

A Geometria do Globo Terrestre

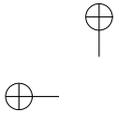
Sérgio Alves

Os Três Problemas Clássicos da Matemática Grega

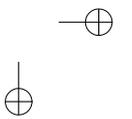
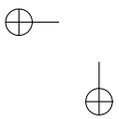
João Pitombeira de Carvalho

A Matemática dos Códigos de Barras

Francisco César Polcino Milies



Texto já revisado pela nova ortografia.



Apresentação

Os textos apresentados a seguir foram produzidos para a II Bienal da Sociedade Brasileira de Matemática, realizada em 2004, em Salvador.

A utilização desses trabalhos pelos bolsistas de Iniciação Científica da OBMEP é motivada pelo fato de tratarem de questões muito interessantes e, a nosso ver, motivadoras.

O primeiro, de Sérgio Alves, aborda o problema do estabelecimento de coordenadas no globo terrestre. O domínio de tais coordenadas e o uso de instrumentos para medi-las, propiciou grandes avanços para a humanidade, desde os tempos das grandes navegações (e o conseqüente aumento do comércio mundial) até o uso do GPS para a construção de mapas e localização.

O segundo texto, de autoria de João Bosco Pitombeira, tem forte sabor histórico pois trata de três problemas clássicos da geometria euclidiana plana: a duplicação do cubo, a quadratura do círculo e a trisseção do ângulo. São perguntas naturais cuja resposta custou muito esforço e fazem parte da bagagem cultural da matemática.

Quem viveu a época das filas nos supermercados, onde o preço de cada produto tinha que ser registrado manualmente no caixa, sabe muito bem o avanço que representa o uso do código de barras. Uma ideia simples e genial, de grande utilidade. Esse é o tema do ter-

ceiro texto, em que Francisco César Polcino Miles expõe os aspectos matemáticos da codificação. Vemos aqui um belo exemplo de como estruturas abstratas ajudam a resolver problemas concretos.

Como será possível notar, os textos não esgotam os respectivos assuntos e possuem referências bibliográficas úteis para aqueles mais curiosos que desejarem aprofundar seus conhecimentos sobre os assuntos. A leitura de um trabalho matemático deve ser pausada e refletida, e aqui não será diferente. O leitor deve se envolver nas questões e ideias discutidas pelo autor e muitas vezes “dialogar” com ele, quem sabe buscando outras soluções.

Esperamos que os bolsistas apreciem tanto a beleza dos problemas e de suas soluções quanto a elegância com que os temas foram tratados.

O Comitê Acadêmico da OBMEP agradece aos autores por terem permitido o uso dos seus trabalhos no Programa de Iniciação Científica da OBMEP.

Comitê Acadêmico

Direção Acadêmica

Sumário

A Geometria do Globo Terrestre	1
Os Três Problemas Clássicos da Matemática Grega	81
A Matemática dos Códigos de Barras	133

“principal1”

2009/8/19

⊕ page iv

Estilo OBME

—⊕

⊕

|

⊕—

⊕—

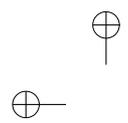
|

⊕

—⊕

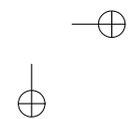
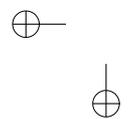
|

⊕



A Geometria do Globo Terrestre

Sérgio Alves





Prefácio

Num encontro de professores que atuam no ensino básico fomos questionados sobre a utilização de aplicações da Matemática em sala de aula. Nossas primeiras considerações foram as esperadas: é sempre bom poder mostrar aos alunos que a Matemática faz parte de nosso cotidiano, ilustrando como ideias e conceitos abstratos levam ao desenvolvimento de conhecimentos e tecnologias que visam o bem-estar da nossa sociedade. Com isso, certamente nossas aulas de Matemática ficam mais interessantes e motivadoras para o aluno.

Porém, existe um preço a se pagar para conseguirmos esse intento e notamos certa surpresa em nossos interlocutores quando passamos a explicar melhor esse lado da questão.

Primeiramente, a chamada *contextualização* da Matemática tem seu lugar e hora para acontecer. Ela deve ser feita de modo que o aluno entenda seu significado e sua necessidade e não se criando situações, por vezes extremamente artificiais, somente pelo gosto de poder dar alguma aplicação prática de um determinado tema. Nem todo tópico da Matemática é passível de contextualização!

Um segundo ponto por nós levantado é que a apresentação de aplicações significativas da Matemática invariavelmente exige do professor certo conhecimento de alguma outra área como, por exemplo, Física, Biologia ou Geografia. Esse professor deve estar ciente que tal

exigência vai requerer dele um maior empenho e dedicação.

Por fim destacamos o fato que, embora a Matemática tenha a marca da ciência exata por excelência, nas suas aplicações frequentemente temos que lidar com a questão das aproximações fazendo uso, ou não, de calculadoras ou programas de computador. Raramente na solução de um problema contextualizado comparecem números como $\sqrt{16}$ ou ainda $\cos 60^\circ$!

Motivado por todas essas considerações escrevemos o texto que aqui se apresenta. Escolhemos como tema a geometria da esfera que encontra na Geografia uma natural contextualização por meio do estudo do globo terrestre e dos vários assuntos a ele relacionados. Incorporamos no texto todos os conceitos geográficos necessários, o que não exige o leitor destas notas de procurar aprofundá-los com suas próprias fontes.

Procuramos explicar, com a maior clareza possível, as ideias matemáticas envolvidas no funcionamento do sistema de posicionamento global (GPS), um dos mais modernos sistemas de localização por satélites, e fazemos ainda uma discussão atualizada do cálculo da distância entre dois pontos da superfície terrestre. Esse cálculo, no passado, era justificado com o uso da trigonometria esférica, tema este que já não frequenta nossos livros didáticos.

Enfim, produzimos este texto na esperança que ele forneça aos professores uma possível resposta àquela pergunta que frequentemente lhes é feita: “Para que serve isso que o senhor está ensinando?”

Sérgio Alves

Sumário

Introdução	7
1 A Superfície Esférica e seus Elementos	9
1.1 A Interseção entre um Plano e uma Superfície Esférica	10
2 Qual é a Forma da Terra?	19
2.1 As Coordenadas Geográficas	25
2.2 Os Movimentos da Terra	31
2.3 Os Fusos Horários	38
3 O Ângulo de Elevação do Sol e da Energia Solar	41
3.1 O Ângulo de Elevação do Sol nos Equinócios	45
3.2 O Ângulo de Elevação do Sol nos Solstícios	47
4 A Superfície Esférica em Coordenadas Cartesianas	54
5 A Matemática do GPS	64
6 A Distância numa Superfície Esférica	71



6

Referências Bibliográficas

79

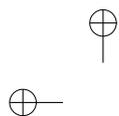


Introdução

Neste curso pretendemos desenvolver o estudo da esfera e seus elementos explorando sua associação com o globo terrestre. Conceitos geográficos como paralelos, meridianos, latitudes, longitudes e fusos horários estão baseados em importantes ideias geométricas que, quando trabalhadas neste contexto, conduzem o aluno a uma melhor compreensão e aprendizagem do tema.

O estudo dos movimentos da Terra nos permite entender, além das quatro estações do ano com seus equinócios e solstícios, porque o Trópico de Capricórnio ou o Círculo Polar Ártico são paralelos notáveis. Veremos ainda que as relações entre longitude e fusos horários bem como entre latitude e o ângulo de elevação do Sol nos levam a problemas geométricos relevantes.

O estudo da posição relativa de duas ou mais esferas e as relações entre as coordenadas geográficas e as coordenadas cartesianas constituem a fundamentação matemática necessária para o entendimento de alguns modernos sistemas de navegação por satélites, em especial do GPS.



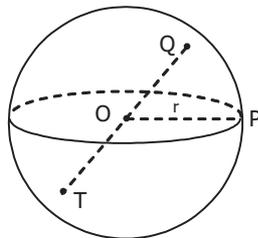
A utilização do globo terrestre, com suas conseqüentes questões envolvendo, por exemplo, cálculo de distâncias e ângulos sobre a esfera, ou ainda, a confecção de mapas por meio de diversas projeções, abre caminho para um interessante trabalho interdisciplinar entre a Matemática e a Geografia.



Capítulo 1

A Superfície Esférica e seus Elementos

Seja O um ponto e r um número real positivo. A **superfície esférica de centro O e raio r** é o conjunto de todos os pontos P do espaço cuja distância a O é igual a r .



Os pontos do espaço cuja distância a O é menor que r são **interiores** à superfície esférica e aqueles cuja distância a O é maior que r são **exteriores** a ela.

A reunião da superfície esférica de centro O e raio r com seus pontos interiores é chamada a **esfera de centro O e raio r** .

O segmento que une o centro a um ponto qualquer da super-

superfície esférica é denominado um raio da superfície esférica enquanto que o segmento que une dois pontos distintos da superfície esférica é chamado uma **corda** da superfície esférica.

Uma corda que contém o centro é chamada um **diâmetro** da superfície esférica.

Evidentemente o comprimento de qualquer diâmetro é o número $2r$ que é chamado o **diâmetro**.

Observamos que a palavra **raio** é usada com dois sentidos: pode ser o número r ou um segmento \overline{OP} . Isto, porém, não causa confusão uma vez que é sempre claro qual o significado utilizado. Um comentário análogo vale para a palavra **diâmetro**.

Neste texto a notação \overline{PQ} representa o segmento de extremos P e Q enquanto que PQ indica o seu comprimento.

Na figura anterior r é o raio da superfície esférica, \overline{OP} , \overline{OQ} e \overline{OT} são raios, \overline{PQ} e \overline{PT} são cordas, \overline{QT} é um diâmetro e $2r = QT$ é o diâmetro da superfície esférica.

1.1 A Interseção entre um Plano e uma Superfície Esférica

A relação entre superfícies esféricas e planos no espaço é similar com a relação entre circunferências e retas no plano.

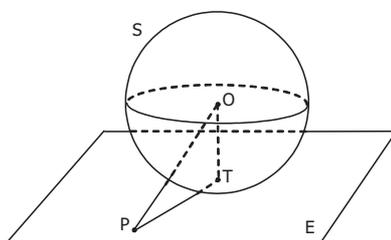
Um plano E é **tangente** a uma superfície esférica S se $E \cap S$ contém exatamente um ponto. Esse ponto é chamado **ponto de tangência**. Dizemos que o plano e a superfície esférica se tangenciam nesse ponto.

Se $E \cap S$ contém mais do que um ponto, então o plano é **secante** à superfície esférica. Podemos caracterizar o plano tangente conforme o seguinte resultado.

▲ SEC. 1.1: A INTERSEÇÃO ENTRE UM PLANO E UMA SUPERFÍCIE ESFÉRICA 11

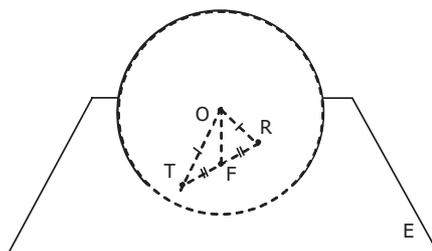
Teorema 1.1. *Um plano perpendicular a um raio na sua extremidade comum com a superfície esférica é tangente à mesma. Reciprocamente, todo plano tangente a uma superfície esférica é perpendicular ao raio que contém o ponto de tangência.*

Demonstração. Sendo E um plano perpendicular ao raio \overline{OT} em T , vamos mostrar que nenhum outro ponto de E está na superfície esférica.



Seja P um ponto qualquer de E , P distinto de T . Como E é perpendicular a \overline{OT} temos que $\triangle OPT$ é um triângulo retângulo com hipotenusa \overline{OP} e catetos \overline{OT} e \overline{PT} . Logo $OP > OT = r$ e, portanto, P não está na superfície esférica.

Reciprocamente seja E um plano tangente à superfície esférica no ponto T . Suponha, por absurdo, que E não seja perpendicular ao raio \overline{OT} . Vamos mostrar que essa hipótese nos leva a uma contradição.



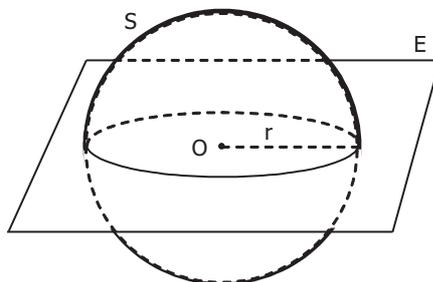
Sendo F o pé da perpendicular ao plano E , traçada a partir de O , temos $F \neq T$, pois E não é perpendicular a \overline{OT} . Seja R o ponto da

reta \overleftrightarrow{FT} tal que $T - F - R$ e $FR = FT$. Então $\triangle OFR \cong \triangle OFT$ (pelo critério LAL de congruência de triângulos) de modo que $OR = OT = r$ e, portanto, R está na superfície esférica. Logo E intersecta a superfície esférica em um ponto R distinto de T . Isso é impossível, pois E é um plano tangente.

□

Vamos agora investigar a figura obtida quando um plano secante corta uma superfície esférica. Se o plano passa pelo centro a resposta é imediata.

Teorema 1.2. *A intersecção de uma superfície esférica com um plano passando pelo seu centro é uma circunferência de mesmo centro e mesmo raio.*



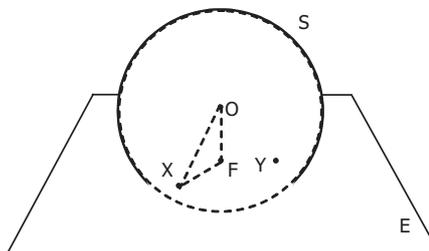
Demonstração. Dada uma superfície esférica S com centro O e raio r e um plano E que passa por O , a intersecção $E \cap S$ é o conjunto de todos os pontos de E cuja distância a O é igual a r . Essa é exatamente a definição de uma circunferência de centro O e raio r .

□

O próximo resultado nos dá a situação mais geral.

▲ SEC. 1.1: A INTERSECÇÃO ENTRE UM PLANO E UMA SUPERFÍCIE ESFÉRICA 13

Teorema 1.3. *Se um plano contém um ponto do interior de uma superfície esférica, então a intersecção do plano com a superfície esférica é uma circunferência. O centro dessa circunferência é o pé da perpendicular ao plano traçada a partir do centro da superfície esférica.*



Demonstração. Seja E um plano que não passa pelo centro O da superfície esférica S e que contém um ponto Y do seu interior. Sendo F o pé da perpendicular a E traçada a partir de O , vamos mostrar que a intersecção $E \cap S$ é uma circunferência de centro F .

Sabemos que $OY < r$, pois Y está no interior de S . No triângulo retângulo $\triangle OFY$ temos que \overline{OY} é hipotenusa, enquanto que \overline{OF} é cateto. Logo $OF < OY < r$, ou seja, F também está no interior de S .

Seja X um ponto qualquer na intersecção $E \cap S$. Então $\triangle OFX$ tem um ângulo reto em F e, pelo teorema de Pitágoras,

$$OF^2 + FX^2 = OX^2 = r^2$$

e, portanto, $FX = \sqrt{r^2 - OF^2}$ (note que $r^2 - OF^2 > 0$).

Logo X está na circunferência de centro F e raio $\sqrt{r^2 - OF^2}$. Provamos assim que a intersecção $E \cap S$ está contida na circunferência de centro F e raio $\sqrt{r^2 - OF^2}$.

Isso ainda não significa que a intersecção é a circunferência. Para completar a demonstração, precisamos mostrar que todo ponto da circunferência pertence à intersecção.

Seja X um ponto qualquer da circunferência, em E , com centro F e raio $\sqrt{r^2 - OF^2}$. Novamente pelo teorema de Pitágoras,

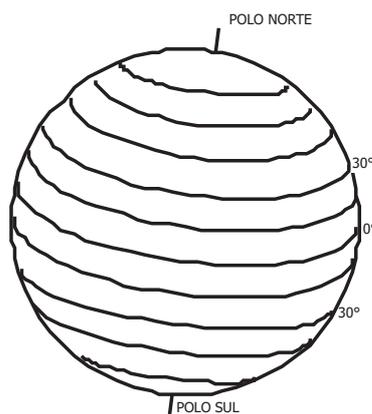
$$\begin{aligned} OX^2 &= OF^2 + FX^2 \\ &= OF^2 + r^2 - OF^2 \\ &= r^2. \end{aligned}$$

Portanto $OX = r$ e X pertence à superfície esférica S .

□

A intersecção da superfície esférica com um plano passando pelo seu centro é chamada uma **circunferência máxima** da superfície esférica.

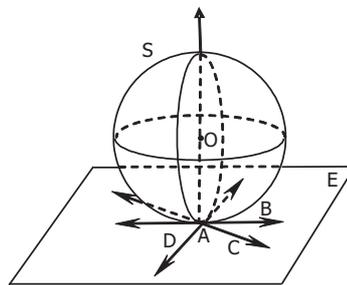
Há uma forte razão para esse nome: as circunferências máximas são as circunferências de maior raio contidas na superfície esférica. Veremos adiante que o Equador é uma circunferência máxima, mas os outros paralelos no globo terrestre não o são. Eles são menores que o Equador, tornando-se muito pequenos perto dos Polos Norte e Sul.



▲ SEC. 1.1: A INTERSEÇÃO ENTRE UM PLANO E UMA SUPERFÍCIE ESFÉRICA 15

Atividades

- 1) A superfície esférica S é tangente ao plano E em A . O ponto O é o centro de S e B, C, D estão em E . Que relação existe entre \overline{OA} e \overline{AB} , \overline{AC} e \overline{AD} ? Explique.



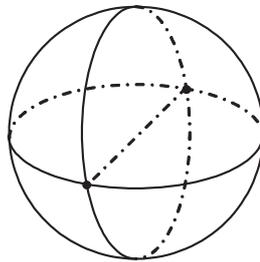
- 2) A perpendicular do centro de uma superfície esférica a uma corda divide-a ao meio. Utilize este resultado na resolução do seguinte problema. Numa superfície esférica de raio 15, a distância de uma corda ao centro é igual a 9. Qual o comprimento da corda? Apresente ainda uma prova do resultado acima.
- 3) Dados um plano E e uma superfície esférica S de centro O e raio r , podemos caracterizar a posição relativa entre E e S por meio da distância $d(O, E)$ de O ao plano E . Faça isso completando a tabela abaixo.

$d(O, E)$	Posição relativa
	E é secante a S
	E é tangente a S
	E é disjunto a S

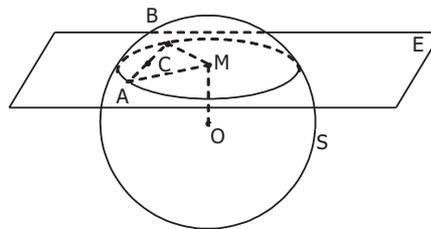
- 4) Sejam A e B dois pontos de uma superfície esférica S que não são extremos de um diâmetro de S . Mostre que existe uma e somente

uma circunferência máxima de S passando por A e B . (SUGESTÃO: Sendo O o centro de S , considere o plano E determinado pelos pontos não-colineares A , B e O .)

- 5) Explique porque duas circunferências máximas quaisquer de uma superfície esférica, se cortam nas extremidades de um diâmetro da superfície esférica.



- 6) Na figura abaixo um plano E intersecta a superfície esférica S de centro O . Os pontos A , B , C e M estão em E sendo que A e B também estão em S . Sabe-se que $\overline{OM} \perp E$, $\overline{AM} \perp \overline{MB}$, C é ponto médio de \overline{AB} , $AM = OM$ e $AB = 5$.



Calcule o raio da superfície esférica S , a medida angular $m(\angle AOB)$ e a medida OC .

- 7) Duas circunferências máximas são ditas perpendiculares se estiverem em planos perpendiculares. Mostre que para cada duas circunferências máximas existe uma terceira circunferência máxima perpendicular a ambas. Se duas circunferências máximas

▲ SEC. 1.1: A INTERSEÇÃO ENTRE UM PLANO E UMA SUPERFÍCIE ESFÉRICA 17

no globo terrestre passam pelos polos, que circunferência máxima é perpendicular a ambas?

8) Na geometria esférica as “linhas retas” são representadas por circunferências máximas. Se puder, encontre cada uma das seguintes figuras em tal geometria.

- (a) Um triângulo equilátero.
- (b) Um triângulo com dois ângulos retos.
- (c) Um triângulo com três ângulos retos.
- (d) Um triângulo cujas medidas de seus ângulos somem 500.
- (e) Um retângulo, isto é, um quadrilátero com quatro ângulos retos.

9) Sejam S e S' duas superfícies esféricas de centros distintos O e O' , respectivamente. Sendo r e r' , $r \geq r'$, seus respectivos raios, verifique que

$$\begin{aligned} S \cap S' \text{ é vazia} &\iff OO' > r + r' \text{ ou } OO' < r - r' \\ S \cap S' \text{ é um ponto} &\iff OO' = r + r' \text{ ou } OO' = r - r' \\ S \cap S' \text{ é uma circunferência} &\iff OO' < r + r' \text{ ou } OO' < r - r'. \end{aligned}$$

Faça desenhos ilustrando todas estas possibilidades.

(SUGESTÃO: Observe inicialmente que $P \in S \cap S' \iff P \in S$ e $PO^2 - PO'^2 = r^2 - r'^2$. A seguir, mostre que o conjunto dos pontos P do espaço que satisfazem $PO^2 - PO'^2 = r^2 - r'^2$ é um plano E perpendicular à reta $\overleftrightarrow{OO'}$ num ponto H definido por $MH = \frac{r^2 - r'^2}{2OO'}$ com $O - M - H$, onde M é o ponto médio do segmento $\overline{OO'}$. Conclua a prova utilizando a atividade 3 para obter as relações desejadas.)

10) Quais as possibilidades para a quantidade de pontos da intersecção de três superfícies esféricas? Faça desenhos ilustrando suas respostas.



Respostas

- 2) 24.
- 6) 5, 60° e $\frac{5\sqrt{3}}{2}$.
- 10) 0, 1, 2 ou infinitos pontos.

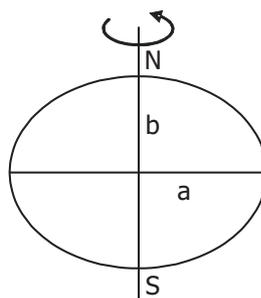


Capítulo 2

Qual é a Forma da Terra?

A Terra é redonda, mas não é uma esfera perfeita, uma vez que é achatada nos polos. Na verdade, a Terra é aproximadamente um **elipsóide**.

A figura abaixo mostra uma secção da superfície terrestre através de um plano que contém a reta que liga os Polos Norte e Sul. Esta secção aproxima-se de uma elipse cujo semieixo maior a é a metade do diâmetro do Equador e o semieixo menor b é a metade da distância entre os polos.



A superfície terrestre é a superfície de revolução gerada por essa elipse quando esta gira em torno da reta que passa pelos polos.

A razão $\varepsilon = \frac{a - b}{b}$ chama-se **achatamento** da Terra.

A tabela seguinte mostra valores dos semieixos a e b e do correspondente achatamento da Terra, obtidos por alguns pesquisadores.

	a (metros)	b (metros)	ε
Bessel (1841)	6 377 397	6 356 078	0,0033541
Clarke (1880)	6 378 249	6 356 515	0,0034191
Helmert (1906)	6 378 200	6 356 940	0,0033443
Hayford (1924)	6 378 388	6 356 911	0,0033785
Krassovski (1940)	6 378 245	6 356 863	0,0033636

O valor extremamente pequeno de ε nos permite, para efeitos didáticos, desprezar esse achatamento e considerar a Terra como se fosse uma esfera.

A superfície da Terra passa a ser considerada daqui por diante como um globo: o globo terrestre. O globo terrestre é ainda o nome do instrumento didático que representa a Terra como uma esfera.

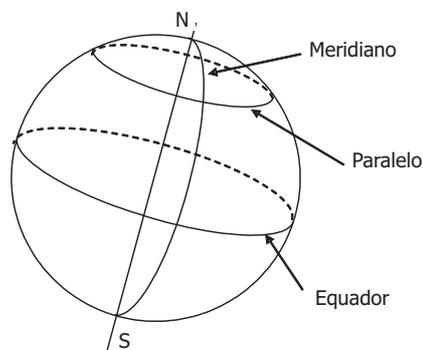
Iniciamos estabelecendo a nomenclatura apropriada. A figura a seguir ilustra um globo terrestre.

O ponto N representa o **Polo Norte** e o ponto S o **Polo Sul**. A reta determinada por N e S é chamada o **eixo polar**. Ela é a reta em torno da qual a Terra efetua seu movimento de rotação.

O plano que passa pelo centro da superfície esférica e é perpendicular ao eixo polar chama-se o **plano do Equador**.

O **Equador** é a intersecção do plano do Equador com a superfície esférica. O Equador é, portanto, uma circunferência máxima.

Um plano que passa pelo centro da superfície esférica divide-a em



duas partes chamadas **hemisférios**.

O plano do Equador divide a superfície terrestre em dois hemisférios: o **Hemisfério Norte** (que contém o Polo Norte) e o **Hemisfério Sul** (que contém o Polo Sul).

Os **paralelos** são as secções da superfície terrestre através de planos paralelos (ou coincidentes) ao plano do Equador. São então circunferências. Os paralelos notáveis são:

- o Equador
- o Trópico de Câncer
- o Trópico de Capricórnio
- o Círculo Polar Ártico
- o Círculo Polar Antártico

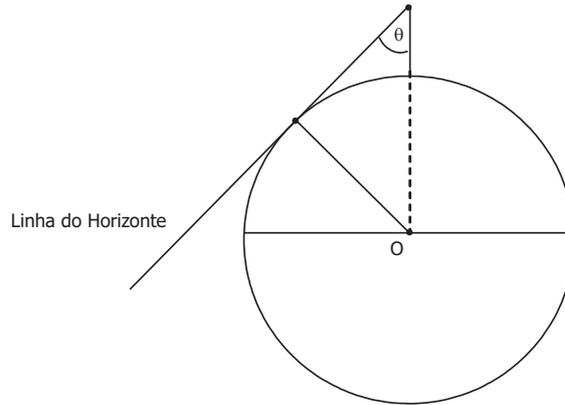
Os **meridianos** são semicircunferências que ligam os Polos Norte e Sul por meio de arcos máximos, isto é, arcos contidos em circunferências máximas que passam pelos polos. Convém ressaltar que os meridianos, ao contrário dos paralelos, não são circunferências. Além disso, eles estão contidos em planos perpendiculares ao plano do

Equador. O meridiano mais notável é o de Greenwich, nome de uma localidade próxima a Londres, onde está instalado um observatório astronômico.

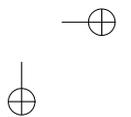
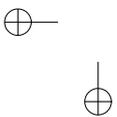
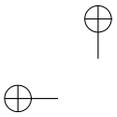
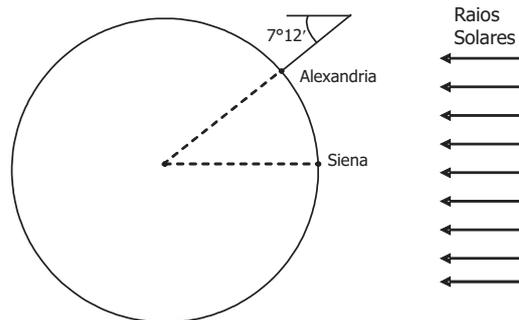
Observação. *Os livros de Geografia utilizam com muita frequência a palavra "círculo" para designar a circunferência. Em particular, descrevem o Equador como um círculo máximo e utilizam nomes como círculo polar. Isto é, na verdade, um abuso de linguagem consagrado pelo uso sistemático ao longo do tempo. A nomenclatura utilizada por profissionais não matemáticos, no caso os geógrafos, não precisa coincidir necessariamente com a usada pelos matemáticos.*

Atividades

- 1) Qual é o comprimento do Equador, supondo que o raio da Terra meça 6 400 km? Considere $\pi = 3,14159$.
- 2) Como se mede o raio da Terra? Uma das maneiras é escalar o topo de uma montanha cuja altitude acima do mar seja conhecida e medir o ângulo entre a vertical e a linha do horizonte. A altura do monte Shasta na Califórnia é 4,3 km. Do seu topo, o horizonte sobre o Oceano Pacífico faz um ângulo de $87^{\circ}53'$ com a vertical. Utilize estes dados para estimar o raio da Terra em quilômetros. Considere $\text{sen } 87^{\circ}53' = 0,99932$.

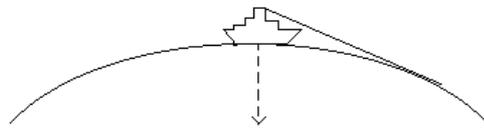


- 3) O cálculo do raio da Terra mais célebre da Antiguidade foi realizado pelo grego Eratóstenes (276-196 a.C.). Consultando as observações astronômicas acumuladas durante séculos na biblioteca de Alexandria, Eratóstenes soube que em Siena, 5 000 estádios (medida grega de comprimento) ao sul de Alexandria e situada aproximadamente no mesmo meridiano, o Sol se refletia no fundo de um poço ao meio-dia de um determinado dia de cada ano. Ao meio-dia deste tal dia, Eratóstenes mediu o ângulo que o raio do Sol fazia com a vertical de Alexandria, achando aproximadamente $7^{\circ}12'$.



Admitindo que os raios solares cheguem ao nosso planeta praticamente paralelos, mostre como obter para o raio da Terra o valor aproximado de $250\,000/2$ estádios. Supondo que a medida de um estádio utilizada por Eratóstenes era de aproximadamente 185 metros, calcule o valor acima em quilômetros.

- 4) Na ponte de um navio em alto-mar, o capitão pediu a um jovem oficial que estava ao seu lado que determinasse a distância ao horizonte. O oficial pegou papel e lápis e, em poucos instantes, deu uma resposta. No papel ele havia escrito a fórmula $d = \frac{8}{5}\sqrt{5h}$. Admitindo-se que o raio da Terra meça 6 400 km, mostre que essa fórmula é uma boa aproximação da distância d , em quilômetros, ao horizonte onde h é a altura, em metros, do observador acima da água. Se a ponte do navio está a 30 metros acima da água, qual é a distância ao horizonte?



- 5) Em um dia claro, até de que distância se poderia enxergar um edifício de 150 metros de altura, supondo-se que não haja obstruções no meio? Utilize a fórmula do exercício anterior.

Respostas

- 1) Cerca de 40 212 km.
- 2) Cerca de 6 319 km.
- 3) Cerca de 7 360 km.

- 4) Cerca de 19 km.
- 5) Cerca de 43 km.

2.1 As Coordenadas Geográficas

O globo terrestre serve para localizar um determinado ponto ou região da Terra. O aluno deve ter acesso a um para efetuar esta localização.

