com efeito, suponhamos que o Nuno conseguia resolver o problema do logaritmo discreto em $\mathcal{F}_q^*$, determinando a partir de $\mathcal{T}$ e $\mathcal{T}^{e_{Cat}}$ o valor de e_{Cat}. Nesse caso, ele rapidamente conseguiria calcular d_{Cat} e, a partir daí, interceptar e decriptar todas as mensagens vindas da Catarina, fossem para ele ou não.

Exemplo: Vamos então, usando o sistema de Massey-Omura ver como se desenrolaria o processo se a Catarina tivesse de enviar a mensagem "LISP" ao Nuno.

Catarina	$\xrightarrow{\mathcal{T}}$	Nuno
	$q = 19$	
$e_{Cat} = 11$		$e_{Nuno} = 7$
$d_{Cat} \equiv 5 \pmod{q-1}$		$d_{Nuno} \equiv 13 \pmod{q-1}$
$\mathcal{T} = \langle 11, 8, 18, 15\rangle$		
$\mathcal{T}^{e_{Cat}} = \langle 7, 12, 18, 3\rangle$	$\longrightarrow$	$(\mathcal{T}^{e_{Cat}})^{e_{Nuno}} = \langle 7, 12, 18, 3\rangle^{e_{Nuno}}$
$(\mathcal{T}^{e_{Cat}e_{Nuno}})^{d_{Cat}} = \langle 7, 12, 18, 2\rangle^{d_{Cat}}$	$\longleftarrow$	$\mathcal{T}^{e_{Cat}e_{Nuno}} = \langle 7, 12, 18, 2\rangle$
$\mathcal{T}^{e_{Nuno}} = \langle 11, 8, 18, 13\rangle$	$\longrightarrow$	$(\mathcal{T}^{e_{Nun}})^{d_{Nun}} = \langle 11, 8, 18, 13\rangle^{d_{Nun}}$
		$\mathcal{T} = \langle 11, 8, 18, 15\rangle$

Cripto-Sistema de ElGamal

Vejamos agora mais um tipo de cripto-sistema, denominado cripto-sistema de ElGamal, onde se começa por fixar um corpo $\mathcal{F}_q$ e um elemento $g \in \mathcal{F}_q^*$ (de preferência, mas não necessariamente, um gerador). Temos uma mensagem $\mathcal{T}$ com equivalentes numéricos $\mathcal{T}_i \in \mathcal{F}_q$ e cada utilizador escolherá um inteiro

a tal que $0 < a < q - 1$ que será a sua chave de decriptação. A sua chave de encriptação será $g^a \in \mathcal{F}_q$, que é tornada pública.

Para enviar a mensagem $\mathcal{T}$ ao Nuno, escolhemos aleatoriamente um inteiro k e enviamos-lhe o par $\langle g^k, \mathcal{T} g^{ak} \rangle$, onde o segundo elemento é uma lista de unidades de mensagens da forma $\langle \mathcal{T}_1 g^{ak}, \mathcal{T}_2 g^{ak}, \cdots, \mathcal{T}_n g^{ak} \rangle$.

É de notar que conseguimos calcular g^{ak} desconhecendo a bastando-nos, para isso, elevar g^a a k. Agora o Nuno, que recebeu a mensagem encriptada $\mathcal{T}$ e conhece a, pode recuperar a mensagem original elevando o primeiro elemento g^k a a e multiplicando o segundo elemento (as unidades de mensagem) por $(g^{ak})^{-1}$.

EXEMPLO: Vamos ver o que se passa quando a Catarina envia ao Nuno a mensagem "LISP", já conhecida de outras aventuras criptográficas.

Catarina	$\xrightarrow{\mathcal{T}}$	Nuno
$a_{Cat} = 5$	$q = 31$	$a_{Nun} = 7$
	$g = 3$	
	$e_{Cat} = g^{a_{Cat}} = 26 \pmod{31}$	
	$e_{Nun} = g^{a_{Nun}} = 17 \pmod{31}$	
$k = 8$		
"LISP" $= \langle 11, 8, 18, 15 \rangle$		
$\langle g^k, \langle \mathcal{T}_1 g^{a_{Nun} k}, \cdots, \mathcal{T}_4 g^{a_{Nun} k} \rangle \rangle$	$\longrightarrow$	$\langle 20, \langle 12, 20, 14, 22 \rangle \rangle$
		$(g^{a_{Nun} k})^{-1} = 19$
		$\langle 12 \times 19, \cdots, 22 \times 19 \rangle$
		$\langle 11, 8, 18, 15 \rangle =$ "LISP"

Se analisarmos detalhadamente o que é enviado, o par $\langle g^k, \mathcal{T} g^{ak} \rangle$ não passa da mensagem com uma máscara, juntamente com uma 'dica' que permite remover a máscara da mensagem (no entanto, a dica só será útil a quem souber o a do receptor).

