

Introdução ao Pensamento Dedutivo

S. C. Coutinho

Agradecimentos

Obrigado a todo mundo que ajudou apontando correções e dando sugestões para melhorar estas notas de aula, especialmente Juliana Vianna Valério, Daniel Kiyoshi Hashimoto Vouzella de Andrade, Pedro Mion Braga Cordeiro, Bernardo Milewski Garcia, Pedro Francisco Staino Santayana, Gabriel da Costa Lins Martins, Rafaela de Carvalho Machado Pinheiro, Luigi Monteiro Santos Rangel, Nicolly Zorzam Moura, Gabriel Vieira do Amaral, Luiz Miguel Viana Barbosa, Demir Novaes Amaral, Pedro Poppolino, Victor Paiva Martins Yao Souto, Pedro Mateus Melo Ormesino Lins, Gustavo Teixeira Breda, João David Jotta Mendona Estorque, Luiz Eduardo Lahm Piccoli,

Contents

Agradecimentos	iii
Part 1. Lógica e conjuntos	1
Chapter 1. Funções	3
1. Um dois, três...funções	3
2. Funções injetivas e sobrejetivas	5
3. Composição de funções	8
Exercícios	11
Part 2. Geometria euclidiana	13
Chapter 2. Comprimentos, áreas e números reais	15
1. Áreas	15
2. A geometria chega à Grécia	19
3. O teorema de Pitágoras e o círculo	22
4. A geometria na Grécia Antiga	25
5. Os elementos de Euclides	26
6. Números reais	30
Exercícios	36
Chapter 3. A geometria axiomatizada	39
1. Introdução	39
2. Incidência e medidas em retas	41
3. Semiplanos e ângulos	46
4. Triângulos	49
5. Paralelas e perpendiculares	56
6. Paralelas	68
7. Medindo comprimento e área	73
8. Semelhança de triângulos	78
Part 3. Funções reais	83
Chapter 4. Funções reais	85
1. Definições e exemplos	85

2. Imagem de uma função	88
3. Gráficos de funções	90
4. Operações com funções	95
 Chapter 5. Funções polinomiais	 105
1. Definições e propriedades básicas	105
2. Divisão com resto e fatoração	108
3. Divisão com resto e raízes	112
4. Fórmulas de Viète	114
 Chapter 6. Trigonometria	 119
1. Seno e cosseno	119
2. Seno e cosseno de soma de ângulos	124
3. Tangente	130
4. Identidades trigonométricas	134
 Chapter 7. Exponenciais e logaritmos	 135
1. Potências	135
2. A função exponencial	137
3. O logaritmo	139
4. Aplicações	144
 Bibliography	 149

Part 1

Lógica e conjuntos

CHAPTER 1

Funções

Como veremos no capítulo 4, a noção matemática de função surgiu no século XVI em mecânica e, eventualmente, espalhou-se por toda a matemática. Contudo, a ideia básica por trás da noção matemática é, provavelmente, tão antiga quanto a humanidade, e é com ela que começaremos este capítulo.

1. Um dois, três...funções

Como ponto de partida investigaremos uma pergunta extremamente básica: em que consiste contar? A versão mais simples do ato de contar é ilustrada quando uma criança pequena conta seus brinquedos, recitando os números em ordem crescente, à medida que aponta para cada um deles. Assim, a cada brinquedo é associada uma palavra (um número), a última das quais designa a quantidade de brinquedos. Portanto, *contar* é um caso especial da ação de *associar*; mais precisamente aquele em que, a cada objeto, associamos certas palavras, recitadas em uma ordem pré-estabelecida. Além disso, para que a contagem seja feita corretamente, é necessário que façamos corresponder a cada objeto uma única destas palavras. Resumindo,

contar pode ser descrito como o ato de associar a cada objeto uma, e uma única, palavra de uma lista padronizada, que chamamos de números.

O processo de associação que está por trás da contagem precede a invenção dos próprios números. Para se certificar, ao fim do dia, de que não perdeu nenhuma de suas ovelhas basta que um pastor associe a cada uma uma pequena pedra que guarda, talvez, em um saco na cintura. O problema é que pedras pesam e, dependendo do tamanho do rebanho, carregá-las à cintura pode tornar-se muito inconveniente. Palavras, contudo, não pesam; mas para que o processo funcione, é necessário que as palavras sejam padronizadas e recitadas sempre na mesma ordem.

Com isto temos o essencial para definir o conceito matemático de função. Para começar, precisamos de dois conjuntos X e Y que vamos relacionar entre si. Nos exemplos acima, X era o conjunto dos brinquedos de uma criança e Y o conjunto \mathbb{N} dos números naturais, ou X era um conjunto de ovelhas e Y o conjunto das pedras que o pastor carrega à cintura. Uma função f de X em Y é uma regra que, a cada elemento de X associa um, e um único,

elemento de Y . No primeiro exemplo, a cada brinquedo era associado um número; no segundo, a cada ovelha uma pedra. Note, porém, que, no primeiro caso, embora a cada brinquedo seja associado um número, há números que não correspondem a nenhum brinquedo, porque \mathbb{N} é um conjunto infinito, o que não ocorre com um conjunto de brinquedos, por mais que isto fosse o que todos desejássemos quando éramos criança.

Se f é uma função que associa, a cada elemento do conjunto X um elemento de Y , dizemos que X é o *domínio* de f e que Y é seu *contradomínio*. Abreviamos isto usando uma das seguintes notações

$$f : X \longrightarrow Y \text{ ou } X \xrightarrow{f} Y.$$

Assim, a afirmação “seja $f : X \longrightarrow Y$ uma função” deve ser interpretada como afirmando que f é uma função cujo domínio é o conjunto X e cujo contradomínio é o conjunto Y .

Se x é um elemento do domínio X de f , então o elemento de Y ao qual f associa x é denotado por $f(x)$. Dizemos que $f(x)$ é a imagem do elemento $x \in X$ pela função f . Por exemplo, seja \mathcal{L} o conjunto das letras maiúsculas no alfabeto latino e $p : \mathcal{L} \longrightarrow \mathbb{N}$ a função que a cada letra associa sua posição no alfabeto. Neste caso, $p(A) = 1$ e $p(M) = 13$; isto é, a imagem de A por p é 1 e a imagem de M por p é 13.

No caso de muitas funções que aparecem em matemática é possível descrever a imagem de um elemento qualquer do domínio usando uma fórmula explícita. Por exemplo, se $d : \mathbb{N} \longrightarrow \mathbb{N}$ é a função que a cada número natural associa seu dobro, a imagem de $n \in \mathbb{N}$ é $f(n) = 2n$. Outro exemplo do mesmo gênero é a função $q : \mathbb{N} \longrightarrow \mathbb{N}$ que a cada número natural associa seu quadrado, de modo que $q(n) = n^2$. Nem toda função pode ser descrita por regras tão simples; por exemplo, uma função cujo comportamento os matemáticos ainda não conseguiram decifrar completamente, é $C : \mathbb{N} \longrightarrow \mathbb{N}$, definida pela regra

$$C(x) = \begin{cases} \frac{x}{2} & \text{quando } x \text{ é par} \\ 3x + 1 & \text{quando } x \text{ é ímpar;} \end{cases}$$

sobre a qual teremos mais a dizer na seção 3. Contudo, a maior parte das funções realmente importantes em matemáticas não pode ser definida por nada que seria razoável chamar de fórmula, mas sim por certos procedimentos; razão pela qual tantas delas têm nomes especiais, como é o caso do seno, do cosseno, da exponencial, do logaritmo e de tantas outras que aparecem de maneira natural em matemática, estatística, física e engenharia, como a função gama, a função de Bessel, a função ζ de Riemann e a função hipergeométrica.

Para encerrar, definimos duas funções f e g , com mesmo domínio X e mesmo contradomínio Y , como sendo *iguais* se

$$f(x) = g(x) \text{ para todo } x \in X.$$

Neste caso, escrevemos $f = g$. Note que duas funções iguais podem ser definidas por regras diferentes; como é o caso de $h_1, h_2 : \{-2, 2\} \rightarrow \mathbb{N}$, definidas por

$$h_1(x) = x^2 \text{ e } h_2(x) = 2|x| = \begin{cases} 2x & \text{quando } x \geq 0 \\ -2x & \text{quando } x < 0. \end{cases}$$

2. Funções injetivas e sobrejetivas

Como vimos no caso da função utilizada na contagem de brinquedos, nem todo elemento do contradomínio de uma função tem que ser imagem do seu domínio. Isto torna conveniente definir a noção de imagem de uma função. Digamos que $f : X \rightarrow Y$ seja uma função. Vimos na seção 1 que o elemento $f(x) \in Y$ ao qual f associa um elemento $x \in X$ é a imagem de x por f . A *imagem* $\text{im}(f)$ da função f é o subconjunto de Y formado por todos os elementos da forma $f(x)$ para algum $x \in X$; isto é,

$$\text{im}(f) = \{f(x) \mid x \in X\}.$$

Portanto, para um dado $y \in Y$,

$$y \in \text{im}(f) \iff \exists x \in X, (y = f(x)).$$

Assim, se uma criança tem 30 brinquedos, a imagem da função que conta os seus brinquedos será

$$\{1, 2, \dots, 30\};$$

ao passo que as imagens das funções d que dobra um número natural e q que calcula seu quadrado serão, respectivamente, o conjunto dos números pares e o conjunto dos quadrados de números inteiros. A imagem da função de Collatz também é fácil de determinar, porque $2n$ é par para todo $n \in \mathbb{N}$, de modo que

$$C(2n) = \frac{2n}{2} = n;$$

o que mostra que $\text{im}(C) = \mathbb{N}$.

Quando a imagem de uma função é igual ao seu contradomínio, dizemos que a função é *sobrejetiva*. Entre as funções citadas no parágrafo anterior, apenas a de Collatz é sobrejetiva, porque nem todo número natural é par ou um quadrado de outro número natural. Vale lembrar que funções são criações humanas e que nada nos impede de redefini-las, sempre que necessário, de modo que se tornem sobrejetivas. Por exemplo, no caso da função d , podemos redefini-la como $d : \mathbb{N} \rightarrow \mathcal{P}$, em que \mathcal{P} é o conjunto dos números pares. Embora isto seja viável, nem sempre é conveniente, como veremos quando tratarmos das operações com funções na seção 4 do capítulo 4.

Voltando à função utilizada na contagem, imaginemos uma criança que tem 15 brinquedos, mas que só sabe contar até dez. Neste caso, a criança pode parar de contar quando chegar em dez, ou pode recomeçar a contagem de um, quando chegar ao décimo primeiro

brinquedo. No primeiro caso, não temos uma função, porque há brinquedos aos quais não estão associados nenhum número. O segundo caso, contudo, é bem mais interessante, porque, ainda que a contagem esteja errada, a associação de números a brinquedos, feita pela criança, continua sendo uma função. Para tornar mais clara a diferença entre as duas situações, observe que dizer que $f : X \rightarrow Y$ é uma função significa que

$$\forall x \in X, (\exists! y \in Y, (y = f(x))).$$

Por outro lado, para que uma função $f : X \rightarrow Y$ possa ser usada como uma contagem é necessário também que

$$(1) \quad \forall y \in Y, (\exists! x \in X, (y = f(x))).$$

Este é um daqueles casos em que a troca de posição dos quantificadores altera completamente o significado do que está sendo dito. Uma maneira menos prolixa de expressar (1) consiste em dizer que se $x, x' \in X$, então

$$x \neq x' \rightarrow f(x) \neq f(x');$$

ou, usando a contrapositiva,

$$f(x) = f(x') \rightarrow x = x'.$$

Uma função que satisfaz (1), ou uma das duas versões equivalentes enunciadas acima, é chamada de *injetiva*. Portanto, apenas funções injetivas podem ser usadas em contagens.

Tanto a função d , quanto a função q , são injetivas; a primeira porque cada número par tem uma única metade, a segunda porque cada quadrado tem uma única raiz positiva. Contudo, se redefirmos q como uma função de \mathbb{Z} em \mathbb{N} , ela deixará de ser injetiva, porque

$$q(-2) = 4 = q(2).$$

Outra função que não é injetiva é a função de Collatz, porque qualquer número natural da forma $3n + 1$ é imagem por C de n e de $6n + 2$; por exemplo, quando $n = 10$, temos que

$$C(10) = 3 \cdot 10 + 1 = 31 = \frac{62}{2} = C(62).$$

As funções que são simultaneamente sobrejetivas e injetivas são conhecidas como *bijetivas*. Estas funções são especialmente importantes, porque se existe uma função bijetiva $f : X \rightarrow Y$, entre dois conjuntos *finitos* X e Y , então estes conjuntos têm a mesma quantidade de elementos. Para entender o porquê desta afirmação, lembre-se que, pela definição da sobrejetividade, $\text{im}(f) = Y$. Logo, para cada elemento $y \in Y$ existe, pelo menos, um elemento $x \in X$ do qual y é a imagem. Assim, Y não pode ter mais elementos do que X . Por outro lado, cada elemento de X tem que ser levado em algum elemento de Y e, além disso, um mesmo elemento $x \in X$ não pode ser levado em mais de um elemento de Y . Portanto, para que a quantidade de elementos de X fosse maior que a quantidade de elementos de Y seria necessário que f aplicasse mais de um elemento de X em um mesmo

elemento de Y , o que não é possível porque f é bijetiva. Mas, se Y não pode ter mais, nem menos, elementos que X , então os dois conjuntos têm a mesma quantidade de elementos.

Uma manifestação simples desta ideia ocorre quando agrupamos elementos aos pares. Por exemplo, para saber se, em um conjunto formado por bolas amarelas e verdes, as quantidades de bolas das duas cores são iguais, podemos formar pares cujo primeiro elemento é uma bola amarela e cujo segundo elemento é uma bola verde. Sempre que a quantidade de bolas verdes for maior ou igual que a de bolas amarelas, teremos uma função que, a cada bola amarela associa a bola verde que é seu par. Naturalmente, se houver mais bolas amarelas que verdes no conjunto, algumas destas últimas não farão parte de nenhum par e a função não será sobrejetiva. Contudo, a função que corresponde a um pareamento tem que injetiva, porque não é possível usar uma mesma bola verde como segundo membro em dois pares distintos. Assim, a função construída a partir de um pareamento será bijetiva quando as quantidades de bolas das duas cores coincidirem.

Podemos utilizar esta estratégia para contar, por exemplo, a quantidade de subconjuntos do conjunto

$$S = \{1, 2, 3, \dots, n\},$$

a partir do produto cartesiano

$$B^n = \underbrace{B \times \cdots \times B}_{n \text{ vezes}}$$

de n -cópias do conjunto $B = \{0, 1\}$. Para isto construímos a função

$$\phi : B^n \longrightarrow S$$

definida pela regra

$$\phi(b_1, \dots, b_n) = \{i \mid b_i = 1\}.$$

Note que, embora o argumento de ϕ seja a n -upla (b_1, \dots, b_n) , escrevemos

$$\phi(b_1, \dots, b_n) \text{ em vez de } \phi((b_1, \dots, b_n)),$$

para não complicar desnecessariamente a notação. Esta função é sobrejetiva, pois o subconjunto

$$X = \{i_1, \dots, i_r\} \subset S$$

é imagem da n -upla

$$(b_1, \dots, b_n) \text{ para a qual } b_k = \begin{cases} 1 & \text{se } i_k \in X \\ 0 & \text{se } i_k \notin X. \end{cases}$$

Por outro lado, para que duas n -uplas

$$(b_1, \dots, b_n) \text{ e } (b'_1, \dots, b'_n)$$

sejam diferentes é necessário que exista pelo menos um índice $1 \leq k \leq n$ tal que $b_k \neq b'_k$. Mas isto significa que $b_k = 1$ e $b'_k = 0$, ou vice-versa: no primeiro caso

$$k \in \phi(b_1, \dots, b_n) \text{ mas } k \notin \phi(b'_1, \dots, b'_n),$$

ao passo que, no segundo,

$$k \notin \phi(b_1, \dots, b_n) \text{ mas } k \in \phi(b'_1, \dots, b'_n).$$

Como, em ambos os casos,

$$\phi(b_1, \dots, b_n) \neq \phi(b'_1, \dots, b'_n),$$

podemos concluir que ϕ é injetiva. Logo, a quantidade de subconjuntos de S é igual à quantidade de elementos de B^n . Como este último conjunto tem 2^n elementos, o mesmo vale para o conjunto dos subconjuntos de S .

3. Composição de funções

Sejam

$$f : X \longrightarrow Y \text{ e } g : Z \longrightarrow W$$

duas funções. Quando a imagem de f está contida no domínio de g , podemos calcular

$$g(f(x)) \text{ para todo } x \in X.$$

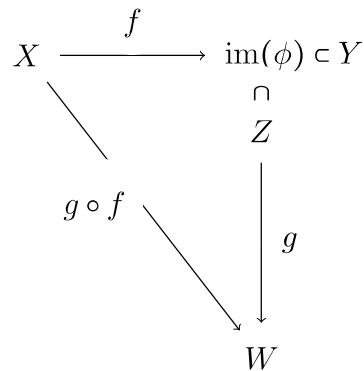
Neste caso a *função composta*

$$g \circ f : X \longrightarrow W$$

é definida pela regra

$$(g \circ f)(x) = g(f(x)) \text{ para todo } x \in X.$$

Podemos ilustrar a definição usando o seguinte diagrama



 Antes de calcular a composta $g \circ f$ é imprescindível verificar se a imagem de f está contida no domínio de g , se isto não ocorrer, a composta não existe. No capítulo 4 veremos que, em muitos casos, é possível redefinir o domínio de f de modo que esta condição seja verificada.

Alguns exemplos ajudarão a entender melhor a definição. Na seção 1, definimos as funções

$$p : \mathcal{L} \longrightarrow \mathbb{N} \text{ e } d : \mathbb{N} \longrightarrow \mathbb{N}$$

que associam, respectivamente, a cada letra maiúscula sua posição no alfabeto e a cada número natural seu dobro. Neste caso,

$$d \circ p : \mathcal{L} \longrightarrow \mathbb{N}$$

é a função definida pela regra

$$(d \circ p)(\ell) = d(p(\ell)) = 2 \cdot p(\ell),$$

para toda letra maiúscula ℓ ; neste modo,

$$(d \circ p)(A) = 2 \cdot p(A) = 2 \text{ e } (d \circ p)(I) = 2 \cdot p(I) = 18.$$

Por outro lado, $p \circ d$ não está definida, porque a imagem de d é o conjunto dos números pares, ao passo que o domínio de p é o conjunto \mathcal{L} das letras maiúsculas. Contudo, como $q : \mathbb{N} \longrightarrow \mathbb{N}$ leva cada número natural em seu quadrado, que também é um número natural, tanto $q \circ d$, quanto $d \circ q$ estão definidas e ambas têm \mathbb{N} como domínio e imagem. As regras que definem estas duas funções são

$$(q \circ d)(n) = q(2n) = 4n^2 \text{ e } (d \circ q)(n) = d(n^2) = 2n^2.$$

Note que isto significa que $q \circ d \neq d \circ q$, porque duas funções são iguais quando têm o mesmo domínio e coincidem para todo valor do seu argumento, o que não ocorre neste caso.

Usando a terminologia introduzida no caso dos conectivos lógicos e das operações com conjuntos, expressamos o fato da composta de funções depender da ordem em que as funções aparecem dizendo simplesmente que não é uma operação *comutativa*. Contudo, a composta é *associativa* sempre que é viável. Para tornar a afirmação mais precisa, sejam

$$f : X \longrightarrow Y, \quad g : Y \longrightarrow Z \text{ e } h : Z \longrightarrow W$$

três funções. Como

$$\text{im}(f) \subset Y = \text{domínio de } g \text{ e } \text{im}(g) \subset Z = \text{domínio de } h$$

as funções $f \circ (g \circ h)$ e $(f \circ g) \circ h$ estão bem-definidas. Além disso, segue da definição de composta que, para todo $x \in X$,

$$(f \circ (g \circ h))(x) = f((g \circ h)(x)) = f(g(h(x)))$$

e que

$$((f \circ g) \circ h)(x) = (f \circ g)(h(x)) = f(g(h(x)))$$

de modo que, pela definição de igualdade de funções, temos

$$f \circ (g \circ h) = (f \circ g) \circ h.$$

O próximo resultado indica como as funções sobrejetivas e injetivas se comportam sob a composição de funções.

TEOREMA 3.1. *Se*

$$f : X \longrightarrow Y \quad e \quad g : Y \longrightarrow Z$$

são funções sobrejetivas, então $g \circ f$ é sobrejetiva e o mesmo vale se substituirmos sobrejetiva por injetiva ou bijetiva.

PROOF. Suponhamos, primeiramente, que ambas as funções sejam sobrejetivas e que $z \in Z$. Como $\text{im}(g) = Z$, existe $y \in Y$ tal que $g(y) = z$. Analogamente, $\text{im}(f) = Y$ garante que existe $x \in X$ tal que $f(x) = y$. Portanto,

$$(g \circ f)(x) = g(f(x)) = g(y) = z.$$

Como isto se aplica a qualquer $z \in Z$, podemos concluir que $\text{im}(g \circ f) = Z$; isto é, que $g \circ f$ é sobrejetiva. Supondo, agora, que f e g são injetivas, digamos que

$$(g \circ f)(x) = (g \circ f)(x'),$$

para alguma escolha de x e x' em X . Como isto equivale a

$$g(f(x)) = g(f(x')),$$

podemos concluir, pela injetividade de g , que

$$f(x) = f(x').$$

Por sua vez, a injetividade de f nos permite deduzir que $x = x'$, mostrando que $g \circ f$ é injetiva. Combinando as duas propriedades que acabamos de provar, podemos concluir que $g \circ f$ é bijetiva quando f e g são bijetivas. \square

Quando a imagem de uma função f está contida em seu domínio X , podemos considerá-la como uma função de X e X . Neste caso a composição pode ser usada para definir potências de f . Há duas maneiras equivalentes de fazer isto. A mais ingênua, consiste em dizer que

$$f^n : X \longrightarrow X,$$

quando $n \in \mathbb{N}$, é dada pela regra

$$f^n(x) = (\underbrace{f \circ \cdots \circ f}_{n \text{ vezes}})(x).$$

Logo,

$$f^1(x) = f(x), \quad f^2(x) = f(f(x)), \quad f^3(x) = f(f(f(x)))$$

e assim por diante. Contudo é preferível utilizar uma definição recursiva,

$$f^n(x) = \begin{cases} f(x) & \text{quando } n = 1 \\ f(f^{n-1}(x)) & \text{quando } n > 1, \end{cases}$$

porque é mais fácil de utilizar em demonstrações. Para calcular $f^3(x)$ usando a definição recursiva, vemos que

$$f^3(x) = f(f^2(x)),$$

ao passo que,

$$f^2(x) = f(f(x)).$$

Substituindo a segunda equação na primeira, obtemos

$$f^3(x) = f(f^2(x)) = f(f(f(x))),$$

que é a mesma resposta encontrada quando utilizamos a primeira definição.

Encerramos considerando o que ocorre quando calculamos potências da função de Collatz, definida na seção 1. Por exemplo,

$$C(3) = 10, \quad C(10) = 5 \quad \text{e} \quad C(5) = 15,$$

onde

$$C^3(3) = 16.$$

Contudo, em vez de parar aqui, continuemos a calcular a imagem de 3 pelas potências consecutivas de C ,

$$C^4(3) = C(16) = 8, \quad C^5(3) = C(8) = 4, \quad C^6(4) = C(4) = 2, \quad \text{e} \quad C^7(2) = 1.$$

Fazendo a mesma coisa com as imagens de 7, obtemos

$C(7) = 22$	$C^2(7) = 11$	$C^3(7) = 34$	$C^4(7) = 17$
$C^5(7) = 52$	$C^6(7) = 26$	$C^7(7) = 13$	$C^8(7) = 40$
$C^9(7) = 20$	$C^{10}(7) = 10$	$C^{11}(7) = 5$	$C^{12}(7) = 16$
$C^{13}(7) = 8$	$C^{14}(7) = 4$	$C^{15}(7) = 2$	$C^{16}(7) = 1,$

e estamos de volta em 1. Todas as potências consecutivas de C , calculadas em um determinado argumento, acabam em 1, ainda que, como nestes dois casos, o expoente da potência que dá 1 dependa do valor escolhido para o argumento. Por exemplo,

$$C^{17}(11) = 1, C^{18}(13) = 1 \quad \text{e} \quad C^{21}(17) = 1.$$

Ninguém sabe ao certo se isto sempre acontece, nem porque deveria ser assim. Segundo o matemático Paul Erdős (1913-1996) “talvez a matemática ainda não esteja madura o suficiente para abordar problemas como este”.

Exercícios

1. Quais das seguintes maneiras de associar um objeto a outro são funções?
 - (a) a cada brasileir(o)a seu cpf;
 - (b) a cada ponto na superfície na Terra sua latitude;
 - (c) a cada tecla um caractere na tela do computador;
 - (d) a cada cartão de crédito VISA um número de 16 algarismos;
 - (e) a cada casa ou prédio de uma rua seu número;

- (f) a cada casa ou prédio da cidade do Rio seu código postal;
 (g) a cada palavra do português sua primeira letra;
 (h) a cada número natural seu consecutivo.
2. Para cada uma das funções encontradas no exercício 1, determine:
- | | |
|------------------------|-----------------------|
| (a) seu domínio; | (d) se é injetiva; |
| (b) seu contradomínio; | (e) se é sobrejetiva; |
| (c) sua imagem; | (f) se é bijetiva. |
3. Determine o domínio, contra-domínio e a imagem de cada uma das funções encontradas no exercício 1.
4. Para cada uma das regras abaixo, determine um subconjunto $D \subset \mathbb{Z}$ sobre o qual a regra define uma função e ache a imagem da função obtida desta maneira.
- | | |
|------------------------------|----------------------------|
| (a) $f_1(x) = 2x;$ | (d) $f_4(x) = \sqrt{x^2};$ |
| (b) $f_2(x) = x - 1 ;$ | (e) $f_5(x) = 2^x;$ |
| (c) $f_3(x) = \sqrt{2 - x};$ | (f) $f_6(x) = 1/x.$ |
- A propósito: a função $f_4(x)$ coincide com a função identidade $\text{id}(x) = x$?
5. Determine a imagem de cada uma das funções de \mathbb{N} em \mathbb{N} definidas abaixo:
- | | |
|-----------------------|---|
| (a) $f_1(x) = x + 1;$ | (d) $f_4(x) = \begin{cases} x - 1 & \text{se } x \text{ é par} \\ x + 1 & \text{se } x \text{ é ímpar.} \end{cases};$ |
| (b) $f_2(x) = 2x;$ | (e) $f_5(x) = x = \begin{cases} x & \text{se } x \geq 0 \\ -x & \text{se } x < 0. \end{cases}$ |
| (c) $f_3(x) = x^2;$ | |
6. Quais das funções do exercício anterior são: sobrejetoras? injetoras? bijetoras?
7. Calcule a $f \circ g$ e $g \circ f$ para cada par de funções de \mathbb{R} em \mathbb{R} descritas abaixo:
- | | |
|--------------------------------------|-----------------------------------|
| (a) $f(x) = 2x$ e $g(x) = 7x;$ | (c) $f(x) = 2x$ e $g(x) = x - 3;$ |
| (b) $f(x) = 2x$ e $g(x) = 9x^2 + 1;$ | (d) $f(x) = x^2$ e $g(x) = -7x.$ |

Part 2

Geometria euclidiana

CHAPTER 2

Comprimentos, áreas e números reais

Neste capítulo investigaremos o método de recorte e colagem que os escribas egípcios provavelmente usaram para chegar às suas regras para cálculos de áreas e veremos como a geometria tornou-se científica na Grécia Antiga. Encerraremos o capítulo enunciando as propriedades dos números reais que serão necessárias no próximo capítulo.

1. Áreas

Toda criança vai-se apropriando, com o passar dos anos, de inúmeras “partículas de informação” que ouve dos adultos. No meu caso, uma destas era a explicação de que *geometria* era o termo grego para “medida da terra”, no sentido de terreno ou região. Os gregos antigos justificavam o uso desta palavra apelando para o que teria acontecido no Antigo Egito. Assim, no início de seu famoso sumário de história da geometria, o filósofo neoplatônico Proclo (410-485 d.C.) nos diz:

como precisamos examinar o início das artes e do conhecimento no presente período, diremos que, de acordo com a maioria dos relatos, os Egípcios foram os primeiros a descobrir a geometria; tendo sua origem na medição de áreas. Isto era necessário, em razão das cheias do Nilo, que, avançando, apagavam os limites [entre os lotes].