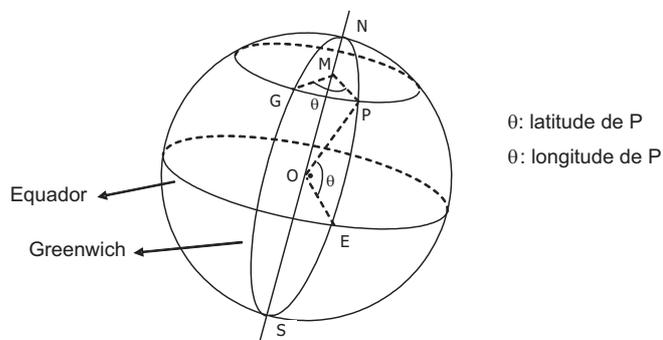
Há razões práticas para se ultrapassar o estágio da localização ingênua (isto é, apontar o dedo para o lugar e dizer: é aqui!) sendo necessário criar um sistema de coordenadas para dar a localização precisa de um ponto no globo.

Para isto utilizamos as chamadas **coordenadas geográficas**: latitude e longitude.

A **latitude** de um ponto P é a medida do arco de meridiano que passa por P situado entre o paralelo que contém P e o Equador. A latitude é expressa em graus, minutos e segundos e se mede de 0° a 90° N (norte) ou de 0° a 90° S (sul).

A **longitude** de um ponto P é a medida do arco de paralelo que passa por P situado entre o meridiano que contém P e o meridiano de Greenwich. A longitude é expressa em graus, minutos e segundos e se mede de 0° a 180° E (leste) ou de 0° a 180° W (oeste).

Na figura a seguir temos que $\theta = m(\angle EOP)$ é a latitude de P enquanto que $\varphi = m(\angle GMP)$ é a longitude de P. Pontos sobre um mesmo paralelo possuem latitudes iguais e pontos sobre um mesmo meridiano possuem longitudes iguais.



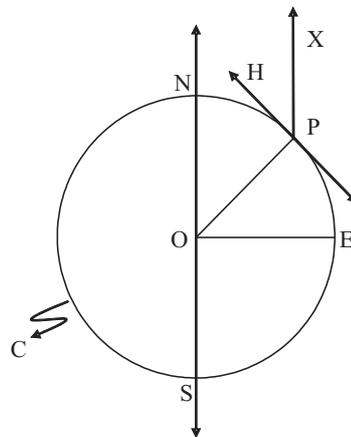
Atividades

- 1) Observando um globo terrestre ou um mapa apropriado estabeleça as coordenadas geográficas de cada uma das cidades abaixo.

Cidade	Latitude	Longitude
São Paulo		
Maceió		
Belo Horizonte		
Nova Orleans		
Chicago		
Roma		
Nova York		
Buenos Aires		
Londres		
Moscou		
Cairo		
Tóquio		
Quito		
Los Angeles		

▲ SEC. 2.1: AS COORDENADAS GEOGRÁFICAS

- 2) Qual a relação entre o raio da superfície terrestre, o raio de um paralelo e a sua respectiva latitude?
- 3) Verifique que a longitude de um ponto P é a medida do arco do Equador situado entre o meridiano que passa P e o meridiano de Greenwich.
- 4) Um dos primeiros fatos que um estudante de astronomia aprende é que a latitude de um ponto da Terra localizado no Hemisfério Norte é igual à medida do ângulo que Polaris (a Estrela do Norte) forma com a linha do horizonte, quando observada daquele ponto. Geometricamente, esta afirmação pode ser interpretada da seguinte maneira. A reta \overleftrightarrow{NS} é o eixo polar da Terra, a circunferência C contém um meridiano, E está no Equador, P é o observador, a reta \overleftrightarrow{PH} é a linha do horizonte e $m(\angle XPH)$ é a elevação de Polaris onde \overleftrightarrow{PX} é paralela a \overleftrightarrow{NS} .



O fato acima equivale a dizer que $m(\angle EOP) = m(\angle XPH)$. Mostre porque isto é verdadeiro.

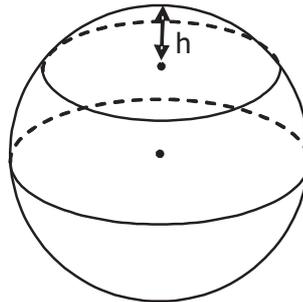
- 5) A aceleração da gravidade, frequentemente denotada por g , é considerada constante para movimentos próximos da superfície ter-

restre. Na verdade, porém, g não é constante, variando ligeiramente com a latitude θ . Uma boa aproximação para g , ao nível do mar, é dada pela expressão

$$g = 9,78049(1 + 0,005288 \operatorname{sen}^2\theta - 0,000006 \operatorname{sen}^2 2\theta) \text{ m/s}^2$$

- (a) Utilizando uma calculadora, determine g para a latitude do lugar onde você mora.
 - (b) Expresse g apenas em função de $\operatorname{sen} \theta$, isto é, elimine o ângulo duplo.
 - (c) Em que lugar da Terra o valor de g é máximo? E mínimo?
- 6) Supondo que o raio da Terra meça 6 400 km, qual o comprimento de um grau de longitude em uma latitude arbitrária θ ? Em particular, qual o comprimento numa latitude de 30° N (aproximadamente a latitude de Nova Orleans), 10° S (aproximadamente a latitude de Maceió) e 20° S (aproximadamente a latitude de Belo Horizonte)?
 - 7) A que latitude um grau de longitude tem comprimento igual a 48 km? A que latitude um paralelo mede 16 000 km?
 - 8) Qual a altura mínima para que um satélite consiga fotografar o Brasil inteiro? Admita que Oiapoque e Chuí estejam sobre o mesmo meridiano, mas em paralelos diferentes: a latitude de Oiapoque é 6° N e a de Chuí é 34° S. Suponha que o raio da Terra meça 6 400 km.
 - 9) Um astronauta encontra-se numa nave espacial que gira numa órbita em torno da Terra. No momento em que a nave está 160 km acima da superfície da Terra, que fração da superfície da Terra é visível para o astronauta? (Esta fração é chamada **calota esférica** e sua área é dada por $2\pi rh$, onde $r = 6\,400$ km é o raio

da Terra e h é a altura da calota). Qual seria a fração visível se a nave estivesse a 20 200 km acima da superfície da Terra? Neste caso, quanto mede o ângulo de visualização sobre a Terra?

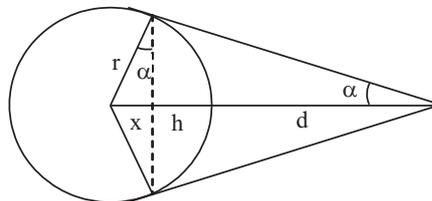


- 10) **Zona** é a região da superfície terrestre compreendida entre dois paralelos. Utilize a expressão da área da calota esférica para deduzir que a área da zona delimitada por um paralelo de latitude e o Equador é dada por $2\pi r d = 2\pi r^2 \sin \theta$, onde r é o raio da Terra e d é a distância entre o plano que contém o paralelo e o plano do Equador. Conclua daí o surpreendente fato: cortando-se uma superfície esférica em fatias de igual espessura, as áreas das zonas são iguais, sendo indiferente que o corte tenha sido feito perto do Equador ou perto do polo.
- 11) Um ponto P do globo terrestre tem latitude θ e longitude φ . Quais são as coordenadas geográficas do ponto Q diametralmente oposto a P ? Discuta cada uma das possibilidades N/S para θ e E/W para φ .

Respostas

- 2) $r_{paralelo} = r_{terra} \cos \theta$.
- 5) (c) máximo nos polos e mínimo no Equador.
- 6) Aproximadamente $111,7 \cos \theta$ km; 96,7 km; 110 km; 104,9 km.
- 7) 65° (N ou S) e 67° (N ou S), ambas aproximadamente.
- 8) Aproximadamente 410 km.
- 9) Na figura abaixo temos $x = r \text{sen } \alpha$ e, portanto, a área A da calota esférica é dada por

$$A = 2\pi r h = 2\pi r(r - x) = 2\pi r^2(1 - \text{sen } \alpha).$$



Sendo F a fração visível temos

$$F = \frac{2\pi r^2(1 - \text{sen } \alpha)}{4\pi r^2} = \frac{1 - \text{sen } \alpha}{2}.$$

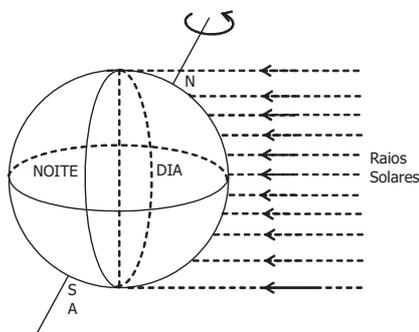
Por outro lado, como $\text{sen } \alpha = \frac{r}{r + d}$ segue que $F = \frac{1}{2} \frac{d}{r + d}$. Para $r = 6400$ e $d = 160$ obtemos $F = 0,012$ de modo que a fração visível é cerca de 1,2%. Para $d = 20200$ obtemos $F = 0,379$ e a fração visível é cerca de 37,9%. Neste caso, a medida 2α do ângulo de visualização é aproximadamente 28° .

2.2 Os Movimentos da Terra

A Terra não está parada. Ela tem dois movimentos principais: o de **rotação**, em torno de seu eixo polar, e o de **translação**, em torno do Sol. As consequências destes dois movimentos são muito importantes em nosso dia-a-dia.

Para dar uma volta completa em torno de seu eixo polar, a Terra leva **24 horas** ou **um dia**. Durante esta rotação, uma metade da Terra está voltada para o Sol, do qual recebe luz e calor: nela então é **dia**. A outra metade não recebe os raios solares: nela é então **noite**. Assim, a sucessão dos dias e das noites é uma consequência da rotação da Terra.

Sob a hipótese de que os raios solares são paralelos temos que, em cada instante, a curva sobre o globo terrestre que separa o dia da noite é uma circunferência máxima.

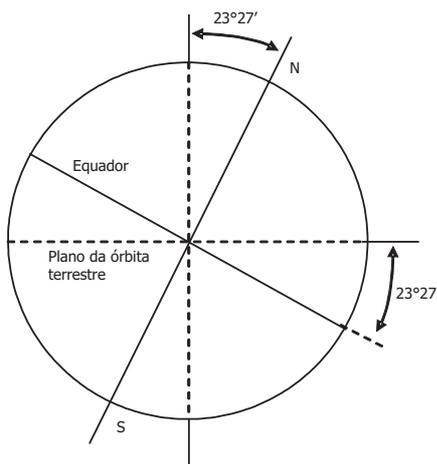


Todos os dias vemos o Sol aparecer no **nascente**, também chamado **oriente** ou **leste**, subir no céu e se deitar no **poente**, também chamado **ocidente** ou **oeste**. Temos a impressão de que é o Sol que caminha no céu. Na realidade, não é isso que acontece. A Terra é que gira em torno de seu eixo polar da esquerda para direita, isto é, do oeste para leste, causando a impressão do movimento do Sol. Chamamos a isto **movimento aparente do Sol**.

O momento em que o Sol, em sua trajetória aparente, está mais alto no céu é o **meio-dia solar**. Neste momento a sombra projetada de uma haste vertical tem o menor comprimento possível. O meio-dia solar para uma particular localidade ocorre quando o meridiano que a contém corta a reta imaginária determinada pelos centros da Terra e do Sol.

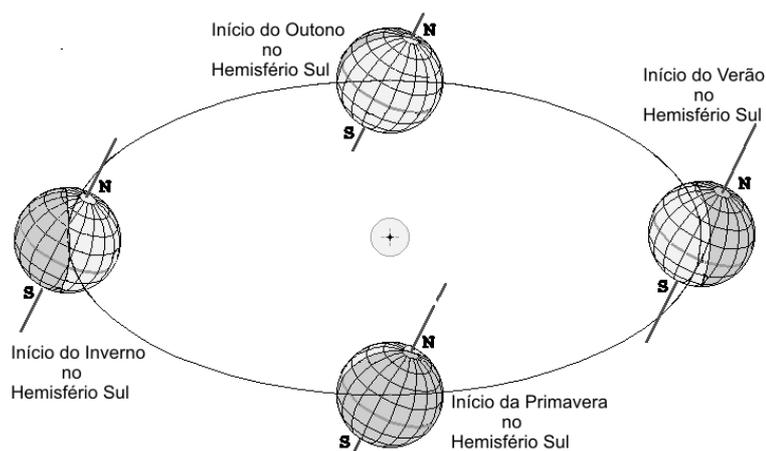
Além de girar em torno de seu eixo, a Terra gira também em torno do Sol. Nesta translação em torno do Sol, a Terra descreve uma curva elíptica denominada **órbita**. A trajetória da órbita da Terra é chamada **eclíptica** e o plano que a contém é referido como o **plano da eclíptica**. Para dar uma volta completa em torno do Sol, a Terra leva 365 dias e 6 horas, isto é, pouco mais que um ano. Para corrigir esta diferença, foi necessário aumentar um dia no ano, a cada período de quatro anos. Nesses anos, chamados bissextos, o mês de fevereiro tem 29 dias.

Um fato extremamente importante deve ser observado: em sua translação em torno do Sol, o eixo da Terra não é perpendicular ao plano da eclíptica, mas **inclinado**, sempre na mesma direção, formando um ângulo de $23^{\circ}27'$ com esta perpendicular.



Se o eixo da Terra fosse perpendicular ao plano da órbita, durante qualquer dia do ano, o Hemisfério Norte receberia a mesma quantidade de luz e calor que o Hemisfério Sul. Mas, como o eixo é inclinado, os dias e as noites não têm a mesma duração em todos os lugares da Terra. Vejamos por que.

Observe a figura abaixo atentamente. Ela mostra as posições da Terra em relação ao Sol ao longo do ano. Como você pode ver, a posição da Terra não é a mesma em todos os meses do ano.



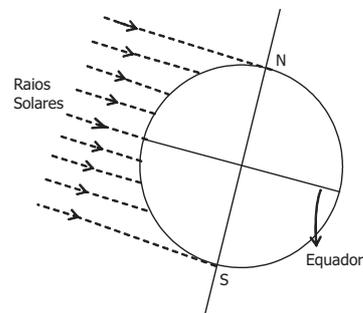
Apenas em duas ocasiões a circunferência máxima que separa a zona iluminada daquela que está no escuro passa rigorosamente pelos polos. Aproximadamente em **21 de março** e em **23 de setembro** a noite e o dia duram, em todos os lugares da Terra, cerca de 12 horas cada um: são os **equinócios** (noites iguais).

Nestas datas a reta imaginária determinada pelos centros da Terra e do Sol corta o globo terrestre num ponto do paralelo de latitude 0° , isto é, **os raios solares incidem perpendicularmente sobre a linha do Equador**, que recebe então mais calor. Para o norte ou para o sul do Equador, o calor vai diminuindo em direção aos polos e

os dois hemisférios são igualmente iluminados e aquecidos.

No dia 21 de março começa para nós (no Hemisfério Sul) o **outono**, enquanto que para os habitantes do Hemisfério Norte é a **primavera** que tem início.

O dia 23 de setembro marca o começo da primavera no Hemisfério Sul e o do outono no Hemisfério Norte.



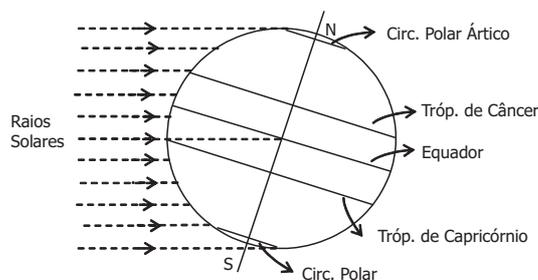
21 de março e 23 de setembro

Vejamos agora o que acontece de setembro a dezembro no Hemisfério Sul, de acordo com o movimento que a Terra faz em torno do Sol.

Durante esta época, os dias vão ficando mais longos e as noites mais curtas, até que, por volta de **21 de dezembro**, temos o dia mais longo e a noite mais curta: é o que chamamos de **solstício de verão**. A Terra inclina o Polo Sul na direção mais próxima do Sol. Nesse dia, começa no Hemisfério Sul o **verão**.

O que se passa no Hemisfério Norte nessa época? Aí a situação se inverte: os dias vão ficando mais curtos e as noites mais longas. A 21 de dezembro, os habitantes daquele hemisfério têm a noite mais longa do ano e o dia mais curto: é o **solstício de inverno**. Para eles, começa o **inverno**.

Você já entendeu por que isso acontece? Observe a figura abaixo. Qual hemisfério está mais exposto aos raios solares em 21 de dezembro? Nessa data a reta imaginária determinada pelos centros da Terra e do Sol corta o globo terrestre num ponto do paralelo de latitude $23^{\circ}27' S$, chamado Trópico de Capricórnio. Assim, no solstício de dezembro, **os raios solares incidem perpendicularmente sobre o Trópico de Capricórnio** que recebe então mais luz e calor fazendo com que o Hemisfério Sul seja mais iluminado e mais quente que o Hemisfério Norte.



21 de dezembro

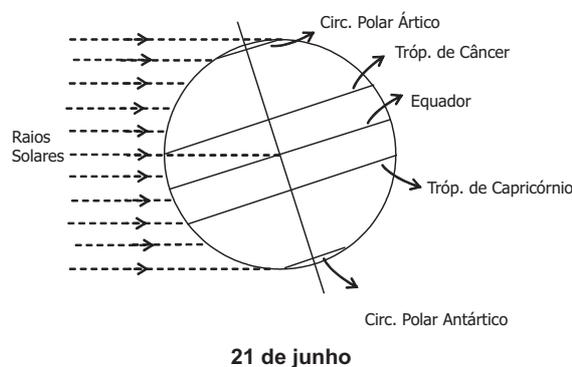
Você já percebeu o que acontece nas regiões polares em 21 de dezembro? Nesse dia, os raios solares não ultrapassam o limite do paralelo de latitude $66^{\circ}33'N$ chamado Círculo Polar Ártico. Os pontos situados entre o Círculo Polar Ártico e o Polo Norte ficam inteiramente no escuro, independentemente da sua longitude. Não há assim dia, e a noite tem duração de 24 horas.

Enquanto isso, no Hemisfério Sul, a zona situada entre o Círculo Polar Antártico, paralelo de latitude $66^{\circ}33'S$, e o Polo Sul fica iluminada durante 24 horas. Não há assim noite, e o dia tem duração de 24 horas.

Vamos agora acompanhar o movimento da Terra de março até

junho. Nessa época, o Hemisfério Sul vai ficando cada vez menos exposto aos raios solares, ao contrário do Hemisfério Norte. Assim, no Hemisfério Sul onde moramos, os dias vão ficando mais curtos e as noites mais longas até que, por volta de **21 de junho**, temos a noite mais longa do ano: é o nosso **solstício de inverno**. Nessa data tem início para nós o **inverno**.

No Hemisfério Norte, a data 21 de junho marca o dia mais longo do ano: é o **solstício de verão**, ou início do **verão**. A Terra inclina seu Polo Norte na direção mais próxima do Sol. A reta imaginária determinada pelos centros da Terra e do Sol corta o globo terrestre num ponto do paralelo de latitude $23^{\circ}27'N$, chamado Trópico de Câncer. Logo, no solstício de junho, **os raios solares incidem perpendicularmente sobre o Trópico de Câncer** e o Hemisfério Norte é, portanto, mais iluminado e mais quente que o Hemisfério Sul.



Vejamos o que ocorre nas regiões polares nesta época. A zona situada entre o Círculo Polar Antártico e o Polo Sul não é atingida pelos raios solares e a noite dura 24 horas, enquanto que a situada entre o Círculo Polar Ártico e o Polo Norte permanece iluminada durante 24 horas. Não há noite e o dia tem duração de 24 horas.

Atividades

- 1) A Terra gira 360° em torno de seu eixo em 24 horas. Quantos graus de longitude a reta determinada pelos centros da Terra e do Sol cruza a cada hora? E a cada minuto?
- 2) Você está numa localidade cujo meio-dia solar ocorreu 6 horas e 8 minutos depois do meio-dia solar em Greenwich. Qual é a sua longitude?
- 3) Quanto tempo depois do meio-dia solar em Greenwich ocorre o meio-dia solar em São Paulo? Qual a diferença de tempo entre o meio-dia solar em São Paulo e em Roma?
- 4) A variação sazonal no tempo de duração do dia pode ser modelada por uma senóide. O dia em Nova Orleans tem cerca de 14 horas no solstício de verão (21 de junho) e cerca de 9 horas e 20 minutos no solstício de inverno (21 de dezembro). Nos itens abaixo desconsidere a possibilidade de um ano bissexto.
 - (a) Esboce o gráfico do número h de horas dos dias de Nova Orleans como função do número x de dias a partir de 21 de março.
 - (b) Encontre uma expressão para h como função de x , determinando A , B e ω de modo que $h(x) = A + B\text{sen}(\omega x)$. (SUGESTÃO: Calcule ω sabendo que h é uma função periódica de período 365. A seguir, obtenha B usando os valores máximos e mínimos da função h .)
 - (c) Seja y o número de dias após 1 de janeiro. Encontre uma expressão para h como função de y .
 - (d) Que dias do ano tem duração aproximada de 13 horas em Nova Orleans?

Respostas

- 1) $15^\circ; 0, 25^\circ$.
- 2) 92°W .
- 3) 3 horas e 8 minutos; 3 horas e 56 minutos.
- 4) (b) $h(x) = 12 + \frac{7}{3}\text{sen}\left(\frac{2\pi x}{365}\right)$.

2.3 Os Fusos Horários

No momento em que for meio-dia em São Paulo, será também meio-dia em todos os lugares situados ao longo do meridiano que passa por São Paulo: é o que chamamos **hora verdadeira**. Todas as localidades situadas sobre um mesmo meridiano têm a mesma hora verdadeira.

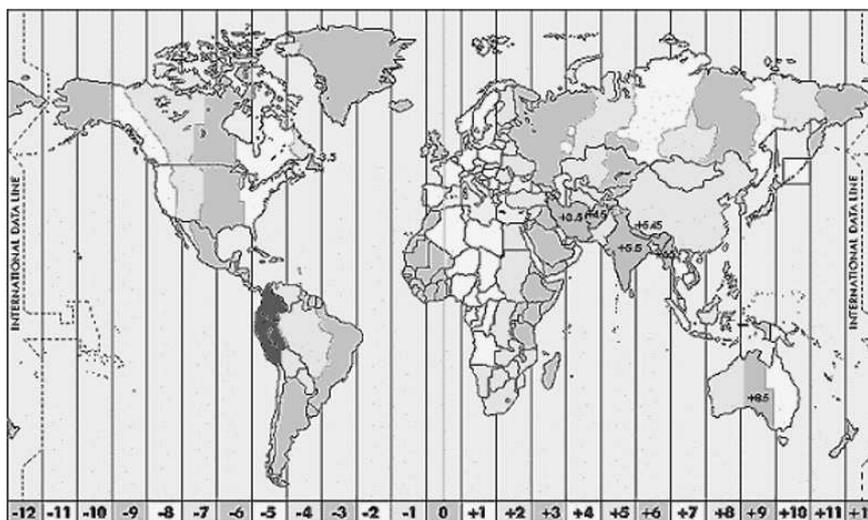
Assim, se é meio-dia em Belo Horizonte, isto é, se o Sol sobre Belo Horizonte está em seu ponto mais alto no céu, o mesmo não acontecerá em Brasília. Isso ocorre porque Belo Horizonte está no meridiano de longitude 44°W , enquanto Brasília encontra-se no meridiano de longitude 48°W . Portanto, meio-dia em Brasília ocorrerá um pouco depois do meio-dia em Belo Horizonte (lembre-se que a Terra gira em torno de seu eixo polar da esquerda para direita). Afinal, se em Belo Horizonte é exatamente meio-dia, qual será a hora verdadeira em Brasília? Para responder essa pergunta basta repetir o argumento utilizado na atividade 3 da página anterior.

Assim, a hora verdadeira não é a mesma em Belo Horizonte, Brasília ou São Paulo. Você já imaginou a confusão que seria se cada lugar tivesse uma hora diferente, mesmo se tratando de lugares próximos uns dos outros? Cada vez que viajássemos para Leste ou para Oeste teríamos que mudar várias vezes os ponteiros do relógio.

▲ SEC. 2.3: OS FUSOS HORÁRIOS

Para tornar as coisas mais simples, foi necessário dividir o globo em 24 partes aproximadamente iguais, uma vez que a Terra leva 24 horas para efetuar seu movimento de rotação. Resultou assim o que chamamos de **fusos horários**. Cada fuso horário corresponde a uma faixa limitada por dois meridianos, distantes 15° um do outro (observe que $360^\circ : 24 = 15^\circ$). Dessa forma, embora haja diferença na hora verdadeira das localidades situadas dentro de um mesmo fuso horário, a hora marcada nos relógios será a mesma para todas essas localidades. É o que chamamos de **hora oficial**. Assim, Rio de Janeiro, Brasília, Belo Horizonte, São Paulo, embora sejam cidades situadas em meridianos diferentes, por encontrarem-se no mesmo fuso horário têm a mesma hora oficial.

Conforme mostra a figura abaixo, países pouco extensos no sentido da longitude, como a Itália e o Paraguai, têm um único fuso horário. Mas países muito extensos possuem vários fusos horários: os Estados Unidos, por exemplo, possuem quatro fusos. Pesquise num atlas geográfico a situação da hora oficial no Brasil.



Atividades

- 1) **Fuso** é a região da superfície terrestre compreendida entre dois meridianos. Calcule a área do fuso limitado pelos meridianos de longitude 20° E e 40° W. (SUGESTÃO: A área de um fuso é proporcional à sua abertura.)
- 2) Indo-se de Oeste para Leste, deve-se adiantar o relógio de uma hora cada vez que se ultrapassa um limite de fuso horário. Utilizando-se um avião bastante rápido, pode-se partir de Greenwich a zero hora do sábado, e chegar ao seu meridiano oposto às 24 horas (tempo local). Nessa altura, serão 12 horas do sábado em Greenwich. Continuando com a mesma velocidade, o viajante completará a volta, podendo chegar a Greenwich às 24 horas do sábado; mas seu relógio, modificado somente em função dos fusos horários marcará 24 horas de domingo! Como resolver este problema?
- 3) Um piloto sai de uma cidade A de latitude 40° N às 12 horas e 15 minutos (hora local verdadeira). Ele voa no rumo Leste a uma velocidade média de 1 000 km/h permanecendo o tempo todo na mesma latitude. Seu destino B é alcançado após percorrer 3 000 km. Supondo que o raio da Terra seja 6 400 km, qual a hora local verdadeira em B?

Respostas

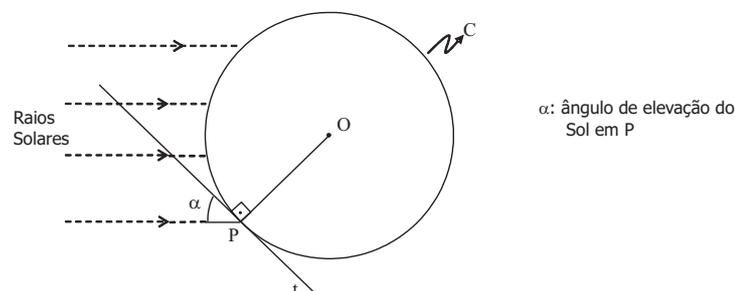
- 1) $\frac{2\pi r^2}{3}$ onde r é o raio da Terra.
- 2) Pesquise sobre a "linha internacional de data."
- 3) 17 horas e 35 minutos.

Capítulo 3

O Ângulo de Elevação do Sol e da Energia Solar

Vimos anteriormente que as estações do ano são consequência da inclinação do eixo da Terra em relação à perpendicular ao plano da eclíptica. A relação entre esta inclinação e a quantidade relativa de energia solar recebida por localidades com diferentes latitudes é uma questão do mundo real que será agora analisada.

A grande maioria da energia solar nunca alcança a superfície da Terra. Porém, a intensidade de radiação que a alcança varia de acordo com o chamado ângulo de elevação do Sol.



42 ■ CAP. 3: O ÂNGULO DE ELEVAÇÃO DO SOL E DA ENERGIA SOLAR

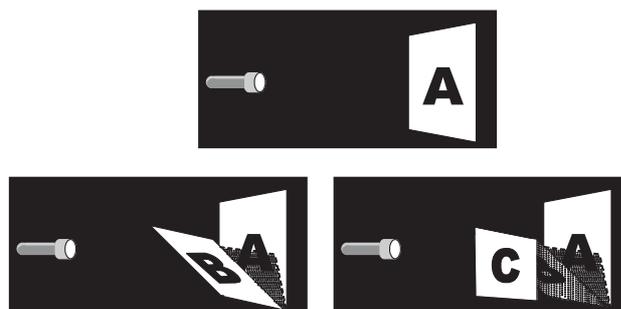
Na figura da página anterior, a circunferência C contém o meridiano que passa por uma localidade P do globo terrestre e t é a reta tangente a C em P . O menor ângulo formado pelos raios solares e pela reta t é chamado o **ângulo de elevação do Sol ao meio-dia solar** em P .

Nesta parte do curso, iremos calcular a intensidade relativa solar que alcança a superfície da Terra em diferentes localidades do mundo e em diferentes épocas do ano. Para nossos propósitos, intensidade relativa é definida como a razão da intensidade de radiação solar incidente na superfície pela intensidade de radiação se os raios incidissem perpendicularmente à superfície (ângulo de elevação com medida igual a 90°).

Você já reparou que a luz solar parece mais quente quando o Sol está a pino? Você é capaz de explicar por que ao meio-dia de um dia claro de verão é muito mais quente que a noite ou a manhã do mesmo dia? O que faz o inverno frio e o verão quente? As respostas a todas estas perguntas estão relacionadas com o ângulo de elevação do Sol. Veremos a seguir que quanto mais próximo de 90° a medida deste ângulo estiver, maior a intensidade da radiação solar. Apesar disso, muitas vezes é mais quente às 15h que ao meio-dia, pois, embora a superfície da Terra receba mais energia solar ao meio-dia, a energia acumulada na atmosfera, desde manhã até este horário, frequentemente torna estas horas as mais quentes do dia.

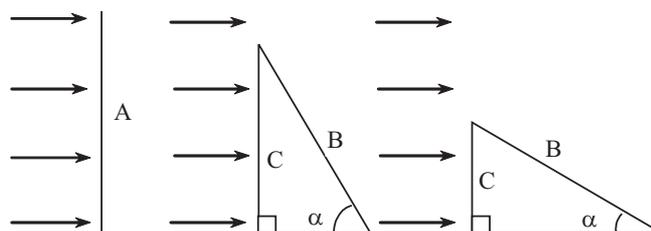
A atividade seguinte oferece um bom modelo para quantificar a relação entre a medida do ângulo de elevação do Sol e a intensidade relativa da radiação solar recebida por uma particular cidade. Escureça a sala e acenda uma lanterna sobre um pedaço de papel branco A . Toda a área do papel será iluminada pela posição perpendicular da lanterna. Coloque outro pedaço de papel idêntico na frente de A , mas incline-o de maneira como mostrada por B . Discuta qual posição do

papel, A ou B,



recebe mais energia luminosa. Observe a sombra que B projeta em A, quando B é mais e mais inclinado. Note que a diminuição da sombra indica que a quantidade de luz sobre B está diminuindo. A quantidade de energia luminosa recebida pelo papel inclinado é a mesma recebida por C, onde C é paralelo a A.