Qualquer pessoa que consiga resolver o problema do logaritmo discreto em $\mathcal{F}_q$ conseguirá quebrar o cripto-sistema obtendo a chave privada a a partir da chave pública g^a. No entanto, e como já foi mencionado anteriormente, é conjecturado que não existe maneira de passar de g^k e g^{ak} para a sem resolver o problema do logaritmo discreto.

4.3 Problema do Knapsack

Nesta secção descreveremos outro tipo de cripto-sistema de chave pública, baseado no chamado "Problema do Knapsack[5]". Suponhamos que temos uma mochila de campismo e que a estamos a encher de coisas para um acampamento no campo. Temos que colocar na mochila um grande número de objectos, digamos k objectos de volume v_i $(i = 0, \ldots, k-1)$ e a mochila tem um volume total V. Se formos experientes neste tipo de situações conseguiremos colocar tudo na mochila sem perder nenhum espaço. Como queremos levar o maior número de coisas possível, queremos encontrar um subconjunto desses k objectos que encha a mochila por completo. Formalmente, queremos encontrar um subconjunto (se existir) $I \subset \{1, \ldots, k\}$ tal que $\sum_{i \in I} v_i = V$. Uma definição rigorosa do problema do Knapsack é a seguinte:

Definição 23. Dado um conjunto $\{v_i\}$ de k inteiros positivos e um inteiro V, queremos encontrar (se existir) um inteiro de k bits $n = (\epsilon_{k-1}\epsilon_{k-2}\cdots\epsilon_1\epsilon_0)_2$ (onde $\epsilon_i \in \{0, 1\}$ são digitos binários de n) tal que $\sum_{i=0}^{k-1} \epsilon_i v_i = V$.

Este problema poderá ter mais que uma solução ou poderá ser impossível. Um caso particular é o chamado "Problema do Knapsack Super-Crescente", no qual todos os v_i, quando escritos por ordem crescente, são maiores que a soma de todos os v_i anteriores. O quíntuplo $\langle 2, 3, 7, 15, 31\rangle$ é um exemplo de uma sequência super-crescente. É sabido que o problema do Knapsack é um problema NP-Completo. No entanto, o Problema super-crescente é muito mais fácil e a sua solução, quando existe, é única. Para o resolver basta olhar por ordem decrescente para cada um dos v_i até encontrar o primeiro menor ou igual a V. Incluimos então o i respectivo no nosso subconjunto I, substituimos V por $V - v_i$ e continuamos a percorrer a nossa lista de v_i até encontrarmos algum que seja menor ou igual ao nosso novo V. Utilizando este método, podemos obter um conjunto de índices I, tal que a soma dos elementos correspondentes é igual a V, ou podemos esgotar os v_i e, nesse caso, o nosso problema é impossível.

Vamos pois analisar o cripto-sistema de Knapsack super-crescente (também conhecido como cripto-sistema de Merkle-Hellman). Começamos, inicialmente, por assumir que as nossas unidades de mensagem são equivalentes a números inteiros de k-bits. Se, como habitualmente, trabalharmos com unidades mono-alfabéticas num alfabeto de 26 letras, necessitamos apenas de 5 bits. De seguida, cada utilizador escolhe aleatoriamente uma sequência super-crescente de k elementos $v = (v_0, \ldots, v_{k-1}), v_i \in \mathbb{N}_0$, um inteiro m maior que $\sum_{i=0}^{k-1} v_i$, e um inteiro a tal que $a \perp m$ e $0 < a < m$. Depois, cada um calcula b tal que $b = a^{-1} \pmod m$ e determina uma sequência w