A tradição grega de que a geometria seria originária do Egito remonta, pelo menos, à *História* de Heródoto (484–425 a.C.). Segundo Heródoto, um faraó (provavelmente Senusret III, que viveu entre 1878 e 1839 a. C.) dividira as margens do Nilo em lotes quadrados de igual tamanho, que foram distribuídos entre a população. Os impostos eram, então, cobrados com base na produção de cada lote. Quando as cheias regulares do Nilo danificavam um lote (ou, na versão de Proclo, apagavam seus limites) um inspetor era enviado para refazer as medidas e determinar a proporção justa do imposto a ser pago. Tendo relatado esta história, Heródoto acrescenta “talvez esta tenha sido a maneira como a geometria foi inventada, tendo depois passado à Grécia”.

Os papiros egípcios que chegaram até nós corroboram a visão da geometria egípcia como sendo motivada por questões práticas. Por exemplo, o papiro Rhind, contém vários problemas relativos ao cálculo de áreas de terrenos e volumes de recipientes. Entre as áreas



FIGURE 1. Medindo com uma corda

calculadas neste papiro está a do círculo; como veremos adiante, a fórmula egípcia para a área do círculo corresponde a tomar $\pi = 3,16$. Seguindo o desenvolvimento histórico, começaremos nosso estudo de geometria com uma introdução ao cálculo de áreas.

Talvez não tenha sido à toa que, na versão de Heródoto, Senusret III tenha dividido as margens do Nilo em lotes quadrados, porque o quadrado é o protótipo de todas as áreas. Assim, os egípcios mediam áreas usando o *setat*, que é a área de um quadrado cujo lado tem um *khet*, que corresponde a, mais ou menos, 5,2 metros. Todas as outras áreas podiam, então, ser medidas em função da quantidade de *setats* que cabem dentro delas. Por exemplo, em um retângulo cujos lados medem 3 *khets* e 5 *khets*, cabem $3 \cdot 5 = 15$ quadrados iguais a um *setat*.

Naturalmente não há nenhuma razão para que os lados de um terreno meçam uma quantidade inteira de *setats*. Os egípcios contornavam o problema dividindo, sucessivamente, o *setat* por dois; deste modo temos, no papiro Rhind, áreas de meio *setat*, um quarto de *setat* e um oitavo de *setat*. Como veremos, os gregos descobriram que representar uma medida a partir de uma dada unidade leva a questões matemáticas bastante profundas, mas isto estava 1400 anos no futuro dos inspetores da época de Senusret III. Seguindo a tradição egípcia, suporemos que sabemos como medir comprimentos com a aproximação desejada, e vamos nos concentrar em como usar estas medidas para calcular áreas.

Como já vimos, se os lados de um retângulo medem um número inteiro de *khets*, sua área em *setats* é igual ao produto do tamanho dos lados, porque esta é a quantidade de quadrados de um *setat* que cabem neste retângulo. Se as medidas dos lados não são exatas, a área será igual a um certo número de *setats* e sobrarão um ou mais retângulos, com lados menores que um *setat*, cuja área deve ser calculada. Os lados do retângulo que sobrou são, então, medidos em unidades de meio *setat*, e assim por diante. Por exemplo, no retângulo

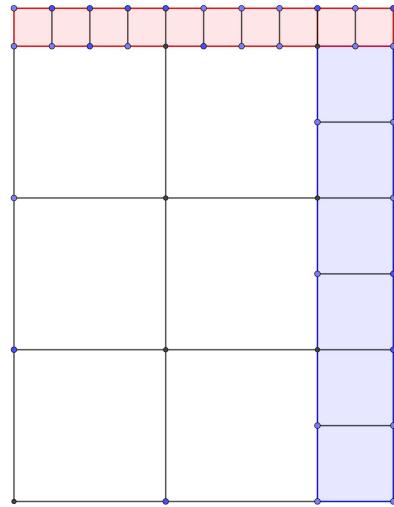


FIGURE 2. Área do retângulo de lados 2 *khets* e meio por 3 *khets* e um quarto.

da figura 2, um dos lado mede 2 *khets* e meio e o outro 3 *khets* e um quarto. Cabem neste retângulo 6 *setats* e a área que sobra pode ser subdividida em dois retângulos, um de meio *khet* por 3 *khets* (em azul na figura), o outro de 2 *khets* e meio por um quarto de *khet* (em vermelho). O primeiro destes retângulos tem área igual a 6 quadrados de um quarto de *setat*, o que corresponde a um *setat* e meio. Já a área do segundo retângulo corresponde a 10 quadrados de um dezesseis avos de *setat*. Portanto, temos um total de

$$6 + 6 \cdot \frac{1}{4} + 10 \cdot \frac{1}{16} = 8\frac{1}{8}$$

ou, como os egípcios provavelmente fariam a conta,

$$\left(3 + \frac{1}{4}\right) \cdot \left(2 + \frac{1}{2}\right) = 8\frac{1}{8}.$$

Portanto, toda esta excursão histórica serve, apenas, para referendar o fato de que *a área de um retângulo é o produto dos comprimentos dos seus lados*.

A partir da área do retângulo é muito fácil achar a área de um paralelogramo. Para isto basta recortar o triângulo à esquerda do paralelogramo, como ilustrado na figura 3, e colá-lo à direita, obtendo assim um retângulo cuja área é igual à do paralelogramo. Observe que um dos lados deste retângulo tem comprimento igual à base do paralelogramo, ao passo que seu outro lado mede o mesmo que a altura do paralelogramo. Dado que as duas figuras têm a mesma área, podemos concluir que *área do paralelogramo é igual ao produto da base pela altura*.

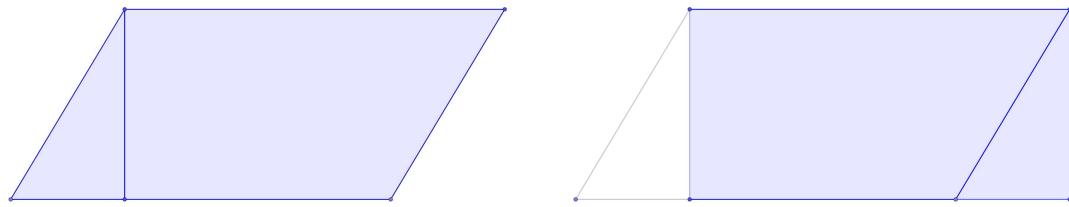


FIGURE 3. Transformando o paralelogramo em um retângulo.

Nosso próximo passo é calcular a área de um triângulo. Desta vez juntamos ao triângulo dado um triângulo igual de modo a formar um paralelogramo. O processo é ilustrado na figura 4, na qual, ao triângulo azul, que foi dado, juntamos o triângulo vermelho para formar um paralelogramo. Portanto, a área do triângulo azul é igual à metade da área do paralelogramo. Como as duas figuras têm mesma base e altura, podemos concluir que *a área de um triângulo é a metade do produto da base pela altura*.

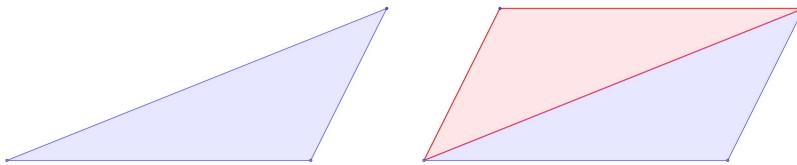


FIGURE 4. Juntando dois triângulos iguais para formar um paralelogramo.

O que fizemos até aqui já era bem conhecido dos antigos egípcios e lhes permitia calcular a área de qualquer figura plana cujos lados fossem segmentos de retas, já que podem ser subdivididas em triângulos. Contudo, os egípcios também sabiam calcular, aproximadamente, a área de um círculo. A regra que utilizavam para isto é dada no papiro Rhind e pode ser formulada da seguinte maneira: *tome o diâmetro do círculo, subtraia sua nona parte e calcule o quadrado do número resultante*. Assim, se o raio do círculo for r , a regra consiste em calcular

$$(2) \quad \left(2r - \frac{2r}{9}\right)^2 = \frac{256}{81}r^2.$$

Como observamos anteriormente isto corresponde a tomar

$$\frac{256}{81} \approx 3,16$$

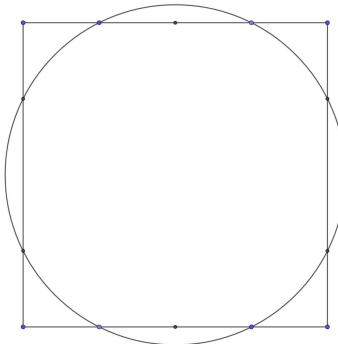


FIGURE 5. Aproximando a área do círculo no Antigo Egito

como aproximação para π . De que forma os escribas do antigo Egito chegaram a esta aproximação é uma pergunta que não sabemos responder. Mas isto não nos impede de especular sobre o que pode ter acontecido. Uma ideia interessante a este respeito é que podem ter chegado a esta aproximação inscrevendo um círculo em um quadrado cujos lados estão divididos em quatro partes iguais, como ilustrado na figura 5. Na próxima seção veremos como passar desta figura à fórmula (2).

2. A geometria chega à Grécia

Segundo o *Sumário* de Proclo, cujo início citamos na seção anterior,

Tales foi o primeiro a ir ao Egito e trazer para a Grécia estes estudos [isto é, a geometria]; encontrando seus próprios resultados, e revelando os princípios àqueles que vieram depois dele.

Em sua *Vida dos Filósofos Ilustres*, Diógenes Laércio (que viveu por volta do século III) é mais explícito quanto ao que Tales teria feito; segundo ele

ninguém o instruíu mas, tendo ido ao Egito, passou um tempo próximo aos sacerdotes de lá. Hieronimus nos diz que mediu as pirâmides a partir de sua sombras, observando quando nossas sombras são iguais a nós mesmos.

De nosso ponto de vista, o que Tales fez foi usar a semelhança entre os triângulos retângulos ABC e ADE ilustrados na figura 6. Entretanto, àquela altura, é quase certo que esta propriedade não fosse conhecida de ninguém. Infelizmente, mais uma vez, nada no registro histórico nos permite determinar como Tales teria chegado à conclusão sobre a igualdade das sombras. O melhor que podemos fazer é tentar criar uma hipótese plausível

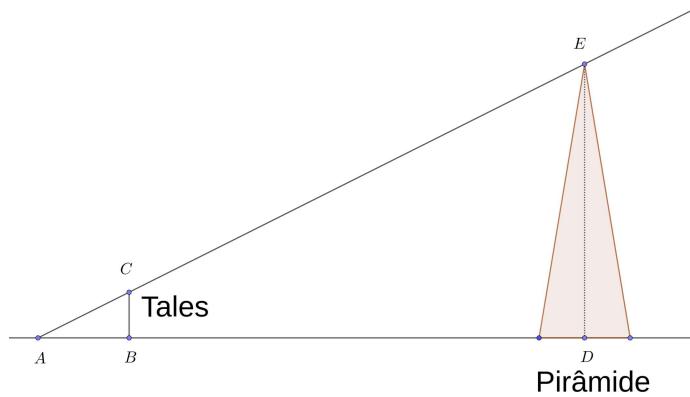


FIGURE 6. Tales e a pirâmide

que, ainda que não seja historicamente justificável, nos ajuda a entender como ele pode ter tido a ideia certa.

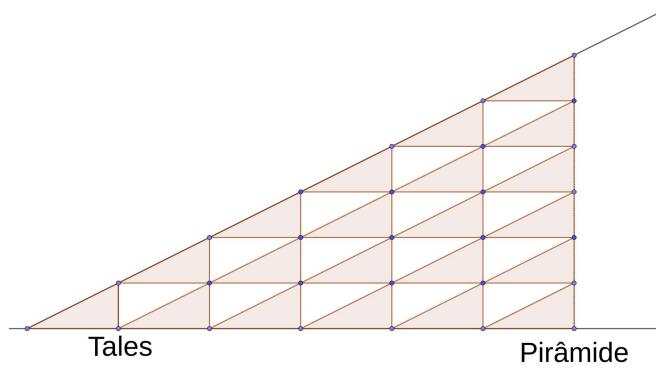


FIGURE 7. Tales e a pirâmide

Uma possibilidade é que Tales tenha usado uma argumentação baseada nas áreas dos dois triângulos pois, como já vimos, os egípcios já sabiam como medir estas áreas. Segundo esta interpretação, Tales teria preenchido o triângulo cuja altura é igual à da pirâmide com várias cópias do triângulo cuja altura coincide com a do próprio Tales, como ilustrado na figura 7. Tendo feito isto, teria observado que a altura da pirâmide, medida em termos da altura de Tales, é igual ao comprimento da sombra da pirâmide, medido em termos da sombra de Tales. Escolhendo o momento em que sua sombra tem o mesmo comprimento de sua altura, este argumento permitiria a Tales concluir que o mesmo tem que valer para a sombra e a altura da pirâmide. Embora esta ideia de como Tales teria procedido seja

bastante atraente, é importante insistir que não há nenhum documento histórico que nos permita determinar se foi mesmo isto que aconteceu.

Voltando ao sumário de Proclo (e deixando de lado “Ameristos, irmão do poeta Estesicorus”, sobre o qual nada sabemos), Pitágoras é o personagem seguinte a ser citado. Segundo Proclo,

Pitágoras transformou o estudo da matemática em uma forma de educação liberal, examinando seus princípios desde o começo, explorando seus teoremas de maneira puramente imaterial e intelectual.

A análise dos princípios da matemática, que Proclo atribui a Pitágoras, é parte das lendas que envolvem a Escola Pitagórica. O fato dos seguidores de Pitágoras atribuírem ao mestre muitas de suas descobertas torna muito difícil determinar o que Pitágoras realmente teria feito. Além disso, é necessário levar em conta que a filosofia neoplatônica, da qual Proclo era um expoente, dava forte ênfase ao imaterial sobre o material, a tal ponto que a biografia de Plotino, o fundador da Escola Neoplatônica, começa com a frase “Plotino, o filósofo nosso contemporâneo, parecia envergonhado de estar em um corpo”.



FIGURE 8. Tabuleta YBC 7289

Em meio a tantas dúvidas, há uma tabuleta de argila, provavelmente escrita por um estudante cerca de mil anos antes de Pitágoras, que sugere que o teorema atribuído a ele talvez já fosse conhecido na Mesopotâmia daquela época. Trata-se de YBC 7289, cuja foto é reproduzida na figura 8. As marcações, em escrita cuneiforme, ao longo do lado e da diagonal do quadrado são números no sistema sexagesimal (base 60) usado na Babilônia. Segundo uma das interpretações mais recentes [6], estes números corresponderiam a $1/2$, ao longo do lado, e a

$$\frac{30547}{43200} \approx 0.7071064814$$

ao longo da diagonal. Este último número é uma excelente aproximação para

$$\frac{1}{\sqrt{2}} \approx 0.7071067811$$

que é o comprimento da diagonal de um quadrado cujo lado é igual a $1/2$.

3. O teorema de Pitágoras e o círculo

Tenham sido os escribas babilônios, Pitágoras, ou outra pessoa, quem descobriu o teorema, é difícil não se perguntar o que levou alguém a perceber a relação entre os quadrados dos catetos e da hipotenusa de um triângulo retângulo. E isto nos leva de volta ao cálculo de áreas. O quadrado de um número é a área de um quadrado. Lembrando disto, o que alguém, em algum momento, percebeu é que o quadrado cujo lado é a hipotenusa tem área igual à soma das áreas dos quadrados sobre os catetos, como ilustrado na figura 9.

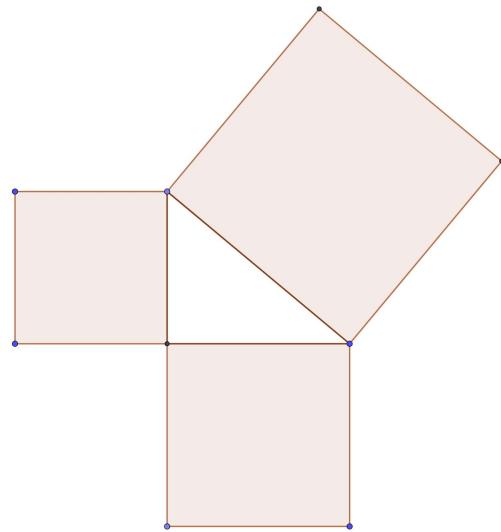


FIGURE 9. O teorema de Pitágoras

Apesar de bem-conhecida esta figura não nos ajuda em nada a entender como alguém teria percebido a relação entre os três quadrados. Para tentar entender o que *pode* ter acontecido, lembre-se que o piso das casas gregas e romanas eram frequentemente recobertos por ladrilhos. Imagine, agora, um grego com inclinação para a matemática que, entediado pela conversa de seus companheiros em um banquete, começa a olhar longamente para os ladrilhos da casa, ilustrados na figura 10.

Talvez você já tenha tido a experiência de olhar longamente para uma figura, de modo que, sem que espere, um detalhe lhe salta aos olhos, ao passo que outros parecem desaparecer. É isso que vamos imaginar ter acontecido a nosso matemático imaginário. Para entender melhor o que teria acontecido, redesenhamos o ladrilhamento na figura 11, mantendo apenas aquilo o que nosso matemático teria visto em seu momento de epifania.

O ponto crucial é que cada um dos quadrados azuis é formado por dois triângulos. Por sua vez, estes quatro triângulos podem ser facilmente rearrumados para formar o quadrado de lados pontilhados sobre a hipotenusa. Reproduzimos assim a figura 9 em um caso muito especial, porque os catetos do triângulo retângulo que estamos considerando são iguais. A

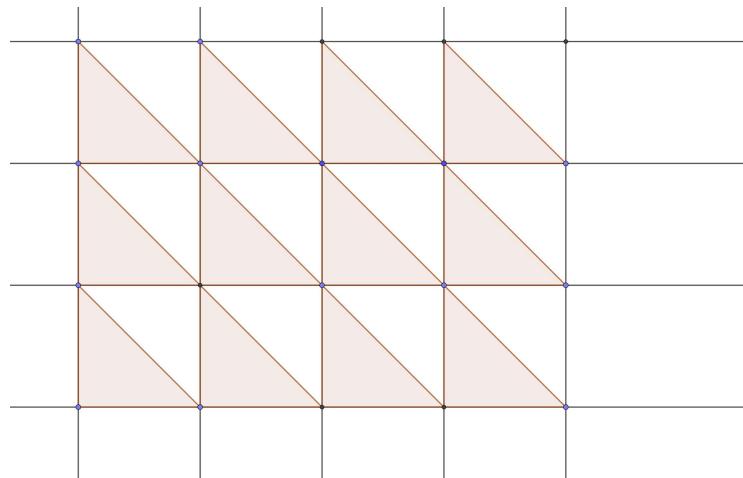


FIGURE 10. O ladrilhamento na casa do banquete

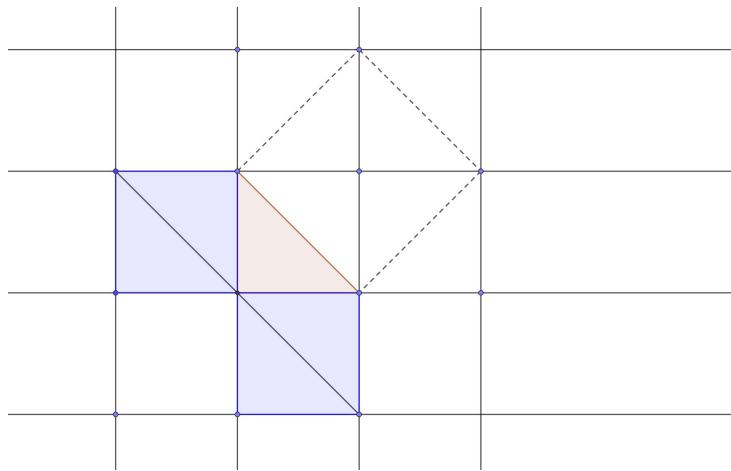


FIGURE 11. O que o matemático viu

beleza deste exemplo é que, para constatar a validade do teorema basta fazer um recorte e colagem do tipo que fizemos na seção 1 e que, como vimos, eram conhecidos dos Egípcios muito antes de Pitágoras. Uma vez que um caso especial tenha sido descoberto é difícil para alguém com propensões matemáticas não se perguntar: dá para generalizar? Um ponto a favor de nosso cenário imaginário, é que a demonstração do caso geral do teorema nos *Elementos* de Euclides é feito por um complicado processo de recorte e colagem. Embora não haja nenhum documento que nos leve a crer que esta história tenha um fundo de verdade, ela nos ajuda a entender como alguém pode ter sido levado a perceber que existe uma relação entre as áreas dos três quadrados.

Para encerrar, vejamos como utilizar o teorema de Pitágoras para encontrar a aproximação do valor de π implicitamente usada nos documentos egípcios. A propósito: embora π seja um letra do alfabeto grego, o primeiro a usá-la para representar o quociente entre o comprimento e o diâmetro da circunferência foi o matemático galês William Jones em 1706.

Na figura 12 construímos um quadrado de lado igual a 8 e dividimos cada um de seus lados em quatro. Em seguida traçamos uma circunferência cujo centro coincide com o do quadrado e que passa pelos pontos que distam duas unidades dos vértices do quadrado. O triângulo de vértices A , B e C é retângulo e seus catetos têm comprimentos iguais a 8 e 4. Portanto, pelo teorema de Pitágoras, a hipotenusa vale $4\sqrt{5}$. Como o diâmetro do círculo coincide com a hipotenusa do triângulo ABC , a área do círculo é igual a

$$\pi \cdot (2\sqrt{5})^2 = \pi \cdot 20.$$

Usando a área do quadrado de lado 8 como aproximação da área do círculo, obtemos

$$\pi \approx \frac{8^2}{20} = \frac{64}{20} = 3,2$$

o que parece contradizer a afirmação feita no final da seção 1 de que a fórmula egípcia dá uma aproximação de π igual a 3,16. A justificativa para a discrepância é que, não tendo conhecimento do teorema de Pitágoras, os egípcios podem ter simplesmente medido o comprimento da hipotenusa em uma figura. Como

$$4\sqrt{5} \approx 8,94$$

o valor que eles encontraram para o comprimento foi, provavelmente, 9. Mas, tomando 9, em vez de $4\sqrt{5}$ como valor do diâmetro, obtemos

$$\pi \approx \frac{8^2}{(9/2)^2} = \frac{16^2}{9^2} \approx 3,1604,$$

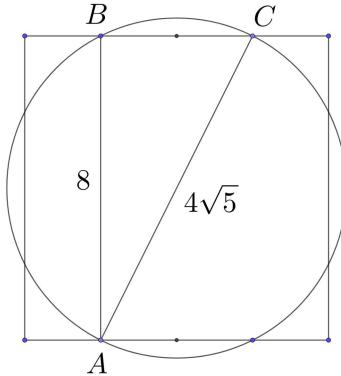


FIGURE 12. Usando o teorema de Pitágoras para aproximar π

de modo que o erro na medida da hipotenusa produz uma aproximação melhor do que o valor exato! Naturalmente temos, mais uma vez, uma história interessante, que é impossível comprovar a partir dos poucos documentos relativos à matemática egípcia que sobreviveram até nossos dias.

4. A geometria na Grécia Antiga

Ainda que não possamos acompanhar Proclo, quando atribui a Pitágoras a transformação que se operou na geometria depois de chegar à Grécia, o fato é que as mudanças que ele descreve de fato ocorreram. Nos 300 anos que separam Tales de Euclides, os gregos transformaram um conjunto de regras práticas para medir terrenos em uma disciplina teórica altamente sofisticada. O que exatamente levou a esta transformação é mais um dos enigmas com que nos deparamos ao contar esta história. Contudo é possível identificar alguns fatores que provavelmente influenciaram essas mudanças.

Um destes fatores foi, certamente, o sistema político vigente em muitas das cidades gregas. A democracia ateniense, por exemplo, permitia que os cidadãos livres (que excluíam mulheres e escravos) debatessem temas políticos importantes na vida da cidade. Esses debates, por sua vez, estimularam a análise crítica dos argumentos propostos por pessoas com interesses diferentes, o aperfeiçoamento da arte da argumentação e a habilidade em justificar, cuidadosamente, a posição pessoal dos participantes.

Estes mesmos fatores levaram ao desenvolvimento paralelo da filosofia. Nos diálogos de Platão, Sócrates, que está sempre em busca do conhecimento, que declara não possuir, interroga seus interlocutores cuidadosamente, expondo toda espécie de falhas e defeitos em argumentos que outros julgavam perfeitamente satisfatórios. Não é a toa que, na *Apologia*, Sócrates descreva a si mesmo como desempenhando, em relação a Atenas, papel semelhante ao da espora que não permite que o cavalo se acomode—a proverbial pedra no sapato.

Essa insistência em tudo contestar teve consequências importantes na filosofia e na matemática. Aristóteles, por exemplo, observa no Livro I dos *Analíticos Posteriores* que, segundo alguns filósofos, é impossível realmente conhecer alguma coisa:

eles afirmam que há uma regressão infinita, porque não é possível conhecer verdades posteriores àquelas que as precedem logicamente, a não ser que estas últimas, por sua vez, dependam de verdades que lhes são anteriores; no que estão certos porque isto requereria percorrer uma série infinita.

Aristóteles, contudo, defende um ponto de vista mais razoável. Embora afirme que

a pessoa que não sabe dar a razão para um fato para o qual existe uma demonstração não possui conhecimento científico,

ele observa que

nem todo conhecimento é obtido através de demonstrações; pois o conhecimento imediato não é demonstrável.

Em outras palavras, embora o que torna seguro nosso conhecimento seja nossa capacidade de demonstrá-lo, isso só é viável se aceitarmos certas afirmações fundamentais como verdadeiras, ainda que não sejam passíveis de demonstração.

As consequências da argumentatividade grega não foram menores na matemática. Na verdade, matemática e filosofia experimentaram um desenvolvimento paralelo e muitos dos exemplos que Aristóteles cita na *Metafísica* e nos *Analíticos posteriores* são de resultados de matemática. Segundo Proclo, Hipócrates de Quios, que não deve ser confundido com o médico Hipócrates de Cós, foi “o primeiro matemático a compilar *elementos*”. O que seus contemporâneos entendiam por esta última palavra é explicado por Aristóteles no capítulo III da *Metafísica*; segundo ele

no caso das proposições geométricas, chamamos de elementos aquelas cujas demonstrações estão presentes nas demonstrações de todas, ou da maioria, das demais.

Como veremos a seguir, o mais famoso livro da matemática grega, ao qual ainda hoje nos referimos como *Os Elementos*, além de respeitar a definição desta palavra dada na *Metafísica*, organiza-se como uma sequência de fatos sistematicamente demonstrados a partir de afirmações cuja veracidade se supõe óbvia. Isto faz justiça à etimologia da palavra grega correspondente a *elementos*, que é derivada de um verbo que significa organizar em uma fila.

5. Os elementos de Euclides

Aristóteles viveu em uma época extremamente conturbada, em que Atenas, derrotada por Filipe II da Macedônia em 338 a.C., começou a perder sua posição como centro por excelência da cultura grega. No século seguinte, esta posição passaria a Alexandria, fundada na foz do Nilo, em 331 a.C., por Alexandre, o filho e herdeiro de Filipe II, do qual Aristóteles havia sido tutor. Com a morte prematura de Alexandre em 323 a.C., o Egito passou ao controle de Ptolomeu, que havia sido um de seus generais, e que estabeleceu sua capital em Alexandria.

Alexandria cresceu rapidamente e logo começou a atrair filósofos, escritores e matemáticos, sobretudo depois da fundação do *Museion* e da famosa biblioteca, que funcionavam em

parte como um instituto de pesquisas moderno. Uma das pessoas a trabalhar em Alexandria foi o matemático Euclides. Em seu livro *Alexandria: a history and a guide*, o escritor inglês E. M. Forster escreve que, em Alexandria

[a] matemática começa com a incrível, mas obscura, carreira de Euclides.
Nada se sabe sobre Euclides: de fato, pensamos nele hoje em dia mais
como sendo um ramo do conhecimento do que um homem.

Publicado em 1922, o guia de Forster foi escrito em uma época em que os *Elementos*, a principal obra de Euclides, ainda era o livro-texto básico de geometria em quase todas as escolas do Reino Unido. Daí a frase de Forster, para quem Euclides era sinônimo de geometria. Apesar disso, os *Elementos* não são a única obra de Euclides que chegou até nós. Entre os outros livros dele dos quais temos cópias estão os *Dados*, a *Óptica* e a *Divisão de Figuras*, este último apenas em uma tradução árabe.