Na figura abaixo, a ilustração da esquerda mostra a visão lateral quando os raios solares atingem A verticalmente, ou perpendicularmente à superfície. As ilustrações do meio e da direita mostram a quantidade relativamente menor de energia que a folha de papel recebe quando o ângulo de elevação do Sol é reduzido.



Quanto mais o papel é inclinado, menos energia luminosa ele recebe. Nas ilustrações anteriores, A e B são na realidade áreas retangu-

lares, mas como as larguras destas áreas são as mesmas, podemos desconsiderá-las e levar em conta apenas os comprimentos mostrados na figura acima.

O comprimento de B é a medida da hipotenusa do triângulo retângulo, α é a medida do ângulo de elevação do Sol e o comprimento de C é a medida do lado oposto a α . A razão entre o comprimento de C e o comprimento de B, que depende do ângulo de elevação e é igual a $\sin \alpha$, nos dá uma medida para a intensidade relativa da luz solar. Se o comprimento de B é tomado como 1, então $\sin \alpha = C/1 = C$. Por exemplo, se o ângulo do papel B mede 90° , isto é, B é perpendicular à direção da luz solar, então $\sin 90^\circ = 1$ e a intensidade relativa da luz solar é 1 ou, em porcentagem, 100%. Como $\sin 80^\circ \approx 0.9848$, temos que quando o ângulo de elevação mede 80° , o papel recebe aproximadamente 98% da radiação solar disponível.

Neste modelo, o pedaço de papel representa uma área da Terra e o ângulo do papel representa o ângulo de elevação do Sol. Quando o Sol se torna visível pela primeira vez de manhã, certa área recebe a luz solar sob um ângulo extremamente pequeno e a intensidade relativa da radiação solar é pequena nesta área. Quando é meio-dia a luz solar é inclinada ao mínimo, ou está perto de se tornar vertical. Ao fim do dia, os raios solares incidem na Terra novamente com o menor ângulo possível. Isto explica por que a luz solar pela manhã ou pela tarde não é tão quente quanto ao meio-dia.

Atividade

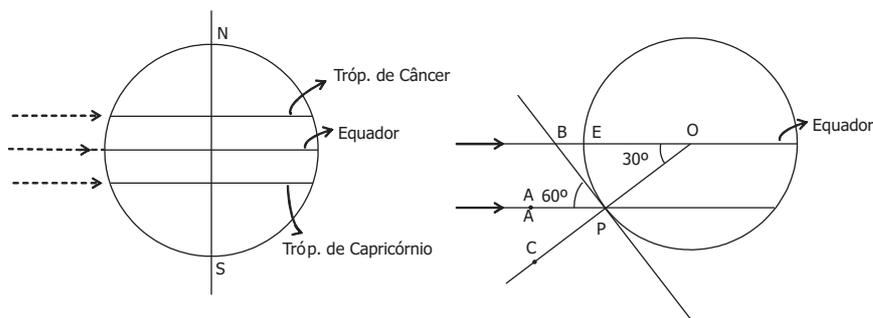
Seja α a medida do ângulo de elevação do Sol ao meio-dia solar em P e β a medida do ângulo de elevação do Sol ao meio-dia solar no Equador, ambas consideradas num mesmo dia. Quais as expressões relacionando α , β e a latitude θ de P? Existem duas possíveis respostas conforme P seja um ponto do Hemisfério Norte ou Sul.

3.1 O Ângulo de Elevação do Sol nos Equinócios

Vamos calcular a medida desse ângulo ao meio-dia solar nos dias de equinócios de primavera e outono. Como vimos anteriormente, nesses dias os raios solares incidem perpendicularmente sobre o Equador. Tomemos, por exemplo, a cidade de Porto Alegre localizada a 30° de latitude sul.

Na figura abaixo, a reta \overleftrightarrow{PB} é tangente a Terra no ponto P, que representa Porto Alegre, O é o centro da Terra e C é um ponto da reta \overleftrightarrow{OP} com P entre C e O. Como $m(\angle BOP) = 30^\circ$ temos $m(\angle APC) = 30^\circ$ (são ângulos correspondentes no contexto de retas paralelas – note que \overleftrightarrow{AP} e \overleftrightarrow{BO} são paralelas aos raios solares). Logo $m(\angle APB) = 60^\circ$ uma vez que $\angle BPC$ é reto.

O ângulo de elevação do Sol ao meio-dia solar mede então 60° durante os equinócios de primavera e outono em Porto Alegre e também em todas as localidades que estão situadas a 30° de latitude sul. Note que o ângulo de elevação do Sol nesses dias e a latitude são ângulos complementares. Para calcular a intensidade relativa da radiação solar encontre $\sin 60^\circ$. Uma calculadora mostrará aproximadamente 0,8660,



indicando que a intensidade de Sol em Porto Alegre nos dias de equinócios é cerca de 86% da que seria se os raios solares incidissem perpendicularmente à superfície, ou ainda, 86% da intensidade do Sol que atinge o ponto E localizado sobre o Equador e que tem a mesma longitude de Porto Alegre.

Pelo mesmo método, podemos calcular a intensidade relativa da radiação solar recebida por qualquer cidade cuja latitude seja conhecida. Por exemplo, Santa Cruz, Argentina, está localizada a 50° de latitude sul. Portanto, o ângulo de elevação do Sol ao meio-dia solar nos dias de equinócios mede $90^\circ - 50^\circ = 40^\circ$.

Para calcular a intensidade relativa da radiação solar que a cidade recebe nesses dias, encontre $\sin 40^\circ$. A calculadora mostrará aproximadamente 0,6427. A intensidade relativa da radiação solar é cerca de 64% ao meio-dia solar em Santa Cruz, nos equinócios de primavera e outono.

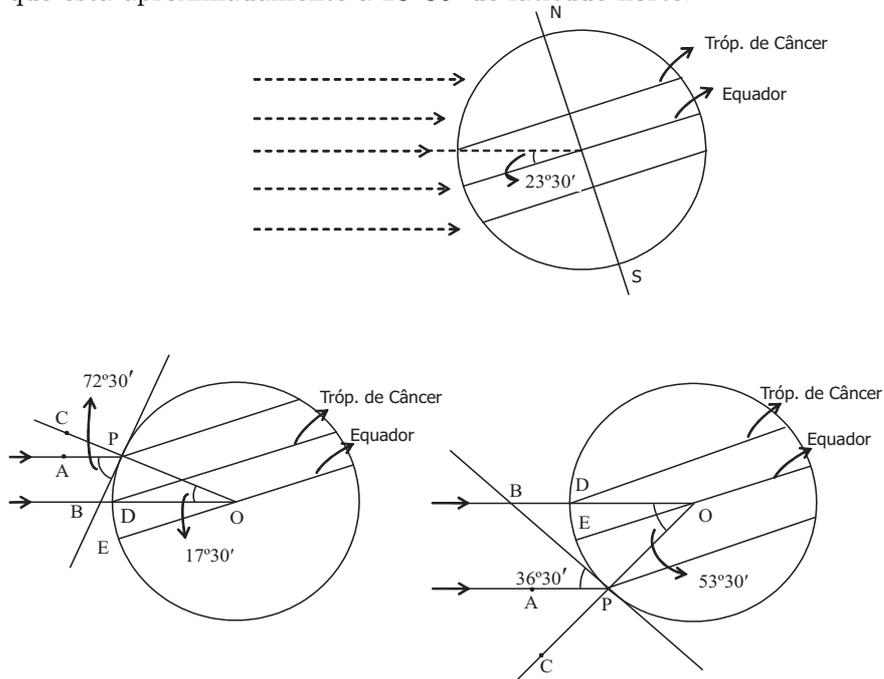
Atividade

Complete a tabela abaixo, obtendo a intensidade relativa da radiação solar ao meio-dia solar nos dias de equinócios.

Cidade	Latitude	Ângulo de Elevação do Sol	Intensidade Relativa
Recife, Brasil	8°S		
Cuiabá, Brasil	16°S		
São Paulo, Brasil	23°S		
Melbourne, Austrália	38°S		
Quito, Equador	0°S		
Miami, EUA	26°N		
Madri, Espanha	41°N		
Oslo, Noruega	60°N		
Barrow, Alaska	71°N		

3.2 O Ângulo de Elevação do Sol nos Solstícios

Vamos agora calcular a medida do ângulo de elevação do Sol ao meio-dia solar no dia de solstício de junho. Sabemos que nesse dia os raios solares incidem perpendicularmente sobre o Trópico de Câncer, que está aproximadamente a $23^{\circ}30'$ de latitude norte.



A ilustração da esquerda na figura acima mostra o ângulo de elevação do Sol em Cleveland (EUA) localizada a 41° de latitude norte. A reta \overleftrightarrow{PB} é tangente a Terra no ponto P que representa Cleveland, O é o centro da Terra e C é um ponto da reta \overleftrightarrow{OP} , com P entre C e O. Como \overleftrightarrow{AP} e \overleftrightarrow{BO} são paralelas aos raios solares temos

$$\begin{aligned} m(\angle APC) &= m(\angle BOP) \\ &= m(\angle EOP) - m(\angle EOB) \\ &= 41^{\circ} - 23^{\circ}30' \\ &= 17^{\circ}30'. \end{aligned}$$

Uma vez que o ângulo BPC é reto segue que APB , que é o ângulo de elevação do Sol ao meio-dia solar, tem medida

$$90^\circ - 17^\circ 30' = 72^\circ 30'.$$

Cleveland e outras cidades localizadas a 41° de latitude norte recebem o maior brilho solar direto do ano durante o dia de solstício de junho, quando o ângulo de elevação do Sol mede $72^\circ 30'$ ao meio-dia solar. Para obter a intensidade relativa da radiação solar que atinge Cleveland nesta hora encontre $\sin 72^\circ 30'$. A calculadora mostrará aproximadamente 0,9537 indicando que Cleveland recebe cerca de 95% da radiação solar que incidiria se o ângulo fosse de 90° ao meio-dia no solstício de junho, ou ainda, 95% da intensidade do Sol que atinge o ponto D localizado sobre o Trópico de Câncer e que tem a mesma longitude de Cleveland.

A ilustração da direita na figura acima mostra o ângulo de elevação do Sol em Porto Alegre. A reta é tangente a Terra no ponto P que representa Porto Alegre, P está C e O com e sendo paralelas aos raios solares. Como P está no Hemisfério Sul, a medida do ângulo BOP é igual a

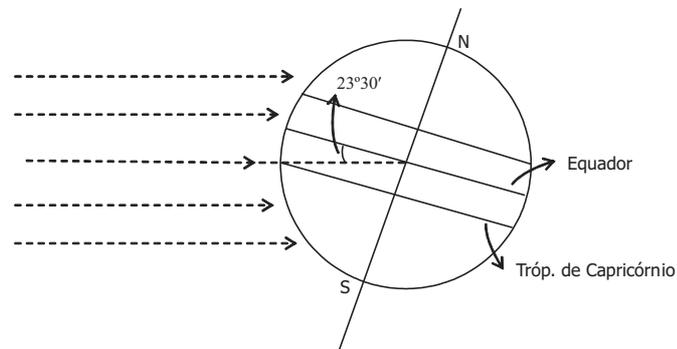
$$m(EOP) + m(EOB) = 30^\circ + 23^\circ 30' = 53^\circ 30'$$

e, pela congruência de ângulos correspondentes no paralelismo, $m(APC) = 53^\circ 30'$. Portanto, $\angle APB$, que é o ângulo de elevação do Sol ao meio-dia solar, tem medida

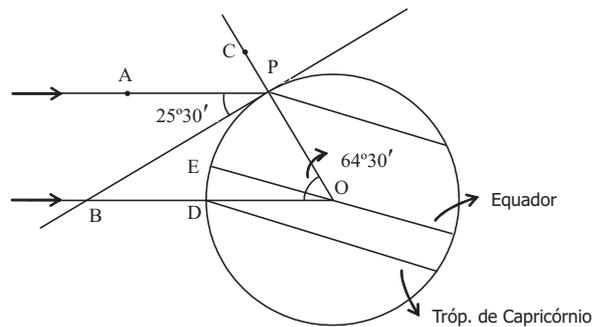
$$90^\circ - 53^\circ 30' = 36^\circ 30'.$$

Porto Alegre e outras cidades localizadas a 30° de latitude sul recebem o menor brilho solar direto do ano durante o dia de solstício de junho, quando o ângulo de elevação do Sol mede $36^\circ 30'$ ao meio-dia solar. Para obter a intensidade relativa da radiação solar que atinge

Porto Alegre nesta hora encontre $\text{sen } 36^\circ 30'$. A calculadora mostrará aproximadamente 0,5948 indicando que Porto Alegre recebe cerca de 59% da radiação solar que incidiria se o ângulo fosse de 90° ao meio-dia no solstício de junho, isto é, 59% da intensidade do Sol que atinge o ponto D localizado sobre o Trópico de Câncer e que tem a mesma longitude de Porto Alegre.



Analogamente podemos calcular a medida do ângulo de elevação do Sol ao meio-dia solar no dia de solstício de dezembro. Nesse dia os raios solares incidem perpendicularmente sobre o Trópico de Capricórnio, que está aproximadamente a $23^\circ 30'$ de latitude sul.



A figura acima mostra o ângulo de elevação do Sol novamente em Cleveland. A reta \overleftrightarrow{PB} é tangente a Terra no ponto P que representa

50 ■ CAP. 3: O ÂNGULO DE ELEVAÇÃO DO SOL E DA ENERGIA SOLAR

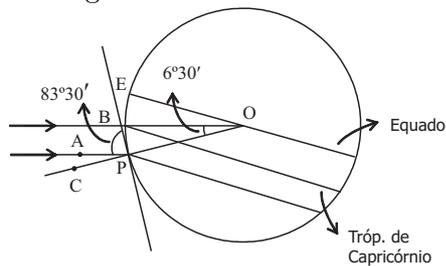
Cleveland, P está entre C e O com \overleftrightarrow{AP} e \overleftrightarrow{BO} sendo paralelas aos raios solares. Como P está no Hemisfério Norte, a medida do ângulo $\angle BOP$ é igual a

$$m(\angle EOP) + m(\angle EOB) = 41^\circ + 23^\circ 30' = 64^\circ 30'$$

e, pela congruência de ângulos correspondentes no paralelismo, $m(\angle APC) = 64^\circ 30'$. Portanto, $\angle APB$, que é o ângulo de elevação do Sol ao meio-dia solar, tem medida

$$90^\circ - 64^\circ 30' = 25^\circ 30'.$$

Cleveland e outras cidades localizadas a 41° de latitude norte recebem o menor brilho solar direto do ano durante o dia de solstício de dezembro, quando o ângulo de elevação do Sol mede $25^\circ 30'$ ao meio-dia solar. Para obter a intensidade relativa da radiação solar que atinge Cleveland nesta hora encontre $\sin 25^\circ 30'$. A calculadora mostrará aproximadamente 0,4305 indicando que Cleveland recebe cerca de 43% da radiação solar que incidiria se o ângulo fosse de 90° ao meio-dia no solstício de dezembro, ou ainda, 43% da intensidade do Sol que atinge o ponto D localizado sobre o Trópico de Capricórnio e que tem a mesma longitude de Cleveland.



A figura anterior mostra o ângulo de elevação do Sol em Porto Alegre. A reta \overleftrightarrow{PB} é tangente a Terra no ponto P que representa Porto Alegre, P está entre C e O com \overleftrightarrow{AP} e \overleftrightarrow{BO} sendo paralelas aos raios solares. Como P está no Hemisfério Sul, a medida do ângulo $\angle BOP$ é igual a

$$m(\angle EOP) - m(\angle EOB) = 30^\circ - 23^\circ 30' = 6^\circ 30'$$

e, pela congruência de ângulos correspondentes no paralelismo, $m(\angle APC) = 6^\circ 30'$. Portanto, $\angle APB$, que é o ângulo de elevação do Sol ao meio-dia solar, tem medida

$$90^\circ - 6^\circ 30' = 83^\circ 30'.$$

Porto Alegre e outras cidades localizadas a 30° de latitude sul recebem o maior brilho solar direto do ano durante o dia de solstício de dezembro, quando o ângulo de elevação do Sol mede $83^\circ 30'$ ao meio-dia solar. Para obter a intensidade relativa da radiação solar que atinge Porto Alegre nesta hora encontre $\sin 83^\circ 30'$. A calculadora mostrará aproximadamente 0,9935 indicando que Porto Alegre recebe cerca de 99% da radiação solar que incidiria se o ângulo fosse de 90° ao meio-dia no solstício de dezembro, isto é, 99% da intensidade do Sol que atinge o ponto localizado sobre Trópico de Capricórnio e que tem a mesma longitude de Porto Alegre.

Vimos nessa seção de que maneira a latitude de um ponto pode ser usada para calcular o ângulo de elevação do Sol ao meio-dia solar naquele ponto em determinadas épocas do ano. Sob certas condições, esse processo pode ser invertido obtendo-se a latitude a partir do ângulo de elevação do Sol (veja atividade 2 a seguir).

Atividades

1) Complete a tabela abaixo, obtendo a intensidade relativa da radiação solar ao meio-dia solar nos dias de solstícios.

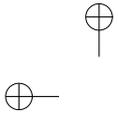
Cidade	Latitude	Solstício de Junho		Solstício de Dezembro	
		Ângulo de elevação do Sol	Intensidade Relativa	Ângulo de elevação do Sol	Intensidade Relativa
Rio Branco, Brasil	10°S				
São Paulo, Brasil	23°S				
Melbourne, Austrália	38°S				
Santa Cruz, Argentina	50°S				
Quito, Equador	0°				
São José, Costa Rica	10°N				
Miami, EUA	26°N				
Paris, França	49°N				
Forel, Groelândia	66°N				

2) No dia 26 de janeiro de 2004 as seguintes medidas foram tomadas numa localidade desconhecida P da região sul do Brasil.

- Altura de uma haste vertical = 35 cm
- Comprimento da menor sombra projetada = 5,4 cm
- Horário local da menor sombra = 12 horas 25 minutos 56 segundos

Consultando um almanaque sabe-se que nesse dia os raios solares incidem perpendicularmente sobre o paralelo de latitude 18°45'S e que o meio-dia solar em Greenwich ocorre às 12 horas 12 minutos 36 segundos. Sabe-se ainda que o horário na localidade P é o de Greenwich diminuído de 3 horas.

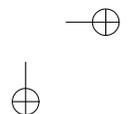
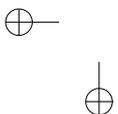
- Determine o tempo decorrido entre o meio-dia solar em Greenwich e o meio-dia solar em P.
- Calcule a longitude de P.



▲ SEC. 3.2: O ÂNGULO DE ELEVAÇÃO DO SOL NOS SOLSTÍCIOS

53

- (c) Determine a medida do ângulo de elevação do Sol ao meio-dia solar naquele dia.
- (d) Calcule a latitude de P.
- (e) Consulte um globo terrestre ou um atlas geográfico para identificar a localidade desconhecida P.

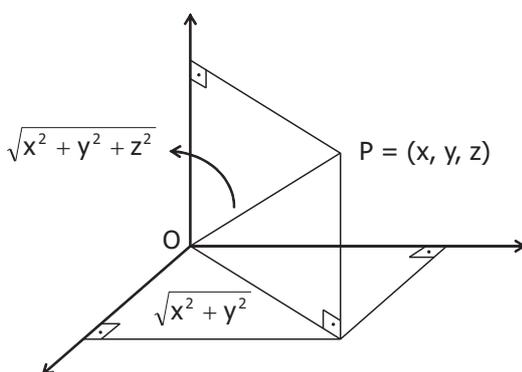


Capítulo 4

A Superfície Esférica em Coordenadas Cartesianas

A partir desta seção trabalharemos num sistema ortogonal de coordenadas cartesianas com origem O . Dado um ponto $P = (x, y, z)$ do espaço, uma dupla aplicação do teorema de Pitágoras mostra que a distância de P a O é expressa por

$$d(P, O) = \sqrt{x^2 + y^2 + z^2}.$$



Mais geralmente, a distância entre os pontos $P = (x, y, z)$ e $C = (u, v, w)$ é dada pela fórmula

$$d(P, C) = \sqrt{(x - u)^2 + (y - v)^2 + (z - w)^2}.$$

Seja r um número real positivo, a superfície esférica S de centro $C = (u, v, w)$ e raio r é o conjunto dos pontos $P = (x, y, z)$ tais que:

$$(x - u)^2 + (y - v)^2 + (z - w)^2 = r^2. \quad (4.1)$$

A equação (4.1) acima é denominada equação reduzida de S . Assim, por exemplo, $(x + 1)^2 + (y - 2)^2 + z^2 = 4$ é a equação reduzida da superfície esférica de centro $C = (-1, 2, 0)$ e raio $r = \sqrt{4} = 2$. Desenvolvendo os quadrados em (4.1), obtemos:

$$x^2 + y^2 + z^2 - 2xu - 2yv - 2zw + u^2 + v^2 + w^2 - r^2 = 0 \quad (4.2)$$

que é uma equação da forma

$$x^2 + y^2 + z^2 + ax + by + cz + d = 0 \quad (4.3)$$

onde a, b, c, d são números reais.

A equação (4.2) é chamada **equação geral** de S . Assim, a superfície esférica de centro $C = (-1, 2, 0)$ e raio $r = 2$ tem como equação geral $x^2 + y^2 + z^2 + 2x - 4y + 1 = 0$.

Dada uma equação da forma (4.3) como decidir se ela é a equação geral de alguma superfície esférica S ? Em caso afirmativo, quais as coordenadas do centro e qual o raio de S ?

56 ■ CAP. 4: A SUPERFÍCIE ESFÉRICA EM COORDENADAS CARTESIANAS

Considerando, por exemplo, a equação

$$x^2 + y^2 + z^2 + 4x - 2y - 6z + 8 = 0,$$

a ideia é **completar os quadrados** e colocar (4.3) na forma (4.1).

Assim:

- $x^2 + 4x = x^2 + 2.2x = x^2 + 2.2x + 2^2 - 2^2 = (x + 2)^2 - 4$
- $y^2 - 2y = y^2 - 2y + 1^2 - 1^2 = (y - 1)^2 - 1$
- $z^2 - 6z = z^2 - 2.3z = z^2 - 2.3z + 3^2 - 3^2 = (z - 3)^2 - 9$

Substituindo na equação dada obtemos

$$(x + 2)^2 - 4 + (y - 1)^2 - 1 + (z - 3)^2 - 9 + 8 = 0,$$

ou seja, $(x + 2)^2 + (y - 1)^2 + (z - 3)^2 = 6$.

Trata-se, portanto, de uma superfície esférica de centro $C = (-2, 1, 3)$ e raio $r = \sqrt{6}$.

Repetindo-se o mesmo argumento para a equação $x^2 + y^2 + z^2 - 2x - 4y + 10 = 0$ obtemos $(x - 1)^2 + (y - 2)^2 + z^2 = -5$ (faça você).

Como uma soma de quadrados nunca é um número negativo concluimos que nenhum ponto do espaço tem coordenadas (x, y, z) satisfazendo a equação acima. Trata-se, portanto, do conjunto vazio.

Já a equação $x^2 + y^2 + z^2 - 2x - 4y - 6z + 14 = 0$ pode ser escrita como $(x - 1)^2 + (y - 2)^2 + (z - 3)^2 = 0$ cuja única solução é o ponto de coordenadas $(x, y, z) = (1, 2, 3)$.

O resultado a seguir desempenhará um papel importante no próximo parágrafo onde discutiremos a fundamentação matemática do funcionamento do GPS.

Teorema 4.1. *Se quatro superfícies esféricas se intersectam e seus centros são não coplanares então essa intersecção consiste de um único ponto.*

Demonstração. Sejam S_1, S_2, S_3 e S_4 superfícies esféricas de centros C_1, C_2, C_3 e C_4 , respectivamente. Mostraremos que se $S_1 \cap S_2 \cap S_3 \cap S_4 \neq \emptyset$ e C_1, C_2, C_3, C_4 são não coplanares então $S_1 \cap S_2 \cap S_3 \cap S_4 = \{P\}$.

Sendo $x^2 + y^2 + z^2 + a_jx + b_jy + c_jz + d_j = 0$ as equações gerais de S_j , onde $j = 1, 2, 3, 4$, ao subtrairmos essas equações duas a duas obtemos equações lineares em x, y e z uma vez que os termos x^2, y^2 e z^2 são eliminados.

Tal equação linear determina o plano que contém a correspondente intersecção. Por exemplo, subtraindo as equações de S_1 e S_2 obtém-se uma equação do plano que contém $S_1 \cap S_2$.

Considerando-se os planos que contém $S_1 \cap S_2, S_1 \cap S_3$ e $S_1 \cap S_4$ temos que se $P = (x, y, z)$ está em $S_1 \cap S_2 \cap S_3 \cap S_4$ então (x, y, z) é a solução do sistema linear

$$(a_1 - a_2)x + (b_1 - b_2)y + (c_1 - c_2)z + (d_1 - d_2) = 0$$

$$(*) \quad (a_1 - a_3)x + (b_1 - b_3)y + (c_1 - c_3)z + (d_1 - d_3) = 0$$

$$(a_1 - a_4)x + (b_1 - b_4)y + (c_1 - c_4)z + (d_1 - d_4) = 0$$

A prova do teorema estará terminada se mostrarmos que o sistema (*) tem uma única solução, pois a existência de dois pontos distintos em $S_1 \cap S_2 \cap S_3 \cap S_4$ acarretariam duas soluções distintas do sistema linear (*).

Sendo $C_j = (u_j, v_j, w_j)$ o centro de $S_j, j = 1, 2, 3, 4$, comparando as equações (4.2) e (4.3) acima temos $a_j = -2u_j, b_j = -2v_j,$

$c_j = -2w_j$ de modo que

$$\begin{vmatrix} a_1 - a_2 & b_1 - b_2 & c_1 - c_2 \\ a_1 - a_3 & b_1 - b_3 & c_1 - c_3 \\ a_1 - a_4 & b_1 - b_4 & c_1 - c_4 \end{vmatrix} = 8 \begin{vmatrix} u_2 - u_1 & v_2 - v_1 & w_2 - w_1 \\ u_3 - u_1 & v_3 - v_1 & w_3 - w_1 \\ u_4 - u_1 & v_4 - v_1 & w_4 - w_1 \end{vmatrix}$$

Como C_1, C_2, C_3, C_4 são não-coplanares segue que o determinante à direita não é nulo e, portanto, (*) é um sistema linear com determinante não nulo tendo assim uma única solução. □

Evidentemente o simples fato do sistema linear (*) ter uma única solução, significando que os centros são não-coplanares, não acarreta necessariamente que a intersecção das quatro superfícies esféricas consiste de um único ponto P. Em outras palavras, a hipótese $S_1 \cap S_2 \cap S_3 \cap S_4 \neq \emptyset$ é essencial para a validade do teorema.

A eventual solução de (*) nos dará o procurado ponto P desde que pertença simultaneamente às quatro superfícies esféricas S_1, S_2, S_3 e S_4 .

Considere, por exemplo, as superfícies esféricas abaixo.

S_1 : centro $(0, 0, 1)$ e raio $\sqrt{2}$

S_2 : centro $(0, 3, 0)$ e raio $\sqrt{10}$

S_3 : centro $(2, 0, 0)$ e raio 1

S_4 : centro $(0, 0, 0)$ e raio 1

Seus centros são não-coplanares e o sistema linear (*), neste caso dado por

$$6y - 2z = 0$$

$$4x - 2z - 4 = 0$$

$$-2z = 0,$$

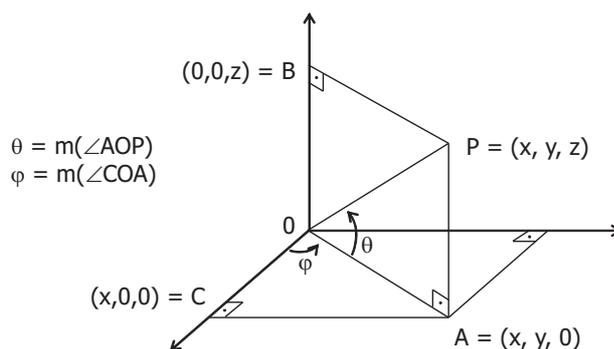
tem como única solução $x = 1, y = 0$ e $z = 0$. Uma verificação

simples mostra que o ponto $P = (1, 0, 0)$ pertence simultaneamente à S_1, S_2, S_3 e S_4 de modo que $S_1 \cap S_2 \cap S_3 \cap S_4 = \{(1, 0, 0)\}$.

Vejam agora, a relação entre as coordenadas geográficas e as coordenadas cartesianas.

Para tanto consideramos um sistema ortogonal de coordenadas cartesianas com origem O no centro da Terra, o eixo Oz positivo apontando na direção do Polo Norte N , o plano Oxy sendo o plano do Equador com o eixo Ox positivo cortando o meridiano de Greenwich e o eixo Oy positivo cortando o meridiano de longitude $90^\circ E$.

Dado um ponto $P = (x,y,z)$ do espaço, sejam θ e φ os ângulos assinalados na figura abaixo.



Quando P está sobre a superfície terrestre os ângulos θ e φ acima indicados correspondem exatamente à latitude e longitude do ponto P como anteriormente definidos na seção 2.1. A diferença entre $OP = \sqrt{x^2 + y^2 + z^2}$ e o raio da Terra é chamada de **elevação** (ou **altitude**) de $P = (x,y,z)$.

No triângulo retângulo $\triangle OPB$ da figura acima temos

$$\cos(90^\circ - \theta) = \frac{OB}{OP} = \frac{z}{\sqrt{x^2 + y^2 + z^2}}$$

60 ■ CAP. 4: A SUPERFÍCIE ESFÉRICA EM COORDENADAS CARTESIANAS

e, como $\cos(90^\circ - \theta) = \sin \theta$, segue que $\sin \theta = \frac{z}{\sqrt{x^2 + y^2 + z^2}}$.

Esta expressão atribui a θ um único valor entre 0 e 90 quando $z > 0$ e um único valor entre -90 e 0 quando $z < 0$. No primeiro caso dizemos que a latitude de P é θ° N enquanto que no segundo a latitude de P é $(-\theta)^\circ$ S.

Por outro lado, no triângulo retângulo $\triangle OAC$ temos

$$\sin \varphi = \frac{AC}{OA} = \frac{y}{\sqrt{x^2 + y^2}} \quad e \quad \cos \varphi = \frac{OC}{OA} = \frac{x}{\sqrt{x^2 + y^2}}.$$

Estas expressões definem um único valor entre 0 e 180 quando $y > 0$ e dizemos que a longitude de P é φ° E. Quando $y < 0$, assume um único valor entre -180 e 0 e, neste caso, a longitude de P é $(-\varphi)^\circ$ W.

Como exemplo, vamos determinar as coordenadas geográficas do ponto P cujas coordenadas cartesianas são dadas por $P = (3\sqrt{3} \times 10^6, -3 \times 10^6, 6\sqrt{3} \times 10^6)$. Considere como unidade de medida o metro.