5 Mochila de campismo.

tal que $w_i = av_i \pmod m$. Cada utilizador mantém $v_i, m, a,$ e b secretos e coloca a público $w = (w_0, \ldots, w_{k-1})$ que será a sua chave pública. A sua chave privada é o tuplo (b, m). As funções de encriptação e decriptação de uma mensagem $\mathcal{T} = (\epsilon_{k-1}\epsilon_{k-2}\cdots\epsilon_1\epsilon_0)$ são então as seguintes:

$$\begin{aligned} \mathcal{C} &= f(\mathcal{T}) = \sum_{i=0}^{k-1} \epsilon_i w_i \\ \mathcal{T}' &\equiv g(\mathcal{C}) \equiv b\mathcal{C} \pmod m. \end{aligned}$$

Utilizando a função g obtemos $\sum \epsilon_i v_i$ (dado que $b\mathcal{C} = \sum \epsilon_i b w_i \equiv \sum \epsilon_i v_i \pmod m$). A partir do somatório obtido é possível obter a mensagem original utilizando o método descrito anteriormente para resolver o problema do Knapsack super-crescente. Note-se que qualquer pessoa que tente decifrar a mensagem, sabendo apenas w, vai enfrentar o problema do Knapsack $\mathcal{C} = \sum \epsilon_i w_i$ que já não é um problema super-crescente, pois esta propriedade é destruída quando v_i é substituído por $av_i \pmod m$. O método descrito acima não pode pois ser utilizado e essa pessoa enfrenta um problema extremamente mais complexo.

EXEMPLO: Não iremos avançar sem consolidar o que foi dito com um exemplo. Vamos mais uma vez simular o envio da nossa conhecida mensagem entre os nossos também conhecidos personagens.

$$\text{“LISP”} = \langle 11, 8, 18, 15\rangle_{10} = \langle 01011, 01000, 10010, 01111\rangle_2$$

Catarina	$\mathcal{T}=\text{“}LISP\text{”}$ $\longrightarrow$	Nuno
	$k = 5$	
$v_{Cat} = \langle 2, 3, 7, 15, 31\rangle$		$v_{Nuno} = \langle 1, 2, 4, 8, 11\rangle$
$m_{Cat} = 61$		$m_{Nuno} = 72$
$a_{Cat} = 17$		$a_{Nuno} = 13$
$b_{Cat} = a_{Cat}^{-1} = 18 \pmod{m_{Cat}}$		$b_{Nuno} = a_{Nuno}^{-1} = 61 \pmod{m_{Nuno}}$
	$w_{Cat} = a_{Cat}v_{Cat} \bmod 61$	
	$w_{Cat} = \langle 34, 51, 58, 11, 39\rangle \bmod 61$	
	$w_{Nuno} = a_{Nuno}v_{Nuno} \bmod 72$	
	$w_{Nuno} = \langle 13, 26, 52, 32, 71\rangle \bmod 72$	
$(b_{Cat}, m_{Cat}) = (18, 61)$		$(b_{Nuno}, m_{Nuno}) = (61, 72)$

$\mathcal{C} = \sum_{i=0}^{k-1} \epsilon_i w_{Nuno_i}$	$\longrightarrow$	$\mathcal{C} = \langle 129, 26, 45, 181 \rangle$
		$\mathcal{T}' = b_{Nuno}\mathcal{C} \pmod{m_{Nuno}}$
		$\mathcal{T}' = \langle 21, 2, 9, 25 \rangle$
		$11 + 8 + 2 = 21 \Rightarrow$ $\mathcal{T}_1 = (01011)_2 = 11$
		$2 = 2 \Rightarrow$ $\mathcal{T}_2 = (01000)_2 = 8$
		$8 + 1 = 9 \Rightarrow$ $\mathcal{T}_3 = (10010)_2 = 18$
		$11 + 8 + 4 + 2 = 25 \Rightarrow$ $\mathcal{T}_4 = (01111)_2 = 15$
		$\mathcal{T}$ = "LISP"

É de notar que o pequeno valor de k e a utilização de unidades de mensagem mono-alfabéticas tornam o sistema bastante inseguro. No entanto, e tal como nos exemplos anteriores, o importante é perceber como este sistema funciona.