Os *Elementos* são um tratado de geometria básica em 13 livros, que, provavelmente, correspondiam aos rolos de papiro em que a obra foi originalmente escrita. Estes livros podem ser agrupados em três grandes blocos, o primeiro dos quais, formado pelos livros I a VI trata de geometria plana. O segundo bloco, cujo tema é a teoria dos números, é formado pelos livros de VII a IX. Já o terceiro bloco, constituído pelos livros de XI a XIII, diz respeito à geometria espacial e inclui a classificação dos poliedros regulares. Não se encaixam nesta subdivisão o livro V, sobre teoria da proporção, e o livro X, sobre a classificação de grandezas incomensuráveis. Teremos um pouco mais a dizer sobre o livro V ao final desta seção.

Como almejamos analisar alguns resultados básicos de geometria plana do ponto de vista dos *Elementos* é conveniente listar os tópicos correspondentes a cada um dos livros que tratam deste tema:

- Livro I:** geometria dos triângulos;
- Livro II:** cálculo com magnitudes;
- Livro III:** geometria do círculo;
- Livro IV:** figuras inscritas e circunscritas no círculo;
- Livro VI:** similaridade de figuras

Contudo, o cálculo da área do círculo só aparece no Livro XII, cujo tema é o volume de pirâmides, cilindros, cones e esferas.

De nosso ponto de vista, o aspecto mais importante dos elementos é a maneira como o material é apresentado. Fazendo eco ao que Aristóteles escrevera nos *Analíticos Posteriores*, Euclides começa o livro definindo os termos técnicos que precisará usar e enunciando o que chama de *postulados* e *noções comuns*. Os primeiros determinam como os objetos que haviam sido definidos se relacionam entre si, ao passo que os segundos enumeram oito

propriedades de caráter mais geral como *coisas iguais a uma mesma coisa são iguais entre si e o todo é maior que a parte*. Como estas últimas são, basicamente, propriedades dos números reais, não será necessário listá-las. As definições e postulados, contudo, requerem uma análise mais cuidadosa.

As definições em Euclides podem ser enquadradas em duas categorias distintas. A primeira, e mais importante, é formada por aquelas que especificam uma figura plana a partir de uma certa propriedade; por exemplo,

um círculo é uma figura plana, compreendida por uma linha, tal que todas as retas que caem sobre ela a partir de um ponto dentro da figura são iguais entre si.

Já as definições na segunda categoria são meros apelos à intuição; por exemplo, *um ponto é o que não tem partes ou uma linha é comprimento sem largura*. Mesmo as definições euclidianas na primeira categoria não são suficientemente precisas do ponto de vista atual; assim, a definição de círculo fala em um ponto dentro da figura, mas “dentro” não aparece em nenhuma das definições que antecedem à do círculo.

Ao contrário das definições, das quais há 23 só no livro I, os postulados são apenas cinco. Os três primeiros afirmam a possibilidade de:

1. desenhar uma reta entre dois pontos quaisquer;
2. alongar um segmento ao longo de uma reta;
3. descrever um círculo com centro e raio dados.

O de número 4 declara a igualdade entre dois ângulos retos quaisquer. Já o quinto postulado, o mais famoso deles, afirma, em uma versão modernizada, que

se uma reta ℓ , ao incidir sobre duas outras, forma ângulos interiores, de um mesmo lado de ℓ , que são menores que dois retos, então estas retas se cortam em um ponto que está do mesmo lado de ℓ em que estão estes ângulos.

Uma figura ajuda a entender melhor o conteúdo deste postulado. Considere uma reta ℓ que, como na figura 13, corta duas retas r_1 e r_2 e sejam α e β os ângulos formados, de um mesmo lado, entre ℓ e r_1 e entre ℓ e r_2 . O postulado afirma que, se $\alpha + \beta$ é menor do que 180° , então r_1 intersecta r_2 do mesmo lado em que estão α e β . Desde a antiguidade clássica este postulado gerou uma enorme polêmica porque é muito menos óbvio que os demais. Séculos de tentativas de deduzi-lo dos outros quatro chegaram ao fim no século XIX, quando N. I. Lobachevsky e J. Bolyai mostraram que é possível construir uma modelo de geometria que respeita os quatro primeiros postulados, mas não o quinto, sem incidir em

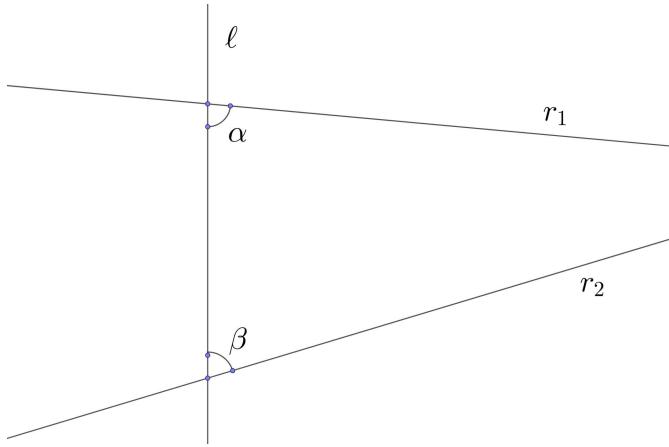


FIGURE 13. Quinto postulado dos *Elementos*

nenhuma contradição, o que nos permite concluir que o quinto postulado é independente dos demais.

Ao contrário de Euclides, não nos referiremos aos princípios básicos da geometria como *postulados*, mas sim como *axiomas*, que se tornou a terminologia padrão a partir dos matemáticos gregos posteriores a Euclides. Este termo é derivado da palavra grega para *digno*, no sentido de que estas hipóteses são dignas de serem tomadas como base de nossa construção.

Um aspecto interessante da exposição de Euclides nos *Elementos* é a ausência de números associados a medições. Em vez de comparar os comprimentos de dois segmentos, os matemáticos gregos comparavam, diretamente, os segmentos entre si. Para explicar como faziam isto, precisamos introduzir algumas das operações que os gregos aplicavam aos segmentos. A primeira é a adição de segmentos. Para somar os segmentos AB e CD , eles executariam as seguintes etapas:

Passo 1: com a agulha do compasso em C , estenderiam a outra ponta até D ;

Passo 2: com o compasso aberto desta maneira, poriam a agulha em B e marcariam, à direita de B , o ponto D' sobre a reta ℓ que contém AB .

Com isso, o segmento AD' é a soma de AB com CD . A segunda operação consiste em comparar segmentos. Para isso, executariam a passo 1 anterior, pondo, em seguida a agulha do compasso em A , com a outra ponta marcando, na reta ℓ , o ponto D' à direita de A . Se $D' \in AB$, então $CD \leq AB$, caso contrário, $AB < CD$.

Com isto podemos explicar a definição de razão entre segmentos, definida por Eudoxo de Cnido (400-350 a.C.), e exposta por Euclides no livro V dos *Elementos*. Suponhamos

pares são iguais a números ímpares. Portanto, argumenta-se até chegar à conclusão de que ímpar é igual a par e prova-se, por hipótese, que a diagonal é incomensurável, porque a proposição contrária produz um resultado falso (*Analíticos anteriores* 41a 23–30).

A afirmação de Aristóteles pressupõe que o leitor está familiarizado com esta demonstração da incomensurabilidade do lado e da diagonal do quadrado, mas um apêndice ao Livro X dos *Elementos* de Euclides dá mais detalhes sobre como o argumento funciona. A demonstração que apresentamos a seguir é uma versão modernizada da que aparece nos *Elementos*.

Começamos traduzindo a noção de incomensurabilidade em termos numéricos. Por definição, os comprimentos, c_1 e c_2 , de dois segmentos de reta são *comensuráveis* se existe uma unidade de medida u e números naturais m_1 e m_2 , tais que

$$c_1 = m_1 u \quad \text{e} \quad c_2 = m_2 u;$$

que equivale a dizer que

$$\frac{c_1}{c_2} = \frac{m_1}{m_2}$$

é uma fração. Portanto,

dizer que a diagonal d e o lado ℓ de um quadrado são incomensuráveis equivale a afirmar que d/ℓ não é uma fração.

Contudo, a versão para triângulos isósceles do teorema de Pitágoras, discutida na seção 3, nos permite concluir que, se um quadrado tem lado de comprimento ℓ , então sua diagonal tem comprimento $d = \ell\sqrt{2}$. Logo, a razão entre estas medidas é igual a

$$\frac{d}{\ell} = \frac{\ell\sqrt{2}}{\ell} = \sqrt{2}.$$

Assim, provar a incomensurabilidade da diagonal e do lado do quadrado equivale a mostrar que $\sqrt{2}$ não é uma fração.

O argumento mencionado no texto de Aristóteles é conhecido como *redução ao absurdo* ou *demonstração por contradição*. Neste tipo de demonstração supomos o contrário do que esperamos provar e argumentamos de modo a chegar a uma afirmação contraditória; no caso citado por Aristóteles, esta afirmação consiste no fato de que um número par seria igual a um número ímpar. Naturalmente, este argumento depende do princípio do meio excluído, segundo o qual, qualquer que seja a proposição P , vale que $\neg P \vee P = \top$.

Passando aos detalhes, suponhamos que

$$(3) \quad \sqrt{2} = \frac{m_1}{m_2}$$

é uma fração. Simplificando os fatores comuns entre o numerador e o denominador desta fração, podemos supor, também, que m_1 e m_2 não são simultaneamente pares. Elevando ao quadrado ambos os lados de (3) e simplificando o denominador, obtemos

$$(4) \quad 2m_2^2 = m_1^2.$$

Em particular, isto mostra que m_1^2 tem que ser um número par. Como o produto de dois números ímpares é um número ímpar, podemos concluir que m_1 tem que ser par, o que nos permite escrever

$$m_1 = 2k,$$

para algum $k \in \mathbb{N}$. Substituindo esta última igualdade em (4) e cancelando os termos comuns, encontramos

$$m_2^2 = 2k^2,$$

de modo que, pelo argumento usado anteriormente, m_2 também tem que ser par. Contudo, tínhamos visto que sempre é possível supor que m_1 ou m_2 é ímpar, o que contradiz a conclusão a que acabamos de chegar. Como a hipótese de que $\sqrt{2}$ é uma fração leva a uma contradição, podemos deduzir que isto é falso e que $\sqrt{2}$ é um número irracional. A propósito, neste contexto, irracional não significa que o número não é compreensível pela razão humana ou, de alguma maneira, ilógico, mas sim que não é igual à razão (ou fração) entre dois números inteiros. Desde o trabalho de Georg Cantor (1845–1918) sabemos que, em um sentido que é possível tornar bastante preciso, quase todos os números reais são irracionais. Apesar disto, há muitos números reais cuja irracionalidade é bastante difícil de provar, como π e a base e dos logaritmos naturais; além de haver outros, como $e + \pi$, $e \cdot \pi$, 2^e ou $\pi^{\sqrt{2}}$ que ainda não sabemos se são ou não racionais. Mais detalhes podem ser encontrados em ????

Já vimos que, trabalhando diretamente com segmentos de retas, os matemáticos gregos precisaram apelar para uma maneira bastante complicada de definir a igualdade entre razões de segmentos. A introdução do sistema decimal na Idade Média reduziu significativamente o contraste entre os números racionais e os irracionais, já que mesmo uma fração simples como $1/3$ é representada por uma expansão decimal infinita. Simon Stévin (1548–1620), por sua vez, foi um dos primeiros a afirmar, categoricamente, que não há diferença básica entre os “números geométricos” usados para medir segmentos, e os “números aritméticos”, usados em contagem; segundo ele,

obtemos, portanto, a conclusão de que não há números absurdos, irrationais, irregulares, inexplicáveis ou surdos, mas que há neles, ao contrário, tal excelência e concordância que deveríamos meditar, dia e noite, em sua admirável perfeição; [1, p.209].

Daí em diante, os números reais foram utilizados sistematicamente pelos matemáticos, ainda que de maneira informal. Contudo, os problemas que surgiram à medida que os matemáticos exploravam a fundo o cálculo infinitesimal, inventado no século XVII por



FIGURE 14. Richard Dedekind

Newton e Leibniz, forçaram-nos a buscar uma definição realmente precisa do que é um número real.

Finalmente, em 1872, Richard Dedekind, inspirado, como ele mesmo admite, pela complicada definição de igualdade de razões utilizada pelos matemáticos gregos, construiu uma das primeiras definições satisfatória do conjunto dos números reais. Partindo dos números inteiros, Dedekind construiu os números racionais e, a partir destes, todos os reais. Tanto sua construção, como outras a ela equivalentes, são bastante sofisticadas e fora do alcance de um curso elementar. Por isso vamos nos contentar apenas em enunciar as propriedades dos números reais que serão necessárias nos próximos capítulos. Começamos com as propriedades algébricas básicas das operações de adição e multiplicação de números reais. Além da soma e multiplicação de números reais ser um número real, temos que, quaisquer que sejam $a, b, c \in \mathbb{R}$

associativa	$a + (b + c) = (a + b) + c$	$a \cdot (b \cdot c) = (a \cdot b) \cdot c$
comutativa	$a + b = b + a$	$a \cdot b = b \cdot a$
elemento neutro	$a + 0 = a$	$a \cdot 1 = a$
distributiva	$a(b + c) = ac + bc$	

Escrevemos as propriedades em colunas para chamar sua atenção de que cada uma das três primeiras propriedades têm versões equivalentes para a adição e a multiplicação. Já a distributividade nos permite relacionar somas a produtos. A última das propriedades básicas afirma que cada elemento $a \in \mathbb{R}$ admite um *simétrico* $-a$ que satisfaz

$$a + (-a) = 0$$

e, se for não nulo, também um *inverso* a^{-1} tal que

$$a \cdot a^{-1} = 1.$$

Destas propriedades decorre que o simétrico de $-a$ é a e que, quando $a \neq 0$, o inverso de a^{-1} também é a .

Antes de passar às demais propriedades, vejamos algumas consequências elementares das que enunciarmos até aqui. A primeira delas diz respeito ao produto de um número por 0. Como $0 + 0 = 0$,

$$a \cdot 0 = a \cdot (0 + 0) = a \cdot 0 + a \cdot 0.$$

Somando o simétrico de $a \cdot 0$ aos dois lados desta equação, obtemos

$$(5) \quad 0 = a \cdot 0.$$

Por sua vez, como $1 + (-1) = 0$, (5) nos dá

$$0 = a \cdot (1 + (-1)) = a \cdot 1 + a \cdot (-1) = a + a \cdot (-1).$$

Somando $-a$ aos dois lados desta equação, concluímos que

$$(6) \quad -a = a \cdot (-1).$$

Em particular, quando $a = -1$, obtemos

$$-(-1) = (-1) \cdot (-1).$$

Como o simétrico de (-1) é 1 , isto nos permite concluir que

$$(7) \quad 1 = (-1) \cdot (-1).$$

As demais propriedades básicas dos números reais de que precisamos dizem respeito a desigualdades. Sejam a e b números reais,

tricotomia $a > 0$, ou $a = 0$, ou $-a > 0$;

fechamento se $a > 0$ e $b > 0$, então $a + b > 0$ e $ab > 0$.

Como consequência da tricotomia, se $a, b \in \mathbb{R}$ e $-a > 0$ e , então

$$(8) \quad a - b > 0, \text{ ou } a - b = 0, \text{ ou } -(a - b) > 0.$$

Mas, combinando (6) à propriedade distributiva,

$$-(a - b) = -(a + (-1) \cdot b) = (-1) \cdot (a + (-1) \cdot b) = (-1) \cdot a + (-1) \cdot (-1) \cdot b.$$

Logo, por (7) e (6),

$$-(a + (-1) \cdot b) = (-1) \cdot a + b = -a + b.$$

Portanto, (8) pode ser reformulada como

$$a - b > 0, \text{ ou } a - b = 0, \text{ ou } -a + b > 0.$$

Definindo $a > b$ como uma maneira abreviada de escrever $a - b > 0$, temos que

$$(9) \quad a > b, \text{ ou } a = b, \text{ ou } b > a.$$

Por outro lado, se $a, b, c \in \mathbb{R}$, então por (6),

$$(a+c) - (b+c) = (a+c) + (-b-c) = (a-b),$$

e podemos concluir que

$$a > b \longrightarrow (a+c) > (b+c).$$

Para deduzir a afirmação análoga relativa à multiplicação, suponhamos que $a > b$ e que $c > 0$. Neste caso, o fechamento nos dá

$$(a-b)c > 0;$$

onde, por (6) e pela distributiva,

$$ac - bc > 0,$$

que equivale a dizer que

$$ac > bc.$$

Denotando, como de costume, $-c > 0$ por $c < 0$, temos neste caso que

$$a(-c) > b(-c).$$

Logo, por (6),

$$-ac > -bc;$$

de modo que

$$ac - bc < 0.$$

Resumindo:

$$a > b \longrightarrow \begin{cases} ac > bc & \text{quando } c > 0 \\ ac < bc & \text{quando } c < 0. \end{cases}$$

Em particular,

$$a < 0 \text{ e } b < 0 \longrightarrow ab > 0.$$

De agora em diante usaremos todas estas propriedades, sistematicamente, sem mencioná-las pelo nome, a não ser que isto seja necessário para tornar o argumento claro. Há, porém, uma última desigualdade que é suficientemente importante para que seu nome seja sempre mencionado quando ela for utilizada. Trata-se da *desigualdade triangular*, segundo a qual

$$(10) \quad |x+y| \leq |x| + |y|,$$

quaisquer que sejam os números reais x e y . Para prová-la, basta lembrar que

$$(x+y)^2 = x^2 + y^2 + 2xy$$

nos permite afirmar que

$$(x+y)^2 \leq x^2 + y^2 + 2|x| \cdot |y|,$$

já que todas as parcelas do lado direito desta última equação são maiores ou iguais a zero. Como

$$x^2 + y^2 + 2|x| \cdot |y| = (|x| + |y|)^2,$$

a desigualdade anterior pode ser escrita na forma

$$|x+y|^2 \leq (|x| + |y|)^2.$$

Logo,

$$0 \leq (|x| + |y|)^2 \cdot |x + y|^2 = (|x| + |y| - |x + y|)(|x| + |y| + |x + y|);$$

que só pode ocorrer quando

$$|x| + |y| - |x + y| \text{ e } |x| + |y| + |x + y|$$

são ambos menores ou iguais a zero, ou ambos maiores ou iguais a zero. No primeiro caso,

$$|x| + |y| \leq |x + y| \text{ e } |x| + |y| \leq -|x + y|$$

o que só pode ocorrer quando $x = y = 0$. Portanto, quando x ou y são não nulos teremos, necessariamente, que

$$|x| + |y| \geq |x + y| \text{ e } |x| + |y| \geq -|x + y|;$$

isto é, que

$$|x| + |y| \geq |x + y| \geq -|x + y|,$$

a primeira das quais é equivalente a (10), concluindo assim a demonstração. Outra maneira de provar a mesma desigualdade consiste em considerar as quatro diferentes possibilidades de escolha de sinal para x e y .

É possível que você esteja se perguntando porque demos ênfase a esta propriedade dos números inteiros em detrimento das demais. A razão é que há versões da desigualdade triangular em muitas outras áreas da matemática, entre elas álgebra linear e análise funcional. Na verdade, a desigualdade tem este nome porque sua primeira aparição ocorre na Proposição ?? dos *Elementos* de Euclides, segundo a qual a soma dos comprimentos de dois lados de um triângulo é sempre maior que o comprimento do outro lado. Esta

Exercícios

- Este exercício tem como objetivo provar que

$$\phi = \frac{1 + \sqrt{5}}{2},$$

conhecido como *razão áurea*, é um número irracional. Para isso você deve seguir o seguinte roteiro.

- Prove que $\phi = 1 + \frac{1}{\phi}$.
 - Supondo que $\phi = m/n$, em que m e n são inteiros primos entre si, use (a) para mostrar que
- $$\frac{m}{n} = \frac{m+n}{m}$$
- Prove que $\text{mdc}(m, n) = \text{mdc}(m+n, n)$.
 - Deduza uma contradição a partir de (b) e (c), usando o fato de uma fração não poder ter mais que uma forma reduzida.

2. Sejam a e b números reais. Prove as seguintes afirmações usando as propriedades algébricas dos números reais:

- (a) se $ab = a$, então $b = 1$;
- (b) se $ab = 0$, então $a = 0$ ou $b = 0$;
- (c) $a^2 - b^2 = (a + b)(a - b)$;
- (d) se $a^2 = b^2$, então $a = b$ ou $a = -b$.

3. Encontre os erros na seguinte “demonstração” de que $2 = 1$:

Se $a = b$, então $a^2 = ab$; donde

$$a^2 - b^2 = ab - b^2.$$

Logo,

$$(a - b)(a + b) = (a - b)b.$$

Assim,

$$a + b = b;$$

mas, lembrando que $a = b$, isto equivale a

$$2a = a,$$

o que nos permite concluir que $2 = 1$.

4. Ache todos os valores de x para os quais as seguintes desigualdades são verdadeiras:

- | | |
|------------------------|-----------------------------|
| (a) $4 - x < 3 - 2x$; | (d) $x^2 + 2x + 9 > 0$; |
| (b) $x^2 + 12 > 8x$; | (e) $x^2 + x \leq -3$; |
| (c) $x^2 + 3 < 4x$; | (f) $(x - 1)/(x + 1) > 0$. |

5. Prove que, se a , b e c são números reais, então as seguintes afirmações são verdadeiras:

- (a) se $a < b$ e $c < d$, então $a + c < b + d$;
- (b) se $a < b$ e $c > d$, então $a - c < b - d$;
- (c) se $a > 1$, então $a^2 > a$;
- (d) se $0 < a < 1$, então $a^2 < a$;
- (e) se $0 \leq a < b$ e $0 \leq c < d$, então $ac < bd$;
- (f) se $0 \leq a < b$, então $a^2 < b^2$;
- (g) se $a, b \geq 0$ e $a^2 < b^2$, então $a < b$.

Para provar (g), argumente por contradição, usando (f) com os papéis de a e b trocados entre si.

6. Prove que se $0 < a < b$, então

$$a < \sqrt{ab} < \frac{a+b}{2} < b.$$

CHAPTER 3

A geometria axiomatizada

Neste capítulo introduzimos uma apresentação axiomática, um tanto simplificada, da geometria, inspirada naquela publicada por George Birkhoff em 1932. Nossa principal propósito é lhe dar uma oportunidade de jogar o *jogo dedutivo* em um ambiente bastante familiar, a tradicional geometria dos triângulos que aparece, em sua maior parte, no Livro I dos *Elementos* de Euclides. A principal característica deste jogo é que

todas as afirmações que você fizer em uma demonstração precisam ser explicitamente justificadas a partir de um axioma ou de um algum resultado previamente demonstrado.

Em outras palavras, trata-se de um jogo em que, a cada jogada, é necessário indicar claramente a regra que permite que a jogada seja executada. Como tudo na vida, a escolha da geometria de triângulos como ambiente para o jogo tem vantagens e desvantagens. A principal vantagem é que o assunto é familiar, de modo que, tanto os axiomas, quanto os resultados que formos provando, serão fáceis de lembrar. A principal desvantagem é que o assunto é tão familiar que você corre o risco de, sem querer, executar um passo da demonstração sem se dar conta de que está usando algo que não foi justificado anteriormente.

1. Introdução

Como vimos no capítulo anterior, a geometria é apresentada nos *Elementos* de Euclides de maneira extremamente sistemática. No início do Livro I, Euclides:

- define noções fundamentais como ponto, reta, plano e circunferência;
- introduz os postulados que determinam como as noções fundamentais estão relacionadas;
- enumera o que chama de *noções comuns*, como “duas coisas iguais a uma terceira são iguais entre si”.

Em princípio, todas as proposições do Livro I dos *Elementos* deveriam ser provadas usando, como ponto de partida, apenas os cinco postulados e as oito noções comuns. Mas, já na antiguidade, vários matemáticos chamaram a atenção de que isto não ocorre. No caso da Proposição 4, por exemplo, Euclides move um triângulo sobre o outro, para provar que são



FIGURE 1. David Hilbert e George D. Birkhoff..

iguais ou, na terminologia usual, congruentes. Contudo, além de nenhum postulado tratar de como estes movimentos podem ser feitos, nenhuma das definições explicita o que significa dizer que dois triângulos são congruentes. É possível que o próprio Euclides tivesse noção do problema, porque argumentos de sobreposição são usados raramente nos *Elementos*.

Há duas maneiras de contornar este problema. A primeira, consiste em definir exatamente o que deve ser entendido pelo movimento de uma figura e pelo que significa que duas figuras sobrepostas são iguais. A segunda, que adotaremos aqui, foi adotada por David Hilbert em um livro, publicado em 1899, no qual dá uma sistematização moderna da geometria euclidiana. Neste caso, a ideia é introduzir como axioma um dos casos de congruência de triângulos.

Outra coisa que a sistematização da geometria euclidiana por Hilbert deixa clara é que precisamos de muito mais axiomas do que os postulados e noções comuns enumerados nos *Elementos*. Por exemplo, seriam necessários oito axiomas diferentes para lidar com a congruência (igualdade) de segmentos, ângulos e triângulos. Além de determinar quando estes objetos são congruentes, estes axiomas estabelecem quais são as propriedades que a congruência satisfaz. Outra categoria de axiomas ausente dos *Elementos* é a daqueles que determinam a ordem em que pontos aparecem sobre uma reta.

Uma característica importante do enfoque de Hilbert é completa ausência das definições das noções mais básicas, como ponto, reta e plano. Como vimos no capítulo 2, as definições destes objetos nos *Elementos* não passam de apelos à intuição, sem nenhum valor formal. Para ele, não importa o que estes objetos *são*, mas sim *como se relacionam entre si*. Otto Blumenthal, que foi aluno de Hilbert, conta, a este respeito, a seguinte história:

numa sala de espera de Berlim, [Hilbert] discutiu com dois geômetras (se não me engano, A. Schoenflies e E. Kötter) sobre os axiomas da geometria e deu sua visão peculiar dizendo: “É preciso sempre que se possa substituir ‘pontos, retas e planos’ por ‘mesas, cadeiras e latas de cerveja’”. Já então sua atitude era de que o substrato intuitivo dos conceitos geométricos era matematicamente irrelevante e de que apenas sua conexão através dos axiomas devia ser considerada.

A axiomatização da geometria euclidiana devida a Hilbert não é a única. Por exemplo, em 1926-27, Alfred Tarski (1901–1983) propôs um sistema de axiomas diferente para a geometria elementar em aulas na Universidade de Varsóvia. O sistema de Tarski tem certas vantagens, de natureza lógica e um tanto técnicas, sobre a sistematização de Hilbert. Entretanto, nenhum dos dois sistemas de axiomas é adequado para um tratamento inicial, porque ambos são bastante abstratos. Por isso apresentaremos neste curso um terceiro sistema, introduzido por George Birkhoff em 1932, com o objetivo explícito de simplificar o tratamento axiomático da geometria elementar, para adequá-lo às necessidades de um curso introdutório.

A principal diferença entre o sistema de Birkhoff, e aqueles propostos por Hilbert e Tarski, está na introdução, entre os axiomas, das medidas de comprimento e ângulo. Hilbert e Tarski introduzem congruências de segmentos e de ângulos sem mencionar medições; em outras palavras, podemos comparar dois segmentos ou ângulos entre si, mas não associar a eles valores numéricos. Além da escolha feita por Birkhoff nos permitir tratar os problemas de geometria de maneira mais natural, ela também permite que trabalhemos com um número bem menor de axiomas. A grande desvantagem deste enfoque é a necessidade de usar números reais para efetuar estas medidas porque, como já vimos no capítulo anterior, descrever os números reais de maneira rigorosa é bastante difícil. Portanto, nosso tratamento da geometria será rigoroso, a menos do fato de que usaremos as propriedades dos números reais enunciadas na seção 6 do capítulo 2.

Para encerrar, uma observação sobre terminologia. Nossos resultados mais importantes serão rotulados como *teoremas*, do verbo grego que significa “contemplar”. Já um resultado cujo papel é apenas o de auxiliar na demonstração de um teorema será rotulado como *lema*, do verbo grego “tomar”, porque um lema é um resultado que se toma como verdadeiro na demonstração de um teorema. Finalmente, diremos que uma consequência facilmente deduzida de um teorema é um *corolário*, do latim “corollarium”, que significa uma coisa que é dada gratuitamente, um presente.

2. Incidência e medidas em retas

Como já dissemos anteriormente, não definiremos pontos, nem retas, nem planos. Segundo Hilbert, vamos nos contentar em estabelecer os axiomas que regulam como estes

objetos se relacionam entre si. Tendo isto em mente, consideraremos o *plano* \mathbb{P} como sendo apenas um conjunto, cujos elementos chamaremos de *pontos*. Estudaremos vários subconjuntos do plano, a começar pelas *retas*, definindo-os a partir de axiomas que enumeram suas propriedades. O primeiro axioma trata da relação entre retas e pontos.