Temos

$$x^2 + y^2 + z^2 = 27 \times 10^{12} + 9 \times 10^{12} + 108 \times 10^{12} = 144 \times 10^{12}$$

e

$$x^2 + y^2 = 27 \times 10^{12} + 9 \times 10^{12} = 36 \times 10^{12}.$$

Logo, $\sin \theta = \frac{6\sqrt{3} \times 10^6}{12 \times 10^6} = \frac{\sqrt{3}}{2}$ e, portanto, $\theta = 60^\circ$.

Como $\sin \varphi = -\frac{3 \times 10^6}{6 \times 10^6} = -\frac{1}{2}$ e $\cos \varphi = \frac{3\sqrt{3} \times 10^6}{6 \times 10^6} = \frac{\sqrt{3}}{2}$ obtemos $\varphi = -30^\circ$.

Assim as coordenadas geográficas de P são $\theta = 60^\circ$ N e $\varphi = 30^\circ$ W. Supondo o raio da Terra igual a $6,4 \times 10^6$ metros temos que a elevação

de P mede $12 \times 10^6 - 6,4 \times 10^6 = 5,6 \times 10^6$ metros.

O processo acima pode ser invertido: conhecendo-se a latitude θ , a longitude φ e a elevação de um ponto P, podemos determinar suas coordenadas cartesianas x , y e z .

Como antes interpretamos as designações N/S para θ e E/W para φ como positivas/negativas, respectivamente. Por exemplo, um ponto com latitude 40° N e longitude 70° W terá $\theta = 40^\circ$ e $\varphi = -70^\circ$ enquanto que um ponto com latitude 40° S e longitude 70° E terá $\theta = -40^\circ$ e $\varphi = 70^\circ$.

A partir da elevação de P obtemos o valor de $\sqrt{x^2 + y^2 + z^2}$ que denotaremos por r . Logo $\text{sen } \theta = \frac{z}{r}$ e, portanto, $z = r \text{sen } \theta$.

Por outro lado, como $\cos \theta = \text{sen}(90 - \theta) = \frac{\sqrt{x^2 + y^2}}{r}$ segue que $x = \sqrt{x^2 + y^2} \cos \varphi = r \cos \theta \cos \varphi$ e $y = \sqrt{x^2 + y^2} \text{sen } \varphi = r \cos \theta \text{sen } \varphi$.

Para referência futura repetimos no quadro abaixo as relações entre as coordenadas geográficas e as coordenadas cartesianas.

$x = r \cos \theta \cos \varphi$ $y = r \cos \theta \text{sen } \varphi$ $z = r \text{sen } \theta$

Atividades

- 1) Determine a equação geral de cada uma das superfícies esféricas abaixo.
 - (a) S_1 : centro $(5,5,15)$ e raio 13.
 - (b) S_2 : centro $(4,2,5)$ e raio 3.

62 ■ CAP. 4: A SUPERFÍCIE ESFÉRICA EM COORDENADAS CARTESIANAS

- (c) S_3 : centro $(5,1,7)$ e raio 5.
- (d) S_4 : centro $(0, -2, -3)$ e raio 7.
- 2) Verifique se as equações dadas abaixo são equações gerais de superfícies esféricas. Em caso afirmativo, determine o centro e o raio.
- (a) $x^2 + y^2 + z^2 - 4x + 6y + 2z - 2 = 0$.
- (b) $x^2 + y^2 + z^2 - 2x + 2y = 0$.
- (c) $x^2 + y^2 + z^2 - 10x + 2y - 6z + 35 = 0$.
- (d) $x^2 + y^2 + z^2 - 2x + 4y + 15 = 0$.
- 3) Obtenha uma equação da superfície esférica de centro $C = (3,-2,5)$ que é:
- (a) tangente ao plano Oxy;
- (b) tangente ao plano Oyz;
- (c) tangente ao plano Oxz.
- 4) Considere as quatro superfícies esféricas dadas na atividade 1. Determine as equações dos planos que contém $S_1 \cap S_2$, $S_1 \cap S_3$ e $S_1 \cap S_4$. Resolva o sistema linear obtido e ache o ponto de intersecção desses três planos. Verifique que esse ponto pertence simultaneamente às quatro superfícies esféricas S_1, S_2, S_3 e S_4 .
- 5) Determine as coordenadas geográficas e a elevação dos seguintes pontos do espaço (considere as coordenadas cartesianas dadas em metros e o raio da Terra medindo $6,4 \times 10^6$ metros):
- (a) $P = (3\sqrt{3} \times 10^6, 3 \times 10^6, -6\sqrt{3} \times 10^6)$;
- (b) $P = (-3\sqrt{3} \times 10^6, -3 \times 10^6, -6\sqrt{3} \times 10^6)$;
- (c) $P = (-3 \times 10^6, 3\sqrt{3} \times 10^6, 6 \times 10^6)$;

(d) $P = (3 \times 10^5, -3\sqrt{3} \times 10^5, 6 \times 10^5)$.

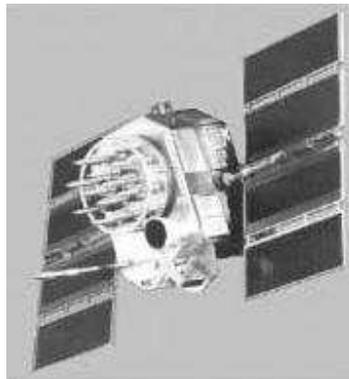
Respostas

- 1) (a) $S_1: x^2 + y^2 + z^2 - 10x - 10y - 30z + 106 = 0$.
(b) $S_2: x^2 + y^2 + z^2 - 8x - 4y - 10z + 36 = 0$.
(c) $S_3: x^2 + y^2 + z^2 - 10x - 2y - 14z + 50 = 0$.
(d) $S_4: x^2 + y^2 + z^2 + 4y + 6z - 36 = 0$.
- 4) $P = (2, 1, 3)$.

Capítulo 5

A Matemática do GPS

A sigla GPS nada mais é do que a abreviatura para **Global Positioning System**. Trata-se de uma constelação de vinte e quatro satélites, orbitando em torno da Terra a uma altura aproximada de 20 200 km acima do nível do mar, permitindo a receptores determinar a sua posição em qualquer lugar sobre a Terra com uma notável precisão.



O projeto foi iniciado em 1973 pelo Departamento de Defesa dos Estados Unidos com o propósito de que aeronaves e navios militares pudessem determinar, em qualquer circunstância de tempo, sua posição exata. Ajuda no lançamento de mísseis e a localização de tropas terrestres em movimento foram outras necessidades que motivaram tal projeto.

Os projetistas do GPS também o planejaram para uso civil, porém, com precisão menor do que para as operações militares.

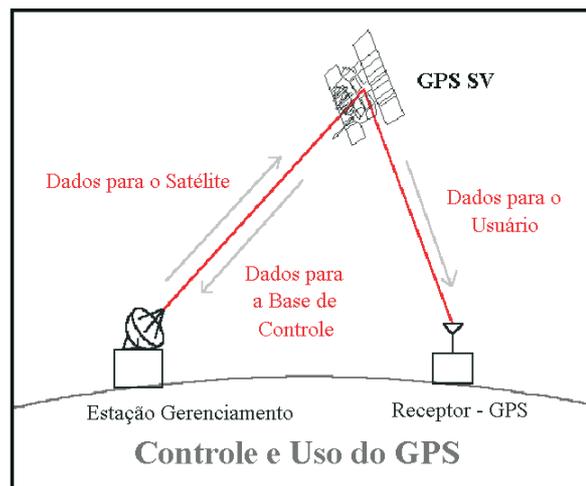
O sistema NAVSTAR (abreviatura para **N**avigation **S**atellite **T**iming and **R**anging), nome oficial dado pelo Departamento de Defesa dos Estados Unidos ao GPS, consiste de um segmento espacial (os satélites), um segmento de controle (as estações terrestres de gerenciamento) e um segmento do usuário.



Os vinte e quatro satélites que formam o segmento espacial do GPS trafegam em torno da Terra em seis órbitas estáveis e predeterminadas com quatro satélites em cada órbita. Os satélites percorrem uma órbita completa a cada 12 horas e cada satélite tem 28° de visualização sobre a Terra (veja atividade 9 da seção 2.1). Isso assegura com que todo ponto da superfície terrestre, em qualquer instante,

esteja visualizado por pelo menos quatro satélites. Várias áreas da Terra são, por alguns momentos, visualizadas por até dez satélites.

Todos os vinte e quatro satélites são controlados pelas estações terrestres de gerenciamento. Existe uma “estação master”, localizada no Colorado (Estados Unidos), que com o auxílio de quatro estações de gerenciamento espalhadas pelo planeta, monitoram o desempenho total do sistema, corrigindo as posições dos satélites e reprogramando o sistema com o padrão necessário. Após o processamento de todos esses dados, as correções e sinais de controle são transferidas de volta para os satélites.



Afinal, de que maneira o GPS determina a localização de um ponto sobre a superfície terrestre?

Cada um dos satélites do GPS transmite por rádio um padrão fixado que é recebido por um receptor na Terra (segmento do usuário) funcionando como um cronômetro extremamente acurado. O receptor mede a diferença entre o tempo que o padrão é recebido e o tempo que foi emitido. Esta diferença, não mais do que um

décimo de segundo, permite que o receptor calcule a distância ao satélite emissor multiplicando-se a velocidade do sinal (aproximadamente $2,99792458 \times 10^8$ m/s – a **velocidade da luz**) pelo tempo que o sinal de rádio levou do satélite ao receptor.

Essa informação localiza uma pessoa sobre uma imaginária superfície esférica com centro no satélite e raio igual à distância acima calculada.

Cada satélite é programado para emitir o que se chama **efeméride**, que informa a sua posição exata, naquele instante, em relação a um sistema ortogonal de coordenadas como o descrito na seção anterior. Tal posição é permanentemente rastreada e conferida pelas estações terrestres de gerenciamento. A unidade receptora processa esses sinais. Com a posição do satélite e a distância acima calculada obtém-se a equação geral da superfície esférica imaginária.

Coletando-se sinais emitidos por quatro satélites, o receptor determina a posição do usuário calculando-a como intersecção das quatro superfícies esféricas obtidas. A localização é dada, não em coordenadas cartesianas, mas por meio das coordenadas geográficas (latitude, longitude) e a elevação.

A precisão do tempo é essencial na operação do GPS. Um erro de um microssegundo (10^{-6} segundos) no registro do lapso de tempo desde a transmissão até a sua recepção resulta num erro de 300 metros. Unidades receptoras do GPS extremamente precisas (e caras!) podem determinar sua posição a menos de um metro.

Com o fim da guerra fria, o sistema GPS passou a oferecer uma precisão muito maior para o usuário civil, disponibilizando a ele a mesma precisão que só os militares tinham a algum tempo atrás. Hoje em dia, com auxílio do piloto automático e do GPS, uma aeronave civil é capaz de percorrer distâncias transatlânticas e pousar sem a interferência do piloto com erro de alguns centímetros com o eixo da

pista.

A navegação é a função primária do GPS sendo usado em aeronaves, navios, veículos e por indivíduos que usam o receptor portátil ("de bolso"). Atualmente o GPS tem se mostrado útil em diversas situações das quais destacamos algumas.

1. Roteirista de viagens: determinam além da sua posição dentro de uma cidade, quais as atrações e pontos turísticos mais próximos, hotéis, postos de emergências, etc.
2. Monitoramento de abalos sísmicos: tais abalos são precedidos por alterações no campo gravitacional que distorcem as ondas de rádio permitindo, através do GPS, tentar prever a ocorrência de um terremoto com algumas horas de antecedência.
3. Meteorologia: o GPS gera informações para a previsão da meteorologia, estudo do clima e outros campos de pesquisa relacionados.
4. Localização para resgate: o serviço usa o GPS para guiar helicópteros de socorro até o lugar do acidente.
5. Aplicações industriais: áreas infectadas por pestes são identificadas por fotografias aéreas e, com uso do GPS, um trator pode ser guiado para aplicações de pesticidas.
6. Uso militar: coordenadas de ataque, orientação e controle para mísseis balísticos, marcação para artilharia, bombardeio de aeronaves, defesa aérea, rastreamento de submarinos, localização de minas e radares inimigos, atos terroristas, etc.

Na atividade a seguir pretendemos discutir, do ponto de vista matemático, o método utilizado pelo GPS na determinação da posição de um ponto sobre a superfície terrestre. As informações transmitidas

no sistema GPS envolvem, por uma questão de precisão, dez ou mais dígitos. Para um aproveitamento mais realista da atividade, sugerimos a utilização de calculadoras ou softwares com capacidade de resolver sistemas lineares com coeficientes dessa ordem. Uma alternativa, abrindo mão eventualmente da precisão, é trabalhar com um número menor de dígitos utilizando a notação científica. Suponha que o raio da Terra meça $6,378164 \times 10^6$ metros e considere a velocidade da luz sendo de $2,99792458 \times 10^8$ m/s.

Atividade

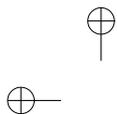
A tabela abaixo indica as efemérides (em metros) de cada satélite.

	x	y	z
Satélite 1	$1,877191188 \times 10^6$	$-1,064608026 \times 10^7$	$2,428036099 \times 10^7$
Satélite 2	$1,098145713 \times 10^7$	$-1,308719098 \times 10^7$	$2,036005484 \times 10^7$
Satélite 3	$2,459587359 \times 10^7$	$-4,336916128 \times 10^6$	$9,090267461 \times 10^6$
Satélite 4	$3,855818937 \times 10^6$	$7,251740720 \times 10^6$	$2,527733606 \times 10^7$

Um receptor GPS registra os seguintes lapsos de tempo (em segundos) entre a transmissão e a recepção do sinal de cada satélite.

Satélite 1	Satélite 2	Satélite 3	Satélite 4
0,08251731391	0,07718558331	0,06890629029	0,07815826940

- Calcule a distância entre o receptor e cada satélite.
- Escreva as equações gerais das imaginárias superfícies esféricas centradas em cada satélite e raios iguais às distâncias calculadas no item anterior.



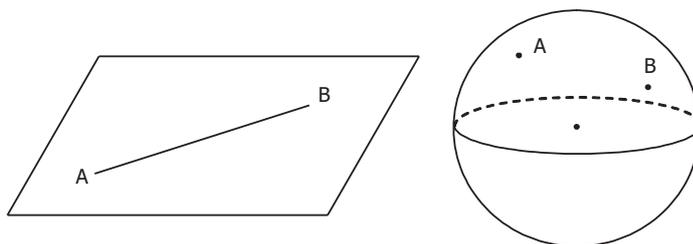
- (c) Determine as coordenadas cartesianas do ponto P que pertence simultaneamente às quatro superfícies esféricas obtidas no item anterior. (SUGESTÃO: Veja atividade 4 do capítulo 4)
- (d) Determine a latitude, a longitude e a elevação do ponto P.
- (e) Consulte um atlas geográfico ou um globo terrestre para identificar a posição desse usuário do GPS.



Capítulo 6

A Distância numa Superfície Esférica

A distância $d(A, B)$ entre dois pontos A e B é, essencialmente, o menor dos comprimentos das trajetórias ligando A a B . No plano, a trajetória de menor comprimento é o segmento de linha reta \overline{AB} e seu comprimento AB é a distância entre A e B . Sobre uma superfície esférica, no entanto, não existe um segmento de linha reta uma vez que ela é curvada em todas as direções e túneis através da Terra não são permitidos. Como medir a distância entre dois pontos A e B neste caso?



Quanto maior o raio de uma circunferência, mais ela se aproxima de ser uma reta. Como as circunferências de maior raio contidas numa superfície esférica S são as circunferências máximas, é razoável esperar que a distância (em S) entre dois pontos A e B seja o comprimento do arco menor \widehat{AB} da circunferência máxima que passa por A e B . Uma prova formal desta afirmação será dada adiante.

O cálculo desse comprimento pode ser feito a partir do conhecimento da medida α do ângulo $\angle AOB$ onde O é o centro da superfície esférica S . Como o comprimento do arco é proporcional à medida do ângulo central correspondente, uma regra de três simples nos dá o valor procurado.

Sendo r o raio da superfície esférica temos

$$\begin{array}{rcl} 360^\circ & \dots\dots & 2\pi r \\ \alpha^\circ & \dots\dots & d(A,B) \end{array}$$

de modo que $d(A,B) = (\alpha/360).(2\pi).r$.

Todos os meridianos estão contidos em circunferências máximas enquanto que, entre os paralelos, apenas o Equador é uma circunferência máxima. Logo quando A e B possuem a mesma longitude, a diferença entre as latitudes pode ser usada para achar a medida α . Analogamente quando A e B estão sobre o Equador é a diferença entre as longitudes que nos permite calcular α . Vejamos como.

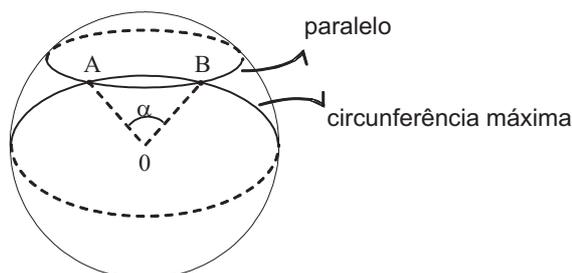
As cidades de Curitiba e Goiânia estão sobre o mesmo meridiano (49° W) e suas latitudes são 26° S e 17° S, respectivamente. Estão assim separadas por 9° de latitude e, tomando o raio da Terra como 6 400 km, segue que a distância entre elas é dada por

$$(9/360).(2\pi).6\,400 \approx 1\,005 \text{ km.}$$

As cidades de Quito, no Equador, e Entebe, em Uganda, estão ambas sobre o Equador. A longitude de Quito é 79° W enquanto que a de Entebe é 32° E. Logo a diferença entre suas longitudes é de 111° de modo que a distância entre elas é igual a

$$(111/360).(2\pi).6\,400 \approx 12\,399 \text{ km.}$$

Quando duas cidades A e B estão sobre um mesmo paralelo, que não seja o Equador, o caminho mais curto possível entre elas, ao contrário do que diz nossa intuição, **não é** o comprimento do arco menor \widehat{AB} daquele paralelo e sim o comprimento do arco menor \widehat{AB} da circunferência máxima que passa por A e B .



Por exemplo, as cidades de Nova York e Nápoles estão praticamente sobre o mesmo paralelo (41° N) e suas longitudes são 74° W e 14° E, respectivamente. O comprimento do arco menor do paralelo entre as duas cidades é cerca de 7 419 km (verifique).

Se A e B representam as cidades de Nova York e Nápoles, respectivamente, vejamos como calcular neste caso o comprimento do arco menor \widehat{AB} da circunferência máxima que passa por A e B , ou seja, como calcular $\alpha = m(\angle AOB)$.

Considerando-se um sistema ortogonal de coordenadas cartesianas como descrito no capítulo 4 e supondo, como de costume, o raio da

Terra medindo 6 400 km, podemos escrever

$$A = 6\,400(\cos 41^\circ \cos(-74^\circ), \cos 41^\circ \sin(-74^\circ), \sin 41^\circ)$$

$$B = 6\,400(\cos 41^\circ \cos 14^\circ, \cos 41^\circ \sin 14^\circ, \sin 41^\circ),$$

ou seja,

$$A = 6\,400(0,20802, -0,72547, 0,65606)$$

$$B = 6\,400(0,73229, 0,18257, 0,65606).$$

A medida procurada α será obtida por meio da relação $\langle \vec{OA}, \vec{OB} \rangle = \|\vec{OA}\| \|\vec{OB}\| \cos \alpha$, onde \langle, \rangle indica o produto interno usual entre os vetores \vec{OA} e \vec{OB} enquanto que $\|\vec{OA}\|$, $\|\vec{OB}\|$ são os módulos desses vetores, neste caso ambos iguais as 6 400.

Como

$$\langle \vec{OA}, \vec{OB} \rangle = 6\,400^2 [0,15233 - 0,13244 + 0,43041] = 6\,400^2 \times 0,4503,$$

segue que $\cos = 0,4503$ e, portanto, $\alpha = 63^\circ$.

O comprimento do arco menor \widehat{AB} da circunferência máxima que passa por A e B , que é a distância entre A e B , é então dada por $(63/360) \cdot 2\pi \cdot 6\,400 \approx 7\,037\text{km}$. Note como esta distância é menor do que aquela calculada ao longo do paralelo.

O argumento acima pode ser utilizado para calcular a distância entre dois pontos quaisquer A e B do globo terrestre. Dadas suas coordenadas geográficas, obtemos suas coordenadas cartesianas e, usando o produto interno $\langle \vec{OA}, \vec{OB} \rangle$, determinamos $\alpha = m(\angle AOB)$. A distância procurada $d(A, B)$ é então dada por $d(A, B) = (\alpha/360) \cdot (2\pi) \cdot 6\,400$.

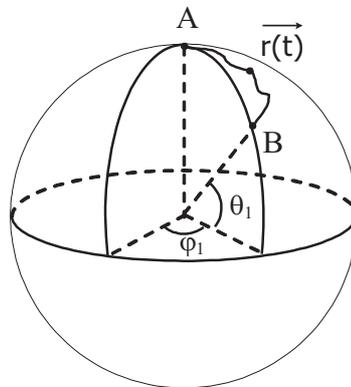
Encerramos a seção apresentando uma prova formal do fato que $d(A, B)$ é o comprimento do arco menor \widehat{AB} da circunferência máxima que passa por A e B . Isso pode ser feito com umas poucas simplificações e a ajuda do cálculo diferencial e integral.

Primeiro, podemos supor sem perda de generalidade que a superfície esférica que estamos trabalhando tem raio igual a 1 unidade (a que você quiser) e que seu centro é a origem do sistema ortogonal de coordenadas cartesianas.

Segundo, pela simetria da superfície esférica podemos assumir que o ponto A é o Polo Norte. O outro ponto, B , será dado pelas suas coordenadas geográficas que vamos supor medidas em **radianos**. Digamos que a latitude de B é θ_1 enquanto que sua longitude é φ_1 .

A circunferência máxima que passa por A e B contém o meridiano por B de modo que o arco menor \widehat{AB} tem por comprimento $\frac{\pi}{2} - \theta_1$ (lembre-se que o raio da superfície esférica é 1).

Devemos agora considerar uma trajetória arbitrária ligando A e B e mostrar que seu comprimento é maior ou igual a $\frac{\pi}{2} - \theta_1$.



Se nos imaginarmos viajando de A até B , em cada instante t , estaremos num ponto do globo terrestre com uma latitude $\theta(t)$ e uma longitude $\varphi(t)$. Iniciando nossa viagem no instante $t = 0$ e a finalizando no instante t_1 teremos $\theta(0) = \frac{\pi}{2}$, $\theta(t_1) = \theta_1$ e $\varphi(t_1) = \varphi_1$.

Uma trajetória arbitrária ligando A e B é então descrita por um

vetor posição $\vec{r}(t)$ dado por

$$\vec{r}(t) = (\cos \theta(t) \cos \varphi(t), \cos \theta(t) \operatorname{sen} \varphi(t), \operatorname{sen} \theta(t)), 0 \leq t \leq t_1.$$

Derivando-se essa função vetorial em relação ao tempo t obtemos o vetor velocidade $\vec{r}'(t)$ da trajetória e conseqüentemente sua velocidade $\|\vec{r}'(t)\|$ no instante t .

Cálculos rotineiros nos dão que

$$\|\vec{r}'(t)\| = \sqrt{\theta'(t)^2 + \varphi'(t)^2 \cos^2 \theta(t)} \geq \sqrt{\theta'(t)^2} = |\theta'(t)| \geq -\theta'(t).$$

Lembrando que o comprimento da trajetória é calculado pela expressão $L = \int_0^{t_1} \|\vec{r}'(t)\| dt$, concluímos que

$$L \geq \int_0^{t_1} [-\theta'(t)] dt = -[\theta(t_1) - \theta(0)] = \frac{\pi}{2} - \theta_1$$

e temos assim provado o resultado desejado.

Atividades

Em todos os exercícios desta série suponha que o raio da Terra meça 6 400 km.

- 1) As cidades de Macapá (Brasil) e Pontianak (Indonésia) estão ambas situadas sobre o Equador. Consulte um atlas geográfico para achar a longitude de cada uma dessas localidades e determine a distância entre elas.
- 2) Qual a distância de Salvador ao Polo Sul? E ao Polo Norte? E a Fortaleza? (SUGESTÃO: Salvador e Fortaleza estão sobre um mesmo meridiano.)
- 3) Chicago e Roma situam-se na mesma latitude (42°N), mas em longitudes diferentes: a de Chicago é aproximadamente 88°W e a de Roma aproximadamente 12°E .

- (a) Suponha que um piloto tenha ido de Chicago a Roma em voo no rumo leste, permanecendo o tempo todo na mesma latitude. (Muita gente acha que esse é o caminho mais curto possível.) Que distância o piloto teve de voar?
- (b) Qual a distância de Chicago a Roma por circunferência máxima?
- (c) Quanto tempo se economizaria voando em uma circunferência máxima num avião a jato capaz de fazer uma velocidade média de 900 km/h?
- 4) Qual é a distância entre Nova York ($40^{\circ}40'N$, $74^{\circ}W$) e Buenos Aires ($34^{\circ}30'S$, $58^{\circ}30'W$)?

Respostas

- 3) (a) Cerca de 8 300 km.
(b) Cerca de 7 752 km.
(c) Aproximadamente 36 minutos.
- 4) Cerca de 8 545 km.



Referências Bibliográficas

- [1] AZEVEDO, G.G.; SANTOS, F.M. *Panorama do Brasil I*. Rio de Janeiro: Editora Atual, 1990.

- [2] FEEMAN, T.G. *Portraits of the Earth*. [S.L.]: American Mathematical Society, 2002.

- [3] HY, Kim. Angled Sunshine, Seasons and Solar Energy. *The Mathematics Teacher*, v. 90, n. 7, October 1997.

- [4] JENNINGS, G.A. *Modern Geometry with Applications*. [S.L.:s.n.], 1997.

- [5] MOISE, E. *Elementary Geometry from an Advanced Standpoint*. Third Edition. [S.L.]: Addison Wesley Publishing Co., 1990.

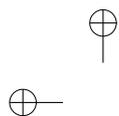
- [6] NORD, G.D.; JABON, D.; NORD, J. The Mathematics of the Global Positioning System. *The Mathematics Teacher*, v. 90, n. 6, September 1997.

- [7] WAIVERIS, C.; CRAINE, T.V. Where are we?. *The Mathematics Teacher*, v. 89, n. 6, September 1996.

Sites Consultados

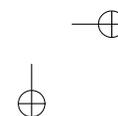
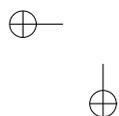
[1] www.garmin.com

[2] <http://astro.if.ufrgs.br/indexhtml>



Os Três Problemas Clássicos da Matemática Grega

João Pitombeira de Carvalho





Prefácio

Este trabalho foi escrito originalmente para um minicurso apresentado na Bienal da Sociedade Brasileira de Matemática, em Salvador, em 2004. Posteriormente, foi feita uma revisão completa do texto, com a correção de erros de impressão, e incorporadas sugestões dos alunos. É esta versão aqui impressa. Mais tarde, foi refeito e publicado como um dos capítulos do livro de João Bosco Pitombeira de Carvalho, *Três excursões pela História da Matemática*.¹

No minicurso, usamos o aplicativo de geometria dinâmica “TABULAE”, para que os alunos pudessem fazer, realmente, as construções geométricas. A mais elaborada, que exige paciência, é a da quadratriz. Acreditamos que a utilização da geometria dinâmica dá vida ao estudo dos problemas clássicos, pois, com ela, é possível realmente ver que as construções feitas pelos gregos funcionam. Isso é instrutivo principalmente no caso das “máquinas” de Platão e de Eratóstenes (ver, 7.1 e 7.2, respectivamente). Esta fusão da história com as tecnologias torna o estudo dos problemas clássicos bem mais interessante. Outra ocasião em que a geometria dinâmica dá vida ao estudo das construções geométricas dos gregos é no estudo de construções por neusis ou ajustamento. Com um aplicativo de geometria dinâmica, o leitor pode ver que a construção é possível.

¹Os dois outros capítulos tratam, respectivamente, da equivalência e aplicação de áreas na matemática grega e da história da equação do segundo grau.

Os tópicos expostos neste texto têm encantado sucessivas gerações de matemáticos profissionais ou amadores. Eles são os problemas da *duplicação do cubo*, *quadratura do círculo* e *trisseção do ângulo*, os quais moldaram muito das atividades de pesquisa matemática dos gregos, e deram origem, inclusive, ao estudo das cônicas. É vastíssima a produção matemática dedicada a estes três problemas, provenientes de erros, por vezes sutis, sobre as exigências feitas para a solução dos problemas. Somente no século XIX é que foi demonstrado que eles não podem ser resolvidos somente com régua (não graduada) e compasso, o que não impede o aparecimento, até hoje, de “provas” de que é possível efetuar as três construções. Essas tentativas não devem ser olhadas com desprezo, mas sim como testemunho de interesse pela Matemática em vários setores na sociedade.

Cabe um pedido de desculpas e, ao mesmo tempo, uma explicação: os leitores notarão por certo a predominância de referências bibliográficas em línguas estrangeiras. Isso se deve ao pouco número de fontes em português sobre o assunto. Embora progressos estejam sendo feitos no sentido de tornar disponíveis, em nossa língua, bons textos sobre os tópicos cobertos neste livro, até hoje a maior parte dos trabalhos sobre o assunto estão escritos em outros idiomas. Em verdade, a motivação para a redação deste material foi tornar disponível exposições introdutórias sobre os tópicos aqui tratados.

Neste trabalho, o livro *Science Awakening*, de Van der Waerden, é simplesmente citado como van der Waerden. A edição padrão, facilmente disponível, dos *Elementos* de Euclides é a de Heath (1956). Para os leitores que não têm acesso a esta edição, recomendamos a tradução da edição de Simson (1773) para o português, perfeitamente suficiente para nossas finalidades. ² Teoremas, figuras, equações ou

²Ela se encontra disponível gratuitamente em www.dominiopublico.gov.br.

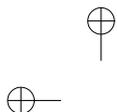


igualdades são numeradas sequencialmente. O símbolo \square indica o fim de uma demonstração ou construção.

Com exceção das Figuras 7.8, 8.1 e 9.4, todas as ilustrações foram feitas com o aplicativo de Geometria dinâmica *TABULAE* [®], desenvolvido no Instituto de Matemática da Universidade Federal do Rio de Janeiro.

Embora o autor tenha procurado corrigir todos os erros tipográficos ou matemáticos do texto, certamente alguns lhe escaparam. Como disse Monteiro Lobato, os erros tipográficos são como sacis. Quando os procuramos, eles se escondem. Uma vez publicado o texto, saltam alegres chamando a atenção dos leitores. Fico, desde já, muito grato pelas indicações dos erros remanescentes (matemáticos ou tipográficos), as quais podem ser encaminhadas para jbpfcarvalho@gmail.com.