Durante algum tempo, muitas pessoas foram optimistas quanto à dificuldade em quebrar um sistema deste tipo, dado que se encontra na classe de problemas NP-Completos. No entanto, existe uma falácia neste raciocínio. O tipo de problema de Knapsack $\mathcal{C} = \sum \epsilon_i w_i$ que tem de ser resolvido para quebrar este sistema, apesar de não ser um problema super-crescente, é obtido de um problema super-crescente através da aplicação de uma simples transformação (multiplicação de cada elemento por a e a sua posterior redução módulo m). Em 1982, Shamir encontrou um algoritmo para resolver este tipo de Knapsack, cujo tempo de execução é polinomial em k. Assim sendo, o sistema original de Merkle-Hellman não pode ser visto como um cripto-sistema de chave pública seguro.

Uma maneira de contornar o algoritmo de Shamir é criar um sistema de Knapsack que utilize uma sequência de transformações do tipo $x \rightarrow ax \mod m$ para diferentes valores de a e de m. Por exemplo, podemos considerar duas transformações correspondentes a (a_1, m_1) e (a_2, m_2) e, a partir de v_i, obter duas sequências w_i e u_i da seguinte forma:

$$w_i = a_1 v_i \mod m_1 \qquad \text{e} \qquad u_i = a_2 w_i \mod m_2.$$

Os números a_1, a_2, m_1 e m_2 estão sujeitos às condições:

$$m_1 > \sum v_i, \quad m_2 > k m_1, \quad a_1 \perp m_1 \text{ e } a_2 \perp m_2.$$

Para encriptar a mensagem utilizamos u_i e, para a decriptar, necessitamos de (b_1, m_1, b_2, m_2) que são calculados como anteriormente. A decriptação faz-se então por "sucessivas" decriptações da mensagem encriptada. Este sistema é conhecido como o sistema de Knapsack iterativo. Actualmente, apesar de não existir um algoritmo que corra em tempo polinomial que solucione um sistema de Knapsack iterativo, já foi possível generalizar o algoritmo de Shamir, mostrando que este tipo de sistemas são bastante vulneráveis a uma cripto-análise eficaz. Seja como for, após a descoberta do algoritmo de Shamir, a maioria dos profissionais desta área perderam a confiança neste tipo de sistemas. Poderíamos ainda descrever outros métodos de Knapsack envolvendo, por exemplo, polinómios em corpos finitos. O leitor interessado poderá procurá-los em [4].

4.4 Protocolos de Conhecimento Nulo

Um pequeno diálogo:

Catarina: "Eu sei a palavra-chave para entrar no computador central do Banco de Portugal, sei os ingredientes do molho especial da McDonalds e também sei todos os segredos da Coca-Cola."

Nuno: "Não sabes, não."

Catarina: "Sei sim."

Nuno: "Não sabes..."

Catarina: "Sei..."

Nuno: "Então prova!"

Catarina: "Ok, eu digo-te o que sei e pronto." Durante algum tempo a Catarina sussurra algo ao Nuno.

Nuno: "Que interessante. Agora eu também sei...e posso dizer a quem quiser."

Catarina: "Oops"

Após a Catarina provar ao Nuno que sabe uma dada informação revelando-lha, o Nuno fica também a conhecê-la. Pode então fazer o que quiser dessa informação e a Catarina nada poderá fazer para o impedir. Nesta secção tentaremos encontrar uma forma de provar algo a alguém sem ter de lhe transmitir a informação.

"Conhecimento Nulo" é o nome de um conceito de criptografia introduzido nos anos 80 para lidar com este problema. Será que isto é realmente possível? Como conseguiremos convencer alguém que temos uma dada informação sem lha revelar? A verdade é que, em várias situações, é possível fazê-lo. A pessoa que tentará provar algo chamar-se-á Catarina e o verificador, que no final ficará convencido que a Catarina tem essa informação, chamar-se-á Nuno.

Prova de Conhecimento Nulo sobre Logaritmos Discretos

Vamos então supor que $\mathcal{G}$ é um grupo finito que contém N elementos, b é um elemento fixo de $\mathcal{G}$ e y é um elemento de $\mathcal{G}$ para o qual a Catarina encontrou o logaritmo discreto de base b, ou seja, resolveu a equação $b^x = y$ para um inteiro positivo x. Ela quer agora demonstrar ao Nuno que sabe x, deixando-o sem qualquer tipo de pista quanto ao valor de x. Vamos supor que o Nuno sabe qual a ordem N do grupo e também os valores de b e y. Eis a sequência de acções feita por ambos:

1. A Catarina gera um inteiro positivo aleatoriamente $e < N$ e envia ao Nuno $b' = b^e$.

2. O Nuno lança uma moeda ao ar. Se sair cara, a Catarina diz-lhe o valor de e e o Nuno verifica que b' é realmente igual a b^e.