AXIOMA 1 (Incidência). *Dois pontos distintos definem uma e somente uma reta e, dada uma reta qualquer, existe um ponto do plano fora desta reta.*

Quando três pontos pertencem a uma mesma reta, dizemos que são *colineares*. Portanto, segundo este axioma, o plano tem, pelo menos duas retas. Na verdade, como você esperaria, o plano contém infinitas retas mas, para provar isto, precisamos do axioma referente à medida de ângulos que será o tema da próxima seção.

Antes de enunciar o próximo axioma, precisamos introduzir mais um objeto indefinido, a função

$$d : \mathbb{P} \times \mathbb{P} \longrightarrow [0, +\infty)$$

que a cada par de pontos (A, B) associa o número real positivo $d(A, B)$, de modo que

$$d(A, B) = d(B, A)$$

quaisquer que sejam os pontos $A, B \in \mathbb{P}$. Diremos que $d(A, B)$ é a *distância* entre A e B . O próximo postulado especifica a relação entre esta função e os pontos de uma reta.

AXIOMA 2 (medidas em retas). *É possível construir funções bijetivas entre os pontos de uma reta e os números reais, de modo que o módulo da diferença entre os números associados a dois pontos da reta é igual à distância entre estes pontos.*

Seja ℓ uma reta. Este axioma prescreve que é possível construir uma função bijetiva

$$(11) \quad \phi : \ell \longrightarrow \mathbb{R},$$

de modo que, se A e B são pontos de ℓ , então

$$(12) \quad |\phi(A) - \phi(B)| = d(A, B).$$

Note que ϕ não está definida em todo o plano, mas apenas para os pontos da reta ℓ . Além disso, há muitas funções possíveis com esta mesma propriedade, ainda que, como veremos abaixo, seja fácil relacioná-las entre si. Este axioma nos permite provar várias propriedades básicas da função distância usando o que aprendemos sobre a função valor absoluto no capítulo 1.

TEOREMA 2.1. *Se $A, B \in \mathbb{P}$, então*

- (a) $d(A, B) = 0$ se, e somente se, $A = B$;
- (b) se A, B e C são três pontos de uma mesma reta, então

$$d(A, C) \leq d(A, B) + d(B, C);$$

(c) se ℓ é uma reta e $\phi, \psi : \ell \rightarrow \mathbb{R}$ são duas funções que satisfazem o axioma 2, então existe um número real r tal que

$$\phi(P) = \psi(P) + r, \text{ para todo } P \in \ell,$$

ou, então,

$$\phi(P) = -\psi(P) + r, \text{ para todo } P \in \ell.$$

PROOF. Sejam ℓ uma reta e A, B e C três pontos de ℓ . Pelo axioma 2 existe uma função bijetiva

$$\phi : \ell \rightarrow \mathbb{R},$$

que satisfaz (12). Logo,

$$|\phi(A) - \phi(B)| = d(A, B) = 0$$

se, e somente se, $\phi(A) = \phi(B)$. Como ϕ é bijetiva, isto equivale a dizer que $A = B$, provando, assim, o item (a). Para provar (b) usamos novamente (12) para escrever

$$d(A, C) = |\phi(A) - \phi(C)| = |(\phi(A) - \phi(B)) + (\phi(B) - \phi(C))|.$$

Mas, pela desigualdade triangular,

$$|(\phi(A) - \phi(B)) + (\phi(B) - \phi(C))| \leq |\phi(A) - \phi(B)| + |\phi(B) - \phi(C)|.$$

Assim,

$$d(A, C) = |\phi(A) - \phi(C)| \leq |\phi(A) - \phi(B)| + |\phi(B) - \phi(C)|.$$

Para obter a desigualdade do item (b) basta lembrar que

$$d(A, B) = |\phi(A) - \phi(B)| \text{ e } d(B, C) = |\phi(B) - \phi(C)|.$$

Resta provar (c). Seja

$$\psi : \ell \rightarrow \mathbb{R}$$

uma função que satisfaz

$$(13) \quad |\psi(A) - \psi(B)| = d(A, B)$$

e seja $O \in \ell$ tal que $\phi(O) = 0$. Supondo, para começar, que

$$\psi(O) = 0,$$

definimos os conjuntos

$$\ell_\phi = \{Q \in \ell \mid \phi(Q) > 0\} \text{ e } \ell_\psi = \{Q \in \ell \mid \psi(Q) > 0\}.$$

Digamos que

$$A \in \ell_\phi \cap \ell_\psi \neq \emptyset.$$

Neste caso, teremos por (12) e (13), que

$$(14) \quad \phi(A) = d(A, 0) = \psi(A).$$

Mas, se $B \neq A$ for outro ponto de ℓ_ϕ , então, por (14),

$$|\psi(A) - \psi(B)| = d(A, B) = |\phi(A) - \phi(B)| = |\psi(A) - \phi(B)|.$$

Logo, pela definição do módulo de um número,

$$\psi(A) - \psi(B) = \psi(A) - \phi(B) \text{ ou } \psi(A) - \psi(B) = -(\psi(A) - \phi(B)).$$

A primeira igualdade nos dá $\psi(B) = \phi(B) > 0$ e a segunda $\psi(A) = \psi(B)$. Contudo, $B \neq A$ e ψ é bijetiva o que invalida a segunda possibilidade. Portanto, admitindo que $\phi(O) = \psi(O) = 0$, provamos

$$\text{se } \ell_\phi \cap \ell_\psi \neq \emptyset, \text{ então } \psi(B) = \phi(B), \text{ para todo } B \in \ell_\phi;$$

o que nos permite concluir que $\ell_\phi = \ell_\psi$. Um argumento análogo mostra que

$$\text{se } \ell_\phi \cap \ell_\psi = \emptyset, \text{ então } \psi(B) = -\phi(B), \text{ para todo } B \in \ell_\psi.$$

Finalmente, precisamos eliminar a hipótese de que $\phi(O) = \psi(O) = 0$. Para isso, note que a função $\tau : \mathbb{R} \rightarrow \mathbb{R}$ definida por $\tau(x) = x - \psi(O)$ é bijetiva, de modo que

$$\psi' = \tau \circ \psi : \ell \rightarrow \mathbb{R}$$

também é bijetiva. Além disso,

$$\psi'(O) = (\tau \circ \psi)(O) = \tau(\psi(O)) = \psi(O) - \psi(O) = 0.$$

Portanto, o resultado provado anteriormente pode ser aplicado a ψ' ; de modo que

$$\psi'(A) = \phi(A) \text{ ou } \psi'(A) = -\phi(A),$$

para todo $A \in \ell$. Logo,

$$\psi(A) = \psi(O) + \phi(A) \text{ ou } \psi(A) = \psi(O) - \phi(A).$$

Para obter o que queríamos provar, basta tomar $r = \psi(O)$. \square

Dada uma reta ℓ , uma função ϕ a ela associada nos permite deduzir que ℓ tem infinitos pontos, já que ϕ é bijetiva e \mathbb{R} é um conjunto infinito. Também podemos usar ϕ para ordenar os pontos de uma reta e, assim, definir quando um ponto está entre outros dois. Mais precisamente, se A, B e C são pontos de ℓ , diremos que B está entre A e C se

$$\phi(A) < \phi(B) < \phi(C) \text{ ou } \phi(C) < \phi(B) < \phi(C).$$

Embora haja uma infinidade de maneiras de associar uma função ϕ a uma reta ℓ de modo que (12) seja satisfeita, o item (c) do teorema 2.1 garante que se B está entre A e C para uma delas, então o mesmo vale para todas as outras. Isto nos permite definir segmentos de retas e semirretas. Se $A \neq B$ são dois pontos da reta ℓ , então o *segmento* AB é o subconjunto de ℓ formado por A, B e os pontos de ℓ que estão entre A e B . Em particular, se $\phi(A) < \phi(B)$, então

$$AB = \{P \in \ell \mid \phi(A) \leq \phi(P) \leq \phi(B)\}.$$

Já uma ponto A na reta ℓ nos permite definir as *semirretas*,

$$A^+ = \{P \in \ell \mid \phi(P) \geq \phi(A)\} \text{ e } A^- = \{P \in \ell \mid \phi(P) \leq \phi(A)\}$$

cuja união é a reta ℓ e cuja interseção é o ponto A . Uma semirreta também pode ser definida a partir de dois pontos A e B de ℓ por

$$\overrightarrow{AB} = \{P \in \ell \mid P \in AB \vee B \in AP\};$$

isto é, como o conjunto dos pontos que estão entre A ou “além de B ”, relativamente a A . Esta semirreta é ilustrada na figura 2.

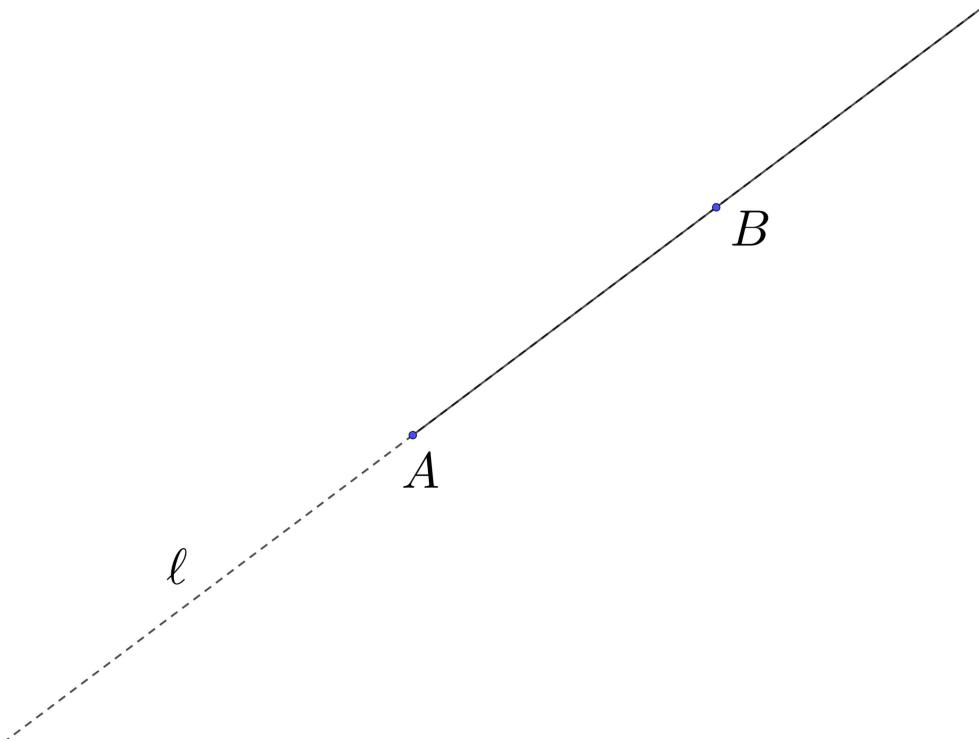


FIGURE 2. Uma semirreta.

Como esta é a primeira figura a aparecer depois que começamos o tratamento axiomático da geometria, é importante explicitar o papel que as figuras desempenharão em nossa apresentação. Afinal, como alguém já disse, *a geometria é a arte de argumentar corretamente a partir de figuras erradas*. Alguns historiadores creem que os matemáticos gregos distorciam deliberadamente algumas figuras, para evitar que a regularidade da figura sugerisse ideias incorretas. Embora, neste curso, todos os nossos argumentos devam, necessariamente, ser baseados nos axiomas e em resultados anteriormente provados, utilizaremos figuras, tanto para ilustrar certos resultados, como para nos inspirar em como fazer uma demonstração. O perigo a evitar, sobretudo no último caso, é o de deixar-nos levar pela figura e fazermos afirmações que não podem ser justificadas com base no que já fomos capazes de provar.

Antes de prosseguirmos, convém fazer uma observação sobre notação. Na medida do possível, denotaremos pontos do plano por letras maiúsculas e retas por letras minúsculas do alfabeto romano. As minúsculas do alfabeto grego serão reservadas para ângulos. Como estamos estudando geometria plana, nosso conjunto universo é o conjunto de pontos que estamos chamando de plano. Portanto, tudo se passa como se conhecêssemos apenas um plano, que denotaremos por \mathbb{P} , e ao qual nos referiremos usando o artigo definido: o plano, em vez de um plano. Encerraremos a seção provando um resultado elementar sobre a interseção de duas retas. Para isto, definimos duas retas como sendo *paralelas* se não têm nenhum ponto em comum.

TEOREMA 2.2. *Se duas retas distintas não são paralelas, então têm exatamente um ponto comum.*

Provaremos este teorema usando a estratégia de demonstração favorita dos geômetras da Grécia Antiga, e que já usamos para provar que $\sqrt{2}$ não é um número racional: a *demonstração por contradição*. Lembre-se que ela se baseia no princípio do meio excluído, segundo o qual, qualquer que seja a proposição P , teremos sempre que $P \vee \neg P = \top$. Desse modo, se formos capazes de mostrar que $\neg P$ é falsa, então P tem que ser verdadeira.

PROOF. Sejam ℓ_1 e ℓ_2 duas retas distintas que não são paralelas. Queremos provar que $\ell_1 \cap \ell_2$ consiste de apenas um ponto. Como, por hipótese, as retas não são paralelas então, por definição, têm que ter *pelo menos* um ponto P em comum. Mas, pelo axioma 2, ℓ_1 tem infinitos pontos. Seja $Q \neq P$ um ponto qualquer de ℓ_1 . Supondo, por contradição, que Q também esteja em ℓ_2 , temos pelo axioma 1, que as retas coincidem, contradizendo a hipótese de que são distintas. Logo, nenhum ponto de ℓ_1 , diferente de P , pode pertencer a ℓ_2 , o que prova o teorema. \square

3. Semiplanos e ângulos

Como definiremos ângulos como subconjuntos de pontos definidos por semiplanos, precisamos de uma axioma que especifique de que maneira uma reta subdivide um plano.

AXIOMA 3 (Semiplanos). *Os pontos do plano fora de uma reta ℓ podem ser agrupados em dois conjuntos disjuntos ℓ^+ e ℓ^- , de modo que se $P, Q \in \ell^+$, então $PQ \subset \ell^+$, e o enunciado análogo vale para ℓ^- . Além disso, se $R \in \ell^+$ e $S \in \ell^-$, então RS intersecta ℓ .*

Os conjuntos ℓ^+ e ℓ^- são os *semiplanos* em que ℓ divide \mathbb{P} . O axioma 3 é menos fácil de digerir do que os anteriores. Segundo ele, que cada reta divide o plano em dois semiplanos com certas propriedades. As situações descritas nos itens (a) e (b) estão ilustradas na figura 3. Note que ℓ não pertence a nenhum dos dois semiplanos, de modo que $\mathbb{P} = \ell^- \cup \ell \cup \ell^+$.

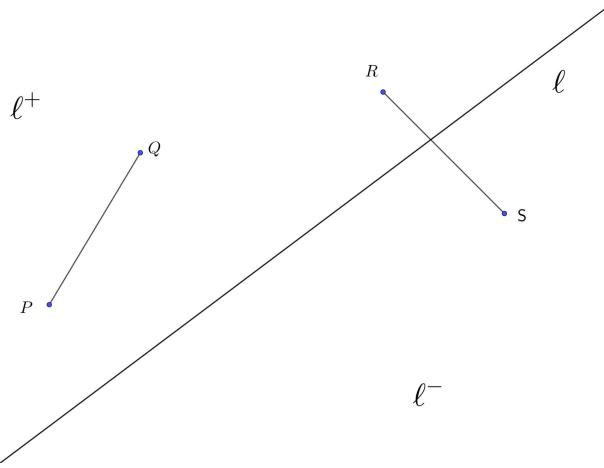
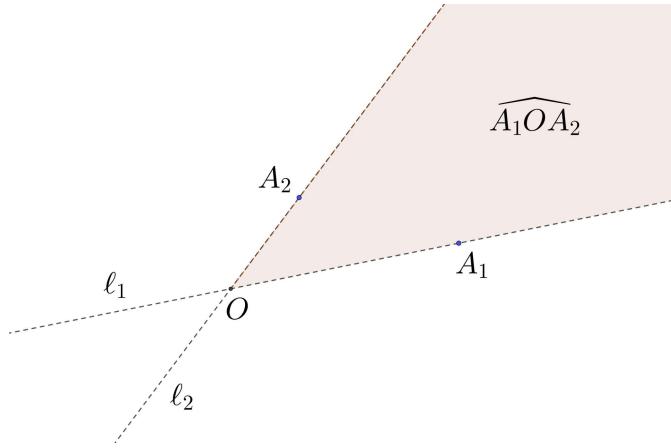


FIGURE 3. Os dois casos do axioma de ordenação 4

O axioma 3 nos permite definir ângulos como subconjuntos do plano. Sejam ℓ_1 e ℓ_2 duas retas que se intersectam no ponto O e sejam A_1 e A_2 pontos do plano tais que

$$A_1 \in \ell_1 \cap \ell_2^- \text{ e } A_2 \in \ell_2 \cap \ell_1^+.$$

O ângulo $\widehat{A_1OA_2}$ é definido como sendo a interseção de $\ell_1^+ \cap \ell_2^-$. Em uma primeira leitura esta definição parece bastante prolixo, mas a figura 4 nos ajuda a pôr os pés no chão. Nela, ℓ_1^+ e ℓ_2^- são os semiplanos situados, respectivamente, acima de ℓ_1 e abaixo de ℓ_2 . A região correspondente ao ângulo \widehat{AOB} foi hachurada para facilitar sua identificação.

FIGURE 4. O ângulo \widehat{AOB} .

A medida de ângulos, que formalizamos no próximo axioma, é feita de maneira muito semelhante à medida de segmentos.

AXIOMA 4 (medida de ângulos). *Existe uma função*

$$\measuredangle : \mathbb{P} \times \mathbb{P} \times \mathbb{P} \longrightarrow [0, \pi],$$

de modo que, dados dois pontos $O \neq A$ em uma reta ℓ , a regra que associa a cada $B \in \ell^+$ o número $\measuredangle(A, O, B) \in [0, \pi]$ define uma bijeção

$$\theta_{\ell, O, A} : \ell^+ \longrightarrow [0, \pi],$$

que satisfaz

$$\theta_{\ell, O}(B) = \theta_{\ell, O}(C),$$

para todo $C \in \overrightarrow{OB}$.

O número $\measuredangle(A, O, B)$, ou simplesmente $\measuredangle AOB$, é a *medida* do ângulo \widehat{AOB} . Como todos os ângulos, por causa da maneira como os definimos, estão sempre contidos em um semiplano, limitar $\measuredangle AOB$ ao intervalo $[0, \pi]$ corresponde à nossa intuição relativa às medidas de ângulos em radianos. A demonstração do seguinte resultado é semelhante à do teorema 2.1 e ficará aos seus cuidados.

TEOREMA 3.1. *Sejam A, O e B pontos do plano e ℓ a reta definida por O e A . Então,*

- (a) $\measuredangle AOB = 0$ se, e somente se, $B \in \overrightarrow{OA}$;
- (b) se $B \in \widehat{AOC}$, então $\measuredangle AOB + \measuredangle BOC = \measuredangle AOC$.
- (c) $\measuredangle AOB = \pi$ se, e somente se, $B \in \ell \setminus \overrightarrow{OA}$.

Encerraremos a seção, provando duas consequências simples do teorema 3.1. A primeira delas diz respeito às retas perpendiculares. Como de costume diremos que \widehat{AOB} é um ângulo *reto* quando $\measuredangle AOB = \pi/2$. Neste caso, as retas \overleftrightarrow{OA} e \overleftrightarrow{OB} são *perpendiculares*.

COROLÁRIO 3.2. *Por um ponto de uma reta ℓ passa uma, e somente uma, reta perpendicular a ℓ .*

PROOF. A existência da perpendicular por um ponto de uma reta é uma consequência imediata do axioma 4. Resta-nos provar que não pode haver mais do que uma destas retas. Sejam $O \in \ell$ e $P, Q \in \ell^+$ dois pontos tais que as retas \overleftrightarrow{OP} e \overleftrightarrow{OQ} são ambas perpendiculares a ℓ . Isto significa, por definição, que se $A \neq O$ é um ponto de ℓ , então

$$\theta_{\ell, O}(P) = \measuredangle AOP = \measuredangle AOQ = \theta_{\ell, O}(Q).$$

Como, pelo axioma 4, a função $\theta_{\ell, O}$ é bijetiva, temos que $P = Q$. Logo, pelo axioma 1, as retas \overleftrightarrow{OP} e \overleftrightarrow{OQ} são iguais, como devíamos mostrar. \square

A segunda consequência do teorema 3.1 de que precisamos é a igualdade de ângulos opostos pelo vértice. Suponhamos que ℓ_1 e ℓ_2 sejam duas retas que se intersectam no ponto O e sejam $A_1, B_1 \in \ell_1$ em semiplanos diferentes relativamente a ℓ_2 e $A_2, B_2 \in \ell_2$

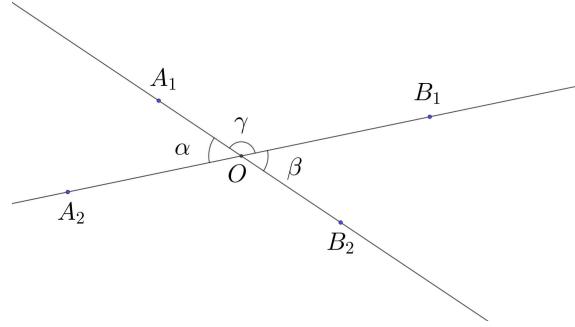


FIGURE 5. Ângulos opostos pelo vértice.

em semiplanos diferentes relativamente a ℓ_1 , como ilustrado na figura 5. Então os ângulos $\widehat{A_1OA_2}$ e $\widehat{B_1OB_2}$ são opostos pelo vértice e o mesmo se aplica a $\widehat{A_1OB_1}$ e $\widehat{A_2OB_2}$. Já $\widehat{A_1OA_2}$ e $\widehat{A_1OB_1}$ são ângulos *adjacentes*, porque têm em comum uma das retas usada para defini-los.

COROLÁRIO 3.3. *Ângulos opostos pelo vértice têm a mesma medida.*

PROOF. Sejam

$$\alpha = \measuredangle A_2OA_1, \quad \beta = \measuredangle B_2OB_1 \quad \text{e} \quad \gamma = \measuredangle A_1OB_1;$$

como na figura 5. Note que $\widehat{A_1OB_1}$ é adjacente, tanto a $\widehat{A_2OA_1}$, quanto a $\widehat{B_2OB_1}$, de modo que

$$\alpha + \gamma = \beta + \gamma.$$

Mas isto implica que

$$\measuredangle A_2OA_1 = \alpha = \beta = \measuredangle B_2OB_1$$

e um argumento análogo mostra que $\measuredangle A_2OB_2 = \measuredangle A_1OB_1$. \square

4. Triângulos

Sejam A , B e C três pontos *não* colineares tais que

$$C \in \overleftrightarrow{AB}^+, \quad B \in \overleftrightarrow{AC}^+ \quad \text{e} \quad A \in \overleftrightarrow{BC}^+.$$

O *triângulo* ABC é o subconjunto de pontos definido por

$$\overleftrightarrow{AB}^+ \cap \overleftrightarrow{AC}^+ \cap \overleftrightarrow{BC}^+ \cup AB \cup AC \cup BC.$$

Note que precisamos incluir os segmentos AB , AC e BC nesta união, porque os semiplanos não contêm pontos sobre as retas que os definem. Os pontos A , B e C são os *vértices* do triângulo ABC e

$$\overleftrightarrow{AB}^+ \cap \overleftrightarrow{AC}^+ \cap \overleftrightarrow{BC}^+$$

é seu *interior*. Diremos também que BC é o lado *oposto* ao ângulo \widehat{A} e vice-versa. Analogamente, AC e AB são os lados opostos aos ângulos \widehat{B} e \widehat{C} , respectivamente.

Os resultados desta seção dizem respeito à comparação de triângulos. Assim, diremos que dois triângulos ABC e $A'B'C'$ são *semelhantes* se

$$(15) \quad \frac{AB}{A'B'} = \frac{AC}{A'C'} = \frac{BC}{B'C'}.$$

Chamamos de *razão de semelhança* ao valor comum destas três frações. Quando este número é igual a um, os triângulos são *congruentes*. Indicaremos que ABC e $A'B'C'$ são triângulos semelhantes, escrevendo

$$\triangle ABC \sim \triangle A'B'C',$$

ao passo que, quando estes triângulos são congruentes, escreveremos

$$\triangle ABC \equiv \triangle A'B'C'.$$

Nosso último axioma nos determina as condições mínimas que dois triângulos devem satisfazer para serem semelhantes.

AXIOMA 5. (*semelhança de triângulos*) *Para que os triângulos ABC e $A'B'C'$ sejam semelhantes, basta que*

$$\frac{AB}{A'B'} = \frac{AC}{A'C'} \text{ e que } \angle BAC = \angle B'A'C'.$$

Quando a razão de semelhança é igual a 1, este axioma coincide com a caso de congruência de triângulos conhecido como ALA.

CASO ALA 1. *Para que os triângulos ABC e $A'B'C'$ sejam congruentes, basta que $d(A, B) = d(A', B')$, $d(A, C) = d(A', C')$ e que $\angle BAC = \angle B'A'C'$.*

Um detalhe que é importante ressaltar é que os ângulos iguais têm que ser, *obrigatoriamente, aqueles que são limitados pelos lados proporcionais*. Por exemplo, na figura 19 os triângulos AOB e AOC têm em comum o lado AO , e o ângulo \widehat{OAB} , ao passo que os OB e OC têm a mesma medida, porque são raios de uma mesma circunferência. No entanto, estes triângulos claramente não são semelhantes. Isto não contradiz o axioma 5 porque o ângulo \widehat{OAB} não está situado entre os lados iguais que, no triângulo AOB , são AO e OB .

Como consequência deste axioma podemos provar facilmente a Proposição 5 dos *Elementos* de Euclides, que diz respeito aos triângulos *isósceles*. Definiremos como sendo *isósceles* aqueles triângulos que têm dois lados de mesmo comprimento. A demonstração de Euclides, que era conhecida na Idade Média como Pons Asinorum (latim para “ponte dos asnos”), é um tanto elaborada; a demonstração que apresentaremos deve-se, segundo

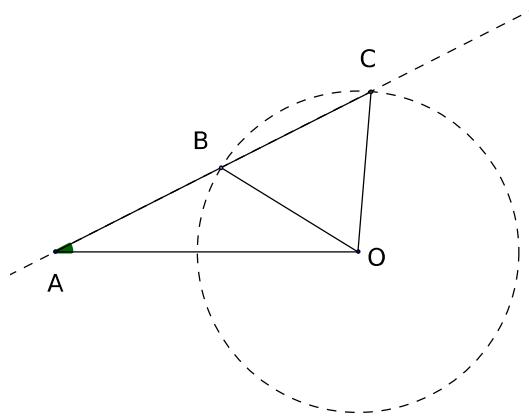


FIGURE 6. Os triângulos AOB e AOC não são congruentes.

Proclus, a Papus (290-350 d.C.). É importante alertá-lo, antes mesmo de ler a demonstração, de que ela trata o mesmo triângulo *como se fossem dois triângulos diferentes, dependendo da ordem em que os vértices são lidos*. A razão para o alerta é que uma demonstração deste gênero às vezes dá a sensação de que estamos “extraíndo algo do nada”.

TEOREMA 4.1. *Em um triângulo isósceles, os ângulos opostos aos lados iguais têm a mesma medida.*

PROOF. Suponhamos que, no triângulo ABC , os lados AB e AC têm a mesma medida. Assim, comparando os triângulos ABC e ACB , temos que:

- (i) o lado AB do triângulo ABC corresponde ao lado AC do triângulo ACB ;
- (ii) o lado AC do triângulo ABC corresponde ao lado AB do triângulo ACB ;
- (iii) \widehat{A} é o ângulo entre os lados correspondentes nos dois triângulos;

e lembrando que (i) e (ii) são consequência da hipótese de que ABC é um triângulo isósceles. Logo, os triângulos ABC e ACB são congruentes, o que nos permite afirmar que todos os ângulos correspondentes são iguais; em particular $\angle B = \angle C$, como devíamos mostrar. \square

Podemos utilizar o axioma 5 para provar dois outros critérios simples de semelhança de triângulos, o primeiro dos quais é o seguinte.