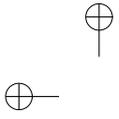
João Bosco Pitombeira de Carvalho



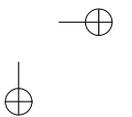
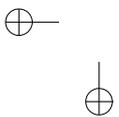


Sumário

Introdução	89
7 A Duplicação do Cubo	97
7.1 A máquina de Platão	99
7.2 A máquina de Eratóstenes	100
7.3 A solução de Nicomedes	104
7.4 A Construção de Árquitas	110
7.5 A solução achada por Menécmo	113
7.6 O método de Diocles	114
7.7 O método de Hierão	117
8 A Quadratura do Círculo	119
8.1 A Quadratriz	119
9 A Trissecção do Ângulo	122
9.1 A Trissecção do Ângulo por Arquimedes	123
9.2 A Trissecção do Ângulo por Nicomedes	124
9.3 A Trissecção do Ângulo Usando a Quadratriz	126



Referências Bibliográficas



Introdução

É indubitável que, na história da Matemática, alguns problemas têm significação especial: agindo como “catalisadores” eles influenciam muito o desenvolvimento da ciência. Tais problemas atraem devido à simplicidade e lucidez de seus enunciados, fascinando muitos matemáticos. Como resultado, são elaborados novos métodos e, até mesmo novas teorias e novas perguntas, profundas e abrangentes, são formuladas (Raigorodski, 2004).

Os matemáticos gregos estudaram três problemas de Geometria que desempenharam papel importante no desenvolvimento da Matemática. Eles são problemas de construção e resistiram a todas as tentativas dos gregos para resolvê-los utilizando somente a régua sem graduação e o compasso, os únicos instrumentos empregados por Euclides nos Elementos.

Os problemas, que ficaram conhecidos como os *três problemas clássicos*, são:

- 1 – A duplicação do cubo.
- 2 – A quadratura do círculo.
- 3 – A trissecação do ângulo.

Sabemos, desde o século XIX, que esses problemas não podem ser resolvidos somente com a régua e o compasso. Referências acessíveis sobre isso são, por exemplo, Courant and Robbins (1996), Hadlock (1978), Klein (1930), Bunt, Jones and Bedient (1988, p. 89-121). Uma discussão de por que os gregos tentavam resolver problemas de construção usando somente a régua e o compasso, pode ser encontrada, por exemplo, em Bkouche et Joëlle (1993).

Para os primeiros geômetras gregos, uma linha era o percurso de um ponto, e a linha reta era um percurso sem asperezas e desvios (Szabó, 2000). No entanto, aos poucos, os matemáticos gregos se distanciaram da realidade palpável, como se vê, por exemplo, em Platão:

[a Geometria] *tem por objeto o conhecimento do que sempre é e não do que nasce e perece.*

Passaram também a dar ao círculo e à reta papéis destacados:

...*Aristóteles – O que não tem nem começo nem fim é portanto ilimitado.*

Parmênides – Ele é ilimitado.

Aristóteles – Portanto ele não tem forma, pois não participa nem do redondo nem do reto.

Além da ideia de perfeição ideal atribuída ao círculo e à linha reta uma outra razão para a restrição à régua e ao compasso pode ter sido a crise devida à descoberta da irracionalidade de $\sqrt{2}$, número que pode, no entanto ser construído com régua e compasso. Esses instrumentos eram a garantia da existência de tais números.

No entanto, é falsa a crença de que os gregos, na resolução de problemas de construções geométricas, trabalhavam somente com a régua e o compasso. Exatamente como os matemáticos de hoje, para resolverem um problema eles usavam todas as ferramentas disponíveis ou criavam novas ferramentas apropriadas. De suas tentativas para achar soluções para os problemas clássicos, surgiram várias curvas e métodos que enriqueceram a Matemática. Encontram-se em Knorr (1986) e Bos (2001) construções geométricas, incluindo soluções dos três problemas clássicos, utilizando várias curvas e outros instrumentos.

O matemático Van der Waerden (p. 263) resumiu a situação como segue:

A ideia por vezes expressa de que os gregos permitiam somente construções com régua e compasso é inadmissível. Ela é negada pelas numerosas construções que nos chegaram para a duplicação do cubo e a trissecção do ângulo. No entanto, é verdade que tais construções eram consideradas mais elementares, e Pappus afirma que sempre que uma construção for possível com régua e compasso métodos mais avançados não deveriam ser usados.

Isso parece seguir a mesma linha que o princípio da *navalha de Ockham*, muito posterior. Na linguagem popular, poderíamos dizer

que não se deve usar um canhão para matar uma mosca.

É impossível superestimar a importância destes problemas. Como diz Yates (1971, p. 5).

Na história da Matemática há três problemas que persistiram com vigor impressionante durante mais de dois mil anos. Eles são a trissecção do ângulo, a duplicação do cubo e a quadratura do círculo, e devido à sua existência robusta eles são atualmente chamados de problemas famosos. (...) Estes três problemas, solidamente inexpugnáveis malgrado todas as tentativas usando geometria plana, o método matemático dos antigos gregos, fizeram com que os matemáticos ficassem fascinados e construíssem novas técnicas e teoremas para sua solução. Por meio deste estímulo surgiu grande parte das estruturas atuais da álgebra e da geometria.

A procura constante de soluções para os três problemas durante tanto tempo forneceu descobertas frutíferas e que lançaram luz sobre tópicos bem distantes. Somente em 1837 é que foi demonstrado, por Wantzel, que um número real é construtível com régua e compasso se, e somente se, ele é um número algébrico, de grau uma potência de dois, sobre os racionais.

Os três problemas são muito naturais para quem tem curiosidade matemática. Sócrates, no diálogo *Meno* de Platão, usando perguntas apropriadas, faz com que um jovem escravo ache um quadrado cuja área é duas vezes a área de um quadrado dado. Isso pode ser feito facilmente usando régua e compasso.

Com efeito, o quadrado cujo lado é a diagonal do quadrado dado é a solução do problema. Se AB é o lado do quadrado dado

(Figura 6.1), então

$$DB = \sqrt{2} \times AB \implies DB^2 = 2 \times AB^2.$$

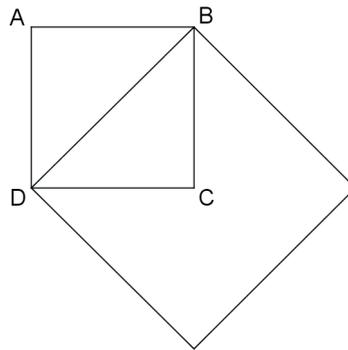


Figura 6.1: Duplicação do quadrado

De $DB^2 = 2 \times AB^2$, vemos imediatamente que

$$\frac{AB}{DB} = \frac{DB}{2AB}.$$

Assim, achar o comprimento de DB é equivalente a inserir uma *meia proporcional* entre AB e $2AB$.

De maneira mais geral, se desejarmos construir um quadrado cuja área seja b vezes a área a do quadrado $ABCD$, devemos ter

$$\frac{AB}{DB} = \frac{DB}{bAB}.$$

Como veremos, a ideia de inserir meias proporcionais entre duas grandezas dadas está por traz da maioria das tentativas de duplicar o cubo.

“Quadrar” uma região plana consiste em traçar, *somente com*

régua e compasso, um quadrado cuja área seja igual à área da região dada. O problema de quadrar qualquer região poligonal está completamente resolvido nos *Elementos* de Euclides, e irá ser estudado no Capítulo 7. O problema da quadratura do círculo é também muito natural. Uma vez resolvido o problema da quadratura de qualquer região poligonal, o próximo passo é tentar “quadrar” regiões limitadas por linhas curvas. Entre estas regiões, o círculo é uma escolha óbvia. Isso levou à investigação das “lúnulas” por Hipócrates de Quios, em torno de 430 a.C. (Van der Waerden, p. 131-132). Curiosamente, somente há pouco tempo, em 1947, usando técnicas muito sofisticadas, é que o problema de achar todas as lúnulas “quadráveis” foi completamente resolvido (Scriba, 1987)!

A primeira menção conhecida do problema da quadratura do círculo encontra-se no problema 50 do papiro Rhind, em torno de 1600 a.C.:

Um campo circular tem 9 khet de diâmetro. Qual é a sua área?

Resolução: Tira 1/9 do diâmetro do seu diâmetro, isto é 1 Khet. O resto é 8 Khet. Multiplica 8 por 8; o que faz 64. Por isso, contém 64 setat de terra.

Nota: 1 setat é khet ao quadrado

Em sua comédia *Os pássaros*, Aristófanes introduz o astrônomo Meton e o ridiculariza por causa de suas tentativas de fazer a quadratura do círculo:

*Farei minhas medições com um esquadro reto [90° graus],
e assim você observa que o círculo se torna quadrangular.*

Para Szabö (2000), o problema de quadratura que deu origem a todos os outros foi o de fazer a quadratura do retângulo. Este problema é facilmente resolvido com régua e compasso usando o resultado que,

em um triângulo retângulo, a altura relativa à hipotenusa é a meia proporcional entre os segmentos que ela determina sobre a hipotenusa.

Aristóteles,³ por sua vez, pensa que este problema surgiu da procura da média geométrica (meia proporcional), mas que isso foi esquecido e só restou o próprio problema:

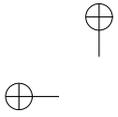
A definição não deve contentar-se em exprimir em que consiste a coisa (...), mas ela deve também incluir e exibir a causa. Ora, as definições são geralmente conclusões. Por exemplo: O que é a quadratura? É a igualdade de um quadrado e de um retângulo. Uma definição como essa é uma conclusão. Mas dizer que a quadratura é a descoberta da meia proporcional é exprimir a causa do que é definido.

Semelhantemente, uma vez que se sabe como bissectar um ângulo (*Elementos*, I-9), é natural perguntar como dividir um ângulo em n partes; em particular, em 3 partes.

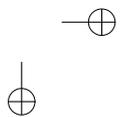
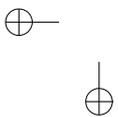
Embora não tenham conseguido resolver estes problemas com os instrumentos especificados, os matemáticos gregos não se deixaram intimidar e, com engenho notável, foram capazes de achar soluções para os três problemas, usando vários outros tipos de instrumentos e construções.

Em verdade, da mesma maneira que a Matemática moderna cresce com respostas aos desafios de novos problemas, muito da Matemática grega se desenvolveu devido a tentativas de resolver os três problemas clássicos. Neste sentido, os matemáticos gregos eram realmente nossos colegas, pois tinham a mesma atitude mental dos matemáticos atuais e tentavam conscientemente atacar novos desafios. Quando os

³*Tratado da alma*, II,2, 413, 13-20.



conceitos e técnicas existentes não conseguiam resolver estes problemas, eles inventavam novos conceitos e técnicas apropriadas para a tarefa.



Capítulo 7

A Duplicação do Cubo

O que sabemos sobre este problema encontra-se principalmente em Eutócio, um comentador de Arquimedes.

Há duas lendas sobre a origem da duplicação do cubo, com detalhes contraditórios. Uma delas se refere à duplicação de um túmulo e a outra à duplicação de um altar (Van der Waerden, p. 160-161).

Segundo a primeira lenda, Minos mandou fazer um túmulo para Glauco. Ao saber que o túmulo era um cubo cuja aresta media 100 pés, ele disse que a residência real tinha sido construída demasiadamente pequena e que ela deveria ser duas vezes maior e ordenou imediatamente que duplicassem cada aresta do túmulo, sem estragar sua bela forma.

De acordo com a segunda lenda, quando um oráculo anunciou aos habitantes de Delos que, para se verem livres da peste, deveriam construir um altar duas vezes maior do que o existente, os arquitetos ficaram muito confusos, pois não sabiam como construir um cubo duas vezes maior do que outro.

Hipócrates de Quios (viveu em torno de 430 a.C.) reduziu este problema ao de achar duas meias proporcionais x e y entre 1 e 2.

Com efeito, se

$$\frac{1}{x} = \frac{x}{y} = \frac{y}{2},$$

vemos que

$$x^2 = y$$

e, portanto, multiplicando ambos os membros por x , obtemos

$$x^3 = xy.$$

Mas como

$$xy = 2$$

temos que

$$x^3 = 2.$$

No caso geral, se x e y são duas meias proporcionais entre a e b , temos

$$\frac{a}{x} = \frac{x}{y} = \frac{y}{b}$$

e disso vemos que

$$x^2 = ay$$

e que

$$xy = ab$$

e daí segue-se que

$$x^3 = axy = a^2b$$

e assim

$$\frac{x^3}{a^3} = \frac{a^2b}{a^3} = \frac{b}{a}.$$

Apresentaremos agora sete soluções do problema da duplicação do cubo, quase todas baseadas em achar duas meias proporcionais entre duas grandezas, usando construções que não podem ser efetuadas somente com régua e compasso e curvas que não podem ser traçadas

usando somente estes dois instrumentos.

7.1 A Máquina de Platão

O filósofo grego Platão (viveu de 429 a 347 a.C.) tinha grande interesse pela Matemática e lhe atribuía importância particular. Gravaram em torno dele excelentes matemáticos, como, por exemplo, Arquitas, Eudoxo, Menécmo, Teeteto, entre outros.

É bem conhecido que Platão desprezava construções mecânicas, materiais (ver Van der Waerden, p. 162-163) em Matemática. Assim, é irônico que a solução discutida a seguir seja conhecida como “máquina de Platão”.

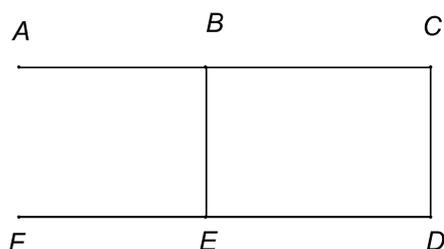


Figura 7.1: A máquina de Platão

A máquina de Platão (Figura 7.1) é um dispositivo, $ACDF$, formado por partes rígidas, com AC e FD paralelas e CD perpendicular a ambas. O segmento BE é paralelo a CD e pode deslizar ao longo de AC e de FD .

Para achar duas meias proporcionais entre $ON = a$ e $OM = b$, movimentamos $ACDF$ de maneira que o segmento CD passe por M , C esteja sobre o eixo horizontal e fazemos BE deslizar até que passe por N e B esteja sobre o eixo vertical, como mostrado na Figura 7.2.

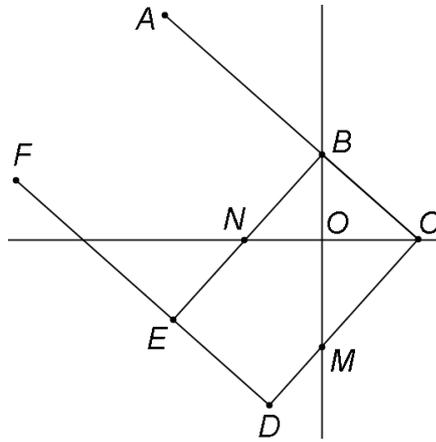


Figura 7.2: Meia proporcional entre $ON = a$ e $OM = b$

Vemos que os triângulos NOB e MOC são semelhantes e, portanto,

$$\frac{ON}{OB} = \frac{OC}{OM}.$$

Como os triângulos NOB e OCB também são semelhantes, temos que

$$\frac{ON}{OB} = \frac{OB}{OC}$$

e assim chegamos a

$$\frac{ON}{OB} = \frac{OB}{OC} = \frac{OC}{OM} \implies \frac{a}{OB} = \frac{OB}{OC} = \frac{OC}{b},$$

o que mostra que OB e OC são de fato meias proporcionais entre a e b .

7.2 A Máquina de Eratóstenes

Considere três placas retangulares $AEZF$, $NMHJ$ e $SGTL$ (Figura 7.3), que podem deslizar sobre uma reta de maneira que a

placa média ($NMHJ$), pode passar por traz da primeira ($AEZF$), e que a última ($SGTL$), deslize por traz da do meio. Suponha que desejamos achar duas meias proporcionais entre $a = AE$ e $b = DT$.

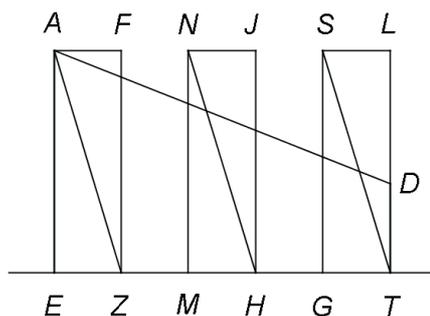


Figura 7.3: A máquina de Eratóstenes

Traçamos o segmento de reta AD e o prolongamos até a intersecção, K , com a reta que passa por ET . Fazemos as placas deslizar, como descrito acima, de maneira que o lado direito ZF da primeira placa ($AEZF$), intercepte a diagonal NH da segunda placa ($NMHJ$), exatamente sobre a reta AD , no ponto B , como mostra na Figura 7.4.

De maneira semelhante, fazemos deslizar a terceira placa ($HGTL$) de tal forma que o lado direito da segunda placa (JH) corte a diagonal (ST) exatamente sobre AD , no ponto C .

Afirmamos então que BZ e HC são meias proporcionais entre AE e DT , ou seja, que

$$\frac{AE}{BZ} = \frac{BZ}{GH} = \frac{GH}{DT}.$$

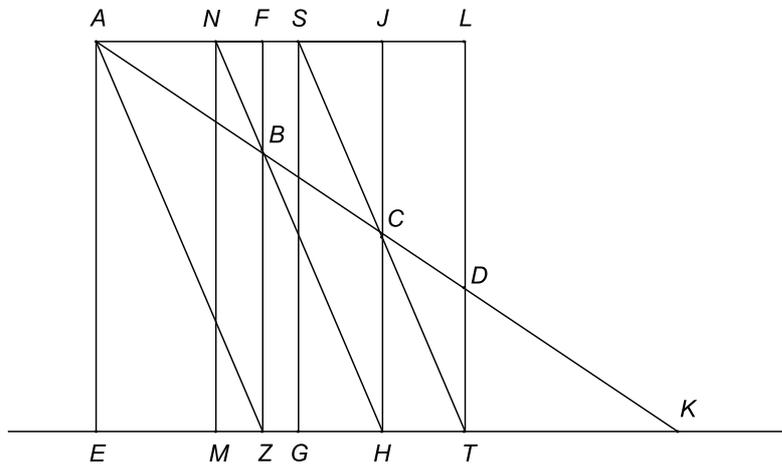


Figura 7.4: Meias proporcionais entre $AE = a$ e $DT = b$

Com efeito, como os triângulos AEK e BZK são semelhantes, temos

$$\frac{AE}{BZ} = \frac{KA}{KB} = \frac{KE}{KZ}. \quad (7.1)$$

Temos também que os triângulos BZK e CHK são semelhantes, e, portanto

$$\frac{BZ}{CH} = \frac{KZ}{KH} = \frac{KB}{KC}. \quad (7.2)$$

Da semelhança dos triângulos CHK e DTK segue-se que

$$\frac{CH}{DT} = \frac{KH}{KT} = \frac{KC}{KD}. \quad (7.3)$$

Da semelhança dos triângulos AZK e BHK decorre

$$\frac{AZ}{BH} = \frac{KZ}{KH} = \frac{KA}{KB}. \quad (7.4)$$

Além disso, como os triângulos BHK e CTK são semelhantes, vemos que

$$\frac{BH}{CT} = \frac{KH}{KT} = \frac{KB}{KC}. \quad (7.5)$$

De 7.3 e 7.5 obtemos

$$\frac{CH}{DT} = \frac{KH}{KT} = \frac{KC}{KD} = \frac{BH}{CT} = \frac{KB}{KC}, \quad (7.6)$$

pois ambos têm $\frac{KH}{KT}$ em comum.

Como 7.2 e 7.6 têm $\frac{KB}{KC}$ em comum, segue-se que

$$\frac{BZ}{CH} = \frac{KZ}{KH} = \frac{BK}{KC} = \frac{BH}{CT} = \frac{KH}{KT} = \frac{CH}{DT} = \frac{KC}{DK}. \quad (7.7)$$

Como 7.1 e 7.4 têm $\frac{KA}{KB}$ em comum, segue-se que

$$\frac{AE}{BZ} = \frac{KA}{KB} = \frac{KE}{KZ} = \frac{AZ}{BH} = \frac{KZ}{KH}. \quad (7.8)$$

De 7.7 mantenhamos somente as razões que nos interessam:

$$\frac{BZ}{CH} = \frac{KZ}{KH} = \frac{CH}{DT}. \quad (7.9)$$

De 7.8 mantenhamos somente as razões que nos interessam:

$$\frac{AE}{BZ} = \frac{KZ}{KH}. \quad (7.10)$$

Como 7.9 e 7.10 têm $\frac{KZ}{KH}$ em comum, obtemos, enfim, que

$$\frac{AE}{BZ} = \frac{BZ}{CH} = \frac{CH}{DT}, \quad (7.11)$$

o que queríamos demonstrar. \square

O processo descoberto por Eratóstenes pode ser usado para inserir qualquer número de meias proporcionais entre a e b . Para inserir n meias proporcionais, é suficiente tomar $n + 1$ retângulos e proceder como acima.

7.3 A Solução de Nicomedes

O método para duplicar o cubo encontrado por Nicomedes, que viveu em torno de 240 a.C., é verdadeiramente muito engenhoso.

Suponha que queremos inserir duas meias proporcionais entre $b = AB$ e $a = BC$ (Figura 7.5).

Construa o retângulo $ABCL$. Seja D o ponto médio de AB e trace a reta que passa por L e D . Seja G seu ponto de intersecção com a reta que passa por C e B . Seja E o ponto médio de BC e trace a perpendicular a BC por E . Chame de F o ponto desta reta para o qual $BF = FC = AD$.

Trace a reta s que passa por F e por G e a paralela a S que passa por C . Por F trace a reta FHK construída de tal maneira que $HK = CF = AD$. Trace a reta que passa pelos pontos K e L e chame de M sua intersecção com a reta definida por A e B .

Afirmamos que AM e CK são meias proporcionais entre a e b , ou seja, que

$$\frac{BC}{MA} = \frac{MA}{CK} = \frac{CK}{AB}. \quad (7.12)$$

Para compreender isso, devemos usar a Proposição II.6, dos

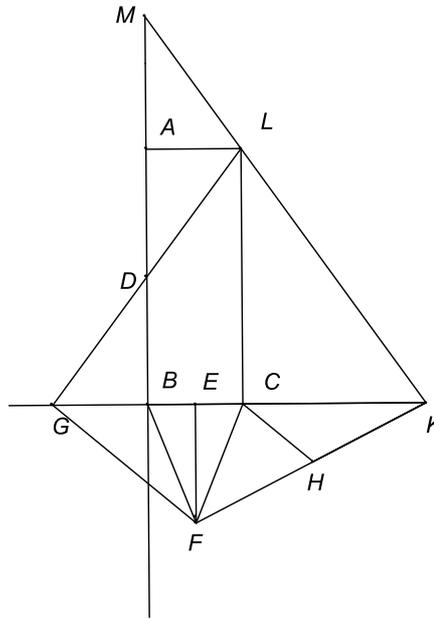


Figura 7.5: Meias proporcionais entre $a = BC$ e $b = AB$

Elementos de Euclides, cujo enunciado é o seguinte:

Elementos, II-6: Se uma linha reta é dividida em duas partes iguais e se uma outra linha reta l he é adicionada, prolongando-a, o retângulo determinado pela linha reta e pela reta adicionada é igual, se l he for adicionado o quadrado sobre a metade da reta, ao quadrado sobre a reta formada pela metade e pela reta adicionada.

Ou seja, na figura, a soma das áreas do retângulo de base AD e altura DB e do quadrado $LEGH$ (que é igual ao quadrado de lado CB) é igual à área do quadrado de lado CD (Ver Figura 7.6).

Voltemos à justificação da construção de Nicomedes, aplicando II-6 ao segmento BC cortado ao meio por E e prolongado até K :

$$BK \cdot KC + CE^2 = EK^2. \quad (7.13)$$

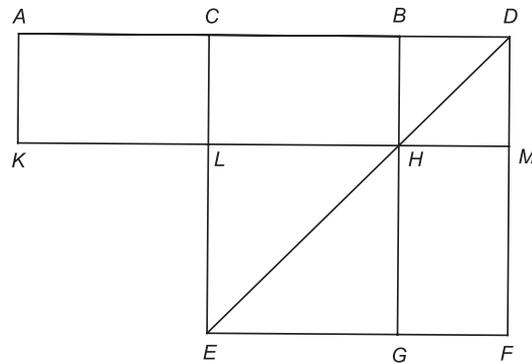


Figura 7.6: *Elementos* II.6

Adicionando EF^2 a ambos os lados desta igualdade e aplicando o teorema de Pitágoras aos triângulos EFK e CEF , temos

$$BK \cdot KC + CF^2 = FK^2. \quad (7.14)$$

Além disso, devido à semelhança dos triângulos AML e MBK temos que

$$\frac{AB}{MA} = \frac{LK}{ML} = \frac{CK}{BC}. \quad (7.15)$$

Observe que $BC = \frac{1}{2}GC$, devido à congruência dos triângulos GBD e DAL .

Como $AB = 2AD$, temos

$$\frac{2 \cdot AD}{MA} = \frac{2 \cdot CK}{GC} \quad (7.16)$$

e assim,

$$\frac{AD}{MA} = \frac{CK}{GC}. \quad (7.17)$$

Como os triângulos GFK e CHK são semelhantes, temos

$$\frac{CK}{GC} = \frac{HK}{FH} \quad (7.18)$$

e daí segue-se que

$$\frac{AD}{MA} = \frac{CK}{GC} = \frac{HK}{FH}. \quad (7.19)$$

Temos então que

$$\frac{AD}{MA + AD} = \frac{HK}{FH + HK}, \quad (7.20)$$

e assim

$$\frac{AD}{MD} = \frac{HK}{FK}. \quad (7.21)$$

Como $HK = AD$ por construção, segue-se de 7.21 que $MD = FK$ e assim, de 7.14, temos

$$BK \cdot KC + CF^2 = FK^2. \quad (7.22)$$

Usaremos mais uma vez Euclides II-6, agora aplicada ao segmento BA , cujo ponto médio é D , e prolongado até M :

$$BM \cdot MA + AD^2 = MD^2. \quad (7.23)$$

Como $MD = FK$, vemos que

$$BM \cdot MA + AD^2 = FK^2. \quad (7.24)$$

Comparando 7.22 e 7.24 vemos que

$$BM \cdot MA = BK \cdot KC \quad (7.25)$$

de maneira que

$$\frac{BK}{BM} = \frac{MA}{CK}. \quad (7.26)$$

Da semelhança dos triângulos MBK , MAL e LCK segue-se que

$$\frac{BK}{BM} = \frac{AL}{MA} = \frac{CK}{LC}. \quad (7.27)$$

Finalmente, usando 7.26 e 7.27, temos que:

$$\frac{CK}{LC} = \frac{MA}{CK} = \frac{AL}{MA}. \quad (7.28)$$

Como $LC = AB$, $AL = BC$, temos, enfim, que

$$\frac{BC}{MA} = \frac{MA}{CK} = \frac{CK}{AB}, \quad (7.29)$$

o que queríamos demonstrar. \square

Esta demonstração encontra-se em Heath (1981). Daremos a seguir uma demonstração analítica para a construção de Nicomedes (veja Bos, 2001).

Na Figura 7.5, faça $x = FH$ e $y = CK$. Então, nos triângulos retângulos EFK e EFC temos que

$$EF^2 + EK^2 = FK^2, \quad (7.30)$$

$$EF^2 + EC^2 = FC^2. \quad (7.31)$$

De 7.30 e de 7.31 vemos imediatamente que

$$FK^2 - EK^2 = FC^2 - EC^2, \quad (7.32)$$

e assim

$$\left(\frac{1}{2}b + x\right)^2 - \left(\frac{1}{2}a + y\right)^2 = \left(\frac{1}{2}b\right)^2 - \left(\frac{1}{2}a\right)^2. \quad (7.33)$$

Disso, segue-se imediatamente que

$$ay + y^2 = bx + x^2. \quad (7.34)$$

Logo

$$\frac{a + y}{b + x} = \frac{x}{y}. \quad (7.35)$$

Os triângulos CHK e GFK são semelhantes e podemos portanto escrever

$$\frac{HK}{CK} = \frac{FH}{GC}, \quad (7.36)$$

ou seja

$$\frac{\frac{1}{2}b}{y} = \frac{x}{2a} \implies \frac{b}{y} = \frac{x}{a} \implies \frac{a}{x} = \frac{y}{b}. \quad (7.37)$$

Mas então

$$\frac{a + y}{y} = \frac{x + b}{b} \implies \frac{a + y}{x + b} = \frac{y}{b}. \quad (7.38)$$

De 7.34 e 7.36 vemos que

$$\frac{x}{y} = \frac{y}{b} \implies \frac{x}{y} = \frac{y}{b} = \frac{a}{x} \quad (7.39)$$

e assim

$$\frac{a}{x} = \frac{x}{y} = \frac{y}{b}. \quad (7.40)$$

Como, pela semelhança dos triângulos AML e LCK , temos que

$$\frac{a}{AM} = \frac{y}{b}, \quad (7.41)$$

vemos imediatamente que

$$AM = x, \quad (7.42)$$

o que conclui a demonstração.

Na construção apresentada por Nicomedes (Figura 7.5), foi necessário construir a reta FHK tal que $HK = CF = AD$. Ou seja, tivemos que *ajustar* o segmento de comprimento AD sobre o segmento FK , a partir de K . Este tipo de construção, que não pode ser efetuado com régua e compasso, é chamado de *construção por neusis* ou *por ajustamento*. Ela é utilizada em várias soluções dos três problemas clássicos, em particular por Arquimedes.

7.4 A Construção de Árqitas

Vamos agora considerar uma das soluções mais engenhosas do problema de duplicar um cubo, devida a Árqitas (viveu em torno de 390 a.C.). Reproduzimos aqui a solução como apresentada em Teixeira (1995, p. 289-290). Embora o raciocínio de Árqitas seja puramente geométrico (veja Van der Waerden, p. 151 ou Heath, 1981,

vol. 1, p. 246-249), usaremos seu equivalente analítico, mais familiar para nós, habituados com a linguagem algébrica.

Sejam a e b os dois segmentos dados, com $b < a$, para os quais desejamos construir duas meias proporcionais. Sejam $OC'A$ uma circunferência cujo diâmetro OA' é igual a a e uma outra circunferência OBA , cujo diâmetro é igual a b e que está contida em um plano perpendicular ao plano da primeira circunferência (veja a Figura 7.7).