3. Se sair coroa, a Catarina terá de dizer ao Nuno o valor do menor inteiro positivo igual a $x + e \mod N$ e, nesse caso, o Nuno poderá confirmar que $yb' = b^{x+e}$.

4. Os passos anteriores deverão ser repetidos até que o Nuno esteja convencido que a Catarina sabe o valor x do logaritmo discreto.

Note-se que, se a Catarina não souber o valor x do logaritmo discreto, não saberá a resposta a mais que um lançamento possível da moeda. Se ela fizer o passo 1 de forma correcta, saberá responder a cara, mas não a coroa senão souber realmente o x. Por outro lado, se ela antecipar coroa durante o passo 1, decidirá dizer ao Nuno $b' = b^e/y$ (logo no passo 3 bastar-lhe-á enviar e em vez de $x + e$) mas, nesse caso, ficará sem saber o que fazer se sair cara (dado que não conhece a potência de b cujo resultado é b').

É também de salientar que a propriedade do Conhecimento Nulo deste protocolo pode ser demonstrada com uma simulação. Suponhamos que o Carlos não sabe o logaritmo discreto de y de base b mas sabe antecipadamente qual o resultado do lançamento da moeda. Então, o Carlos pode simular os mesmos passos que a Catarina (o envio de $b' = b^e$ para cara e $b' = b^e/y$ para

coroa), dando ao Nuno a mesma informação que a Catarina lhe daria. O Carlos, ao transmitir esta informação, não estará a dizer nada de útil para descobrir o logaritmo discreto, dado que ele próprio não faz ideia de qual o seu valor.

Transferência Desmemoriada

Um canal de transferência desmemoriada da Catarina para o Nuno é um sistema em que a Catarina envia ao Nuno dois pacotes de informação encriptada nas seguintes condições:

1. O Nuno consegue decriptar e ler apenas um dos dois pacotes;

2. A Catarina não sabe qual dos pacotes o Nuno consegue ler;

3. Tanto a Catarina como o Nuno conhecem as duas condições anteriores.

Apesar de este sistema ser um pouco estranho à primeira vista, é muito importante em criptografia. Por exemplo, podemos utilizar uma transferência desmemoriada em vez do lançamento da moeda visto na secção anterior, bastando para isso colocar e no primeiro pacote e $x + e$ no segundo.

Suponhamos que temos um corpo finito $\mathcal{F}_q$ e um elemento fixo b do grupo multiplicativo $\mathcal{F}_q^*$ para o qual, dados b^x e b^y, não é possível calcular b^{xy}. Esta é a suposição de Diffie-Hellman, já mencionada anteriormente, que se conjectura ser verdadeira caso o problema do logaritmo discreto seja impossível em $\mathcal{F}_q^*$. Seja ψ uma transformação do nosso corpo finito $\mathcal{F}_q$ para o espaço vectorial $\mathcal{F}_2^n$ (onde $\mathcal{F}_2^n$ é o espaço de vectores com n bits), facilmente calculável e invertível e tal que a sua imagem contenha $\mathcal{F}_2^{n-1}$ (ou seja, todos os vectores cujo primeiro bit é 0). Por exemplo, se q for um primo, podemos escolher n tal que $2^{n-1} < q < 2^n$ e transformar qualquer elemento de $\mathcal{F}_q$ na sua representação vectorial em dígitos binários. Suponhamos que a nossa unidade de mensagem são vectores de n bits (ou seja elementos $m \in \mathcal{F}_2^n$) e que escolhemos um elemento $C \in \mathcal{F}_q^*$ fixo (sempre!) cujo logaritmo discreto ninguém conhece. A transferência desmemoriada prossegue da seguinte forma. O Nuno escolhe um inteiro x tal que $0 < x < q - 1$ e também um elemento $i \in \{1, 2\}$. Depois faz $\beta_i = b^x$ e $\beta_{3-i} = C/b^x$, construindo a sua chave pública (β_1, β_2) e mantendo x e i secretos. Note-se que assumimos que o Nuno não conhece o logaritmo discreto de β_{3-i} (a que adiante chamaremos x') pois, caso contrário, também saberia o logaritmo discreto de $C = \beta_i\beta_{3-i}$. Vamos agora supor que a Catarina possui a unidade de mensagem $m_1 \in \mathcal{F}_2^n$ do primeiro pacote de informação e $m_2 \in \mathcal{F}_2^n$ do segundo pacote. Ela escolhe