TEOREMA 4.2. *Dois triângulos são semelhantes quando dois ângulos de um deles são iguais a seus correspondentes no outro triângulo.*

PROOF. Sejam ABC e $A'B'C'$ os dois triângulos tais que

$$\measuredangle B = \measuredangle B' \text{ e } \measuredangle A = \measuredangle A'$$

e digamos que $d(A', B') \geq d(A, B)$, donde

$$k = \frac{A'B'}{AB} \geq 1.$$

Marque, na semirreta $\overrightarrow{B'C'}$ o ponto C'' cuja distância de B' é igual a $kd(B, C)$; como ilustrado na figura 7. Por causa da maneira como o ponto C'' foi construído, temos que

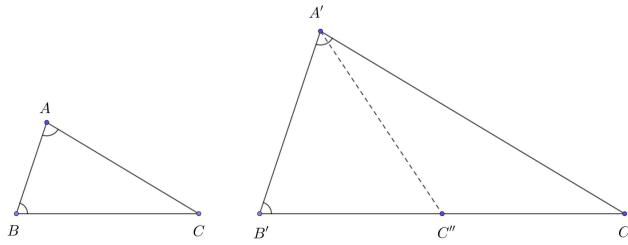


FIGURE 7. Os triângulos ABC e $A'B'C'$ com o ponto C'' .

$$k = \frac{A'B'}{AB} = \frac{B'C''}{BC};$$

de modo que, pelo axioma 5, os triângulos ABC e $A'B'C'$ são semelhantes. Suponhamos que

$$C' \in \overleftrightarrow{A'B'}^+ \text{ e } B' \in \overleftrightarrow{A'C'}^-$$

e que o ponto C'' está situado entre B' e C' . Segue-se, da definição de triângulo, que

$$C'' \in \overleftrightarrow{A'B'}^+ \cap \overleftrightarrow{A'C'}^- = \widehat{B'A'C'}.$$

Logo, pelo teorema 3.1,

$$\measuredangle A = \measuredangle A' = \measuredangle B'A'C' = \measuredangle B'A'C'' + \measuredangle C''A'C',$$

em que a primeira igualdade é uma das hipóteses do teorema. Contudo, como $\triangle ABC \sim \triangle A'B'C'$, temos também que

$$\measuredangle A = \measuredangle BAC = \measuredangle B'A'C''.$$

Substituindo esta última igualdade na anterior, obtemos

$$\measuredangle A = \measuredangle A + \measuredangle C''A'C',$$

o que nos permite concluir que $\angle C''A'C' = 0$. Por sua vez, isto implica que

$$C'' = C' \text{ e } \frac{A'B'}{AB} = \frac{B'C'}{BC}.$$

Mas, combinando a hipótese de que $\angle A = \angle A'$ ao que acabamos de mostrar, podemos deduzir do axioma 5, que os triângulos ABC e $A'B'C'$ são semelhantes. A demonstração no caso em que $C' \in B'C''$ é análoga e fica por sua conta. \square

Quando, sob as hipóteses do teorema, os dois triângulos também têm um lado igual, a razão de semelhança é, necessariamente, igual a 1, de modo que os dois triângulos serão congruentes. Portanto, o caso de congruência conhecido como LAL é uma consequência direta do teorema 4.2. Contudo, ao contrário do que acontecia no caso LAL, o lado não tem que estar situado entre os dois ângulos.

CASO ALA 2. *Dois triângulos são congruentes se têm um lado e dois ângulos iguais.*

Comparando os enunciados do axioma 5 e do teorema 4.2, vemos que um pode ser obtido do outro trocando os papéis dos ângulos com os lados e vice-versa. Já o terceiro critério de semelhança, que apresentamos a seguir, tem um teor bastante diferente, já que apenas impõe condições sobre os lados.

TEOREMA 4.3. *Dois triângulos são semelhantes quando seus lados correspondentes são proporcionais.*

PROOF. Sejam ABC e $A'B'C'$ os dois triângulos. Temos, por hipótese, que

$$(16) \quad \frac{AB}{A'B'} = \frac{AC}{A'C'} = \frac{BC}{B'C'}.$$

No que segue, denotaremos este número por k . Supondo que $C' \in \overleftrightarrow{AB}^+$, construa a semirreta $\vec{s} \subset \overleftrightarrow{AB}^-$ que forma com \overleftrightarrow{AB} o ângulo $\angle CAB$ e marque o ponto $C'' \in \vec{s}$ tal que $d(A', C') = kd(A, C)$; como ilustrado na figura 8. Pelo axioma 5, os triângulos ABC e $A'B'C''$ são semelhantes com razão de semelhança igual a k . Podemos concluir disto e de (16) que

$$\frac{A''B''}{A'B'} = \frac{A''C''}{A'C'} = \frac{B''C''}{B'C'} = 1,$$

de modo que os triângulos $A'B'C'$ e $A'B'C''$ são congruentes. Em particular, os segmentos $A'C'$ e $A'C''$, assim como $B'C'$ e $B'C''$, têm o mesmo comprimento. Logo, pelo teorema 5.6,

$$(17) \quad \angle C''C'A' = \angle C'C''A' \text{ e } \angle C''C'B' = \angle C'C''B'.$$

Mas, pelo teorema 3.1,

$$\angle A'C'B' = \angle C''C'A' + \angle C''C'B' \text{ e } \angle A'C''B' = \angle C'C''A' + \angle C'C''B'.$$

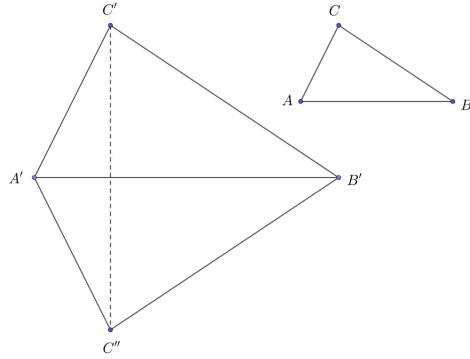


FIGURE 8. Os triângulos ABC , $A'B'C'$ e $A'B'C''$.

Combinando isto com (17), verificamos que

$$\not\propto A'C'B' = \not\propto A'C''B'.$$

Levando em conta, também, que

$$A'C' = A'C'' \text{ e } B'C' = B'C''$$

podemos concluir, pelo axioma 5, que os triângulos $A'B'C'$ e $A'B'C''$ são congruentes. Como ABC e $A'B'C''$ são semelhantes por construção, segue-se que ABC e $A'B'C'$ também são semelhantes. \square

Quando a razão de semelhança é igual a 1, o teorema 4.3 corresponde ao caso de congruência de triângulos conhecidos como LLL.

CASO LLL 1. *Dois triângulos que têm lados correspondentes de mesmo tamanho são congruentes.*

Seguindo a tradição, usaremos estas abreviações *apenas* para os casos de congruência. Encerramos esta seção provando a Proposição 32 dos *Elementos*.

TEOREMA 4.4. *A soma das medidas dos ângulos internos de um triângulo é igual a π .*

PROOF. Seja ABC o triângulo e sejam K , L e M os pontos médios dos lados AC , BC e AB , respectivamente, como ilustrado na figura 11. Como

$$\frac{1}{2} = \frac{AK}{AC} = \frac{AM}{AB}$$

e \widehat{A} é um ângulo comum aos triângulos ABC e KAM , o axioma 5 nos permite concluir que estes triângulos são semelhantes. Como os ângulos \widehat{LBM} e \widehat{AMK} e os lados KM e

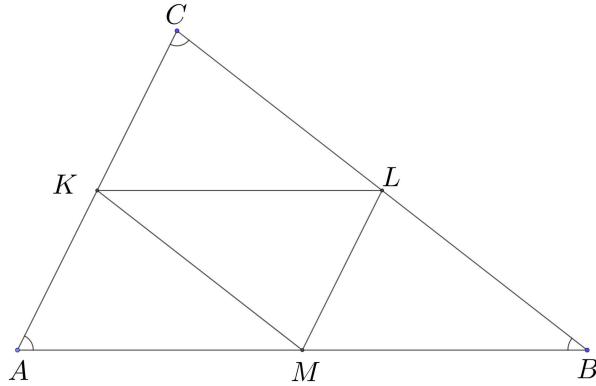


FIGURE 9. O triângulo ABC com os pontos médios dos lados.

BC , se correspondem nestes dois triângulos, concluímos pela definição de semelhança de triângulos que

$$(18) \quad \angle LBM = \angle AMK \text{ e } \frac{KM}{BC} = \frac{1}{2}.$$

Argumentos análogos mostram que

$$\triangle ABC \sim \triangle BML \text{ e } \triangle ABC \sim \triangle CKL$$

de modo que

$$(19) \quad \angle CBA = \angle CAB \text{ e } \frac{KL}{AB} = \frac{1}{2};$$

assim como

$$(20) \quad \frac{KL}{AB} = \frac{1}{2}.$$

Contudo, levando em conta que

$$\frac{KM}{BC} = \frac{LM}{AC} = \frac{KL}{AB} = \frac{1}{2}$$

podemos concluir pelo teorema 4.3 que os triângulos ABC e KLM também são semelhantes. Como os ângulos \widehat{KML} e \widehat{ACB} são correspondentes nestes dois triângulos, temos que

$$\angle KML = \angle ACB.$$

Combinando esta última equação com (18) e (19), obtemos

$$\angle CBA + \angle CAB + \angle ACB = \angle LMB + \angle KMA + \angle KML = \angle AMB = \pi,$$

em que as últimas duas igualdades seguem do teorema 3.1 porque os pontos A , M e B são colineares, o que completa a demonstração deste teorema. \square

Encerramos com uma consequência simples deste teorema, que será usada mais de uma vez nos argumentos da próxima seção.

COROLÁRIO 4.5. *Duas retas perpendiculares a uma terceira são paralelas entre si.*

PROOF. Sejam ℓ uma reta e A e A' pontos distintos em ℓ . Se r e r' são retas perpendiculares a ℓ em A e A' , respectivamente, então r é paralela a r' . Suponhamos, por contradição, que

$$(21) \quad r \cap r' = \{C\},$$

de modo que ABC é um triângulo, como na figura 10. Então, pelo teorema 4.4,

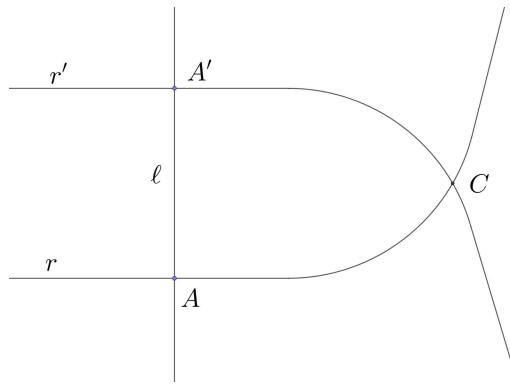


FIGURE 10. Triângulo formado por duas perpendiculares a uma mesma reta.

$$\angle CAB + \angle CBA + \angle ACB = \pi.$$

Como, por hipótese,

$$\angle CAB = \angle CBA = \frac{\pi}{2},$$

podemos concluir que $\angle ACB = 0$. Mas, pelo teorema 3.1, isto implica que $A = C$. Logo, por (21), $A = B = C$, contradizendo o fato de A e B serem pontos distintos de ℓ . \square

5. Paralelas e perpendiculares

Nesta seção provaremos vários resultados, extremamente úteis, relativos a retas paralelas e perpendiculares, entre eles, a existência e unicidade de tais retas sob certas hipóteses adicionais. Nossa primeiro resultado, que é de natureza mais geral, diz respeito à *mediatriz* de um segmento de reta, que é a reta perpendicular a este segmento em seu ponto médio.

LEMA 5.1. *Um ponto pertence à mediatriz de um segmento de reta se, e somente se, é equidistante das extremidades do segmento.*

PROOF. Seja AB o segmento de reta e seja P um ponto da mediatrix ℓ do segmento AB , que é perpendicular a AB e intersecta este segmento em seu ponto médio M . Dado um ponto P de ℓ , precisamos mostrar que PA e PB têm a mesma medida. Como PM é

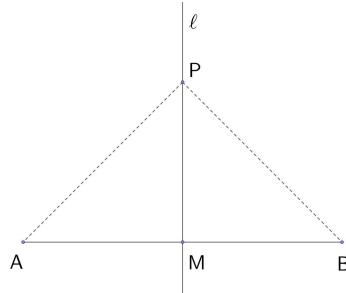


FIGURE 11. Mediatrix de um segmento de reta.

um lado comum aos triângulos AMP e BMP e, por definição,

$$d(A, M) = d(B, M) = \frac{1}{2}d(A, B) \text{ e } \angle AMP = \angle BMP = \frac{\pi}{2},$$

podemos concluir, pelo axioma 5, que estes triângulos são congruentes. Em particular, isto implica que

$$d(A, P) = d(B, P)$$

o que mostra a primeira metade da equivalência. Para mostrar a outra metade, suporemos que P é um ponto do plano tal que

$$(22) \quad d(A, P) = d(B, P)$$

e provaremos que P pertence à mediatrix ℓ . Mais uma vez o argumento se baseia na congruência entre os triângulos APM e BPM . Contudo, precisamos de um argumento diferente daquele usado na primeira parte da demonstração, porque já não temos mais a hipótese de que PM é perpendicular a AB . Contudo, continua sendo verdade que PM é um lado comum aos dois triângulos e que

$$d(A, M) = d(B, M) = \frac{1}{2}d(A, B).$$

Como, nesta segunda parte da demonstração temos também que (22) vale por hipótese, podemos utilizar o teorema 4.3 para concluir que APM e BPM são congruentes. Em particular, $\angle AMP = \angle BMP$. Lembrando que A, M e B são colineares, obtemos

$$\pi = \angle AMB = \angle AMP + \angle BMP = 2\angle AMP = 2\angle BMP;$$

que equivale a dizer que PM é perpendicular a AB . Como, pelo corolário 6.5, a perpendicular a uma reta dada, por um de seus pontos, é única, podemos concluir que $P \in \ell$, como desejávamos mostrar. \square

Embora a existência da perpendicular a uma reta ℓ por um ponto de ℓ seja consequência imediata do axioma 4, muitas construções geométricas requerem que seja construída uma perpendicular a ℓ por um ponto fora de ℓ . A existência de uma tal reta é garantida pelo seguinte teorema.

TEOREMA 5.2. *Dada uma reta ℓ e um ponto $P \notin \ell$, existe uma, e uma única, reta perpendicular a ℓ que contém P .*

PROOF. Sejam A e B pontos de ℓ e suponhamos que $P \in \ell^+$. Pelo axioma 4 existe um ponto $Q \in \ell^-$, tal que $\angle BAQ = \angle BAP$. Usando, por sua vez, o axioma 2, podemos encontrar $P' \in \overrightarrow{AQ}$ tal que $d(A, P) = d(A, P')$. Como o lado AB é comum aos triângulos

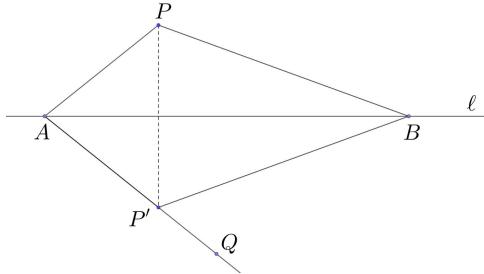


FIGURE 12. Perpendicular por um ponto P fora da reta.

ABP e ABP' e, pela construção acima,

$$AP = AP' \text{ e } \angle BAP = \angle BAP',$$

temos, pelo axioma 5, que estes triângulos são congruentes. Por sua vez, isto implica que $BP = BP'$. Logo, tanto A , quanto B , são equidistantes de P e P' . Portanto, pelo lema 5.1, a reta $\overleftrightarrow{AB} = \ell$ é a mediatrix do segmento PP' . Em particular, ℓ e $\overleftrightarrow{PP'}$ são perpendiculares. \square

Para provar a unicidade da perpendicular, suponhamos, por contradição, que $r \neq r'$ são retas perpendiculares a ℓ por P que intersectam esta última reta, respectivamente, nos pontos Q e Q' . As posições relativas destas retas e pontos estão ilustradas na figura 13, na qual os ângulos $\overline{PQQ'}$ e $\overline{PQ'Q}$ são ambos retos. Logo, pelo corolário 4.5, as retas r e r' são paralelas, o que não é possível porque têm P como ponto em comum. Portanto, existe uma única reta perpendicular a ℓ por P . \square

O teorema 5.2 corresponde à Proposição 12 dos *Elementos*, porém a construção de Euclides utiliza régua, sem marcações, e compasso, ao passo que a que descrevemos acima é baseada no uso de uma régua, com marcações, e de um transferidor.

Nosso próximo resultado é o famoso *quinto postulado* dos *Elementos*. Na seção 5 vimos que, durante séculos, os matemáticos tentaram, sem sucesso, provar este postulado a

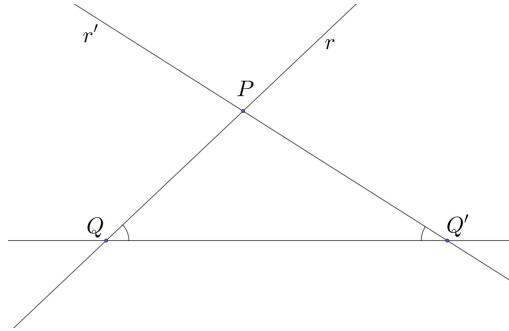


FIGURE 13. Unicidade da perpendicular por um ponto fora de uma reta.

partir dos demais e que, na verdade, isto não é possível. Por isso, talvez você se surprenda que estejamos prestes a prová-lo. A razão porque isto é possível está em nossa escolha de axiomas; mais precisamente, no fato de que nossos axiomas nos permitem medir segmentos e ângulos, ao passo que Euclides admitia apenas construções com uma reta, sem marcações, e um compasso. Assim, para achar o ponto médio de um segmento basta-nos medi-lo e encontrar, na régua, onde está sua metade; Euclides só podia fazer isto às custas da construção geométrica que apresenta na Proposição 10 dos *Elementos*.

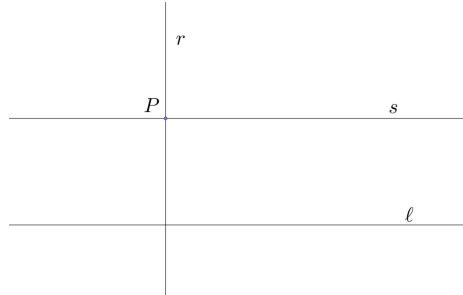


FIGURE 14. Paralela por um ponto P fora da reta.

TEOREMA 5.3. *Por um ponto, fora de uma reta, passa uma, e somente uma, reta paralela à reta dada.*

PROOF. Sejam ℓ uma reta e P um ponto fora de ℓ . Pelo teorema 5.2, existe uma reta r perpendicular a ℓ e que passa pelo ponto P . Já, pelo corolário 6.5, existe uma reta s perpendicular a r e passando por P . A situação relativa das retas é ilustrada na figura 14. Como s e ℓ são ambas perpendiculares a r , o corolário 4.5 nos permite concluir que s é paralela a ℓ .

A demonstração não está completa, porque ainda falta mostrar que não pode haver duas retas paralelas a ℓ por P . Mais uma vez, argumentaremos por contradição. Suponhamos, então, que s e s' sejam duas retas *distintas* que são paralelas a ℓ e passam por P . Pelo axioma 1, P é o único ponto comum a s e s' . Para os passos seguintes precisamos construir várias retas e segmentos, todos ilustrados na figura 15. Começamos escolhendo um ponto

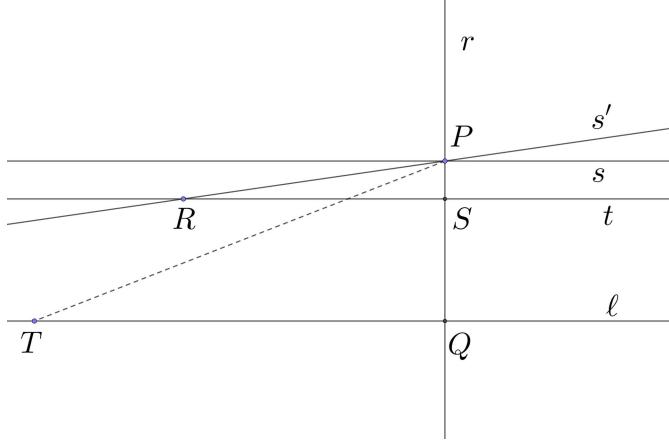


FIGURE 15. Provando a unicidade da paralela por um ponto.

qualquer $R \neq P$ na reta s' . Em seguida construímos a reta t que é perpendicular a r por R . Digamos que S é o ponto de interseção de t e r . Finalmente, e esta é a parte menos óbvia da construção, usamos o axioma 2 para marcar o ponto $T \in \ell$ tal que

$$(23) \quad d(S, T) = \frac{d(S, R) \cdot d(P, Q)}{d(P, S)},$$

em que, como antes, Q é o ponto em que ℓ e r se intersectam. Logo, PRS e PTQ são triângulos nos quais os ângulos nos vértices S e Q são retos e os segmentos que se encontram nestes vértices satisfazem

$$\frac{SR}{PS} = \frac{ST}{PQ},$$

pois esta relação é equivalente a (23). Portanto, pelo axioma 5, os triângulos são semelhantes. Em particular,

$$\sphericalangle RPQ = \sphericalangle TPQ.$$

Mas, pela bijetividade da função $\theta_{s,P}$, isto implica que as semirretas \overrightarrow{PR} e \overrightarrow{PT} coincidem; de modo que,

$$\overrightarrow{PR} \cap \ell = \overrightarrow{PT} \cap \ell = \{T\}.$$

Contudo, $\overrightarrow{PR} \subset s'$, que é uma reta paralela a ℓ por hipótese, o que nos dá a desejada contradição. \square

Encerraremos investigando as relações entre os ângulos formados por uma reta r , que intersecta outras duas retas ℓ e ℓ' . Como a descrição formal dos ângulos que nos interessam é bastante prolixo, vamos nos contentar em localizá-los usando a figura 16. Assim, α e β são ângulos *alternos internos*, ao passo que α e α' são ângulos *correspondentes*. Esta terminologia facilmente se estende aos demais ângulos formados por estas três retas.

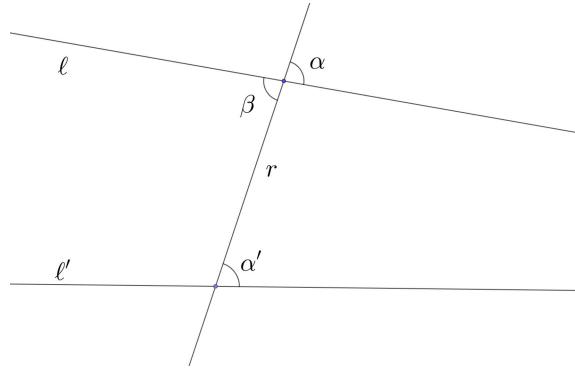


FIGURE 16. Ângulos alternos internos e correspondentes.

TEOREMA 5.4. *Seja r uma reta que intersecta duas outras retas $\ell \neq \ell'$ em pontos diferentes de $\ell \cap \ell'$. Os ângulos alternos internos formados pela interseção de r com ℓ e ℓ' são iguais se, e somente se, estas retas são paralelas.*

PROOF. Começaremos provando que, se os ângulos alternos internos formados por uma reta que intersecta outras duas são iguais, então as retas são paralelas. Argumentando por contradição, suponhamos que, embora os ângulos alternos internos sejam iguais, as retas não sejam paralelas, e digamos que

$$\ell_1 \cap \ell_2 = \{C\}.$$

Denotando por A e A' os pontos de interseção de r com ℓ e ℓ' , respectivamente, temos que $AA'C$ é um triângulo, como na figura 17.

Porém, pelo teorema 4.4,

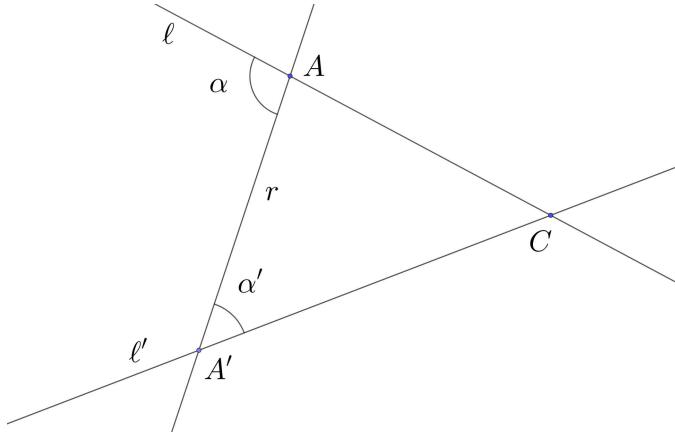
$$\angle CAA' + \angle CA'A + \angle ACA' = \pi.$$

Contudo,

$$\angle CA'A = \alpha' \text{ e } \angle CAA' = \pi - \alpha = \pi - \alpha',$$

em que a última igualdade segue da hipótese de que $\alpha = \alpha'$, pois são medidas de ângulos alternos internos. Substituindo estas igualdades na fórmula para a soma dos ângulos internos do triângulo $AA'C$, obtemos

$$\alpha' + (\pi - \alpha') + \angle ACA' = \pi.$$

FIGURE 17. O triângulo $AA'C$.

Mas, simplificando esta equação, encontramos $\angle ACA' = 0$ que, pelo teorema 3.1 só é possível se $A = A'$, contradizendo o fato de que $l \neq l'$. Logo, l e l' não podem ter nenhum ponto em comum, o que mostra que são paralelas.

Passando, agora, à recíproca, suponhamos que as retas paralelas l e l' são intersectadas por uma reta r . Sejam A e A' os pontos de interseção de r com l e l' , respectivamente e seja $B \in l \cap r^+$. O axioma 4 nos permite construir uma reta ℓ'' , pelo ponto A' , de modo que se $B \in \ell'' \cap r^-$, então

$$(24) \quad \angle BAA' = \angle B''A'A,$$

como na figura 18. Isto significa que os ângulos alternos internos formados quando r

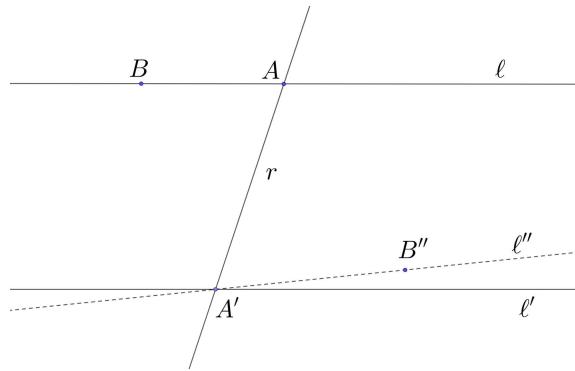


FIGURE 18. Igualdade dos ângulos alternos internos.

intersecta l e l'' são iguais. Mas, pela primeira parte desta demonstração, sabemos que a igualdade destes ângulos implica que ℓ'' e l são paralelas. Portanto, pelo teorema 5.3, $l = \ell''$. Assim, $B'' \in l'$ e (24) nos dá a desejada igualdade entre os ângulos alternos internos formados quando r corta l e l' . \square

Combinando o teorema 6.2 ao Corolário 3.3, que estabelece a igualdade dos ângulos opostos pelo vértice, temos o seguinte resultado.

COROLÁRIO 5.5. *Seja r uma reta que intersecta duas outras retas $\ell_1 \neq \ell_2$ em pontos diferentes de $\ell \cap \ell'$. Os ângulos correspondentes formados pela interseção de r com ℓ e ℓ' são iguais se, e somente se, estas retas são paralelas.*

O ângulo definido pelas semirretas \overrightarrow{AB} e \overrightarrow{AC} no triângulo ABC será denotado por \widehat{CAB} ou simplesmente por \widehat{A} . Diremos que dois triângulos ABC e $A'B'C'$ são *congruentes* se

$$\overline{AB} \equiv \overline{A'B'}, \quad \overline{AC} \equiv \overline{A'C'} \quad \text{e} \quad \overline{BC} \equiv \overline{B'C'}$$

e também

$$\widehat{CAB} \equiv \widehat{C'A'B'}, \quad \widehat{ABC} \equiv \widehat{A'B'C'} \quad \text{e} \quad \widehat{ACB} \equiv \widehat{A'C'B'}.$$

Em outras palavras, dois triângulos são congruentes quando é possível estabelecer uma correspondência entre seus lados e seus ângulos de modo que os lados e os ângulos correspondentes sejam congruentes.

AXIOMAS DE CONGRUÊNCIA.