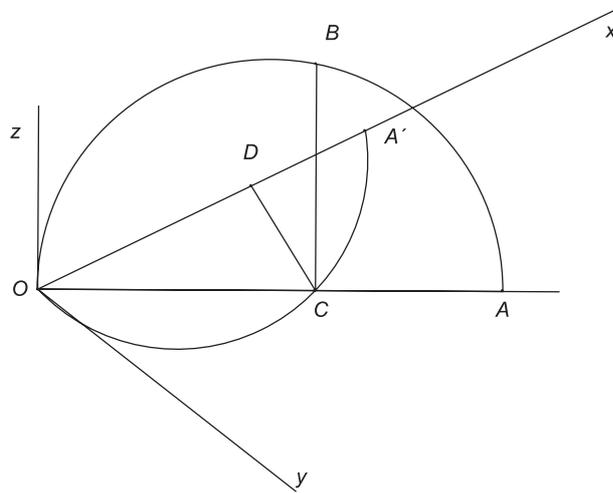


Figura 7.7: Duplicação do cubo por Árquitas

Considere o cilindro circular reto gerado pela circunferência OCA' e o toro gerado pela circunferência OBA ao girar em torno da reta OZ , perpendicular ao plano de OCA' .

A intersecção destas duas superfícies define uma curva, conhecida como *curva de Árquitas*, dada pelas equações

$$\begin{aligned} x^2 + y^2 &= ax, \\ (x^2 + y^2 + z^2)^2 &= a^2(x^2 + y^2). \end{aligned}$$

Seja $\Theta = \frac{b}{a}$ e considere o cone cujo eixo é a reta OA' e cuja geratriz forma com o eixo o ângulo Θ . A equação deste cone é

$$(x^2 + y^2 + z^2) = \frac{a^2}{b^2}x^2. \quad (7.43)$$

Seja D a projeção de C sobre Ox . Então a curva de Árquitas corta o cone no ponto B cuja projeção sobre o plano de OCA' é o ponto C da circunferência tal que $\frac{OC}{OB} = \frac{b}{a}$.

Afirmamos que OC e OB são duas meias proporcionais entre a e b . Com efeito,

$$\begin{aligned} OC^2 &= a \cdot OD, \\ OB^2 &= a \cdot OC, \\ b \cdot OB &= a \cdot OD. \end{aligned}$$

Disso, decorre que

$$\begin{aligned} OC^2 &= b \cdot OD, \\ OB^2 &= a \cdot OC, \end{aligned}$$

e temos enfim que

$$\frac{a}{OB} = \frac{OB}{OC} = \frac{OC}{b}.$$

□

Teixeira (1995, p. 290) comenta que,

esta solução é muito engenhosa e tem grande interesse histórico, porque é o mais antigo exemplo de solução de um problema de geometria plana usando geometria

espacial, e a curva usada é a mais antiga curva reversa conhecida.

Comentando a versão original, geométrica, desta solução, Van der Waerden (p. 151) afirma que Árqitas deve ter tido uma inspiração divina para achar esta construção.

7.5 A Solução Achada por Menécmo

Menécmo viveu em torno de 350 a.C. Em seu *Comentário sobre o primeiro livro de Euclides*, Proclus, no *Sumário de Eudemo*, afirma que Menécmo foi um aluno de Eudoxo e um membro dos filósofos e matemáticos em torno de Platão. Ele era irmão de Dinóstrato o qual, segundo Proclus, “aperfeiçoou ainda mais a geometria”.

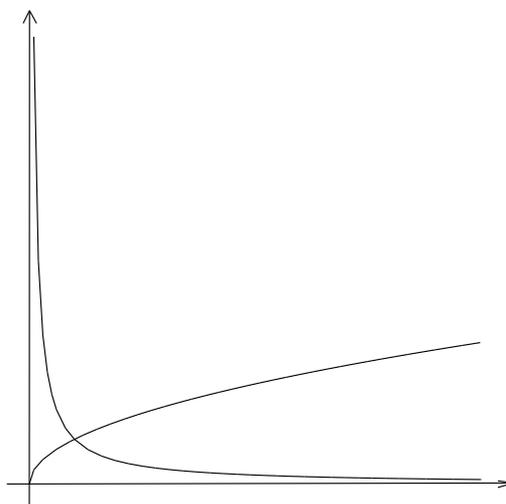


Figura 7.8: Duplicação do cubo por Menécmo

Se x e y são duas meias proporcionais entre a e b , temos que

$$\frac{a}{x} = \frac{x}{y} = \frac{y}{b}.$$

Isso é inteiramente equivalente às equações simultâneas,

$$y^2 = bx, \quad xy = ab.$$

Assim, para resolver o problema, é suficiente achar a intersecção das parábolas e hipérbolas definidas por estas equações, respectivamente (veja a Figura 7.8). As coordenadas deste ponto são as meias proporcionais que estamos procurando.

É óbvio que 7.5 é também equivalente a

$$y^2 = bx, \quad x^2 = ay. \tag{7.44}$$

Desta maneira, o problema pode também ser resolvido usando duas parábolas cujos vértices coincidem e cujos eixos são ortogonais. Estas duas soluções são descritas por Eutócio em seu comentário do *Tratado sobre a esfera e o cilindro*, de Arquimedes.

Nem todas as soluções deste problema consistiam em inserir duas meias proporcionais entre duas grandezas dadas. Isso pode ser verificado pela solução devida a Diocles. É claro que Menécmo não formulou sua solução usando a representação analítica das parábolas ou hipérbolas, mas seu raciocínio é equivalente a isso.

7.6 O Método de Diocles

A solução de Diocles se baseia na cissóide, uma curva definida como segue.

Seja uma circunferência que passa pela origem O do sistema de

coordenadas, tem seu centro sobre o eixo dos x e diâmetro igual a 1. Sejam $T = (1, 0)$ e r a reta vertical que passa por T . Seja P um ponto qualquer sobre a circunferência. A reta que passa por O e por P intercepta r no ponto W . Tome o ponto M sobre OW tal que $OP = MW$. A *cissóide* é o lugar geométrico do ponto M quando P percorre a circunferência (veja a Figura 7.9, que não mostra toda a cissóide, mas somente a parte que nos interessa para a resolução do problema da duplicação do cubo).

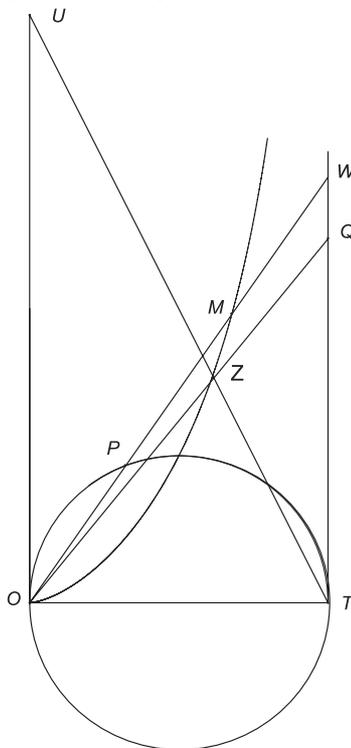


Figura 7.9: Duplicação do cubo usando a cissóide

A equação polar de uma circunferência que passa pela origem, tem raio R , e centro sobre o eixo dos x é

$$\frac{\rho}{2R} = \cos \theta.$$

Como em nosso caso $R = \frac{1}{2}$, a equação 7.6 se reduz a $\rho = \cos \theta$.

Por outro lado, $OW = \frac{1}{\cos \theta}$.

Como $OM = OW - OP$, a equação polar da cissóide é

$$\rho = \frac{1}{\cos \theta} - \cos \theta = \frac{\text{sen}^2 \theta}{\cos \theta}.$$

Para achar a equação cartesiana da cissóide temos que:

$$x = \rho \cdot \cos \theta = \text{sen}^2 \theta = \frac{\text{tg}^2 \theta}{1 + \text{tg}^2 \theta} = \frac{\frac{y^2}{x^2}}{\frac{x^2 + y^2}{x^2}} = \frac{y^2}{x^2 + y^2},$$

ou seja,

$$x(x^2 + y^2) - y^2 = 0.$$

Seja $U(0, 2)$ e considere o ponto de intersecção, $Z = (a, b)$, da reta TU com a cissóide. A equação cartesiana da reta TU é $y = 2(1 - x)$, de maneira que $b = 2(1 - a)$, e assim, como Z pertence à cissóide,

$$\begin{aligned} a(a^2 + b^2) - b^2 = 0 &\implies a^3 + ab^2 - b^2 = 0 \implies \\ a^3 = b^2(1 - a) = \frac{b^3}{2} &\implies 2a^3 = b^3. \end{aligned}$$

É imediato então que $\frac{b}{a} = \sqrt[3]{2}$. Assim, a equação cartesiana da reta que passa por O e por Z é $y = \sqrt[3]{2}x$. Seja Q o ponto de intersecção dessa reta com a reta $x = 1$. Temos, então, que as coordenadas de Q são $(1, \sqrt[3]{2})$. Se temos um cubo cuja aresta mede 1, para duplicar seu volume, devemos achar a aresta y de um cubo que tem volume 2. Ou seja, devemos ter $y = \sqrt[3]{2}$. Vemos portanto que a ordenada do ponto Q , achada acima, resolve o problema.

7.7 O Método de Hierão

Descrevemos a seguir o método proposto por Hierão ¹ para achar duas meias proporcionais entre os segmentos a e b .

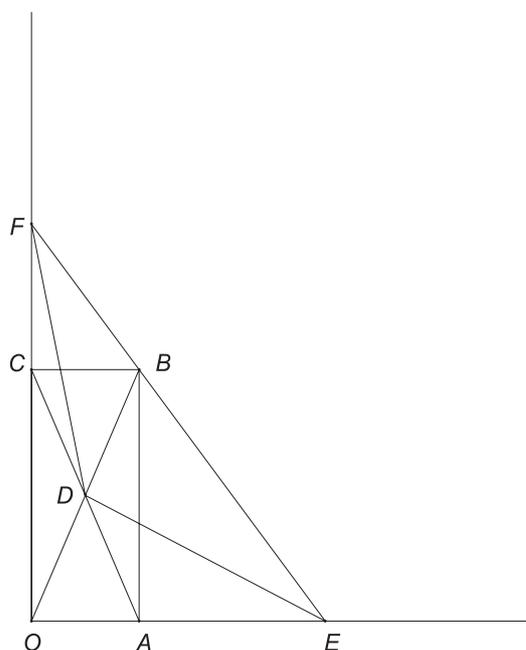


Figura 7.10: Método de Hierão para achar meias proporcionais

Construa o retângulo $OACB$, no qual $AO = a$ e $OB = b$ (Figura 7.10). Seja D o centro do retângulo. Tome uma régua que passa por C e sejam E e F seus pontos de intersecção com as retas definidas por AO e OB respectivamente. Faça a régua girar até que $DF = DE$. Afirmamos então que $BF = x$ e $AE = y$ são as duas meias proporcionais entre a e b .

¹Matemático e físico grego, que viveu em torno de 60 d.C., em Alexandria.

Com efeito, usando a semelhança dos triângulos FBC , CAE e FOE , temos que

$$\frac{a}{x} = \frac{y}{b} = \frac{a+y}{b+x}.$$

Como $DE^2 = DF^2$, segue-se que

$$\left(y + \frac{1}{2}a\right)^2 + \frac{1}{4}b^2 = \left(x + \frac{1}{2}b\right)^2 + \frac{1}{4}a^2,$$

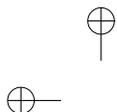
de que obtemos $y(a+y) = x(b+x)$. Desta última igualdade decorre que

$$\frac{z}{y} = \frac{a+y}{b+x} = \frac{a}{x} = \frac{y}{b},$$

e daí vemos imediatamente que

$$\frac{a}{x} = \frac{x}{y} = \frac{y}{b}.$$

Mais detalhes sobre a duplicação do cubo podem ser achados em Heath (1981, vol. 1, p. 244-270). Uma explicação lúcida da solução de Árqitas encontra-se em Van der Waerden (p. 150-152), Heath (1981, vol. 1, p. 246-249) e Teixeira (1995, p. 285-326), o qual apresenta um total de 17 soluções para a duplicação do cubo, incluindo os métodos propostos por Viète, Descartes, Fermat, Newton e Clairaut.



Capítulo 8

A Quadratura do Círculo

Como já dissemos, “quadrar” o círculo, ou seja, construir, com régua e compasso, um quadrado com área igual à área do círculo é um problema bem natural, uma vez resolvido o problema de fazer a quadratura de polígonos! Veremos agora como os matemáticos gregos encontraram maneiras de resolver este problema usando curvas e construções que não podem ser obtidas somente com régua e compasso.

A origem do interesse grego nos problemas de quadratura é pouco conhecida. Segundo Zsabó (2000), o problema primitivo do qual se originaram todos os outros foi o da quadratura do retângulo. Aristóteles afirma que a origem deste problema foi a procura da média geométrica, mas que isso foi esquecido e que só foi preservado o problema.

8.1 A Quadratriz

Esta curva notável resolve dois dos problemas clássicos: a quadratura do círculo e a trisseção de um ângulo arbitrário. Para construí-la, suponhamos que no quadrado $ABCD$ o lado AD gira com



movimento circular uniforme em torno de A até que coincide com o lado AB . Ao mesmo tempo, o lado DC desce com velocidade constante até coincidir com AB . Os dois movimentos estão sincronizados de maneira que ambos os lados, DC e AD coincidam com AB no mesmo instante.

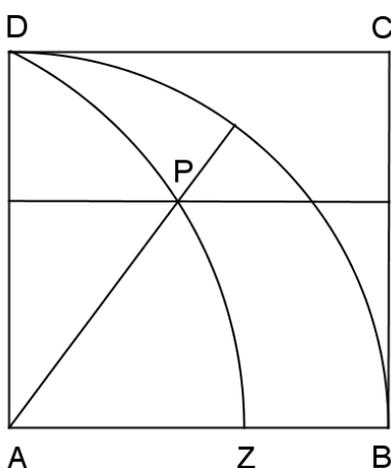


Figura 8.1: Duplicação do cubo com a quadratriz

A *quadratriz* é o lugar geométrico gerado pelas intersecções destes dois lados móveis. É a curva DPZ da Figura 8.1. Ela foi inventada por Hípias de Elis (viveu em torno de 420 a.C.), originariamente em suas tentativas para trissectar o ângulo. Tudo indica que foi Dinóstrato (viveu em torno de 350 a.C.) quem pela primeira vez usou esta curva para fazer a quadratura do círculo.

Afirmamos que $AZ = \frac{2a}{\pi}$, com a o comprimento do lado do quadrado. Com efeito, sejam θ o ângulo PAZ , $x = MP$, $y = AM$ e $AB = AD = DC = a$. Então, devido à proporcionalidade dos dois movimentos, temos que $\frac{y}{\theta} = k$, com k a constante de proporcionali-

dade. Quando $\theta = \frac{\pi}{2}$, temos que

$$\frac{a}{\frac{\pi}{2}} = k,$$

de maneira que

$$k = \frac{2a}{\pi}$$

e podemos concluir que

$$\theta = \frac{\pi y}{2a} \implies y = \frac{2a\theta}{\pi}.$$

Assim,

$$\frac{y}{\rho} = \text{sen } \theta \implies \rho = \frac{y}{\text{sen } \theta} = \frac{2a\theta}{\pi \text{sen } \theta}.$$

Temos então que

$$AZ = \lim_{\theta \rightarrow 0} \rho = \lim_{\theta \rightarrow 0} \frac{2a\theta}{\pi}.$$

Quando $\theta \rightarrow 0$,

$$\lim_{\theta \rightarrow 0} \frac{\theta}{\text{sen } \theta} = 1,$$

e assim vemos que

$$AZ = \rho = \frac{2a}{\pi}.$$

Após obter um segmento de comprimento $\frac{2a}{\pi}$ é imediato construir π para fazer a quadratura do círculo. Com efeito, é fácil dividir, usando somente régua e compasso, $\frac{2a}{\pi}$ por $2a$ e, em seguida, tomar o inverso de $\frac{1}{\pi}$.

Um tratamento mais completo do problema da quadratura do círculo pode ser encontrado em Heath (1981, vol. I, p. 220-235). Uma boa exposição encontra-se em Teixeira (1995, p. 362-384). A história detalhada do número π pode ser lida em Beckmann (1977).

Capítulo 9

A Trissecção do Ângulo

Voltamo-nos agora para o terceiro dos problemas clássicos, ou seja, a trissecção de um ângulo qualquer. Este problema pode ser resolvido de várias maneiras. Acredita-se que Hípias de Elis, que viveu no século V a.C. foi um dos primeiros a tentar resolver este problema, utilizando curvas e construções que não podem ser efetuadas somente com régua e compasso.

Faremos, em primeiro lugar, uma digressão sobre as *construções por ajustamento* ou *por neusis*.

Em uma construção por neusis deve-se ajustar um segmento dado entre duas curvas dadas, com a exigência de que o segmento passe por um ponto dado. Nas palavras de Heath (1953, p. c):

Assim, uma linha reta tem que colocada entre duas linhas ou curvas de maneira que passe por um ponto dado e o segmento determinado sobre ela pelas intersecções com as linhas ou curvas seja igual a um comprimento dado.

Mostremos um exemplo de uma construção por neusis.

9.1 A Trissecção do Ângulo por Arquimedes

A construção que mostraremos a seguir, que utiliza *neusis*, é um exemplo das várias soluções do problema da trissecção do ângulo propostas por Arquimedes.

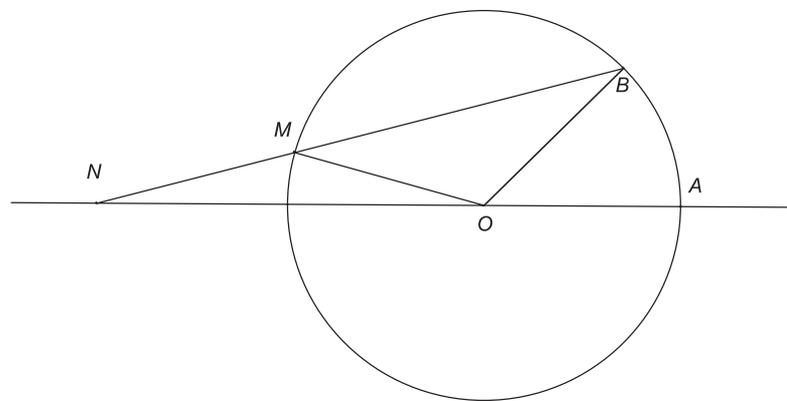


Figura 9.1: Trissecção do ângulo por Arquimedes

Suponha que desejamos trissectar o ângulo BOA . Tome uma reta r que passa por B e, tendo o cuidado para que ela sempre passe por B , movimente-a para que o segmento MN seja igual ao raio OM do círculo. Isso é exatamente o que se denomina uma construção por *neusis*: *Ajustamos* um segmento (o raio OM) entre o círculo e a linha reta que passa por C e por A .

Observe que os triângulos NMO e MOB são isósceles, de maneira que $x = \widehat{MNO} = \widehat{MON}$. De maneira semelhante, $\widehat{BMO} = \widehat{OBM}$.

No triângulo NOM , o teorema do ângulo externo permite escrever que $\widehat{BMO} = \widehat{MBO} = 2 \times \widehat{MNO}$. Aplicando o teorema do ângulo

externo ao triângulo BNO , vemos que

$$\widehat{BOA} = \widehat{BNO} + \widehat{MBO} = 3 \times \widehat{BNO},$$

e vemos assim que com esta construção é possível dividir o ângulo \widehat{BOA} em três partes iguais. \square

9.2 A Trissecção do Ângulo por Nicomedes

Examinemos agora outro exemplo de trissecção do ângulo, devido a Nicomedes, no qual ele utiliza uma construção por ajustamento.

Suponha que desejamos trissectar o ângulo $\alpha = \widehat{AOB}$. Por B levante uma reta perpendicular a OB e por A uma reta paralela a OB (Figura 9.2). Trace uma reta por O e sejam P e C suas intersecções com AB e AC respectivamente. Mova-a de maneira que $PC = 2OA$. Afirmanos que $\widehat{POB} = \widehat{AOB}/3$.

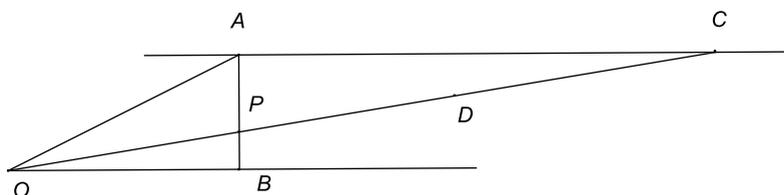


Figura 9.2: Trissecção do ângulo por Nicomedes

Com efeito, seja D o ponto médio de PC . Então o triângulo APC está inscrito em um círculo de centro D e raio PD , e assim $PD = AD = DC = OA$. Sejam $\beta = \widehat{AOD} = \widehat{ADO}$ e $\phi = \widehat{DAC} = \widehat{ACD}$. Aplicando o teorema do ângulo externo ao triângulo ADC vemos que $\beta = 2\phi$. Como AC e OB são paralelas, segue-se que $\phi = \widehat{POB}$ e portanto $\alpha = 3\phi$.

A construção neusis usada neste problema pode ser efetuada usando a *conchóide de Nicomedes*, cuja definição é dada a seguir.

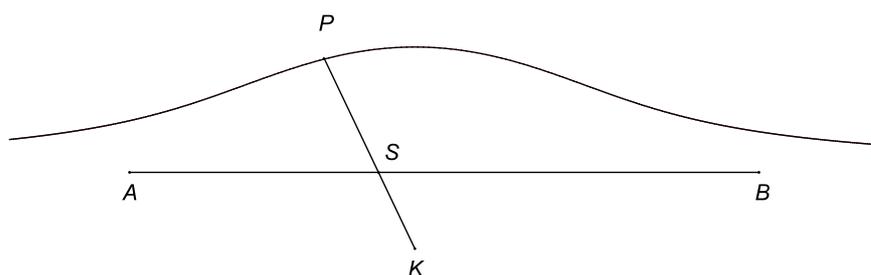


Figura 9.3: Conchóide de Nicomedes

Sejam K um ponto fixo, o *polo* da conchóide, e uma linha reta SA , denominada *diretriz* da conchóide, e um comprimento fixo d , chamado de *distância* da conchóide.

A conchóide é a curva gerada por P quando S se desloca sobre SA e o comprimento do segmento PS permanece sempre igual a d (Figura 9.3).

Para vermos como a conchóide pode ser usada para a trisseção do ângulo proposta por Nicomedes, trace a conchóide com polo C , diretriz AB e distância igual a duas vezes OA . A intersecção C desta conchóide com a paralela a OB que passa por A é o ponto procurado, como já foi mostrado.

A conchóide pode também ser usada na duplicação do cubo feita por Nicomedes feita na página 104. Com efeito, a intersecção K da conchóide cujo polo é F , diretriz CH e distância AD com a linha reta definida por G e C é exatamente o ponto que achamos anteriormente (compare com a Figura 7.5).

9.3 A Trissecção do Ângulo Usando a Quadratriz

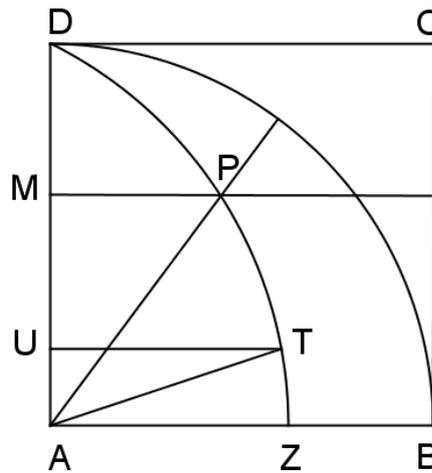


Figura 9.4: Trissecção do ângulo com a quadratriz

Como já dissemos, a quadratriz foi utilizada em primeiro lugar para resolver o problema da trissecção do ângulo. Como acontece frequentemente em Matemática, às vezes uma ideia que permite atacar com sucesso um problema mostra, posteriormente, ser capaz de resolver outros problemas. Isso aconteceu com a quadratriz, pois foi logo depois observado que ela também permite resolver o problema da quadratura do círculo.

Suponha que desejamos dividir o ângulo PAZ em três partes iguais. Por P , trace a paralela a AZ que intercepta AD em M (Figura 9.4). Divida o segmento AM em três partes iguais e seja AU uma dessas partes. Por U trace a paralela a AZ que corta a quadratriz no ponto T . Por proporcionalidade, o tempo gasto por P para chegar a T é igual ao tempo gasto por M para chegar a U , e é igual a $2/3$ do tempo que P gasta para chegar a Z . Assim, $P\hat{A}Z = 3 \times T\hat{A}Z$.



▲ SEC. 9.3: A TRISSECÇÃO DO ÂNGULO USANDO AQUADRATRIZ

127

Esta é a solução mais antiga que conhecemos para o problema da trissecção do ângulo. Um tratamento mais amplo deste problema pode ser encontrado em Heath (1981, vol. I, p. 235-244). Para uma discussão geral de problemas resolvidos usando construções neusis, ver Heath (1953, Capítulo V, p. c-cxxii) e um tratamento mais didático em Yates (1971). Várias outras construções podem ser achadas em Teixeira (1995, p. 327-358). Consultar também Dudley (1987) em que se encontram muitos processos aproximados ou exatos para trissectar um ângulo.





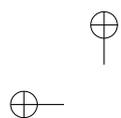
Referências Bibliográficas

- [1] BECKMANN, Peter. *A history of pi*. 4th ed. Boulder. Colorado: Golem Press, 1977.
- [2] BOS, Henk J. M. *Redefining geometrical exactness: Descartes' transformation of the early modern concept of construction*. New York: Springer, 2001.
- [3] BUNT, Lucas N. H.; Phillip S. Jones; Jack D. Bedient. *The historical roots of elementary mathematics*. New York: Dover, 1988.
- [4] CARVALHO, João Pitombeira de. *Três excursões pela história da Matemática*. Rio de Janeiro: Intermat, 2008.
- [5] COURANT, Richard; ROBBINS, Herbert. *What is mathematics?* New York: Oxford University Press, 1996.
- [6] BKOUICHE, Rudolf; DELATTRE, Joëlle. “Pourquoi la règle et le compas”. Comissão Inter-IREM. *Histoire de problèmes, Histoire des Mathématiques*. Paris: Ellipses, 1993.
- [7] DUDLEY, Underwood. *A budget of trisections*. New York: Springer, 1987.
- [8] EVES, Howard. *Introdução à História da Matemática*. Campinas: Editora da UNICAMP, 1995.

- [9] HADLOCK, Charles Robert. *Field theory and its classical problems*. Washington, DC: Mathematical Association of America, 1978. (Carus Mathematical Monographs 18.)
- [10] HEATH, Thomas L. *The works of Archimedes*. New York: Dover, 1953.
- [11] HEATH, Thomas L. *The Thirteen books of Euclid's Elements*, vols. 1, 2, 3. New York: Dover, 1956.
- [12] HEATH, Thomas L. *A history of Greek mathematics*, 2 v. New York: Dover, 1981.
- [13] KATZ, Victor J. *A History of Mathematics – an introduction*. New York: Harper Collins, 1993.
- [14] KLEIN, Felix. *Famous problems of elementary geometry*, translated by Wooster Woodruff Beman e David Eugene Smith. New York: G. E. Stechert & Co. 1930.
- [15] KNORR, Wilbur Richard. *The evolution of the euclidean elements: a study of the theory of incommensurable magnitudes and its significance for early Greek geometry*. Dordrecht [u.a.], Holanda: Reidel, 1975.
- [16] KNORR, Wilbur Richard. *The ancient tradition of geometric problems*. Boston, Basel, Stuttgart: Birkhäuser, 1986.
- [17] RAIGORODSKI, Andrei M. “The Borsuk partition problem: the seventieth anniversary”. *The mathematical intelligencer*, vol. 26, n. 3, 2004, p. 4-12.
- [18] SANTOS, Christovam dos. “Trissecção de um ângulo”. *Revista brasileira de Matemática*, ano 2, n. 5, jan./1931, p. 43-47.

- [19] SCRIBA, Christoph J. On the so-called “Classical Problems” in the History of Mathematics. Ivor Grattan-Guinness (Ed.) *Cahiers d’Histoire & de Philosophie des Sciences*, n. 21, 1987, *History in mathematics education*. Proceedings of a Workshop held at the University of Toronto, Canada, July – August 1983. Paris: Belin, 1987.
- [20] SIMSON, Robert. *Elementos de Euclides. Dos seis primeiros livros, do undécimo e duodécimo, da versão latina de Frederico Commandino, adicionados e ilustrados por Roberto Simson, Professor de Mathematica na Academia de Glasgow*. Coimbra: Real Imprensa da Universidade, 1773. Com privilégio real.
- [21] SZABÓ, Arpad. *L’aube des mathématiques grecques*. Paris: Vrin, 2000.
- [22] TEIXEIRA, Francisco Gomes. *Traité des courbes spéciales planes et gauches*, vol III. Paris: Jacques Gabay, 1995. 1. ed. Coimbra: Universidade de Coimbra, 1909.
- [23] VAN DER WAERDEN, B. L. *Science Awakening I*. Third edition. Grönigen: Wolters Noordhoff, s/d.
- [24] VAN DER WAERDEN, B. L. *Geometry and algebra in ancient civilizations*. New York: Springer Verlag, 1983.
- [25] VAN DER WAERDEN, B. L. *A History of algebra, from al-Khwarizmi to Emmy Noether*. New York: Springer Verlag, 1985.
- [26] WANTZEL, Pierre. Recherches sur les moyens de reconnaître si un problème de géométrie peut se résoudre avec règle et compas. *Journal de Mathématiques*, 2, 366-372, 1837.
- [27] YATES, Robert C. *The trisection problem*. Reston, Virginia: The National Council of Teachers of Mathematics, 1971. *Classics in mathematics education*, vol. 4.





A Matemática dos Códigos de Barras

Francisco César Polcino Milies





Prefácio

Em 1977 fiz meu primeiro estágio de pesquisa no exterior, no Canadá. Foi lá, na biblioteca da universidade em que estagiava, que vi pela primeira vez a utilização de códigos de barras, usados para catalogar os livros do acervo. Logo depois, pude testemunhar o seu emprego nos caixas de supermercados e grandes magazines. Parecia coisa de mágica. Fiquei admirado de ver como agilizavam os atendimentos e diminuíam o tamanho das filas.

Hoje em dia, os códigos de barras são de uso rotineiro e nos deparamos com eles continuamente no nosso dia a dia. Não somente estão presentes nos diversos artigos que compramos, como aparecem também, em versões bem mais longas, em recibos, contas a pagar e inúmeras outras aplicações.

Mesmo assim, eles estão completamente ausentes das nossas salas de aula. Sua estrutura é muito simples e certamente eles poderiam ser usados para motivar o estudo de alguns aspectos da teoria da divisibilidade e para plantear questões instigantes.

A exposição que se segue tem por objetivo apresentar os conceitos e ideias básicas ligadas a estes códigos. Vamos descrever como é atribuído um número a cada produto, como é feita a tradução de números para barras de espessura variável e, finalmente, vamos discutir um dos aspectos mais interessantes da teoria: como se pode

estruturar o código de modo que um cálculo simples (em geral efetuado automaticamente pela caixa que registra a compra) pode detectar se foi cometido algum erro de leitura. Partes deste trabalho já foram publicadas na forma de artigos, na Revista do Professor de Matemática.