dois inteiros aleatoriamente $y_1 > 0, \quad y_2 < q-1$ e envia ao Nuno os seguintes elementos pertencentes a $\mathcal{F}_2^*$ e $\mathcal{F}_2^n$ dois a dois,

$$b^{y_1}, \quad b^{y_2}, \quad \alpha_1 = m_1 + \psi(\beta_1^{y_1}), \quad \alpha_2 = m_2 + \psi(\beta_2^{y_2})$$

(onde a operação "+" em $\mathcal{F}_2^n$ é também conhecida como "ou exclusivo") mantendo y_1 e y_2 secretos. Como $\beta_i^{y_i} = (b^{y_i})^x$, e o Nuno conhece y_i e x, ele pode facilmente determinar $\psi(\beta_i^{y_i})$ e encontrar $m_i = \alpha_i - \psi(\beta_i^{y_i})$. No entanto, se ele quisesse descobrir m_{3-i}, teria de calcular $\beta_{3-i}^{y_{3-i}} = b^{x'y_{3-i}}$, sabendo $b^{y_{3-i}}$ e $b^{x'}$ mas não y_{3-i} ou x'. No entanto, pela suposição de Diffie-Hellman isto é impossível.

Note-se que a Catarina pode facilmente verificar que $\beta_1\beta_2 = C$, pelo que poderá ter a certeza que o Nuno não conhece os logaritmos discretos de ambos os elementos da sua chave pública (β_1, β_2). Como é do interesse do Nuno ter o máximo de informação possível, a Catarina terá a certeza que ele desconhece o logaritmo discreto de um dos elementos. Não existe, no entanto, maneira da Catarina distinguir β_1 de β_2 para determinar qual foi o obtido pelo Nuno como b^x e qual foi o obtido como C/b^x. Logo, tanto a Catarina como o Nuno podem estar certos que as condições 1 e 2 acima são satisfeitas.

Se uma sequência de pares (m_1, m_2) é enviada usando o mesmo (β_1, β_2), então a Catarina sabe que o elemento do par (m_1, m_2) que o Nuno está a decriptar (m_i) permanece igual para todos os pares da mensagem da sequência que está a ser enviada. Se quisermos enviar outra sequência de unidades de mensagem independente das anteriores o Nuno terá de aleatoriamente seleccionar novos valores para x e i e enviar uma nova chave pública (β_1, β_2).

É também possível realizar uma transferência não interactiva, ou seja, sem comunicação entre o Nuno e a Catarina. O leitor interessado poderá ler [4].

5 Agradecimentos

Gostaria de agradecer a excelente oportunidade que me foi dada de participar no seminário diagonal pela Prof.^{a} Leonor Godinho e pela organização. Agradeço também os contributos, a enorme disponibilidade e o apoio indispensável da Prof.^{a} Leonor Godinho e do Prof. Carlos Florentino. Quero também agradecer a uma pessoa, especial e muito importante para mim, pela paciência, ajuda, apoio, motivação e carinho sempre que necessário, à minha namorada M^a Catarina Dias.

Referências

[1] B. Engquist, W. Schmid, *Mathematics Unlimited – 2001 and Beyond*, Springer, 2000.

[2] R.L. Graham, D. Knuth, O. Patashnik, *Concrete Mathematics – A Foundation for Computer Science*, 2ª edição, Addison-Wesley, 1994.

[3] D. Knuth, *The Art of Computer Programming*, Vol. 1, 3ª edição, Addison-Wesley, 1997.

[4] N. Koblitz, *A Course in Number Theory and Cryptography*, Springer, 1994.

[5] T. Reis, *Criptografia e jogos por telefone*, Seminário Diagonal, 2001.

[6] B. Schneier, *Applied Cryptography – Protocols, Algorithms, and Source Code in C*, 2ª edição, Wiley Computer Publishing, 1996.

"It is insufficient to protect ourselves with laws;
we need to protect ourselves with mathematics."
— B. Schneier