8. *Sejam ABC e $A'B'C'$ dois triângulos. Se*

$$\overline{AB} \equiv \overline{A'B'}, \quad \overline{AC} \equiv \overline{A'C'} \quad \text{e} \quad \widehat{CAB} \equiv \widehat{C'A'B'}$$

então os triângulos ABC e $A'B'C'$ são congruentes.

Um detalhe que é importante ressaltar é que o ângulo no axioma de congruência 8 *tem que ser aquele que é definido pelos dois lados*, caso contrário, podemos facilmente construir triângulos que *não* são congruentes, mas têm dois lados e um ângulo congruentes. Por exemplo, na figura 19 os triângulos AOB e AOC têm em comum o lado \overline{AO} , e o ângulo \widehat{OAB} , ao passo que os \overline{OB} e \overline{OC} são congruentes porque são raios de uma mesma circunferência. No entanto, estes triângulos claramente não são congruentes.

O axioma 8 corresponde ao caso de congruência de triângulos conhecido como LAL (abreviação de Lado-Ângulo-Lado) e é desta forma que vamos nos referir a ele pelo resto do livro. Os outros dois casos de congruência de triângulos serão provados a partir do LAL. Antes, porém, precisamos de um resultado sobre triângulos *isósceles*, que são aqueles que têm dois ângulos iguais. Mais precisamente, necessitamos mostrar que, nestes triângulos, os ângulos da base também são iguais. É importante que você seja alertado, antes mesmo de ler a demonstração, de que ela trata o mesmo triângulo *como se fossem dois triângulos diferentes, dependendo da ordem em que os vértices são lidos*. A razão para o alerta é que uma demonstração deste gênero às vezes dá a sensação de que estamos “extraíndo algo do nada”.

TEOREMA 5.6. *Se ABC é um triângulo no qual $\overline{AB} \equiv \overline{AC}$, então $\widehat{B} \equiv \widehat{C}$.*

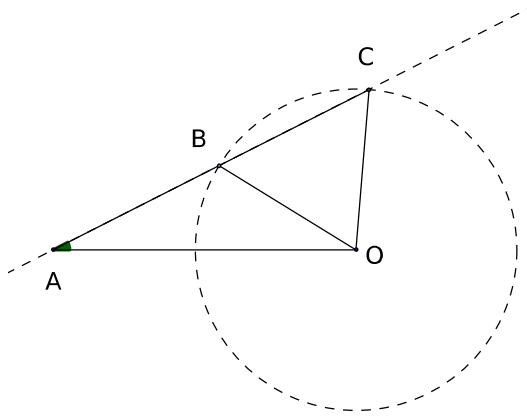


FIGURE 19. Os triângulos AOB e AOC não são congruentes.

PROOF. Como $\overline{AB} \equiv \overline{AC}$, os triângulos ABC e ACB têm os lados correspondentes iguais, assim como o ângulo \widehat{A} entre eles, logo são congruentes por LAL. Portanto, os ângulos correspondentes na base \overline{BC} são iguais, o que nos dá $\widehat{B} \equiv \widehat{C}$. \square

Na demonstração do próximo teorema precisaremos “copiar” um triângulo em uma determinada posição; mas, para fazer isto de maneira legítima, é necessário mostrar que o procedimento de cópia pode ser deduzido a partir dos axiomas que já estabelecemos.

LEMA 5.7. *Sejam ABC um triângulo, ℓ uma reta, ℓ^+ um dos semiplanos definidos por ℓ e $A', Q \in \ell$. Então existem $B' \in \overrightarrow{A'Q}$ e $C' \in \ell^+$ tais que o triângulo $A'B'C'$ é congruente a ABC .*

PROOF. Pelo axioma de congruência 1, existe um ponto $B' \in \overrightarrow{A'Q}$ tal que

$$\overline{AB} \equiv \overline{A'B'};$$

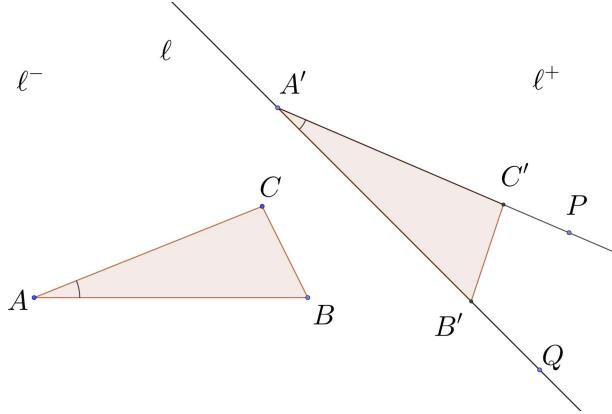
ao passo que, pelo axioma de congruência 4 existe uma semirreta $\overrightarrow{A'P} \in \ell^+$ para a qual

$$\widehat{B'A'P} \equiv \widehat{BAC},$$

como ilustrado na figura 20. Finalmente, pelo axioma de congruência 1 podemos escolher C' na semirreta $\overrightarrow{A'P}$, de forma que

$$\overline{A'C'} \equiv \overline{AC}.$$

Portanto, pelo caso LAL, os triângulos ABC e $A'B'C'$ são congruentes, como desejado. \square

FIGURE 20. Copiando o triângulo ABC

Note que usamos ℓ^+ , no enunciado do lema, apenas como uma maneira de especificar um dos dois semiplanos nos quais um plano é dividido pela reta ℓ . Com isso, podemos provar os outros casos de congruência de triângulos.

TEOREMA 5.8. *Dois triângulos ABC e $A'B'C'$ são congruentes se qualquer uma das alternativas abaixo é verificada:*

- ALA:** $\overline{AB} \equiv \overline{A'B'}$, $\widehat{CAB} \equiv \widehat{C'A'B'}$ e $\widehat{ABC} \equiv \widehat{A'B'C'}$;
- LLL:** $\overline{AB} \equiv \overline{A'B'}$, $\overline{AC} \equiv \overline{A'C'}$ e $\overline{BC} \equiv \overline{B'C'}$.

PROOF. Vamos começar provando ALA. Sejam ABC e $A'B'C'$ dois triângulos nos quais

$$\overline{AB} \equiv \overline{A'B'}, \quad \widehat{CAB} \equiv \widehat{C'A'B'} \quad \text{e} \quad \widehat{ABC} \equiv \widehat{A'B'C'}.$$

Suponhamos, por contradição, que $\overline{B'C'} \neq \overline{BC}$. Pelo axioma de congruência 1 podemos escolher um ponto D , do mesmo lado de B em que está C , tal que

$$\overline{BD} \equiv \overline{B'C'}.$$

Como estamos supondo que $\overline{B'C'} \neq \overline{BC}$, sabemos que D não pode coincidir com C . Temos, portanto, dois casos: ou D está entre B e C , ou C está entre B e D . Vamos fazer a demonstração no primeiro caso e deixar o outro por sua conta. A posição dos vários pontos no triângulo ABC é ilustrada na figura 21. Com isto, os triângulos ABD e $A'B'C'$ têm

$$\overline{AB} \equiv \overline{A'B'}, \quad \widehat{ABC} \equiv \widehat{A'B'C'} \quad \text{e} \quad \overline{BD} \equiv \overline{B'C'}.$$

podemos concluir, pelo caso LAL, que estes triângulos são iguais. Em particular,

$$\widehat{BAD} \equiv \widehat{B'A'C'}.$$

Como, por hipótese, $\widehat{BAC} \equiv \widehat{B'A'C'}$, segue-se do axioma de congruência 5 que

$$(25) \quad \widehat{BAD} \equiv \widehat{BAC},$$

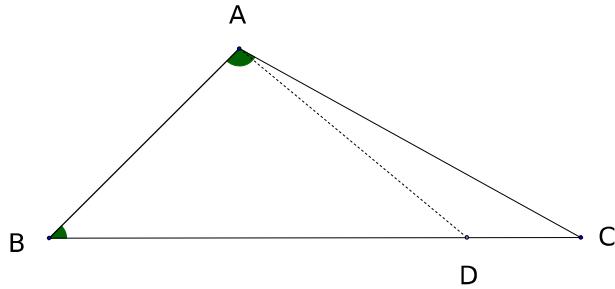


FIGURE 21. O triângulo ABC com o ponto $D \in \overline{BC}$.

Logo,

$$(26) \quad \widehat{BAC} \equiv \widehat{BAD} + \widehat{DAC}.$$

Substituindo (25) em (26), obtemos

$$\widehat{BAC} \equiv \widehat{BAC} + \widehat{DAC},$$

o que só é possível se $D = C$, contradizendo nossa suposição de que

$$\overline{BD} \equiv \overline{B'C'} \neq \overline{BC},$$

e concluindo a demonstração do caso ALA.

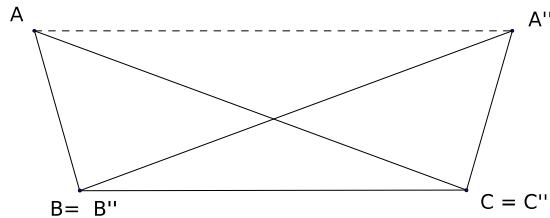


FIGURE 22. O triângulo $A'B'C'$ copiado como $A''BC$.

Passando, agora, ao caso LLL, suponhamos que os triângulos ABC e $A'B'C'$ tenham os três lados iguais. Pelo lema 5.7 fazer uma cópia $A''B''C''$ de $A'B'C'$, do mesmo lado da reta BC no qual está o ponto A , de modo que $\overline{BC} \equiv \overline{B''C''}$.

Como, por construção, $\overline{A''B''} \equiv \overline{A'B'}$ e, por hipótese, $\overline{AB} \equiv \overline{A'B'}$, o axioma de congruência 2 nos permite concluir que

$$\overline{AB} \equiv \overline{A''B''}.$$

Logo,

$$\overline{AB} \equiv \overline{A''B}.$$

pois construímos $B'' = B$. Portanto, o triângulo $AA''B$ é isósceles e o teorema 5.6 nos permite concluir que

$$\widehat{BAA''} \equiv \widehat{BA''A}.$$

Um argumento análogo mostra que o triângulo $AA''C$ é isósceles, de modo que $\widehat{CAA''} \equiv \widehat{CA''A}$. Resumindo, mostramos até aqui que

$$(27) \quad \widehat{BAA''} \equiv \widehat{BA''A} \quad \text{e que} \quad \widehat{CAA''} \equiv \widehat{CA''A}.$$

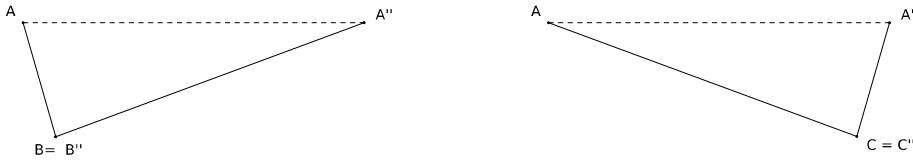


FIGURE 23. Os triângulos isósceles BAA'' e CAA'' .

Por outro lado,

$$\widehat{BAA''} = \widehat{BAC} + \widehat{CAA''} \quad \text{e} \quad \widehat{CA''A} = \widehat{CA''B} + \widehat{BA''A};$$

donde,

$$\widehat{BAC} = \widehat{BAA''} - \widehat{CAA''} \quad \text{e} \quad \widehat{CA''B} = \widehat{CA''A} - \widehat{BA''A};$$

Logo,

$$\widehat{BAC} + \widehat{CA''B} = (\widehat{BAA''} - \widehat{CAA''}) + (\widehat{CA''A} - \widehat{BA''A}).$$

Rearrumando, obtemos, de (27) que

$$\widehat{BAC} + \widehat{CA''B} \equiv (\widehat{BAA''} - \widehat{BA''A}) + (\widehat{CA''A} - \widehat{CAA''}) \equiv 0.$$

Mas isto só é possível se

$$\widehat{BAC} = -\widehat{BA''C},$$

que equivale a dizer que $A'' = A$, como queríamos mostrar. \square

6. Paralelas

Nesta seção provaremos o bem conhecido fato sobre a soma dos ângulos de um triângulo. Lembre-se que um ângulo é *raso* quando as duas semirretas que o definem são colineares e apontam em direções opostas.

TEOREMA 6.1. *A soma de dois ângulos de um triângulo não pode ser igual a um ângulo raso.*

PROOF. Suponhamos, por contradição, que ABC é um triângulo, no qual a soma dos ângulos \widehat{BAC} e \widehat{BCA} é igual a um ângulo raso. Pelo axioma de congruência 1 existe um ponto D sobre a reta BC tal que C está entre B e D e os segmentos \overline{AB} e \overline{CD} são congruentes, como ilustrado na figura 24.

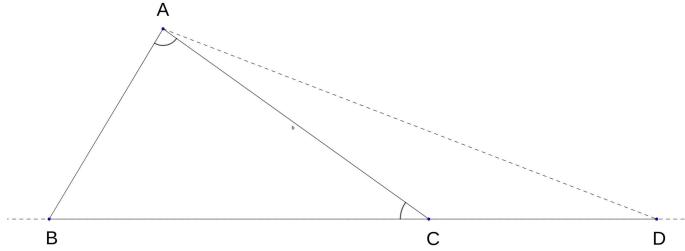


FIGURE 24. No triângulo ABC o ângulo $\widehat{BAC} + \widehat{BCA}$ é raso.

Contudo,

$$\widehat{BAC} + \widehat{BCA} = \text{raso} = \widehat{BCA} + \widehat{ACD}$$

pois os pontos B , C e D pertencem a uma mesma reta. Logo, $\widehat{BAC} \equiv \widehat{ACD}$. Mas isto significa que, nos triângulos ABC e ACD , os lados \overline{AB} e \overline{CD} são congruentes e os ângulos \widehat{BAC} e \widehat{ACD} são congruentes; como este ângulo está situado entre os lados congruentes \overline{AC} , que é comum aos dois triângulos, podemos deduzir do axioma LAL que ABC e ACD são triângulos congruentes. Em particular, os ângulos \widehat{ACB} e \widehat{CAD} têm que ser congruentes. Entretanto,

$$\widehat{BAD} \equiv \widehat{BAC} + \widehat{CAD} \equiv \widehat{BAC} + \widehat{ACB}$$

que estamos supondo ser igual a dois ângulos rasos, o que equivale a dizer que B , A e D são colineares. Contudo, isto implica que as retas AB e BD se intersectam em B e D . Porém, pelo teorema 2.2, isto só pode acontecer se $AB = BD$. Como $C \in BD = AB$, os pontos A , B e C não podem definir um triângulo, gerando, assim, uma contradição.

Portanto, a suposição de que $\widehat{BAC} + \widehat{BCA}$ é raso é falsa, o que conclui a demonstração do teorema. \square

Os próximos dois teoremas dizem respeito à situação ilustrada na figura 25. Nela as retas paralelas AB e $A'B'$ são cortadas pela reta AA' , e os pontos B e B' estão situados em lados opostos da reta AA' . Neste caso, dizemos que os ângulos $\widehat{BAA'}$ e $\widehat{B'A'A}$ são *alternos internos*.

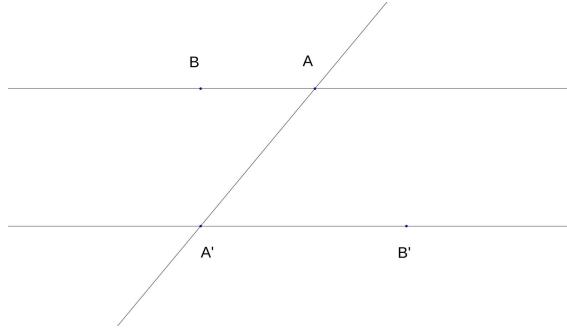


FIGURE 25. Posição das retas e pontos no enunciado dos teoremas 6.2 e 6.7.

TEOREMA 6.2. *Sejam AB e $A'B'$ duas retas e seja AA' uma reta transversal às outras duas, de modo que B e B' estejam em semiplanos diferentes relativamente a AA' . Se $\widehat{BAA'} \equiv \widehat{B'A'A}$, então as retas AB e $A'B'$ são paralelas.*

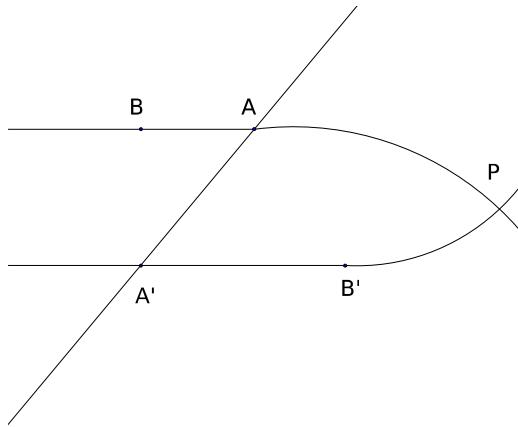


FIGURE 26. Posição dos pontos A, B, A', B' e P .

PROOF. Suponhamos, por contradição, que AB e $A'B'$ não são paralelas. Pelo teorema 2.2 estas retas têm que se encontrar em um único ponto P . Sem perda de generalidade podemos supor que P pertence ao mesmo semiplano que B' , relativamente à reta

AA' . Se for este o caso, a figura que ilustra o teorema 6.2 é mais parecida com a figura 26 do que com a figura 25. Em particular, os pontos A , A' e P formam um triângulo. Contudo, sabemos que o ângulo $\widehat{BAA'} \equiv \widehat{B'A'A}$ é suplementar de $\widehat{A'AP}$, pois B , A e P são colineares. Portanto, a soma $\widehat{B'A'A} + \widehat{A'AP}$ dá um ângulo raso, contradizendo o teorema 6.1. Logo, as retas AB e $A'B'$ não podem se encontrar em um ponto, o que significa que são paralelas. \square

Nosso próximo passo é provar a recíproca do teorema acima. como isto requer que tracemos uma reta paralela a uma reta dada por um ponto fora dela, precisamos mostrar primeiro que é possível justificar a existência de tal reta a partir dos axiomas.

LEMA 6.3. *Sejam ℓ uma reta e P um ponto fora de ℓ . Existe uma reta ℓ' à qual P pertence e que é perpendicular a ℓ .*

PROOF. Escolha dois pontos O e A em ℓ . Pelo axioma de incidência 1 podemos construir a reta OP e, pelo axioma de congruência 4 podemos construir uma semirreta \overrightarrow{OR} , do lado oposto de P relativamente a ℓ tal que $\widehat{POA} \equiv \widehat{AOR}$. Finalmente, pelo axioma de congruência 1, existe um ponto Q na semirreta \overrightarrow{OR} tal que $\widehat{OP} \equiv \widehat{OQ}$. Pelo axioma de ordenação 4, as retas OA e PQ se intersectam em um ponto, que chamaremos de C . A posição de todos estes pontos é ilustrada na figura 27.

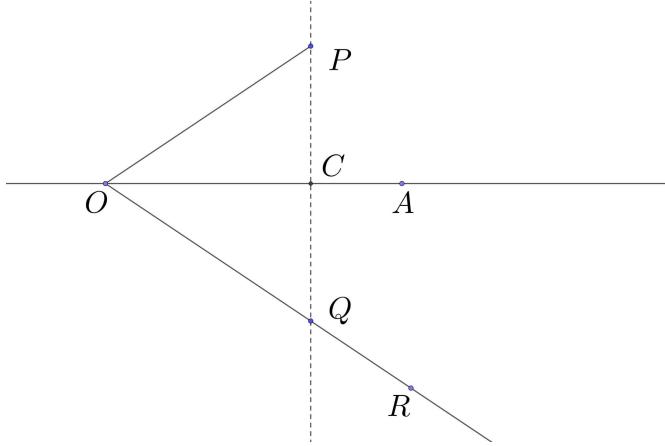


FIGURE 27. Construindo a perpendicular por um ponto

Nosso objetivo é mostrar que o ângulo \widehat{OCP} é reto. Contudo, como $C \in PQ$, o ângulo \widehat{PCQ} é raso, de modo que basta provar que $\widehat{OCP} \equiv \widehat{OCQ}$. Mas, por construção,

$$\widehat{POA} \equiv \widehat{AOR} \text{ e } \widehat{OP} \equiv \widehat{OQ},$$

ao passo que \overline{OC} é um lado comum aos triângulos POC e COQ . Portanto, pelo caso LAL estes triângulos são congruentes. Em particular, isto significa que os ângulos \widehat{OCP} e \widehat{OCQ} têm que ser congruentes, como queríamos provar. \square

Este lema tem várias consequências importantes, a primeira das quais, provavelmente, não passaria pela sua cabeça que teríamos que provar.

COROLÁRIO 6.4. *Ângulos retos existem.*

As outras duas consequências não são tão imediatas, por isso explicaremos em linhas gerais como proceder para prová-las.

COROLÁRIO 6.5. *Sejam ℓ uma reta e $P \in \ell$. Existe uma reta ℓ' à qual P pertence e que é perpendicular a ℓ .*

COROLÁRIO 6.6. *Sejam ℓ uma reta e P um ponto fora de ℓ . Existe uma reta ℓ' à qual P pertence e que é paralela a ℓ .*

Para provar o corolário 6.5 basta usar o Lema 6.3 para construir um ângulo perpendicular a ℓ em algum lugar e transferi-lo para o ponto P usando o axioma de congruência 4. Já a demonstração do corolário 6.6 é feita construindo uma perpendicular ℓ' a ℓ por P e, depois, uma perpendicular ℓ'' a ℓ' por P , como ilustrado na figura 28. As retas ℓ e ℓ'' são paralelas pelo teorema 6.2.

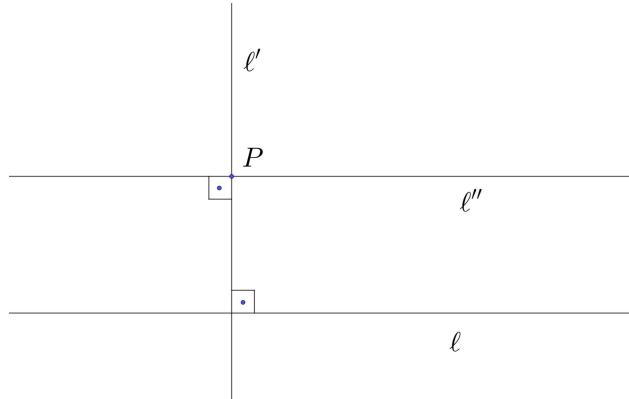
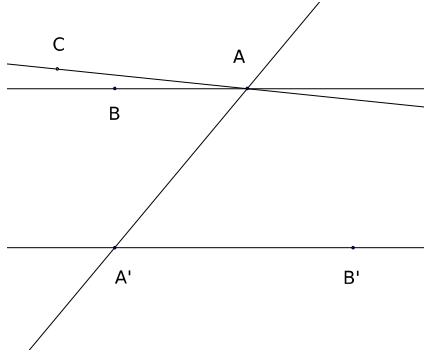


FIGURE 28. Construindo uma paralela

Com isso estamos prontos para provar a recíproca do teorema 6.2. A posição relativa dos pontos no enunciado do próximo teorema é a mesma da figura 25.

TEOREMA 6.7. *Sejam AB e $A'B'$ duas retas e seja AA' uma reta transversal às outras duas, de modo que B e B' estejam em semiplanos diferentes relativamente à reta AA' . Se as retas AB e $A'B'$ são paralelas então $\widehat{BAA'} \equiv \widehat{B'A'A}$.*

PROOF. Suponhamos, por contradição, que as retas AB e $A'B'$ são paralelas, mas que os ângulos alternos internos $\widehat{BAA'}$ e $\widehat{AA'B'}$ não são congruentes. Pelo axioma de congruência 4, existe uma reta AC de modo que o ângulo $\widehat{CAA'} \equiv \widehat{AA'B'}$, como ilustrado na figura 29.

FIGURE 29. Construção da reta AC .

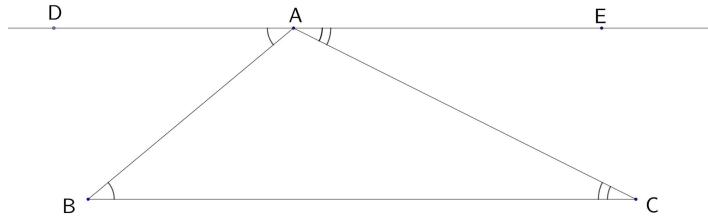
Entretanto, pelo teorema 6.2, a igualdade entre os ângulos alternos internos implica que as retas $A'B'$ e AC têm que ser paralelas. Como AB é paralela a $A'B'$ por hipótese, o axioma de paralelismo implica que $AC = AB$. Portanto, podemos deduzir da maneira como AC foi construída que

$$\widehat{AA'B} \equiv \widehat{CAA'} = \widehat{BAA'},$$

o que contradiz a suposição que fizemos no início da demonstração; provando, assim, o teorema. \square

TEOREMA 6.8. *A soma dos ângulos internos de um triângulo é um ângulo raso.*

PROOF. Seja ABC um triângulo. Pelo axioma de paralelismo existe uma única reta DE paralela ao lado BC pelo vértice A do triângulo, como ilustrado na figura 30.

FIGURE 30. A reta DE , paralela ao lado \overline{BC} .

Mas os ângulos \widehat{DAB} e \widehat{ABC} são alternos internos relativamente à reta transversal AB . Portanto, pelo teorema 6.7 temos que $\widehat{DAB} \equiv \widehat{ABC}$. Analogamente, $\widehat{EAC} \equiv \widehat{ACB}$



FIGURE 31. Richard Dedekind

são alternos internos relativamente à reta AC , de modo que $\widehat{EAC} \equiv \widehat{ACB}$. Logo, a soma dos ângulos internos do triângulo ABC é

$$\widehat{BAC} + \widehat{ACB} + \widehat{ABC} \equiv \widehat{EAC} + \widehat{ACB} + \widehat{DAB}$$

que é claramente um ângulo raso porque os ponto D , A e E são colineares. \square

7. Medindo comprimento e área

Será que você percebeu que nossa exposição do enfoque dedutivo da geometria não envolveu nenhuma medição? Ainda que você tenha presumido que dois segmentos são congruentes quando têm o mesmo tamanho, isto não faz, até aqui, parte das regras do jogo dedutivo. Na verdade, no caso específico dos *Elementos*, medidas de comprimento, ângulo ou área não aparecem em nenhum lugar.

Contudo, para podermos estudar triângulos que são proporcionais (ou semelhantes, para usar a terminologia tradicional), precisamos ser capazes de medir os comprimentos dos seus lados e suas áreas. Infelizmente, fazer isto com o mesmo grau de detalhe com que vínhamos tratando a geometria plana é muito difícil. O problema está na própria definição dos números reais. Por isso, teremos que nos contentar em listar as propriedades dos comprimentos de segmentos, sem investigar o tema mais a fundo. Os gregos contornaram este problema definindo diretamente proporções entre segmentos e áreas. Esta teoria, devida a Eudoxo de Cnido, é exposta em detalhes no Livro V dos *Elementos*. Este é um dos livros mais difíceis de todo o tratado de Euclides e foi a fonte de inspiração para uma das primeiras construções satisfatórias dos números reais, dada por R. Dedekind em 1872.

Levando tudo isto em consideração, você pode tomar o conteúdo deste parágrafo como um novo conjunto de axiomas, sem, contudo, esquecer que, em troca de muito esforço,

poderíamos provar a maior parte das propriedades abaixo a partir de apenas dois axiomas. A cada segmento de reta \overline{AB} associaremos um número real não negativo, o *comprimento* deste segmento. Este número satisfaz as seguintes propriedades:

Propriedade 1: \overline{PP} tem comprimento zero, qualquer que seja o ponto P ;

Propriedade 2: dois segmentos têm os mesmos comprimentos se, e somente se, são congruentes;

Propriedade 3: B é um ponto entre A e C se, e somente se, o comprimento de \overline{AC} é igual à soma dos comprimentos de \overline{AB} e \overline{BC} ;

Propriedade 4: todo número real é o comprimento de algum segmento e conhecemos um segmento \overline{OI} de comprimento igual a 1.

Para não complicar demais as fórmulas, denotaremos o comprimento de um segmento pelo mesmo símbolo usado para denotar o próprio segmento. Por exemplo, para indicar que o segmento \overline{AB} tem comprimento 5, escreveremos simplesmente $\overline{AB} = 5$. Em geral, o contexto torna claro se o símbolo se refere ao segmento ou a seu comprimento.

As regras para medir segmentos que estabelecemos acima nos permitem provar, de maneira rigorosa, as fórmulas para áreas de polígonos que investigamos, à maneira egípcia, no início deste capítulo. Antes, porém, precisamos definir o que queremos dizer quando falamos em polígonos. Sejam P_0, P_1, \dots, P_m pontos do plano tais que

1. $P_j \in (P_i P_{i+1})^+$ para todo $i = 0, \dots, m - 1$ e todo $j \neq i, i + 1$;
2. $P_j \in (P_m P_0)^+$ para todo $j \neq 0, m$.