É claro que este é um tópico em contínuo desenvolvimento, que está relacionado com outras áreas da matemática atual e suas aplicações, tais como a moderna criptografia ou a teoria de códigos corretores de erros. Esperamos que a leitura destas notas possa servir de estímulo para que o leitor continue seus estudos, aprofundando-se nos assuntos aqui tratados.

César Polcino Milies

Instituto de Matemática e Estatística

Universidade de São Paulo.

Sumário

Introdução	139
Um Pouco de História	141
10 Código de Barras	149
10.1 Escrevendo com Barras	149
10.2 A Detecção de Erros	155
10.3 Outros Códigos Numéricos	160
11 Códigos sobre Grupos	171
11.1 O Grupo Dihedral	171
Referências Bibliográficas	181



Introdução

Hoje em dia, muitos produtos são identificados com um código numérico. O progresso da tecnologia, que tornou relativamente baratos e acessíveis aparelhos de leitura óptica e computadores, tornou também o uso deste tipo de códigos bastante frequente. Por exemplo, os produtos que compramos num supermercado estão identificados por um código de barras, como o que mostramos na Figura 9.5. Ele não é mais do que um número, assignado ao produto para sua identificação, escrito de forma a permitir uma leitura rápida no caixa. Note que imediatamente abaixo das barras, aparece o mesmo número escrito em algarismos correntes, de forma que o leitor humano também possa ler o número.



Figura 9.5: Código de barras

Porém, algumas vezes acontece que, ao passar um produto pela leitora ótica (por exemplo, quando a embalagem está úmida ou enrugada), esta não consegue realizar a leitura. O que vemos então é a pessoa que está no caixa tenta passar o produto em sentido contrário, ou inverte o produto, de modo que o código de barras fique de cabeça para baixo, e tentar passá-lo mais uma vez. Se nem assim der certo, então ele próprio lê o código e o digita rapidamente.

Naturalmente, estas atitudes sugerem algumas perguntas. Em primeiro lugar, uma vez que o desenho das barras é totalmente simétrico para a máquina, que o lê usando um feixe de luz transversal, ao passá-lo “de ponta cabeça”, ela não deveria ler o número na ordem contrária? E, o que é pior, o operador do caixa, ao digitar o número rapidamente, não poderia cometer um erro e nós acabarmos pagando por um produto muito mais caro do que aquele que estamos comprando?

Na verdade, isso não ocorre. Tanto quando lido numa ordem, como na ordem contrária, o código sempre é interpretado de forma correta. Mais ainda, quando o operador comete algum erro de digitação – e todos nós já vimos isso acontecer algumas vezes – a máquina simplesmente emite um som, para avisar que houve um erro!

O objetivo destas notas é explicar como e porquê isso acontece. Começaremos contando um pouco da história destas ideias na próxima seção. No capítulo seguinte vamos explicar como é feita a tradução dos números em termos de barras e como a leitora distingue entre esquerda e direita; na seção 10.2 vamos explicar as ideias matemáticas que fazem com que eventuais erros possam ser detectados. No capítulo final, mostraremos brevemente como estas ideias podem ser extendidas para se obter códigos mais sofisticados.

Um Pouco de História

A ideia de utilizar máquinas para realizar cálculos é, na verdade, bastante antiga. Já em 1642, **Blaise Pascal** construiu a primeira máquina de calcular de que se tem notícia. Ela funcionava com engrenagens mecânicas e era capaz de realizar apenas somas. Em 1694, **Wilhelm Leibniz** aprimorou o invento de Pascal e criou uma máquina capaz de realizar também multiplicações. Do ponto de vista destas notas, o que é realmente interessante é a forma de transmitir dados à máquina já que, praticamente desde o começo da automação, isto foi feito com cartões perfurados que são antepassados diretos dos códigos de barras.

De forma muito esquemática, os estágios do processo que nos interessa são os seguintes:

- Em 1728, **B. Bouchon** concebeu a ideia de cifrar informações em folhas de papel perfurado. A descoberta verdadeiramente importante veio em 1801, quando **Joseph-Marie Jacquard** (1752-1834) construiu um tear que era comandado por cartões perfurados e que foi, talvez, a primeira máquina programável. O tear de Jacquard era capaz de realizar todos os movimentos necessários e ele foi o primeiro a produzir tecidos com padrões figurativos. Essa máquina deu ímpeto à revolução tecnológica da era industrial e foi a base para o desenvolvimento do moderno tear automático.
- Em 1857, **Sir Charles Wheatstone** utilizou fitas de papel para armazenar dados, seguindo o mesmo princípio básico dos cartões perfurados, mas com a vantagem de poder alimentar dados de forma contínua.

- Em 1822, **Charles P. Babbage** (1792-1871), um professor de matemática de Cambridge e um dos fundadores de *Analytical Society* (grupo de professores que modificaria o ensino da matemática na Inglaterra e que tiveram influência fundamental na criação da álgebra abstrata) inventou um instrumento de cálculo mais sofisticado, que denominou *Máquina Diferencial*. Nessa época, ele observou que “as operações matemáticas repetitivas poderiam ser desenvolvidas com mais agilidade e confiabilidade pelas máquinas que pelos homens”. Mais adiante, em 1833, projetou uma *Máquina Analítica*, isto é, uma máquina capaz de executar todas as operações aritméticas, de fazer comparações e analisar seus próprios resultados, que era programada através de *cartões perfurados*. Ele foi o primeiro a perceber que uma máquina de computar deveria ter um dispositivo de entrada, uma memória (que ele chamou de *mill = moinho*) e um dispositivo de saída. Sua máquina, em particular, seria alimentada por duas séries de cartões perfurados: uma com os dados e outra com as operações a serem executadas. Por causa disto, ele é considerado o pai do computador digital.

Suas ideias despertaram o interesse de Ada August, condessa de Lovelace e filha de Lord Byron, que foi a primeira programadora da história. No período de 1842-1843 ela traduziu do italiano um artigo de Luigi Menebrea sobre a máquina da Babbage e, num apêndice, detalhou um método completo para calcular números de Bernoulli com a máquina.

Babbage conseguiu convencer o governo britânico a financiar seu projeto mas, apesar dos esforços de anos e de vários investimentos governamentais, a máquina jamais chegou a ser construída. De acordo com o plano original, ela seria movida a vapor e de tamanho maior do que uma locomotiva.

- A próxima personagem importante nesta história é **Hermann Hollerith** (1860-1951) que obteve um doutorado em estatística na Columbia University em 1879 e, logo em seguida, foi empregado pelo Bureau de Censos dos EUA para trabalhar com seu professor, William P. Trowbridge, no censo de 1880. Foram necessários dez anos de trabalho para tabular completamente todos os dados recolhidos. Durante este período, Hollerith deu também aulas durante algum tempo no Massachusetts Institute of Technology e trabalhou no Escritório de Patentes de Washington. Sobretudo, ele empregou seu tempo projetando uma máquina que pudesse tabular dados automaticamente. Para isso, ele utilizou novamente a ideia dos cartões perfurados de Jacquard, escrevendo dados em oito colunas que utilizavam o sistema de numeração binária. Esses cartões eram então lidos por sua máquina que utilizava sensores elétricos. Quando foi realizado um novo censo, em 1890, apesar dele ser mais sofisticado e coletar mais dados que o anterior, a invenção de Hollerith pode tabular todos os dados em apenas seis semanas.

Sua criação teve sucesso imediato e ele deixou seu emprego para fundar uma companhia dedicada ao desenvolvimento de máquinas semelhantes, a Tabulating Machine Company que, com o decorrer dos anos, se transformou na atual IBM.

- O desenvolvimento de computadores eletrônicos ganhou mais força a partir da segunda guerra mundial, quando foi percebido seu potencial estratégico.

Muitos autores consideram que o Atanasoff-Berry Computer (ABC), desenvolvido nos EUA no período 1937-42 é o primeiro computador eletrônico digital. Porém, ele não tinha capacidade de programação geral e servia apenas para resolver sistemas li-

neares, além de ter outras limitações técnicas.

Em 1941 os alemães desenvolveram o Z3, desenhado por Konrad Zuse. Era eletromecânico, mas com objetivos mais gerais e totalmente programável mediante fitas perfuradas.

Em 1944, Tommy Flowers desenhou, na Inglaterra, o computador Colossus, totalmente eletrônico, concebido para decodificar mensagens interceptadas aos alemães.

O computador mais famoso deste período foi o ENIAC (Electronic Numerical Integrator and Computer), desenhado por John Mauchly e J. Presper Eckert, da Universidade de Pennsylvania e construído na Escola Penn Moore de Engenharia Eletrônica daquela universidade. Embora ele fosse um computador de propósitos gerais, ele foi desenhado originalmente para calcular tabelas de fogo de artilharia para o Laboratório de Pesquisas Balísticas, durante a guerra, mas sua construção só foi completada três meses após o fim da guerra. Os primeiros problemas computados pelo ENIAC foram relacionados à construção da bomba de hidrogênio. Tanto a entrada quanto a saída de dados era feita através de cartões perfurados.

O aparelho pesava 27 toneladas, usava 17 468 válvulas e precisou, para sua construção, de mais de cinco milhões de soldas feitas a mão. Ocupava todo um galpão e consumia 150 kW de energia elétrica. Costuma-se dizer que, quando estava em uso, provocava apagões na cidade de Pennsylvania mas isto não deve ser verdade pois possuía alimentação independente da rede elétrica. Quando estava em operação, elevava a temperatura do local a 50 graus. Ele foi desativado em 2 de outubro de 1955.

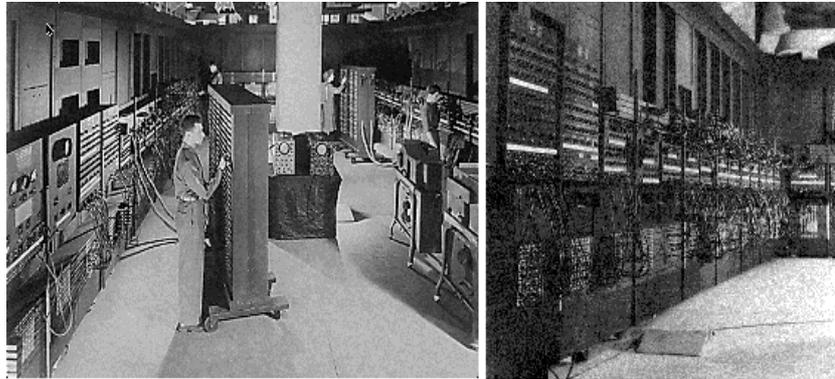


Figura 9.6: O computador ENIAC

Outro computador de grande porte construído nesse período foi o Mark I. O projeto, concebido por Howard Aiken da Universidade de Harvard, iniciou-se em 1939 mas foi concluído apenas em 1943, na IBM. Foi trasladado a Harvard onde foi mostrado publicamente, pela primeira vez, em 1944 e foi batizado oficialmente como Harvard-IBM Automatic Sequence Controlled Calculator (ASCC). Media 15.5 m de comprimento, 2.40 m de altura e aproximadamente 60 cm de largura.

A partir de então, os progressos da tecnologia permitiram diminuir gradativamente o tamanho (e o custo!) dos computadores até popularizá-los definitivamente. Foi também a tecnologia que permitiu usar feixes de luz e *scanners* para transmitir dados direta e rapidamente aos computadores, criando assim condições para a utilização da codificação que nos interessa estudar aqui.

Códigos de Barras

A primeira patente de um código de barras foi atribuída em 1952 a Joseph Woodland e Bernard Silver. Seu código consistia num padrão de circunferências concêntricas de espessura variável. Ao dar entrada ao pedido de patentes, eles descreviam seu invento como *uma classificação de artigos através de identificação de padrões*.

Em torno de 1970, uma firma de assessoria, a McKinsey & Co., junto com a Uniform Grocery Product Code Council¹ definiu um formato numérico para identificar produtos e pediu a diversas companhias que elaborassem um código adequado. Dentre as firmas contactadas, a que acabou apresentando a proposta vencedora foi a IBM e o código foi criado por George J. Laurer.²



Figura 9.7: George J. Laurer

O código proposto, formalmente aceito em maio de 1973, passou a

¹A palavra *Grocery* em inglês, equivale aproximadamente a nossa expressão *secos e molhados*.

²Estes dados, bem como a foto do autor, foram obtidos no site do próprio Laurer, que diz ter tido a colaboração de outras duas pessoas, mas não cita os nomes. O endereço do site é:

http://www.bellsouthpwp.net/1/a/laurergj/upc_work.html

ser conhecido como código UPC (Universal Product Code) e foi adotado nos Estados Unidos e Canadá. Ele consistia de uma sequência de 12 dígitos, traduzidos para barras da forma que analizaremos detalhadamente no próximo capítulo. Existem várias versões sucessivas do UPC, com pequenas modificações. Posteriormente foi solicitado a Laurer que ampliasse o código, para permitir uma maior difusão do sistema, de modo a identificar também o país de origem de cada produto classificado. Baseado no UPC-A, ele acabou criando um novo código, com 13 dígitos, que foi adotado em dezembro de 1976 com o nome EAN (European Article Numbering system). Alguns países adotam este mesmo sistema, dando-lhe outro nome. Por exemplo, no Japão o sistema é conhecido como JAN (Japanese Article Numbering system).



Capítulo 10

Códigos de Barras

10.1 Escrevendo com Barras

Vamos estudar inicialmente o código UPC, que é mais simples. Se observamos o código de barras da Figura 10.1, notamos imediatamente que ele é formado por listras brancas e pretas alternadas, de espessura variável. Há, na verdade, quatro espessuras possíveis para essas listras, que podem ser classificadas como finas, médias, grossas ou muito grossas.



Figura 10.1: Código UPC

Vamos utilizar o símbolo 0 para indicar uma listra branca fina, o símbolo 00 para uma listra branca média, 000 para uma listra branca grossa e 0000 para uma listra muito grossa. Da mesma forma, vamos representar por 1, 11, 111 e 1111, uma listra preta fina, média, grossa ou muito grossa, respectivamente.

Assim, as primeiras quatro listras da figura (sem contar, é claro as listras que servem de limite e que aparecem mais compridas na figura), que são uma listra branca grossa, uma preta média, uma branca fina e uma preta fina respectivamente, podem ser representadas pela sequência 0001101.

Como já dissemos, o código de barras representa uma série de números. A cada número lhe é assignado um espaço de espessura fixa, que corresponde sempre a uma sequência de sete dígitos iguais a 1 ou 0. Por exemplo, a sequência 0001101 que achamos acima representa o número 0, o primeiro do código da figura. O seguinte número do código, o 7, é representado pela sequência 0111011.

Agora que começamos a compreender a forma de escrever com barras, já podemos responder à primeira das nossas perguntas: como a leitora distingue à direita da esquerda, quando o artigo pode ser passado em uma ou outra direção?

A resposta é muito engenhosa e também bastante simples. Os dígitos são codificados de maneira *diferente* quando estão do lado direito ou do esquerdo do código de barras. Isto é feito conforme a seguinte tabela:

dígito	do lado esquerdo	do lado direito
0	0001101	1110010
1	0011001	1100110
2	0010011	1101100
3	0111101	1000010
4	0100011	1011100
5	0110001	1001110
6	0101111	1010000
7	0111011	1000100
8	0110111	1001000
9	0001011	1110100

Note que a codificação de um dado número, à direita, se obtém da sua codificação à esquerda, trocando cada 0 por 1 e reciprocamente. Agora, o mecanismo de reconhecimento fica claro se notamos que *cada sequência do lado esquerdo tem um número ímpar de dígitos iguais a 1 e, conseqüentemente, cada uma das que estão à direita tem um número par*. Assim, verificando a paridade de cada sequência de sete dígitos, a máquina “sabe” imediatamente de que lado está lendo o código.

A elaboração do código EAN se deparou com um problema bastante delicado. Era necessário adicionar um dígito à cada código, de modo a permitir a identificação do país de origem do produto, mas se desejava fazer isto de uma forma tal que a mesma máquina leitora pudesse ler indistintamente códigos UPC e EAN.

Se observamos a figura a seguir, que representa o mesmo código numérico escrito em ambos sistemas, veremos que, a primeira vista, parecem diferentes pois, no número escrito para o leitor humano, vemos que há um 0 a mais, escrito no início da sequência. Porém, se

observamos o código de barras, vemos que é exatamente o mesmo.



Figura 10.2: Os códigos UPC-A e EAN-13

O que foi feito é o seguinte. Os países que utilizavam o código UPC antigo, EUA e Canadá, são identificados com um 0, na frente, e o resto da codificação é feita utilizando-se o sistema anterior.

Para outros países, os primeiros dois ou três dígitos, identificam o país. Por exemplo, o código de barras de todos os produtos produzidos no Brasil começa com a sequência 789, que é a que identifica o país.¹ Como era necessário adicionar um dígito e também manter o mesmo padrão de tamanho do código de barras, para não ter que modificar todas as leitoras, a ideia utilizada foi fazer com que o novo dígito *estivesse implícito na forma de escrita de todos os outros*. Para isso, não foi modificada a codificação do lado direito (permitindo assim que as leitoras continuassem a identificar o lado correspondente) mas *a codificação do lado esquerdo varia, dependendo do dígito inicial*.

¹Uma tabela completa, com os números identificatórios de cada país, pode ser encontrada na página internet <http://www.barcodeisland.com/ean13.phtml>

Um dígito do lado esquerdo pode ser agora codificado com um número par ou ímpar de dígitos iguais a 1, de acordo com a seguinte tabela:

dígito	lado esquerdo ímpar	lado esquerdo par	lado direito
0	0001101	0100111	1110010
1	0011001	0110011	1100110
2	0010011	0011011	1101100
3	0111101	0100001	1000010
4	0100011	0011101	1011100
5	0110001	0111001	1001110
6	0101111	0000101	1010000
7	0111011	0010001	1000100
8	0110111	0001001	1001000
9	0001011	0010111	1110100

Finalmente, para cada dígito inicial escolhe-se uma alternância diferente de pares e ímpares de acordo com o seguinte critério:

Dígito inicial	1º	2º	3º	4º	5º	6º
0	ímpar	ímpar	ímpar	ímpar	ímpar	ímpar
1	ímpar	ímpar	par	ímpar	par	par
2	ímpar	ímpar	par	par	ímpar	par
3	ímpar	ímpar	par	par	par	ímpar
4	ímpar	par	ímpar	ímpar	par	par
5	ímpar	par	par	ímpar	ímpar	par
6	ímpar	par	par	par	ímpar	ímpar
7	ímpar	par	ímpar	par	ímpar	par
8	ímpar	par	ímpar	par	par	ímpar
9	ímpar	par	par	ímpar	par	ímpar

Vamos ver um exemplo. Uma barra de cereais produzida no Brasil é identificada pelo código 7895000266241. Como corresponde, começa com a sequência 789, de modo que o primeiro dígito, que estará implícito na codificação dos demais, é sete. Consequentemente, deve-se usar, do lado esquerdo, a seguinte ordem de codificação (obtida na tabela acima):

ímpar, par, ímpar, par, ímpar, par.

Consultando então a tabela de codificação do EAN-13 obtemos:

8 \mapsto 0110111 9 \mapsto 0010111 5 \mapsto 0110001

0 \mapsto 0100111 0 \mapsto 0001101 0 \mapsto 0100111

Para os dígitos do lado direito não temos que nos preocupar com paridade, e obtemos, diretamente da tabela, a seguinte codificação:

2 \mapsto 1101100 6 \mapsto 1010000 6 \mapsto 1010000

2 \mapsto 1101100 4 \mapsto 1011100 1 \mapsto 1100110

Por tanto, o código de barras correspondente é:



Figura 10.3:

Um último comentário à respeito deste código. Como já dissemos, os primeiros dois ou três dígitos do código de barras (dependendo do caso) servem para identificar o país de origem do produto. Os cinco ou quatro dígitos que restam, até as barras centrais, identificam o fabricante. Os primeiros cinco dígitos do lado direito identificam o produto específico, desse fabricante. O último dígito, chamado **dígito de verificação**, é adicionado no final do processo de elaboração do código, de acordo com um método que veremos adiante.

Falta ainda responder a nossa segunda pergunta: como faz a máquina para detectar quando um operador apressado comete um erro de digitação? Isto será o assunto da nossa próxima seção.

10.2 A Detecção de Erros

Para compreender como funciona o processo de detecção de erros precisamos entender, inicialmente, como se atribui a cada produto, o dígito de verificação.

Suponhamos que um determinado produto está identificado, no sistema EAN-13, por uma dada sequência de dígitos $a_1a_2 \dots a_{12}a_{13}$. Como os primeiros dígitos identificam o país de origem, o fabricante e o produto específico, os primeiros doze dígitos da sequência, estão determinados naturalmente, por um método-padrão, a cargo de uma autoridade classificadora em cada país. Denotaremos o décimo terceiro dígito, de verificação, por x .

Para facilitar nossa exposição, vamos escrever esta sequência como um vetor

$$\alpha = (a_1, a_2, \dots, a_{11}, a_{12}, x).$$

O sistema EAN-13, se utiliza de um vetor fixo, que chamaremos, **vetor de pesos** que é:

$$w = (1, 3, 1, 3, 1, 3, 1, 3, 1, 3, 1, 3, 1).$$

Calcula-se, então o “produto escalar” de ambos vetores:

$$\begin{aligned} \alpha \cdot \omega &= (a_1, \dots, a_{12}, x) \cdot (1, 3, 1, 3, 1, 3, 1, 3, 1, 3, 1, 3, 1) = \\ &= a_1 + 3a_2 + a_3 + 3a_4 + a_5 + 3a_6 + a_7 + 3a_8 + a_9 + 3a_{10} + a_{11} + 3a_{12} + x. \end{aligned}$$

Agora, o dígito de verificação x se escolhe de forma tal que a soma acima seja múltiplo de 10, isto é, tal que

$$\alpha \cdot \omega = 0 \pmod{10}.$$

Por exemplo, no caso do código da Figura 10.3, os números que indicam o país de origem, o fabricante e o produto são 789500026624. Vamos ver como foi determinado o dígito de verificação. Chamando este dígito de x e fazendo o “produto escalar” com o vetor de pesos, temos:

$$\begin{aligned} 7 + (3 \times 8) + 9 + (3 \times 5) + 0 + (3 \times 0) + 0 + (3 \times 2) + 6 + (3 \times 6) + 2 + (3 \times 4) + x = \\ = 99 + x. \end{aligned}$$

Conseqüentemente, deve-se tomar $x = 1$.

Vejamos agora um exemplo de como funciona a detecção de erros. Um livro do autor [14] recebeu o código de barras 9781402002380. Suponhamos que, por um erro de digitação no quarto dígito, este número é transmitido como $\alpha = 9782402002380$. Ao fazer a verificação de leitura, o computador que recebeu a informação faz a operação $\alpha \cdot \omega$

e obtém:

$$9+(3 \times 7)+8+(3 \times 2)+4+(3 \times 0)+2+(3 \times 0)+0+(3 \times 2)+3+(3 \times 8)+0 = 73.$$

Como o resultado não é um múltiplo de 10, o computador avisa que foi cometido algum erro.

O código UPC é muito semelhante. Como utiliza apenas 12 dígitos (pois usa apenas um para identificar o país de origem do artigo, enquanto o EAN utiliza-se de dois), e o vetor de pesos utilizado pelo UPC também tem um dígito a menos; ele é:

$$\omega = (3, 1, 3, 1, 3, 1, 3, 1, 3, 1, 3, 1).$$

O leitor notará que, se o digitador comete apenas UM erro de digitação, trocando um dos dígitos a_i por um outro valor, então necessariamente o produto $\alpha \cdot \omega$ não será congruente a 0 em módulo 10 e assim será possível detectar que o erro foi cometido. Se mais de um erro for cometido na digitação, o fato provavelmente ainda será detectado, mas já não podemos ter certeza, pois eles poderiam se “compensar” mutuamente e a soma poderia ainda continuar sendo um múltiplo de 10.

O leitor pode-se perguntar qual é a função do vetor de pesos ω . De fato, se a escolha do dígito de verificação x fosse feita simplesmente de modo que

$$a_1 + a_2 + \dots + a_{12} + x \equiv 0 \pmod{10},$$

ainda assim UM erro de digitação seria detectado. Acontece que há um outro tipo de erro de digitação muito comum, que consiste em digitar todos os números corretamente, mas trocar a ordem de dois

dígitos consecutivos.

Suponha que, ao digitar o número 9 788531 404580 do nosso primeiro exemplo, tenha se cometido esse tipo de erro, e que o número de fato digitado fosse 9 788351 404580. Ao efetuar a verificação ter-se-ia:

$$\begin{aligned} & (9, 7, 8, 8, 5, 3, 1, 4, 0, 4, 5, 8, 0)(1, 3, 1, 3, 1, 3, 1, 3, 1, 3, 1, 3, 1) = \\ & = 9 + 21 + 8 + 24 + 3 + 15 + 1 + 12 + 12 + 5 + 24 \\ & = 134 \not\equiv 0 \pmod{10}. \end{aligned}$$

Desta forma, o erro seria detectado.

Suponha agora que, ao digitar o número 9 781402 002380 do nosso segundo exemplo, tenha se cometido um erro desse mesmo tipo, e que o número de fato digitado fosse 9 781402 002830. Ao efetuar a verificação ter-se-ia:

$$\begin{aligned} & (9, 7, 8, 1, 4, 0, 2, 0, 0, 2, 8, 3, 0) \cdot (1, 3, 1, 3, 1, 3, 1, 3, 1, 3, 1, 3, 1) = \\ & = 9 + 21 + 8 + 3 + 4 + 2 + 6 + 3 + 24 \\ & = 80 \equiv 0 \pmod{10}. \end{aligned}$$

Este exemplo mostra que o sistema de detecção adotado acima *não tem a capacidade de detectar todo erro de transposição* cometido. É fácil ver que a transposição de dois dígitos consecutivos a_i e a_{i+1} não é detectada, neste sistema de codificação, se, e somente se, $|a_i - a_{i+1}| = 5$ (veja a atividade 4).

Atividades

- 1) (i) Um determinado produto deve ser identificado, no código UPC pelo número 7 897595 90071. Determine o dígito de segurança correspondente.
(ii) Faça o mesmo para um produto cujo número é 7 894900 01152.
- 2) Vamos definir um sistema de detecção de erros da seguinte maneira. A cada número de 12 dígitos $a_0 \dots a_{11}$ vamos asignar um dígito de verificação a_{12} de forma tal que

$$\sum_{i=0}^{12} a_i \equiv 0 \pmod{10}.$$

- (i) Achar o dígito de verificação que deve se adicionar ao número
723443501297.
(ii) Provar que toda vez que apenas um número é alterado na digitação, este sistema é capaz de detectar o erro.
(iii) Mostrar que este sistema não é capaz de detectar qualquer erro de transposição.
- 3) Nós afirmamos no texto que se apenas um erro de digitação for cometido, alterando um dos dígitos a_i para um outro valor b_i , então o sistema UPC sempre será capaz de detectar o erro. Dê uma demonstração cuidadosa deste fato.

- 4) (i) Mostrar que uma *transposição adjacente*; isto é, um erro do tipo

$$\dots a_i a_{i+1} \dots \mapsto \dots a_{i+1} a_i \dots$$

é detectada pelo sistema EAN-13 se, e somente se, $|a_i - a_{i+1}| \neq 5$. (SUGESTÃO: note que $|w_i - w_{i+1}| = 2$).

- (ii) Mostre que um erro de *transposição não adjacente* do tipo

$$\dots a_i a_{i+1} a_{i+2} \dots \mapsto \dots a_{i+2} a_{i+1} a_i \dots$$

não pode ser detectado pelo sistema EAN-13.

- (iii) Mostrar que um erro de transposição em que dois dígitos não adjacentes a_i e a_j são trocados não pode ser detectado pelo sistema se a diferença $i - j$ é par.
- (iv) Mostrar que, num erro como o descrito no item anterior, se a diferença $i - j$ é ímpar, então o erro pode ser detectado pelo sistema EAN-13 se, e somente se, $|a_i - a_j| \neq 5$.

10.3 Outros Códigos Numéricos

Como observamos na seção anterior, existem diversos tipos de erros que podem ser cometidos ao digitar um *vetor de identificação*. Os erros num único dígito e as transposições são, de longe, os mais frequentes. Autores como D.F. Beckley [1] e J. Verhoeff [16] investigaram sistematicamente os erros cometidos por operadores humanos. No quadro abaixo damos as frequências relativas obtidas por Verhoeff, que citamos abreviando quadro publicado por H.H. Schulz [15] e também por G.B. Belyavskaya, V.I. Izbash e V.A. Shcherbacov [2].

Tipo de erro		Frequência relativa %
erro único	$\dots a \dots \mapsto \dots b \dots$	79
transposição adjacente	$\dots ab \dots \mapsto \dots ba \dots$	10.2
transposição alterna	$\dots abc \dots \mapsto \dots cba \dots$	0.8
erro gêmeo	$\dots aa \dots \mapsto \dots bb \dots$	0.6
erro gêmeo alternado	$\dots aba \dots \mapsto \dots cbc \dots$	0.3
outros		9.1

Tabela 10.1: Tipos de erros e suas frequências segundo Verhoeff

Para descrever vários dos métodos existentes, vamos introduzir um pouco de linguagem geral. Denotaremos por \mathcal{A} o conjunto de valores que podem assumir os dígitos utilizados na codificação. Por exemplo, no caso do código UPC da seção anterior, esse conjunto é

$$\mathcal{A} = \{x \in \mathbf{Z} \mid 0 \leq x \leq m - 1\}.$$

O vetor com os dados $\alpha' = (a_1, \dots, a_{n-1})$ será chamado de **vetor de informação** e o vetor, já acrescido do dígito de verificação será chamado de **número ou vetor de indentificação**.

Definição 10.1. *Sejam $\omega = (w_1, \dots, w_n)$, com $w_i \in \mathcal{A}$, $1 \leq i \leq n$ um vetor de pesos e $c \in \mathcal{A}$ um inteiro fixado. Dados dois inteiros positivos m e n e um conjunto de números a_1, \dots, a_{n-1} tais que $a_i \in \mathcal{A}$, $1 \leq i \leq n - 1$, define-se o **número de verificação** a_n como o único elemento de \mathcal{A} que verifica a equação:*

$$\sum_{i=1}^n a_i w_i \equiv c \pmod{m}.$$

Um sistema de codificação assim definido será denotado por $\mathcal{C} = (\mathcal{A}, m, n, c, \omega)$.

Note que frequentemente $\mathcal{A} = \{0, 1, \dots, m - 1\}$. Neste caso, tomando classes módulo m , temos que a_n é o único elemento de \mathcal{A}

que verifica:

$$\bar{a}_n = \bar{w}_n^{-1} \left(\bar{c} - \sum_{i=1}^{n-1} \bar{a}_i \bar{w}_i \right).$$

Exemplo 10.3.1.