O *polígono* definido pelos pontos P_0, P_1, \dots, P_m é a região

$$\mathcal{P} = (P_0 P_1)^+ \cap (P_1 P_2)^+ \cap \dots \cap (P_{m-1} P_m)^+ \cap (P_m P_0)^+.$$

Na figura 32 temos um polígono à esquerda e uma figura que não é um polígono à direita. Os vértices P_i e P_j do polígono P_0, P_1, \dots, P_m são *consecutivos* se

- $i < m$ e $j = i + 1$, ou
- $i = m$ e $j = 0$.

Os *lados* do polígono P_0, P_1, \dots, P_m são os segmentos que ligam vértices consecutivos. Para evitar uma multiplicidade de casos nas demonstrações, suporemos, de agora em diante, que se $k = m$, então $k + 1 = 0$. Em outras palavras, usaremos aritmética módulo m nos índices dos pontos que definem um polígono.

O próximo resultado é uma consequência do axioma de ordenação 8 e será movido para a seção onde está o teorema em uma versão futura da apostila.

COROLÁRIO 7.1. *Seja \widehat{AOB} um ângulo e $C \in \widehat{AOB}$. A semirreta \overrightarrow{OC} intersecta \overline{AB} .*

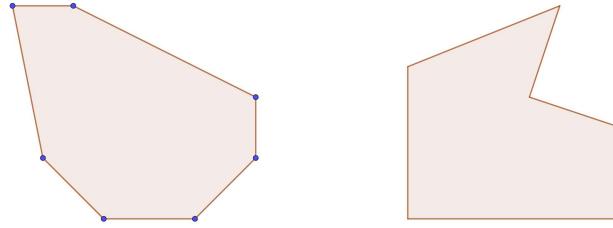


FIGURE 32. A figura da esquerda é um polígono, mas não a da direita.

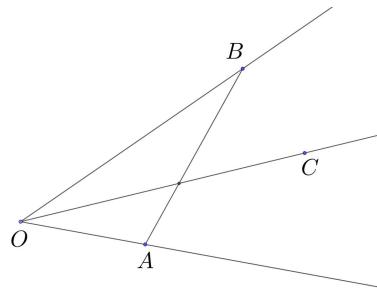


FIGURE 33. Pontos e semirretas no Corolário 7.1

PROOF. Seja

$$\widehat{AOB} = (OB)^+ \cap (OA)^+$$

Pelo axioma de ordenação 1 existe um ponto $D \in OA \setminus \{O, A\}$ tal que $O \in \overline{DA}$. Como, por construção, OC corta \overline{DA} em um ponto interior, temos, pelo lema A, que OC corta \overline{AB} ou \overline{DB} . Contudo, como $C \in \widehat{AOB}$, podemos concluir que $\overrightarrow{OC} \subset (OA)^+$. Analogamente, como $D \in (OB)^-$, segue-se que $\overrightarrow{DB} \subset (OB)^-$. Logo, \overrightarrow{OC} não intersecta \overrightarrow{DB} . Como B e \overrightarrow{OC} estão do mesmo lado de OA , temos que OC não corta \overline{DB} . Portanto, pelo axioma de ordenação 4, OC corta \overline{AB} .

□

LEMA 7.2. *Dois vértices não consecutivos dividem um polígono em dois polígonos com um lado comum.*

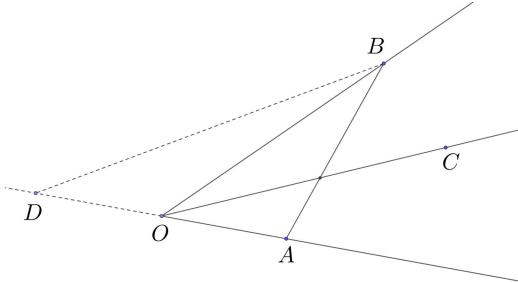


FIGURE 34. Pontos e semirretas na demonstração do Corolário 7.1

PROOF. Sejam P_i e P_j dois vértices não consecutivos do polígono \mathcal{P} definido pelos pontos

$$P_0, P_1, \dots, P_m$$

e seja ℓ a reta P_iP_j . Definimos

$$\mathcal{P}^+ = \mathcal{P} \cap \ell^+ \text{ e } \mathcal{P}^- = \mathcal{P} \cap \ell^-$$

de modo que $\mathcal{P} = \mathcal{P}^+ \cup \mathcal{P}^-$. Para provar o lema basta mostrar que \mathcal{P}^+ e \mathcal{P}^- são ambos polígonos não vazios.

Mostraremos, primeiramente, que \mathcal{P}^+ e \mathcal{P}^- não são ambos vazios. Para isto basta considerar o triângulo $P_{i-1}P_iP_{i+1}$. Como $P_j \in \widehat{P_{i-1}P_iP_{i+1}}$, segue-se do corolário 7.1 que a reta P_iP_j intersecta o segmento $\overline{P_{i-1}P_{i+1}}$ em um ponto interior. Logo, P_{i-1} e P_{i+1} estão situados em lados opostos de P_iP_j ; portanto, um deles está em \mathcal{P}^+ e o outro em \mathcal{P}^- , mostrando que nenhum dos dois é vazio.

Para mostrar que \mathcal{P}^- e \mathcal{P}^+ são ambos polígonos basta provar que os únicos vértices de \mathcal{P} na reta ℓ são P_i e P_j . Suponhamos, por contradição, que ℓ também contém o vértice P_k , com $k \neq i, j$. Logo, um destes três vértices estará entre os outros dois. Em primeiro lugar,

$$k \notin \{i-1, i+1, j-1, j+1\};$$

caso contrário, haveria três vértices em um mesmo lado de \mathcal{P} , o que contradiz a definição de polígono. Sem perda de generalidade, podemos supor que P_k está entre P_i e P_j . Como isto implica que P_i e P_j estão de lados opostos da reta P_kP_{k+1} obtemos, novamente, uma contradição com a definição de polígono, o que completa a demonstração. \square

Para definir rigorosamente a área de polígonos precisamos de mais três axiomas. A *altura* de um triângulo ABC relativamente ao lado \overline{BC} é o segmento de reta perpendicular a \overline{BC} pelo ponto A .

AXIOMAS DE ÁREA.

- 1.** A área de um triângulo é igual à metade do produto dos comprimentos de um de seus lados e da altura sobre este lado;
- 2.** Dois triângulos congruentes têm a mesma área.
- 3.** A área de um polígono é igual à soma das áreas dos dois polígonos em que fica dividido pelo segmento que une dois vértices que não são adjacentes.

Ao contrário do que você talvez estivesse esperando, tendo em vista a maneira como introduzimos áreas na seção 1, o primeiro axioma estabelece como medir áreas de triângulos, e não de quadrados. A razão para isso é que qualquer polígono pode ser facilmente subdividido em triângulos, mas não em quadrados. Esta inversão é típica de tratamentos axiomáticos, porque, frequentemente, a maneira mais natural de introduzir um conceito não é a mais adequada quando se trata de um tratamento puramente dedutivo.

A partir destes axiomas podemos deduzir todas as fórmulas de áreas de polígonos que havíamos obtido usando o método de copiar e colar. Por exemplo, um paralelogramo $ABCD$ é um polígono para o qual

$$\overline{AB} \equiv \overline{CD} \text{ e } \overline{AD} \equiv \overline{BC}.$$

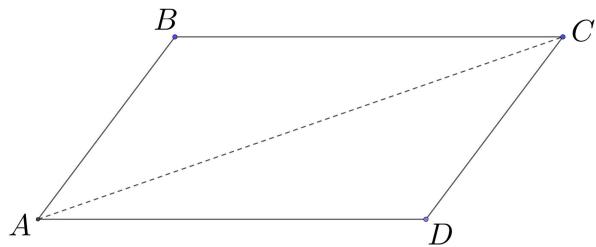


FIGURE 35. Demonstrando a fórmula da área do paralelogramo

Podemos dividir $ABCD$ nos triângulos ABC e ACD , que têm em comum o segmento \overline{AC} . Pelo caso LLL, estes triângulos são congruentes. Logo, pelo axioma de área 2, ABC e ACD têm a mesma área. Portanto, se h for a altura de ACD , temos, pelo axioma de área 3, que

$$\text{área}(ABCD) = 2 \times \left(\frac{\overline{AD} \cdot h}{2} \right) = \overline{AD} \cdot h.$$

Como $ABC \equiv ACD$, temos, pelo teorema 6.2, que as retas BC e AD são paralelas, de modo que h é o comprimento de qualquer segmento que liga um ponto de AD perpendicularmente a um ponto de BC . A demonstração da fórmula da área do trapézio é semelhante e ficará como exercício.

8. Semelhança de triângulos

Triângulos que são congruentes são essencialmente iguais, apenas estão localizados em lugares diferentes. Mas em muitas questões práticas, nos deparamos com triângulos que não são congruentes, mas sim semelhantes; isto é, diferem apenas na escala. Por exemplo, vimos na seção 2 que Tales teria utilizado triângulos semelhantes para medir a altura de uma pirâmide. A semelhança de figuras é o tema do Livro VI dos *Elementos* e é de lá que vêm os resultados que estudaremos nesta seção.

Começaremos dando uma definição precisa da noção de semelhança de triângulos. Dois triângulos ABC e $A'B'C'$ são *semelhantes* se

$$(28) \quad \frac{\overline{AB}}{\overline{A'B'}} = \frac{\overline{AC}}{\overline{A'C'}} = \frac{\overline{BC}}{\overline{B'C'}}.$$

Provaremos primeiro o teorema que Tales teria utilizado para determinar a altura de uma pirâmide; veja seção 2. A demonstração fará uso de resultados sobre áreas de triângulos que provamos anteriormente.

TEOREMA 8.1. *Seja ABC um triângulo e seja B' um ponto do lado \overline{AB} , diferente dos vértices A e B . Se C' for o ponto de interseção da paralela a BC por B' , então*

$$\frac{\overline{AB}}{\overline{AB'}} = \frac{\overline{AC}}{\overline{AC'}}.$$

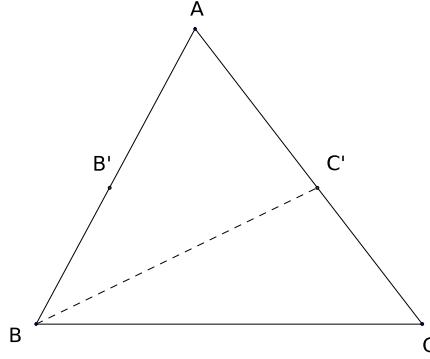
A construção no teorema 8.1 requer que a reta paralela a BC por B' corte o lado \overline{AC} do triângulo, que é uma consequência do lema A, na página ??.

PROOF. A posição relativa dos triângulos ABC e $AB'C'$ é ilustrada na figura 36. Como os triângulos ABC e BCC' , têm o vértice B em comum e os lados \overline{AC} e $\overline{CC'}$ são colineares, segue-se que

$$\text{altura}(ABC) = \text{altura}(BCC').$$

Portanto,

$$(29) \quad \frac{\text{área}(BCC')}{\text{área}(ABC)} = \frac{\overline{CC'}}{\overline{AC}}.$$

FIGURE 36. Os triângulos ABC e BCC' .

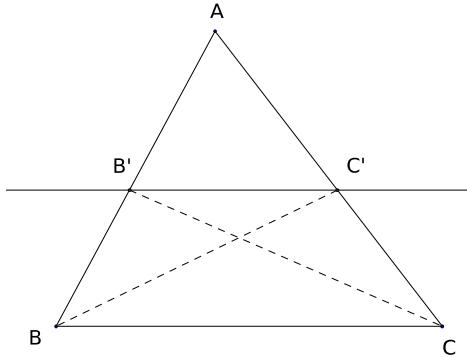
Repetindo o mesmo argumento, desta vez com os triângulos ABC e $BB'C$, obtemos

$$(30) \quad \frac{\text{área}(BB'C)}{\text{área}(ABC)} = \frac{\overline{BB'}}{\overline{AB}}.$$

Por outro lado, como os triângulos $BB'C$ e $BC'C$ têm a mesma base e a mesma altura, então

$$(31) \quad \text{área}(BB'C) = \text{área}(BC'C).$$

Combinando (29) com (31),

FIGURE 37. Os triângulos $BB'C$ e $BC'C$.

$$\frac{\overline{CC'}}{\overline{AC}} = \frac{\text{área}(BCC')}{\text{área}(ABC)} = \frac{\text{área}(BB'C)}{\text{área}(ABC)}.$$

Mas, levando em conta (30),

$$(32) \quad \frac{\overline{CC'}}{\overline{AC}} = \frac{\text{área}(BB'C)}{\text{área}(ABC)} = \frac{\overline{BB'}}{\overline{AB}}.$$

Contudo, $\overline{AB} = \overline{AB'} + \overline{B'B}$, de modo que

$$1 = \frac{\overline{AB'}}{\overline{AB}} + \frac{\overline{B'B}}{\overline{AB}}.$$

Substituindo a igualdade entre o primeiro e o terceiro termos de (32) nesta equação, obtemos

$$1 = \frac{\overline{AB'}}{\overline{AB}} + \frac{\overline{CC'}}{\overline{AC}}.$$

Analogamente, $\overline{AC} = \overline{AC'} + \overline{C'C}$ nos dá

$$1 = \frac{\overline{AC'}}{\overline{AC}} + \frac{\overline{C'C}}{\overline{AC}},$$

o que nos permite concluir que

$$\frac{\overline{AC'}}{\overline{AC}} = \frac{\overline{AB'}}{\overline{AB}},$$

completando, assim, a demonstração. \square

 Há uma pequena sutileza na demonstração do teorema anterior que merece um comentário a parte. O argumento começa com a observação de que os triângulos ABC e $BC'C$ têm a mesma altura, mas o mesmo vale para os triângulos ABC e ABC' , de modo que também é verdade que

$$\frac{\text{área}(ABC')}{\text{área}(ABC)} = \frac{\overline{AC'}}{\overline{AC}}.$$

Aplicando o mesmo argumento aos triângulos ABC e $AB'C$, obteríamos

$$\frac{\text{área}(AB'C)}{\text{área}(ABC)} = \frac{\overline{AB'}}{\overline{AB}}.$$

À primeira vista estas duas relações parecem muito melhores que (29) e (30), porque envolvem diretamente as frações

$$\frac{\overline{AC'}}{\overline{AC}} \text{ e } \frac{\overline{AB'}}{\overline{AB}},$$

cuja igualdade queremos provar. Porém, para que o argumento pudesse continuar, precisaríamos mostrar que as áreas dos triângulos ABC' e $AB'C$ são iguais. Contudo, enquanto os triângulos $BC'C$ e $CB'B$ têm, claramente, mesma base e altura e, portanto, mesma área, nenhuma relação semelhante vale para os triângulos ABC' e $AB'C$. Com isso, essa versão do argumento encalha, ao passo que a versão mais complicada, usada na demonstração do teorema 8.1 funciona perfeitamente bem.

O resultado que realmente desejamos é o caso AAA de semelhança de triângulos que provaremos no próximo teorema.

TEOREMA 8.2. *Dois triângulos ABC e $A'B'C'$ nos quais*

$$\widehat{A} = \widehat{A}', \quad \widehat{B} = \widehat{B}' \quad \text{e} \quad \widehat{C} = \widehat{C}'$$

são semelhantes.

PROOF. Sem perda de generalidade, podemos supor que $\overline{A'B'} < \overline{AB}$. Começamos fazendo uma cópia ADE do triângulo $A'B'C'$ de modo que A' vá parar sobre A , $\overline{AD} = \overline{A'B'}$ caia sobre o segmento \overline{AB} e o ponto E , correspondente a C' , esteja no mesmo semiplano da reta AB que o ponto C . Como, por hipótese,

$$\widehat{B'A'C'} = \widehat{BAC}$$

e, por construção,

$$\widehat{B'A'C'} = \widehat{DAE},$$

temos que

$$\widehat{BAC} = \widehat{DAE}.$$

Logo a semirreta \overrightarrow{AE} coincide com \overrightarrow{AC} , como ilustrado na figura 38.

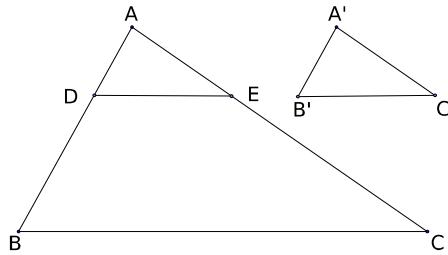


FIGURE 38. O triângulo DAE é uma cópia de $A'B'C'$.

Mas, com isto, estamos exatamente na situação do teorema de Tales, o que nos permite concluir que

$$\frac{\overline{AB}}{\overline{A'B'}} = \frac{\overline{AB}}{\overline{AD}} = \frac{\overline{AC}}{\overline{AE}} = \frac{\overline{AC}}{\overline{A'C'}}.$$

Com isto provamos a primeira igualdade na equação (28). Para provar a segunda igualdade em (28) faremos a figura, desta vez copiando $A'B'C'$ em BFG , como ilustrado na figura 39.

Aplicando o teorema de Tales aos triângulos ABC e FBG , concluímos que

$$\frac{BA}{B'A'} = \frac{BA}{BF} = \frac{BC}{BG} = \frac{BC}{B'C'},$$

que nos dá a igualdade que faltava, porque $\overline{BA} = \overline{AB}$ e $\overline{B'A'} = \overline{A'B'}$. \square

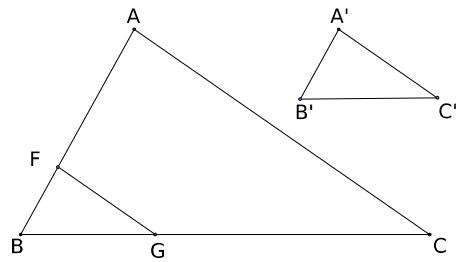


FIGURE 39. O triângulo FBG é uma cópia de $A'B'C'$.

Part 3

Funções reais

CHAPTER 4

Funções reais

Neste capítulo começamos nosso estudo das funções reais e suas propriedades básicas. Nos capítulos seguintes cobriremos as funções polinomiais, as funções trigonométricas e as funções exponencial e logaritmo, a partir das quais é possível construir as chamadas funções elementares, que são aquelas que aparecem nos cursos básicos de cálculo.

1. Definições e exemplos

A única parte da mecânica explorada pelos matemáticos gregos foi a estática, que é o estudo dos corpos em equilíbrio. Mesmo a matemática utilizada na investigação da queda dos corpos, iniciado na Idade Média e que culminou nos trabalhos de Galileu, baseava-se apenas em resultados conhecidos desde Apolônio de Perga. O conceito moderno de função surgiu com o desenvolvimento da dinâmica por Huygens e Newton e a criação do cálculo infinitesimal por Leibniz e pelo próprio Newton. A posição e a velocidade de um corpo em movimento variam, inevitavelmente, com o tempo e estavam entre as primeiras funções a serem estudadas.



FIGURE 1. Leonhard Euler e Augustin-Louis Cauchy

O primeiro a utilizar o conceito de função como fundamento do cálculo infinitesimal parece ter sido Euler, em seu livro *Introductio in analysin infinitorum* de 1748 [7, p. 405].

Contudo, a definição moderna de função só se estabeleceu no século XIX. Em seu *Cours d'Analyse*, publicado em 1821, Cauchy escreve

para que uma função de uma só variável esteja completamente determinada, é necessário e suficiente que cada valor atribuído à variável seja possível deduzir um valor correspondente da função; [5, p. 244].

Note que não há, nesta definição, nenhuma menção à forma como a cada valor da variável é associado o valor da função. A generalidade desta definição levou, ainda no século XIX, à definição de funções cada vez mais exóticas, causando uma cascata de problemas que só vieram a ser completamente esclarecidos no século XX.

Com isso estamos prontos para nossa definição oficial de função. Sejam X e Y conjuntos, uma função $f : X \rightarrow Y$ é uma regra que, a cada $x \in X$ associa um único elemento $f(x) \in Y$. Dizemos que $f(x)$ é a *imagem* do argumento x pela função f . O conjunto X , ao qual pertencem os argumentos de f , é o *domínio* de f e o conjunto Y , ao qual pertencem as imagens dos seus argumentos, é o *contradomínio* de f .

As funções que estudaremos nesta parte do curso terão sempre subconjuntos de \mathbb{R} como domínio e contradomínio. Por exemplo, dado $c \in \mathbb{R}$ podemos definir a função constante $\widehat{c} : \mathbb{R} \rightarrow \mathbb{R}$ pela regra $\widehat{c}(x) = c$. Um exemplo importante de função constante é a função nula, que a cada número real associa o valor 0. Outra família muito importante de funções é a das funções lineares. Digamos que $a, b \in \mathbb{R}$ e que $a \neq 0$, definimos a função linear $\ell : \mathbb{R} \rightarrow \mathbb{R}$ pela regra

$$\ell(x) = ax + b.$$

Como no caso da função constante, há uma função linear especialmente importante: conhecida como função identidade e denotada por id , esta função é definida pela regra $\text{id}(x) = x$. Nosso terceiro exemplo é a função quadrática $q : \mathbb{R} \rightarrow \mathbb{R}$ definida por

$$q(x) = ax^2 + bx + c,$$

em que $a \neq 0$, b e c são números reais. Estas três funções são exemplos de *funções polinomiais*, que estudaremos em mais detalhe no capítulo 5.

Até aqui consideramos apenas funções cujas regras são definidas usando apenas as operações de adição, subtração e multiplicação. Introduzindo a divisão, podemos definir as funções r_1 , r_2 e r_3 , cujas regras são

$$r_1(x) = \frac{1}{x}, \quad r_2(x) = \frac{x+1}{x-2} \quad \text{e} \quad r_3(x) = \frac{x+1}{x^2-2}.$$

Ao contrário das funções constantes, lineares e quadráticas, as regras que definem estas funções não fazem sentido quando o denominador se anula, de modo que estes números precisam ser excluídos dos seus domínios. Portanto, os domínios de r_1 , r_2 e r_3 são, respectivamente,

$$\mathbb{R} \setminus \{0\}, \quad \mathbb{R} \setminus \{2\} \quad \text{e} \quad \mathbb{R} \setminus \{-\sqrt{2}, \sqrt{2}\}.$$

Agora que já temos vários exemplos de funções, é conveniente chamar sua atenção para a diferença entre o nome de uma função e seu valor no argumento. Assim, nos exemplos acima, r_1 , r_2 e r_3 são os nomes das funções, ao passo que $r_1(x)$, $r_2(x)$ e $r_3(x)$ são aos valores destas funções no argumento x . De maneira geral os matemáticos não são muito consistentes em fazer esta diferença e frequentemente se referem à “função $r_1(x)$ ” quando deveriam dizer a “função r_1 ”. Contudo, é útil ter isto em mente, porque muitas vezes é necessário diferenciar entre o nome da função e seu valor em x . Resumindo r_1 é o nome da função, $r_1(x)$ é o valor de r_1 no argumento x e $1/x$ é a regra que define como calcular $r_1(x)$.

Outra função que encontraremos com frequência é

$$\sqrt{} : [0, +\infty] \rightarrow \mathbb{R}$$

cujo valor em um número real positivo é a raiz quadrada *positiva* deste número. Note que não podemos definir $\sqrt{}$ como sendo a raiz positiva e a raiz negativa porque uma função associa a cada valor do seu domínio um *único* número real. Até aqui, todas as funções que consideramos foram definidas por uma única regra em todo o seu domínio, mas muitas funções importantes requerem regras diferentes para diferentes valores do domínio. O exemplo mais simples é a função *valor absoluto* $\text{abs} : \mathbb{R} \rightarrow \mathbb{R}$, também conhecida como *função módulo*, que é definida por

$$\text{abs}(x) = \begin{cases} x & \text{se } x \geq 0 \\ -x & \text{se } x < 0. \end{cases}$$

Esta função também é, frequentemente, denotada por $|x|$. Uma função bem mais exótica, mas que também requer duas regras, é a *função de Dirichlet* \mathcal{D} definida por

$$\mathcal{D}(x) = \begin{cases} 1 & \text{se } x \text{ é racional} \\ 0 & \text{se } x \text{ é irracional.} \end{cases}$$

Tanto a função de Dirichlet, quanto a valor absoluto, estão definidas em todo o conjunto \mathbb{R} . Naturalmente nada nos impede de descrever funções por três ou mais regras. Por exemplo,

$$f(x) = \begin{cases} 0 & \text{se } x \leq -1 \\ -x + 1 & \text{se } -1 \leq x \leq 0 \\ 1 - x & \text{se } 0 \leq x \leq 1 \\ 0 & \text{se } x \geq 1, \end{cases}$$

define uma função cujo domínio e contradomínio são iguais a \mathbb{R} .

As funções que aparecem com mais frequência nos cursos de cálculo, são as chamadas *funções elementares*, que são obtidas combinando funções racionais, funções trigonométricas, exponenciais e logaritmos usando as operações que estudaremos na seção 4. Antes disso, porém, precisamos estudar o conjunto imagem de uma função.

2. Imagem de uma função

Ao contrário do que você possa ter, implicitamente, admitido ao ler a seção anterior, não é necessário que todo elemento do contradomínio de uma função seja imagem de algum elemento em seu domínio. Considere, por exemplo, a função quadrática $q_1 : \mathbb{R} \rightarrow \mathbb{R}$ definida por $q_1(x) = x^2 + 1$ e digamos que $a \in \mathbb{R}$. Para que exista x satisfazendo

$$q_1(x) = a$$

é necessário que

$$x^2 + 1 = a$$

tenha solução real. Mas esta equação equivale a

$$x^2 = a - 1$$

que só tem solução, quando $a - 1 \geq 0$; isto é, quando $a \geq 1$. Portanto, nenhum $a < 1$ é imagem de algum número real por q .

Este e outros exemplos da mesma natureza sugerem a seguinte definição. Sejam $X, Y \subset \mathbb{R}$. O *conjunto imagem*, ou simplesmente *a imagem* $\text{im}(f)$ de uma função $f : X \rightarrow Y$ é o subconjunto de Y formado pelos números reais da forma $f(x)$ para $x \in X$. Usando quantificadores, podemos escrever

$$\text{im}(f) = \{y \in Y \mid \exists x \in X, (y = f(x))\}.$$

Portanto, no exemplo acima,

$$\text{im}(q) = [1, +\infty).$$

A imagem da função de Dirichlet é, obviamente, $\{0, 1\}$. Igualmente fácil de determinar é a imagem da função valor absoluto. Como $\text{abs}(x) \geq 0$, para todo $x \in \mathbb{R}$, podemos concluir que

$$\text{im}(\text{abs}) \subset [0, +\infty).$$

Por outro lado, $\text{abs}(x) = x$ quando $x \geq 0$, de modo que

$$[0, +\infty) \subset \text{im}(\text{abs}).$$

Combinando as duas inclusões, concluímos que $\text{im}(\text{abs}) = [0, +\infty)$.