Um sistema usado em alguns bancos (mas não todos) é o seguinte: o número de conta de um cliente é composto de 9 dígitos, sendo que o último é o dígito de verificação. Na nossa notação, o sistema pode ser descrito como $\mathcal{C} = (\mathcal{A}, 10, 2, 0, \omega)$ onde \mathcal{A} é o conjunto dos dígitos de 0 a 9 e $\omega = (7, 3, 9, 7, 3, 9, 7, 3, 9)$. Por exemplo, o número de uma conta num certo banco é 95-005541-9. Podemos verificar que

$$\begin{aligned} (9, 5, 0, 0, 5, 5, 4, 1, 9) \cdot (7, 3, 9, 7, 3, 9, 7, 3, 9) &= \\ &= 63 + 15 + 15 + 45 + 28 + 3 + 81 \\ &= 250 \equiv 10 \pmod{10}. \end{aligned}$$

Nosso próximo Teorema descreve a capacidade que tem um sistema definido desta forma, para detectar os diversos tipos de erros mais frequentes.

Teorema 10.2. *(Capacidade de detecção) Sejam m um inteiro positivo e $\omega = (w_1, \dots, w_n)$ um vetor de pesos. Suponhamos que um vetor de identificação $\alpha = (a_1, \dots, a_n)$ (onde assumimos que $0 \leq a_i < m$, para todo índice i , $1 \leq i \leq n$) satisfaz a condição*

$$\alpha \cdot \omega = a_1 w_1 + \dots + a_n w_n \equiv c \pmod{m}.$$

Então:

1. *Todo erro consistente numa única alteração na posição i -ésima será detectado se e somente se $\text{mdc}(w_i, m) = 1$.*

2. *Todo erro de transposição da forma*

$$\dots a_i \dots a_j \dots \mapsto \dots a_j \dots a_i \dots$$

será detectado se e somente se $\text{mdc}(w_i - w_j, m) = 1$.

Demonstração. Suponhamos inicialmente que o dígito a_i , na posição i , foi trocado por um outro valor b_i e vamos denotar por β o vetor resultante deste erro. É claro que o erro não será detectado se, e somente se,

$$\alpha \cdot \omega - \beta \cdot \omega \equiv 0 \pmod{m}.$$

Mas $\alpha \cdot \omega - \beta \cdot \omega = (a_i - b_i)w_i$, de modo que o erro não será detectado se, e somente se, $m \mid (a_i - b_i)w_i \equiv 0 \pmod{m}$ ou, se denotamos por \bar{x} a classe de um inteiro x em \mathbf{Z}_m , se, e somente se, $(\bar{a}_i - \bar{b}_i)\bar{w}_i = \bar{0}$ em \mathbf{Z}_m .

Se $\text{mdc}(w_i, m) = 1$ tem-se que \bar{w}_i é inversível em \mathbf{Z}_m , donde a condição acima implica que $\bar{a}_i = \bar{b}_i$, logo $a_i \equiv b_i \pmod{m}$ e, como ambos os números são menores que m , isto só aconteceria se $a_i = b_i$. Logo, o erro será detectado.

Por outro lado, se $\text{mdc}(w_i, m) = d \neq 1$, dado a_i tem-se que dentre os números $b_i = a_i + m/d$ e $b_i = a_i - m/d$, um deles verifica a condição $0 \leq b_i < m$ e o erro que substitui a_i por esse número não pode ser detectado. Isto completa a demonstração de (i).

Suponhamos agora que foi cometido um erro do tipo

$$\alpha = \dots a_i \dots a_j \dots \mapsto \alpha' = \dots a_j \dots a_i \dots$$

Note que, neste caso podemos calcular a diferença

$$\alpha \cdot \omega - \alpha' \cdot \omega = (a_i w_i + a_j w_j) - (a_j w_i - a_i w_j) = (a_i - a_j)(w_i - w_j).$$

Assim, este erro não será detectado se, e somente se,

$$(a_i - a_j)(w_i - w_j) \equiv 0 \pmod{m}.$$

Daqui em diante, o argumento é inteiramente análogo ao anterior. \square

A luz deste teorema, resulta claro que a melhor forma de ter certeza que o sistema de codificação será capaz de detectar todos os erros únicos e todos os erros de transposição (contigua ou não) é tomar, para o valor do módulo m , um número primo. De fato, existem vários sistemas em uso que procedem desta forma.

Exemplo 10.3.2.

Um sistema universalmente adotado para a classificação de livros é o ISBN (**International Standard Book Number**). Ele trabalha módulo 11, mas para facilitar a notação, utiliza também como conjunto de valores \mathcal{A} os dígitos de 0 a 9 e os vetores de identificação tem 10 componentes. Ele pode ser descrito, na nossa notação, por $(\mathcal{A}, 11, 10, 0, \omega)$ com $\omega = (10, 9, 8, 7, 6, 5, 4, 3, 2, 1)$.

Por exemplo, o livro do autor [14] mencionado anteriormente tem o número ISBN 1-4020-0238-6. O dígito final, de verificação é 6 porque

$$\begin{aligned} (1, 4, 0, 2, 0, 0, 2, 3, 8, 6) \cdot (10, 9, 8, 7, 6, 5, 4, 3, 2, 1) &= \\ &= 10 + 36 + 14 + 8 + 9 + 16 + 6 \\ &= 99 \equiv 0 \pmod{11}. \end{aligned}$$

Este método, porém, tem um pequeno inconveniente, que se compreenderá melhor analisando mais um exemplo. Um conhecido livro de álgebra [13] tem como vetor de informação de seu código ISBN o número 0-387-96035. Vamos tentar determinar o dígito de verificação. Ele será um número $a \in \mathcal{A}$ tal que

$$(0, 3, 8, 7, 9, 6, 0, 3, 5, a) \cdot (10, 9, 8, 7, 6, 5, 4, 3, 2, 1) \equiv 0 \pmod{11}.$$

Efetuando as operações necessárias, obtemos que

$$243 + a \equiv 0 \pmod{11}$$

ou

$$a \equiv -243 \pmod{11}.$$

Como $243 \equiv 1 \pmod{11}$ e $-1 \equiv 10 \pmod{11}$ temos que $a = 10$. Porém, no conjunto dos dígitos de 0 a 9, não temos nenhum que represente o número 10. Devemos introduzir então mais um símbolo para representar este número. A convenção usual é utilizar o símbolo X e assim, o código ISBN que aparece neste livro é ISBN 0-387-96035-X.

Finalmente observamos que, se tomamos o número m de modo que seja primo e o conjunto \mathcal{A} é formado por inteiros menores do que m – como em todos os exemplos acima – como cada componente w_i do vetor de pesos é prima com m , resulta que multiplicar por w_i , em módulo m , equivale a definir uma permutação do conjunto \mathcal{A} (isto é, uma bijeção de \mathcal{A} em si mesmo). Isto sugere um método mais geral de definir o vetor de pesos.

Dado um vetor de informação $\alpha' = (a_1, \dots, a_{n-1})$ podemos escolher n permutações $\delta_1, \dots, \delta_n$ do conjunto \mathcal{A} , definir um “vetor de pesos” por $\gamma = (\delta_1, \dots, \delta_n)$, fixar um número $c \in \mathcal{A}$ e escolher o dígito

de verificação a_n de modo que verifique a equação:

$$\gamma(\alpha) = \delta_1(a_1) + \cdots + \delta_n(a_n) \equiv c \pmod{m}.$$

Neste caso, o dígito de verificação fica definido por:

$$a_n = \delta_n^{-1} \left(c - \sum_{i=1}^{n-1} \delta_i(a_i) \right).$$

Este tipo de codificação também é usada na prática, como mostrado a seguir.

Exemplo 10.3.3.

Um código usado pela IBM utiliza como conjunto \mathcal{A} os dígitos de 0 a 9; os valores $m = 10$, um valor qualquer $c \in \mathcal{A}$ e a permutação

$$\delta = \begin{pmatrix} 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ 0 & 2 & 4 & 6 & 8 & 1 & 3 & 5 & 7 & 9 \end{pmatrix}.$$

No caso de um vetor de identificação com um número ímpar de componentes n , se utiliza o vetor de pesos

$$\gamma = (I, \delta, I, \delta, \dots, \delta, I).$$

Assim, a equação de verificação resulta:

$$a_n + \delta(a_{n-1}) + a_{n-2} + \delta(a_{n-3}) \cdots \equiv c \pmod{10}$$

e o dígito de verificação é dado pela fórmula:

$$a_n \equiv c - (\delta(a_{n-1}) - a_{n-2} - \delta(a_{n-3}) - \cdots) \pmod{10}.$$

No caso em que o número de componentes do vetor de identificação é par utiliza-se o vetor de pesos

$$\gamma = (\delta, I, \delta, I, \dots, \delta, I),$$

e resulta uma equação de verificação análoga, com uma fórmula similar para o cálculo do dígito de verificação.

Este é o sistema utilizado para determinar os números nos cartões de crédito. Por exemplo, um determinado cartão tem o número 5745 5195 0431 5412. Vamos a aplicar a ele o sistema de verificação IBM:

5	7	4	5	5	1	9	5	0	4	3	1	5	4	1	2
↓ δ	↓ I	↓ δ	↓ I	↓ δ	↓ I	↓ δ	↓ I	↓ δ	↓ I	↓ δ	↓ I	↓ δ	↓ I	↓ δ	↓ I
1	7	8	7	1	1	9	1	0	4	6	1	1	9	2	2

Agora, calculamos:

$$1+7+8+7+1+1+9+1+0+4+6+1+1+9+2+2 = 60 \equiv 0 \pmod{10}$$

Note que esta codificação permite detectar todo erro único de digitação e toda transposição adjacente exceto no caso em que a_i e a_j assumem os valores 0 e 9 ou 9 e 0 respectivamente (veja a atividade 3). Observe, porém, que este sistema não detecta transposições do tipo

$$\dots a_i \dots a_j \dots \mapsto \dots a_j \dots a_i \dots$$

quando a diferença $i - j$ é par e também não permite detectar erros gêmeos.

O código IBM generalizado utiliza a mesma permutação, com o vetor de pesos:

$$\gamma = (\delta^{n-1}, \delta^{n-2}, \dots, \delta, \delta^0)$$

onde $\delta^0 = I$. A equação é então:

$$\sum_{i=1}^n \delta^{i-1}(a_{n+1-i}) \equiv c \pmod{10}$$

e

$$a_n \equiv c - \sum_{i=2}^n \delta^{i-1}(a_{n+1-i}) \pmod{10}.$$

Este sistema detecta todo erro único de digitação, toda transposição, adjacente ou não, (exceto no caso já apontado, quando os dígitos envolvidos são 0 e 9) e todo erro gêmeo.

Vimos acima que o código ISBN detecta todo erro único de digitação e todo erro de transposição. É natural se perguntar, então, se existe um código com essa capacidade de detecção, trabalhando na base 10. Infelizmente, a resposta é negativa, como mostra o seguinte teorema.

Teorema 10.3. *(Gumm [10]) Se um sistema numérico de detecção de erros, com um módulo par, detecta todo erro único de digitação, então, para todo par de índices i, j existe um erro de transposição entre as posições i e j que não é detectada pelo sistema.*

Demonstração. Como vamos trabalhar com os números de 0 a $2m - 1$ e tomar congruências em módulo $2m$, vamos considerar nossos dígitos como elementos de \mathbf{Z}_{2m} , para simplificar nossos argumentos. Suponhamos que o sistema transforma o vetor (a_1, \dots, a_n) num outro vetor, que denotaremos por $(\sigma_1(a_1), \dots, \sigma_n(a_n))$. Claramente, se o sistema é capaz de detectar todo erro único de digitação, então a aplicação na posição i -ésima $x \mapsto \sigma_i(x)$ deve ser uma permutação de \mathbf{Z}_{2m} .

Para que o sistema detecte todo erro de transposição entre as posições i e j é necessário que $\sigma_i(a) + \sigma_j(b) \neq \sigma_j(a) + \sigma_i(b)$, todo par

de elementos diferentes $a, b \in \mathbf{Z}_{2m}$. Isto é equivalente a dizer que a aplicação $\sigma = \sigma_i - \sigma_j$ é uma permutação \mathbf{Z}_{2m} .

Mas $m \in [0, 2m - 1]$ e temos que

$$0 + 1 + 2 + \dots + 2m - 1 = \frac{(2m - 1) \cdot 2m}{2} = 2m \cdot m - m \equiv m \pmod{2m},$$

ou seja,

$$\overline{0} + \overline{1} + \dots + \overline{2m - 1} = \overline{m} \quad \text{em } \mathbf{Z}_{2m}.$$

Logo:

$$\begin{aligned} \overline{m} &= \sum_{\overline{x} \in \mathbf{Z}_{2m}} \overline{x} = \sum_{\overline{x} \in \mathbf{Z}_{2m}} \sigma(\overline{x}) \\ &= \sum_{\overline{x} \in \mathbf{Z}_{2m}} (\sigma_i(\overline{x}) - \sigma_j(\overline{x})) = \sum_{\overline{x} \in \mathbf{Z}_{2m}} \sigma_i(\overline{x}) - \sum_{\overline{x} \in \mathbf{Z}_{2m}} \sigma_j(\overline{x}) \\ &= \overline{m} - \overline{m} = \overline{0}, \quad \text{em } \mathbf{Z}_{2m}. \end{aligned}$$

uma contradição. □

Atividades

- 1) Calcular o dígito de verificação para um livro cujo número ISBN tem, como vetor de informação o número 85-314-0458.
- 2) Use o sistema do exemplo 10.3.1 para determinar o dígito de verificação do número 13-010765.
- 3) Demonstre que o código IBM e o código IBM generalizado tem a capacidade de detecção de erros mencionada no texto. (SUGESTÃO: Para discutir erros de transposição adjacente, note que $\delta(x) = 2x$, se $x \in [0, 4]$ e que $d(x) = 2x - 9$ se $x \in [5, 9]$. Considere separadamente três casos: (i) $a, b \in [0, 4]$, (ii) $a, b \in [5, 9]$ e



(iii) $a \in [0, 4], b \in [5, 9]$ ou vice-versa. Mostre que nos casos (i) e (ii) o erro é sempre detectado e que, no caso (iii) o erro só não é detectado se $a = 0$ e $b = 9$ ou $a = 9$ e $b = 0$.

- 4) Mostre que no código UPC podem ocorrer 90 erros de transposição adjacente e que o código é capaz de detectar todos eles, exceto quando os pares de números adjacentes são 05, 16, 27, 38, 49 ou aqueles que se obtém invertindo estes.



Capítulo 11

Códigos Sobre Grupos

11.1 O Grupo Dihedral

Na seção anterior apresentamos diversos métodos de detectar erros usando um dígito de verificação. Dentre estes, só o sistema ISBN para livros era capaz de detectar todo erro único de digitação e todo erro de transposição. Ele tinha, porém, o inconveniente de precisar da introdução de um dígito extra, para representar o número 10, que denotamos por X .

Em 1969 Verhoeff, na sua tese de doutoramento [16], desenvolveu um método simples, baseado não em cálculos com números inteiros, mas com os elementos de um certo grupo, que também detecta erros únicos de digitação e todos os erros de transposição adjacentes, sem necessidade de símbolos extras. A exposição elementar deste método que damos a seguir aparece em [6] e num texto básico de álgebra, do mesmo autor [5, Capítulo V].

Consideremos o **grupo dihedral** D_5 , que pode ser definido como o grupo das isometrias do plano que deixam fixo um pentágono regular dado. Este grupo contém dez elementos. Cinco deles são rotações: a

identidade R_0 , a rotação R_1 de ângulo $2\pi/5$ em sentido anti-horário, e as rotações R_2 , R_3 e R_4 de ângulos $2(2\pi/5)$, $3(2\pi/5)$ e $4(2\pi/5)$ respectivamente. Contém ainda cinco reflexões, em relação aos seus eixos de simetria, que passam por cada um dos vértices e o ponto médio do lado oposto: S_5 , S_6 , S_7 , S_8 , S_9 e S_{10} .

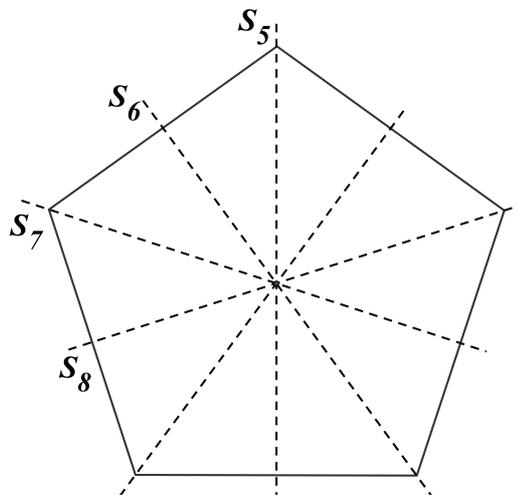


Figura 11.1: Simetrias do pentágono

Vamos usar simplesmente os subíndices 0, 1, 2, 3, 4 para designar as rotações correspondentes e os subíndices 5, 6, 7, 8, 9 para as respectivas reflexões.

Assim, por exemplo, a composição das rotações R_3 e R_4 seria a rotação de ângulo $3(2\pi/5) + 4(2\pi/5) = 7(2\pi/5) = 2(2\pi/5)$; isto é, a rotação R_2 . Em vez de escrever então que $R_3 \circ R_4 = R_2$ escreveremos simplesmente que $3 \cdot 4 = 2$.

Da mesma forma, como a composição $S_6 \circ S_5 = R_1$ (lembre que, como se trata de composição de funções, aplicamos primeiro S_1 e depois S_2) escrevemos $6 \cdot 5 = 1$. Por outro lado, é fácil verificar que $S_5 \circ S_6 = R_0$ donde escrevemos $6 \cdot 5 = 0$.

Procedendo desta forma, obtemos a seguinte tabela de multiplicação para D_5 .

·	0	1	2	3	4	5	6	7	8	9
0	0	1	2	3	4	5	6	7	8	9
1	1	2	3	4	0	6	7	8	9	5
2	2	3	4	0	1	7	8	9	5	6
3	3	4	0	1	2	8	9	5	6	7
4	4	0	1	2	3	9	5	6	7	8
5	5	9	8	7	6	0	4	3	2	1
6	6	5	9	8	7	1	0	4	3	2
7	7	6	5	9	8	2	1	0	4	3
8	8	7	6	5	9	3	2	1	0	4
9	9	8	7	6	5	4	3	2	1	0

Tabela 11.1: A tabela de multiplicação do grupo D_5

Consideremos a permutação:

$$\sigma = \begin{pmatrix} 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ 1 & 5 & 7 & 6 & 2 & 8 & 3 & 0 & 9 & 4 \end{pmatrix}.$$

A ideia de Verhoeff consiste em transformar um vetor de informação (a_1, \dots, a_{n-1}) num vetor de codificação, adicionando um dígito de verificação a_n de forma tal que

$$\sigma(a_1) \cdot \sigma^2(a_2) \cdot \dots \cdot \sigma^{n-1}(a_{n-1}) \cdot a_n = 0 \text{ em } D_5.$$

Note que σ é uma permutação de D_5 . Ela foi escolhida para desenvolver este sistema porque pode-se verificar diretamente que

$$a \cdot \sigma(b) \neq b \cdot \sigma(a) \text{ para todo } a, b \in D_5. \tag{11.1}$$

Logo abaixo veremos a importância deste fato.

Como σ^i também é uma permutação de D_5 , para todo inteiro positivo i , resulta claro que este sistema de codificação detecta todo erro único de digitação.

Note que um erro de transposição adjacente $\dots a_i a_{i+1} \dots \mapsto \dots a_{i+1} a_i \dots$ será detectado se, e somente se, $\sigma^i(a_i) \cdot \sigma^{i+1}(a_{i+1}) \neq \sigma^i(a_{i+1}) \cdot \sigma^{i+1}(a_i)$. Mas, vimos na equação 11.1 que $a \cdot \sigma(b) \neq b \cdot \sigma(a)$ para todo $a, b \in D_5$ o que implica, aplicando a permutação σ^i a ambos os membros desta equação, que

$$\sigma^i(a) \cdot \sigma^{i+1}(b) \neq \sigma^i(b) \cdot \sigma^{i+1}(a) \text{ para todo } a, b \in D_5$$

como queríamos demonstrar.

Uma variante do método de Verhoeff¹ foi usada pelo órgão emissor de dinheiro da Alemanha, o Deutsche Bundesbank. As notas de dinheiro são numeradas num código alfanumérico; isto é, se utilizam tanto de letras quanto de números. Na Figura 11.2 reproduzimos uma nota de 10 marcos (que já está fora de circulação desde o advento da moeda unificada da Europa, o Euro).



Figura 11.2:

¹Citado por Gallian [7].

O código utilizado pelo banco segue uma variante do método anterior. Eles numeram as notas usando os dígitos de 0 a 9 e também dez letras: A, D, G, K, L, N, U, V e Z. Também se utiliza da tabela da operação do grupo D_5 , mas ao invés de usar uma permutação e suas potências, utiliza dez permutações diferentes. Para determiná-las damos, na tabela abaixo, em cada fila i os valores da função σ_i . Em outras palavras, na posição i, j está o valor de $\sigma_i(j)$.

	0	1	2	3	4	5	6	7	8	9
σ_1	1	5	7	6	2	8	3	0	9	4
σ_2	5	8	0	3	7	9	6	1	4	2
σ_3	8	9	1	6	0	4	3	5	2	7
σ_4	9	4	5	3	1	2	6	8	7	0
σ_5	4	2	8	6	5	7	3	9	0	1
σ_6	2	7	9	3	8	0	6	4	1	5
σ_7	7	0	4	6	9	1	3	2	5	8
σ_8	0	1	2	3	4	5	6	7	8	9
σ_9	1	5	7	6	2	8	3	0	9	4
σ_{10}	5	8	0	3	7	9	6	1	4	2

Isto significa, por exemplo, que a permutação σ_5 é:

$$\sigma_5 = \begin{pmatrix} 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ 4 & 2 & 8 & 6 & 5 & 7 & 3 & 9 & 0 & 1 \end{pmatrix}.$$

O número de série da nota da Figura 11.2 é DU7124458G6. Vamos verificar que esta é uma numeração válida. Para poder usar o método de Verhoeff devemos trabalhar apenas com os dígitos de 0 a 9; por causa disso, a cada letra das usadas pelo Deutch Bundesbank deve-se lhe assignar um valor numérico. Isto é feito de acordo com a seguinte tabela.

A	D	G	K	L	N	S	U	Y	Z
0	1	2	3	4	5	6	7	8	9

Usando esta tabela, o número da nota em questão se transforma em 17712445826. Aplicamos então ordenadamente as permutações dadas:

1	7	7	1	2	4	4	5	8	2	6
$\downarrow \sigma_1$	$\downarrow \sigma_2$	$\downarrow \sigma_3$	$\downarrow \sigma_4$	$\downarrow \sigma_5$	$\downarrow \sigma_6$	$\downarrow \sigma_7$	$\downarrow \sigma_8$	$\downarrow \sigma_9$	$\downarrow \sigma_{10}$	\downarrow
5	1	5	4	8	8	9	5	9	0	6

Agora precisamos “multiplicar” estes números, com ajuda da tabela 11.1. Temos:

$$5 \cdot 1 = 9, \quad 9 \cdot 5 = 4, \quad 4 \cdot 4 = 3,$$

$$3 \cdot 8 = 6, \quad 6 \cdot 8 = 3, \quad 3 \cdot 9 = 7,$$

$$7 \cdot 5 = 2, \quad 2 \cdot 9 = 6, \quad 6 \cdot 0 = 6,$$

e finalmente

$$6 \cdot 6 = 0.$$

Este método, porém, tem um inconveniente. Nos cálculos, ele não distingue entre uma letra e o número que lhe é assignado. Assim por exemplo, se a letra K for trocada pelo número 3, o sistema será incapaz de detectar o erro. O mesmo acontece se ocorre uma transposição de 3 e K, ou vice-versa. Para evitar este problema, poder-se-ia usar o grupo D_{18} , que tem 36 elementos (e portanto os vinte símbolos usado no código alfanumérico das notas corresponderiam a elementos diferentes em D_5), com uma permutação adequada. Sugestões nesse sentido apareceram, por exemplo, em trabalhos de Winters, em 1990 [17] e de Gallian e Mullin, em 1995 [8].

Note que a equação 11.1 foi essencial para podermos mostrar que o método de Verhoeff permite detectar erros de transposição adjacentes. Isto justifica a seguinte definição.

Definição 11.1. *Uma permutação σ de um grupo G diz-se uma **aplicação antissimétrica** se verifica a seguinte condição:*

$$x\sigma(y) \neq y\sigma(x), \quad \text{para todo par de elementos } x, y \in G.$$

O grupo D_5 desempenha um papel importante na elaboração de códigos detectores de erros porque pode-se mostrar que ele é o único grupo de ordem 10 que possui uma aplicação antissimétrica.

Como as transposições adjacentes se encontram entre os erros mais comuns e como códigos que detectam estes erros podem-se elaborar a partir de grupos com aplicações antissimétricas, houve vários trabalhos que dedicaram especial atenção a este tipo de grupos.

No caso dos grupos abelianos, há um tipo de permutação que é também importante.

Definição 11.2. *Uma permutação σ de um grupo G diz-se uma **aplicação completa** se a função $x \mapsto x\sigma(x)$, para todo $x \in G$, é uma permutação de G .*

Pode-se demonstrar que se G é um grupo abeliano, então G possui uma aplicação antissimétrica se, e somente se, G possui uma aplicação completa.

Um grupo abeliano de ordem $2m$, com m ímpar não possui aplicações completas [3]. Como consequência imediata, temos que *Um grupo abeliano de ordem $2m$, com m ímpar não possui aplicações antissimétricas.*

Outros resultados relativos a este tipo de aplicações são os seguintes:

- Um grupo cíclico admite uma aplicação antissimétrica se, e somente se, é de ordem ímpar [15].

- Todo grupo solúvel não abeliano admite uma aplicação antisimétrica [11].
- Todo grupo simples, exceto \mathbf{Z}_2 , admite uma aplicação antisimétrica [8].

Foi anunciado por Heiss que todo grupo finito não abeliano também admite uma aplicação antissimétrica [12].

Para os grupos dihedrais, diversas classes de aplicações antisimétricas foram achadas em [4] e [9]. Como vimos, isto implica que estes grupos podem ser usados para construir códigos que detectam erros únicos de digitação ou transposições adjacentes. Porém, eles não podem detectar outros erros frequentes, como mostra o seguinte teorema, devido a Damm [3, Teorema 5].

Teorema 11.3. *Seja $m > 2$ um inteiro ímpar. Não existe um sistema de dígito de controle sobre D_m que seja capaz de detectar todas as transposições alternadas, todas os erros gêmeos ou todos os erros gêmeos alternados.*

Atividades

- 1) Determine todos os subgrupos cíclicos de D_5 .
- 2) Ache o centro $\mathcal{Z}(D_5)$ de D_5 e determine o quociente $D_5/\mathcal{Z}(D_5)$.
- 3) Determine a decomposição em produto de ciclos disjuntos e a paridade da permutação σ do método de Verhoeff.
- 4) Determine o valor de x para que o número 3572498x seja um número válido no método de Verhoeff.



- 5) Mostre que tomando a permutação $\tau = (1\ 4)(2\ 3)(5\ 8\ 6\ 9\ 7)$ no método de Verhoeff, também é possível detectar todo erro único de digitação e toda transposição adjacente.
- 6) Determine o valor do dígito x para que o número $AD377345654Kx$ seja um número válido para uma nota emitida pelo Deustche Bank.
- 7) Idem, para o número $NZ357x29477L2$.
- 8) (Gallian [5]) Seja $\sigma = (1\ 2\ 4\ 8\ 7\ 5)(3\ 6)$. A cada número da forma $a_1a_2 \dots a_n$ (com n ímpar) atribuímos o dígito de controle $-(\sigma(a_1) + a_2 + \sigma(a_3) + a_4 + \dots + \sigma(a_n)) \pmod{10}$. Calcule o valor do dígito de controle para o número 3125600196431. Prove que este método detecta todo erro único de digitação. Determine quais transposições adjacentes não podem ser detectadas por este método.





Referências Bibliográficas

- [1] BECKLEY, D.F. An optimum system with modulo 11. *The Computer Bulletin*, n. 11, p. 213-215, 1967.
- [2] BELYAVSKAYA, G.B.; IZBASH V.I.; MULLEN, G.L. Check character systems over quasegroups and loops. *Quasigroups and related systems*, n. 10, p. 1-28, 2003.
- [3] DAMN, M. Check digit over groups and anti-symmetric mappings. *Archiv der Math.*, n. 75, p. 413-421, 2000.
- [4] ECKER, A.; POCH, G. Check character systems. *Computing*, n. 37, p. 277-301, 1986.
- [5] GALLIAN, J.A. *Contemporary Abstract Algebra*, D.C. Heath and Co., Lexington, 1990.
- [6] GALLIAN, J.A. The Mathematics of Identification Numbers, *The College Math. J.*, v. 22, n. 3, p. 194-202, 1991.
- [7] GALLIAN, J.A. Error detecting methods, *ACM Computing Surveys*, v. 28, n. 3, p. 504-517, 1996.
- [8] GALLIAN, J.A.; MULLIN, M. Groups with antisymmetric mappings, *Archiv der Math.*, n. 65, p. 273-280, 1995.
- [9] GUMM, H.P. A new class of check-digit methods for arbitrary number systems, *IEEE Trans, Inf. Th.*, n. 31, p. 102-105, 1985.

- [10] GUMM, H.P. Encoding of numbers to detect typing errors, *Inter. J. Applied Eng. Ed.*, n. 2, p. 61-65, 1986.
- [11] HEISS, S. Anti-symmetric mappings for finite solvable groups, *Archiv der Math.*, n. 69, p. 445-454, 1997.
- [12] HEISS, S. Anti-symmetric mappings for finite groups, *preprint*, 1999.
- [13] LIDL, R.; PILZ, G. *Applied Abstract Algebra*. Undergraduate Texts in Math. New York: , Springer Verlag, 1984.
- [14] POLCINO MILIES, C.; SEHGAL, S.K. *An introduction to Group Rings*, Kluwer Acad. Publ., Dordrecht, 2002.
- [15] SCHULZ, R.H. On check digit systems using anti-symmetric mappings. In *Numbers, Information and Complexity*, p. 295-310, Dordrecht: Kluwer Acad. Publ., 2000.
- [16] VERHOEFF, J. Error detecting decimal codes. *Math. Centre Tracts*. Amsterdam: Mathematische Centrum, 1969.
- [17] WINTERS, S. Error detecting codes using dihedral groups. *UMAP J.*, n. 11, p. 299-308, 1990.

Há também muitas páginas na internet com informações sobre o assunto. Veja, por exemplo:

<http://en.wikipedia.org/wiki/ENIAC>

http://en.wikipedia.org/wiki/Harvard_Mark-I



REFERÊNCIAS BIBLIOGRÁFICAS

183

http://en.wikipedia.org/wiki/Ada_Lovelace

<http://www-etsi2.ugr.es/alumnos/mili/Harvard20I.htm>

http://www.bellsouthpwp.net/1/a/laurergj/upc_work.html

<http://www.barcodeisland.com/ean13.phtml>

<http://www.adams1.com/pub/russadam/barcode1.html>