Vejamos como determinar a imagem das funções r_1 , r_2 e r_3 da seção anterior. Pela definição do conjunto imagem, temos que $a \in \text{im}(r_1)$ se, e somente se, a equação

$$r_1(x) = \frac{1}{x} = a$$

tiver solução. Mas isto ocorre sempre que $a \neq 0$ porque, neste caso, basta tomar $x = 1/a$. Isto nos permite concluir que

$$\text{im}(r_1) = \mathbb{R} \setminus \{0\}.$$

No caso da função r_2 , precisamos considerar as soluções da equação

$$r_2(x) = \frac{x+1}{x-2} = a.$$

Mas esta igualdade nos dá,

$$x+1 = a(x-2),$$

onde

$$x = \frac{2a+1}{a-1},$$

que só faz sentido se $a \neq 1$. Assim,

$$\text{im}(r_2) = \mathbb{R} \setminus \{1\}.$$

Finalmente,

$$r_3(x) = \frac{1}{x^2 - 2} = a$$

equivale a

$$ax^2 - 2a - 1 = 0$$

de modo que

$$x = \pm \sqrt{\frac{2a+1}{a}},$$

só está definido quando

$$\frac{2a+1}{a} \geq 0.$$

Mas isto só acontece quando $a > 0$ e $2a+1 \geq 0$ ou, então, $a < 0$ e $2a+1 \leq 0$. Como $2a+1 \geq 0$ ocorre se, e somente se, $a \geq -1/2$, as condições acima equivalem, respectivamente, a

$$a \in (0, +\infty) \quad \text{e} \quad a \in \left(-\infty, -\frac{1}{2}\right),$$

o que nos permite concluir que

$$\text{im}(r_3) = (0, +\infty) \cup \left(-\infty, -\frac{1}{2}\right).$$

Encerraremos a seção analisando, com todo cuidado, como determinar a imagem de uma função quadrática geral $q(x) = ax^2 + bx + c$, em que $a \neq 0$, b e c são números reais. Pela definição de imagem, isto consiste em determinar os valores de y para os quais

$$q(x) = ax^2 + bx + c = y,$$

que, por sua vez, equivale a encontrar os valores de y para os quais equação quadrática

$$ax^2 + bx + (c - y) = 0$$

tem solução. Mas, para que isto ocorra é necessário que o discriminante desta equação seja maior ou igual a zero; isto é, que

$$b^2 - 4a(c - y) \geq 0.$$

Rearrumando a desigualdade, obtemos

$$4ay \geq -(b^2 - 4ac).$$

A partir deste ponto, a análise requer um cuidado extra. Naturalmente, o que queremos é dividir tudo por $4a$, mas a desigualdade resultante depende do sinal de a . Assim,

$$\begin{aligned} y &\geq \frac{4ac - b^2}{4a} \text{ se } a > 0 \\ y &\leq \frac{4ac - b^2}{4a} \text{ se } a < 0. \end{aligned}$$

Portanto,

$$\text{im}(q) = \begin{cases} \left[\frac{4ac - b^2}{4a}, +\infty \right) & \text{quando } a > 0 \\ \left[-\infty, \frac{4ac - b^2}{4a} \right) & \text{quando } a < 0. \end{cases}$$

3. Gráficos de funções

Na seção 8 do capítulo 2 introduzimos números reais positivos para podermos medir o comprimento de segmentos. Contudo, para podermos definir gráficos de funções precisamos ir mais longe e introduzir coordenadas no plano.

Os sistemas de coordenadas tiveram origem na geografia, possivelmente introduzidos por Eratóstenes (275–195 a.C.), ao qual também devemos o crivo usado para listar números primos. Sobre seu mapa do mundo helenístico, Eratóstenes desenhou linhas horizontais e verticais imaginárias, precursoras de nossos paralelos e meridianos. Um avanço expressivo veio com Cláudio Ptolomeu (100-170 d.C.) que, em sua *Geografia*, aperfeiçoou uma maneira de projetar corretamente o globo terrestre sobre o plano e descreveu métodos para fixar a posição de um local da Terra, baseado em um sistema de latitudes e longitudes semelhante ao nosso.

Em matemática os sistemas de coordenadas do plano foram introduzidos no século XVII, independentemente, por Pierre de Fermat e René Descartes. Embora o sistema de dois eixos perpendiculares que usamos hoje em dia seja chamado de cartesiano em homenagem Descartes (Cartesius é a versão latina de Descartes), nem ele, nem Fermat usaram o sistema de dois eixos que se tornou padrão. Ambos utilizavam um único eixo a partir do qual traçavam linhas, geralmente perpendiculares, até os pontos cuja posição queriam determinar.

O primeiro matemático a utilizar sistematicamente dois eixos de coordenadas foi Isaac Newton, embora, frequentemente, escolhesse eixos que não eram perpendiculares. Em sua *Enumeratio Linearum Tertii Ordinis*, publicado em 1704, ele também usa sistematicamente coordenadas negativas para esboçar os gráficos das 72 “espécies” diferentes de



FIGURE 2. Pierre de Fermat e René Descartes

curvas definidas por equações de grau três que identifica em seu artigo. Dada a grande quantidade de casos a serem identificados, não é surpreendente descobrir que, nem mesmo Newton, escapou de cometer erros. Nos anos que se seguiram à publicação da *Enumeratio*, James Stirling, François Nicole e Nicolaus I Bernoulli descobriram mais seis tipos de cúbicas que não haviam sido identificadas por Newton.

O sistema de coordenadas planas usado atualmente utiliza quase sempre dois eixos perpendiculares. O ponto onde estes eixos se cruzam é tomado como o marco a partir do qual são feitas as medições de distância ao longo dos eixos. Cada ponto pode, então, ser identificado por um par de números reais, suas *coordenadas*. A primeira coordenada de um ponto do plano é conhecida como sua *abscissa* e a segunda, como sua *ordenada*. Os pontos à esquerda do eixo vertical têm abscissas negativas, ao passo que aqueles que estão abaixo horizontal têm ordenadas negativas. Com isso os eixos dividem o plano em quatro partes, conhecidas como *quadrantes*. A posição de cada quadrante e os sinais das coordenadas dos seus pontos são indicados na tabela 1.

		Quadrante	Sinal da abscissa	Sinal da ordenada
segundo	primeiro	Primeiro	positivo	positivo
		Segundo	negativo	positivo
terceiro	quarto	Terceiro	negativo	negativo
		Quarto	positivo	negativo

TABLE 1. Os quadrantes e os sinais das coordenadas dos seus pontos

Nossa próxima definição requer uma operação entre conjuntos que ainda não consideramos. Sejam A e B dois conjuntos. O *produto cartesiano* de A e B , denotado por $A \times B$, é o conjunto formado pelos pares ordenados cujo primeiro elemento pertence a A e

o segundo a B . Assim,

$$A \times B = \{(a, b) \mid a \in A \wedge b \in B\}.$$

Note que, quando A e B são conjuntos finitos, a quantidade de pares no produto cartesiano $A \times B$ é igual ao produto das quantidades de elementos em A e em B . Quando $A = B$, o produto cartesiano $A \times B$ é abreviado como A^2 . Como estamos associando a cada ponto um par ordenado formado pelos números reais que correspondem à sua abscissa e ordenada, podemos escrever

$$\mathbb{R}^2 = \{(x, y) \mid x \in \mathbb{R} \wedge y \in \mathbb{R}\}.$$

Por isso, frequentemente nos referiremos ao plano como “o \mathbb{R}^2 ”, quando estivermos expressando seus pontos em termos de coordenadas.

Com isto, podemos definir o *gráfico* de uma função $f : X \rightarrow Y$ como o conjunto

$$(33) \quad \Gamma_f = \{(x, f(x)) \in \mathbb{R}^2 \mid x \in X\}.$$

Note que, pela definição de função, há apenas um ponto em Γ_f para cada $x \in X$. Se você está acostumado a pensar no gráfico de uma função como um desenho, talvez tenha ficado um pouco surpreso com a definição acima; mas, na verdade, o que chamamos usualmente de gráfico de f é apenas a figura formada pelos pontos do conjunto Γ_f . Por exemplo, o gráfico da função constante $\widehat{1}$ é o conjunto formado pelos elementos da forma $(x, 1)$, para todo $x \in \mathbb{R}$. Marcando estes pontos no plano coordenado, obtemos a figura 3.

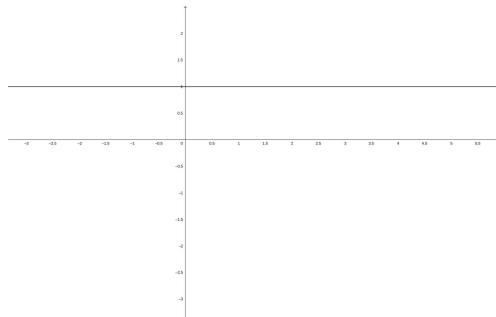


FIGURE 3. Gráfico
de $\widehat{1}$

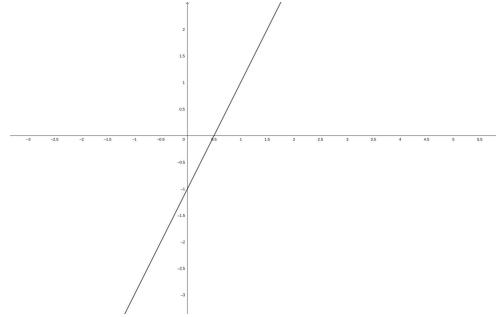


FIGURE 4. Gráfico
de $\ell(x) = 2x - 1$

Já o gráfico da função linear $\ell(x) = 2x - 1$ é o conjunto

$$\Gamma_\ell = \{(x, 2x - 1) \mid x \in \mathbb{R}\},$$

ilustrado na figura 4. É fácil desenhar ambos os gráficos porque, pelo axioma de incidência 1, dois pontos bastam para determinar uma reta.

Ao contrário da função linear, o gráfico da função quadrática

$$q(x) = ax^2 + bx + c, \text{ em que } a, b, c \in \mathbb{R} \text{ e } a \neq 0,$$

é bem mais difícil de esboçar. Mas, se você sabe que o gráfico é uma parábola, uma ideia mais ou menos razoável da curva pode ser obtida determinando o vértice da parábola e dois pontos em que o gráfico intersecta uma reta horizontal. Começaremos determinando as coordenadas do vértice, que é o ponto mais baixo, ou mais alto, da parábola, dependendo se $a > 0$ ou $a < 0$. Levando em conta que, quando $a > 0$, a imagem da função quadrática q é o intervalo

$$\left[\frac{b^2 - 4ac}{4a}, +\infty \right),$$

podemos concluir que esta é a ordenada do vértice da parábola que representa o gráfico de $q_{a,b,c}$. A abscissa x_0 do vértice será, então, o número real que satisfaz

$$q(x_0) = ax_0^2 + bx_0 + c = -\frac{b^2 - 4ac}{4a}.$$

A equação quadrática correspondente tem como única raiz

$$x_0 = \frac{-b}{2a},$$

de modo que as coordenadas do vértice da parábola serão

$$\left(-\frac{b}{2a}, -\frac{b^2 - 4ac}{4a} \right).$$

Por exemplo, o vértice do gráfico da função quadrática

$$q_2(x) = x^2 - x - 2$$

será

$$\left(\frac{1}{2}, -\frac{9}{4} \right).$$

Como as raízes de $x^2 - x - 2 = 0$ são -1 e 2 , o gráfico

$$\Gamma_{q_2} = \{(x, x^2 - x - 2) \in \mathbb{R}^2 \mid x \in \mathbb{R}\}$$

será como na figura 5.

Considere, agora, a função quadrática

$$q_3(x) = -4x^2 + 4x - 2.$$

Desta vez o vértice será o ponto

$$\left(\frac{1}{2}, -1 \right).$$

Como o coeficiente do termo em x^2 é negativo, a parábola que representa o gráfico desta função terá concavidade para baixo. Logo, esta parábola não corta o eixo das abscissas. Para termos outros pontos para nos ajudar a fazer o esboço, podemos intersectar o gráfico com a reta $y = -2$. Para obter as abscissas dos pontos de interseção precisamos resolver a equação

$$-4x^2 + 4x - 2 = -2,$$

cujas soluções são 0 e 1 . O gráfico desta função é ilustrado na figura 6.

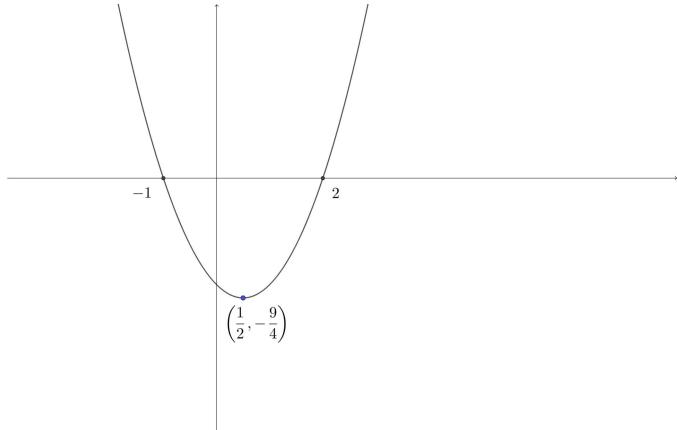


FIGURE 5. Gráfico da função quadrática $q_2(x) = x^2 - x - 2$.

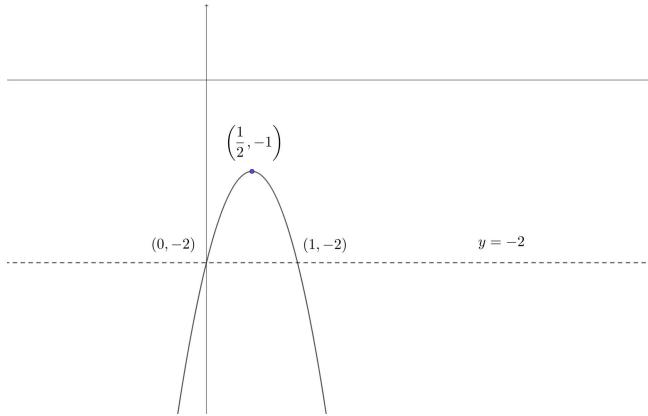


FIGURE 6. Gráfico da função $q_3(x) = -4x^2 + 4x - 2$.

Das funções que estudamos até agora, há duas cujos gráficos ainda não esboçamos. A primeira é a função valor absoluto. Por definição, temos que

$$\Gamma_{\text{abs}} = \{(x, \text{abs}(x)) \mid x \in \mathbb{R}\} = \{(x, -x) \mid x \in (-\infty, 0]\} \cup \{(x, x) \mid x \in [0, +\infty)\},$$

que é a união de duas semiretas que se tocam na origem e cuja representação geométrica é dada na figura 7

O outro gráfico que ainda não investigamos é o da função de Dirichlet. Pela definição desta função,

$$\Gamma_D = \{(x, 1) \mid x \in \mathbb{Q}\} \cup \{(x, 0) \mid x \notin \mathbb{Q}\}.$$

Entretanto, entre quaisquer dois números irracionais existe um número racional, e entre dois racionais, sempre existe um número irracional. Portanto, entre dois pontos de ordenada 1, sempre há pontos de ordenada 0 e vice-versa. Nem adiantaria dar um zoom no gráfico porque, se fizéssemos isto, continuaríamos tendo que enfrentar o mesmo problema.

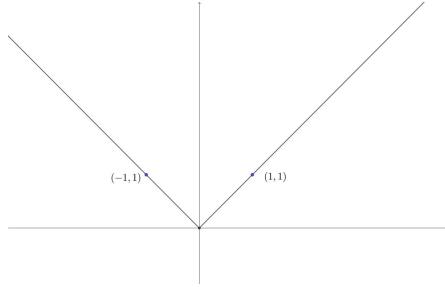


FIGURE 7. Gráfico da função valor absoluto

Em outras palavras, se insistíssemos em, mesmo assim, desenhar o gráfico, veríamos o que pareceria serem duas retas paralelas horizontais, uma com ordenada 0 e a outra com ordenada 1. Porém sabemos que tal figura não pode representar o gráfico de uma função, porque teríamos duas ordenadas diferentes para cada abscissa, o que não é permitido pela definição de função. Note que o problema não é que o gráfico não exista, mas sim que não é possível fazer um esboço dele que seja matematicamente correto.

4. Operações com funções

Nesta seção temos como meta definir as operações básicas com funções reais. Antes disso, porém, precisamos esclarecer quando duas funções são iguais. Ao longo de toda a seção suporemos que X seja um subconjunto de \mathbb{R} e que f_1 e f_2 sejam funções de X em \mathbb{R} . Dizemos que $f_1 = f_2$ se

$$f_1(x) = f_2(x) \text{ para todo } x \in X.$$

Não é, de forma alguma, o caso que duas funções só sejam iguais quando são definidas pela mesma regra; por exemplo, a função $h : \mathbb{R} \rightarrow \mathbb{R}$, definida por

$$h(x) = \sqrt{x^2}$$

é igual a $\text{abs}(x) = |x|$ porque, como vimos na seção 1, \sqrt{x} sempre representa a raiz quadrada positiva de x .

Podemos definir a soma $f_1 + f_2$ e a multiplicação $f_1 \cdot f_2$ como sendo as funções de X em \mathbb{R} dadas pelas fórmulas

$$(f_1 + f_2)(x) = f_1(x) + f_2(x) \text{ e } (f_1 \cdot f_2)(x) = f_1(x) \cdot f_2(x).$$

É importante observar que, para que a soma e a multiplicação de funções estejam definidas, as funções têm que ter o mesmo domínio. Por exemplo, se $g_1, g_2 : \mathbb{R} \rightarrow \mathbb{R}$ são as funções

$$(34) \quad g_1(x) = x^2 + 2 \text{ e } g_2(x) = -x^2 + 2x - 1$$

então $g_1 + g_2 : \mathbb{R} \rightarrow \mathbb{R}$ é definida por

$$(g_1 + g_2)(x) = g_1(x) + g_2(x) = (x^2 + 2) + (-x^2 + 2x - 1) = 2x + 1,$$

e $g_1 \cdot g_2 : \mathbb{R} \rightarrow \mathbb{R}$ por

$$(g_1 \cdot g_2)(x) = g_1(x) \cdot g_2(x) = (x^2 + 2) \cdot (-x^2 + 2x - 1) = -x^4 + 2x^3 - 3x^2 + 4x - 2.$$

Outras operações entre funções que podem ser definidas a partir destas incluem a multiplicação por um número real e a subtração. Supondo que $a \in \mathbb{R}$ e que $\widehat{a} : X \rightarrow \mathbb{R}$ é a função constante $\widehat{a}(x) = a$, para todo $x \in \mathbb{R}$, definimos

$$af_1 = \widehat{a} \cdot f_1 \quad \text{e} \quad f_1 - f_2 = f_1 + (\widehat{(-1)} \cdot f_2).$$

Como seria de esperar,

$$(af_1)(x) = af_1(x) \quad \text{e} \quad (f_1 - f_2)(x) = f_1(x) - f_2(x).$$

Inspirados pelas definições da soma e multiplicação de funções, gostaríamos de definir a divisão pela fórmula

$$(35) \quad \left(\frac{f_1}{f_2} \right)(x) = \frac{f_1(x)}{f_2(x)}.$$

Contudo, para que esta fórmula faça sentido, é necessário que f_2 não se anule em X . Diante disto temos duas opções. A primeira consiste em declarar que f_1/f_2 só está definida quando $f_2(x) \neq 0$ para todo $x \in X$. Na prática, contudo, é preferível adotar uma opção mais flexível, restringindo o domínio das duas funções para permitir que a divisão possa ser efetuada. Mais precisamente, se

$$\mathcal{Z}(f_2) = \{x \in X \mid f_2(x) = 0\},$$

usamos a fórmula (35) para definir f_1/f_2 como uma função cujo domínio é $X \setminus \mathcal{Z}(f_2)$ e não X . Vejamos o que acontece se aplicarmos estas considerações às funções $g_1(x) = x^2 + 2$ e $g_2(x) = -x^2 + 2x - 1$. Como $g_2(x) = -(x-1)^2$, a função g_2 só se anula quando $x = 1$. Portanto,

$$\mathcal{Z}(g_2) = \{1\},$$

de modo que g_1/g_2 é a função de domínio $\mathbb{R} \setminus \{1\}$, que é definida por

$$\left(\frac{g_1}{g_2} \right)(x) = \frac{g_1(x)}{g_2(x)} = \frac{x^2 + 2}{-x^2 + 2x - 1}.$$

Por outro lado, como $g_1(x) = x^2 + 2$ não tem raiz real,

$$\mathcal{Z}(g_1) = \emptyset.$$

Logo,

$$\left(\frac{g_2}{g_1} \right)(x) = \frac{g_2(x)}{g_1(x)} = \frac{-x^2 + 2x - 1}{x^2 + 2}$$

tem como domínio todo o conjunto \mathbb{R} . Outros dois exemplos interessantes de divisão são dados pelas funções abs/\mathcal{D} e \mathcal{D}/abs , cujos domínios são respectivamente, \mathbb{Q} e $\mathbb{R} \setminus \{0\}$, porque a função de Dirichlet se anula sobre os números irracionais.

A quarta operação com funções que estudaremos neste curso tem uma definição um pouco mais sutil que as anteriores, mas é tão, ou até mais, importante que elas. Desta vez suporemos que X e Y são subconjuntos de \mathbb{R} e que as funções são

$$h_1 : X \rightarrow \mathbb{R} \quad \text{e} \quad h_2 : Y \rightarrow \mathbb{R}.$$

Se $\text{im}(h_1) \subset Y$, então a função composta $h_2 \circ h_1 : X \rightarrow \mathbb{R}$ como sendo a função de X em \mathbb{R} é definida pela regra

$$(36) \quad (h_2 \circ h_1)(x) = h_2(h_1(x)).$$

Por exemplo, se $p_1 : \mathbb{R} \setminus \{0\} \rightarrow \mathbb{R}$ e $p_2 : \mathbb{R} \rightarrow \mathbb{R}$ são definidas por

$$p_1(x) = \frac{1}{x} \quad \text{e} \quad p_2(x) = x - 1$$

então

$$\text{im}(p_1) = \mathbb{R} \setminus \{0\} \quad \text{e} \quad \text{im}(p_2) = \mathbb{R}.$$

Como o domínio de p_2 é igual a \mathbb{R} , a função

$$(p_2 \circ p_1)(x) = p_2(p_1(x)) = p_2\left(\frac{1}{x}\right) = \frac{1}{x} - 1$$

está bem definida. Contudo, $\text{im}(p_2) = \mathbb{R}$ não está contida no domínio $\mathbb{R} \setminus \{0\}$ de p_1 , de modo que a composta $p_1 \circ p_2$ não está definida. Entretanto, como no caso da divisão de funções, adotaremos o ponto de vista, mais flexível, de que $p_1 \circ p_2$ está definida pela fórmula

$$(p_1 \circ p_2)(x) = p_1(p_2(x)) = p_1(x - 1) = \frac{1}{x - 1}$$

no domínio $\mathbb{R} \setminus \{1\}$. Em geral, se $\text{im}(h_2) \notin Y$, para que a função composta $h_2 \circ h_1$ possa ser definida é necessário considerar seu domínio como sendo o conjunto

$$X' = \{x \in X \mid h_1(x) \in Y\}.$$

Observe que, segundo os cálculos que fizemos acima

$$(p_2 \circ p_1)(x) = \frac{2}{x} - 1 \neq \frac{1}{x - 1} = (p_1 \circ p_2)(x),$$

mesmo que $x \in \mathbb{R} \setminus \{0\}$, que é o conjunto onde ambas as compostas estão definidas. Portanto, a composta de funções não é uma operação comutativa, embora seja fácil verificar que é associativa e tenha a função identidade $\text{id}(x) = x$ como elemento neutro.

Sejam X e Y subconjuntos de \mathbb{R} e $f : X \rightarrow Y$ uma função. A inversa de f é a função $f^{-1} : Y \rightarrow X$ que satisfaz

$$(37) \quad (f^{-1} \circ f)(x) = x \quad \text{para todo } x \in X \quad \text{e} \quad (f \circ f^{-1})(y) = y \quad \text{para todo } y \in Y.$$

A ideia por trás desta definição é que se $f(x) = y$, então sua inversa deve ser a função que leva y em x . Por exemplo, como

$$p_1(x) = \frac{1}{x} = y,$$

então $x = 1/y$. Logo, a inversa de p_1 é a própria função p_1 . Podemos confirmar isto facilmente, calculando

$$(p_1 \circ p_1)(x) = p_1\left(\frac{1}{x}\right) = \frac{1}{1/x} = x,$$

o que mostra que as fórmulas em (37) são verificadas neste exemplo. Por outro lado, a partir de

$$p_2(x) = x - 1 = y,$$

obtemos $x = y + 1$, de modo que p_2^{-1} é a definida por $p_2^{-1}(y) = y + 1$. De fato,

$$(p_2 \circ p_2^{-1})(x) = p_2(p_2^{-1}(x)) = p_2(x + 1) = (x + 1) - 1 = x$$

e um cálculo semelhante mostra que $(p_2^{-1} \circ p_2)(x) = x$.

Infelizmente estes exemplos não são nem um pouco representativos da dificuldade de calcular a inversa de uma função. Uma ideia mais clara dos problemas a serem enfrentados é dada pelas funções quadráticas. Digamos, por exemplo, que $q(x) = x^2 - 4x + 3$ e vamos tentar aplicar a esta função a mesma estratégia utilizada para encontrar as inversas de p_1 e p_2 . Assim, escrevemos

$$q(x) = x^2 - 4x + 3 = y$$

e resolvemos a equação quadrática

$$x^2 - 4x + (3 - y) = 0$$

para descobrir de que forma x depende de y . O problema é que, ao fazer isto, obtemos duas soluções possíveis,

$$x = 2 + \sqrt{y+1} \text{ ou } x = 2 - \sqrt{y+1}.$$

Qual das duas escolher? Como se isto não bastasse, nenhuma destas duas expressões está definida quando $y < -1$. Isto, porém, é fácil de explicar, porque um tal y não pertence à imagem de $\text{im}(q) = [-1, +\infty)$ e, portanto, não há nenhum elemento do domínio onde levar estes números.

Sistematizaremos isto para futura referência. Voltando ao caso em que $X \subset \mathbb{R}$ e $f : X \rightarrow Y$ precisamos, para que a inversa exista que

- (i) todo elemento de Y seja imagem de um elemento em X ;
- (ii) não haja dois elementos em X que sejam levados em um mesmo elemento de Y .

A condição (i) equivale a dizer que $\text{im}(q) = Y$; já a condição (ii) equivale à proposição

$$\forall x_1, x_2 \in X, ((x_1 \neq x_2) \rightarrow f(x_1) \neq f(x_2)).$$

As funções que satisfazem (i) são chamadas de *sobrejetivas* (ou *sobrejetoras*), e as que satisfazem (ii) de *injetivas* (ou *injetoras*). Podemos enunciar os problemas que enfrentamos ao tentar inverter q dizendo que não é sobrejetora, nem injetora. Uma função que é, simultaneamente, injetiva e sobretiva é chamada de *bijetiva* (ou *bijetora*). Nossa análise mostra que uma função tem inversa se, e somente se, é bijetiva.

À primeira vista nada mais há a falar sobre a inversa de q ; afinal a função não é sobrejetiva, nem injetiva. Contudo, refinando nossa análise, podemos mostrar que é possível restringir seu domínio e contradomínio de modo que a obter funções inversíveis. Antes, porém, precisamos definir exatamente o que é a restrição de uma função. Seja f a função em (37). A *restrição* de f a $S \subset X$ é a função

$$f|_S : S \rightarrow \mathbb{R}$$

definida pela mesma regra de f ; isto é,

$$f|_S(x) = f(x) \text{ para todo } x \in S.$$

Por exemplo, as restrições de q aos intervalos $(-\infty, 2]$ e $[2, +\infty)$ são as funções

$$(38) \quad q_- : (-\infty, 2] \rightarrow [1, +\infty) \quad \text{e} \quad q_+ : [2, +\infty) \rightarrow [1, +\infty)$$

ambas definidas pela mesma regra que $q(x)$; isto é,

$$q_-(x) = x^2 - 4x + 3 \quad \text{e} \quad q_+(x) = x^2 - 4x + 3.$$

Como 2 é a abscissa do vértice da parábola que representa o gráfico de q , tanto q_- , quanto q_+ são injetivas. De fato,

$$q_-(x_1) = q_-(x_2)$$

equivale a

$$x_1^2 - 4x_1 = x_2^2 - 4x_2;$$

onde

$$0 = (x_1^2 - x_2^2) - 4(x_1 - x_2) = (x_1 - x_2)(x_1 + x_2 - 4).$$

Mas, se $x_1 \neq x_2$, então $x_1 - x_2 \neq 0$ e podemos concluir, da equação acima, que

$$x_2 = 4 - x_1.$$

Logo, se $x_1 \in (-\infty, 2]$, então

$$x_2 = 4 - x_1 \geq 4 - 2 = 2,$$

de forma que

$$x_1 \in (-\infty, 2] \longrightarrow x_2 \in [2, +\infty).$$

Note que além dos domínios de q_- e q_+ terem sido restrinidos para produzir funções injetivas, limitamos seus contradomínios a $\text{im}(q) = [1, +\infty)$ para que também sejam sobrejetivas. As inversas

$$q_-^{-1} : [-1, +\infty) \rightarrow (-\infty, 1/2] \quad \text{e} \quad q_+^{-1} : [-1, +\infty) \rightarrow [1/2, +\infty)$$

das funções bijetivas resultantes destas restrições são definidas pelas regras

$$q_-^{-1}(x) = 2 - \sqrt{y+1} \quad \text{e} \quad q_+^{-1}(x) = 2 + \sqrt{y+1}.$$

No capítulo 6 veremos que restrições ainda mais drásticas são necessárias para podermos determinar as inversas das funções trigonométricas.

Uma função cuja inversa requer ainda mais cuidado é $r : \mathbb{R} \setminus \{2, 3\} \rightarrow \mathbb{R}$, definida pela regra

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

Como a função q analisada acima, esta também não é injetiva. De fato,

$$r\left(\frac{5}{2} + x\right) = \frac{1}{x^2 - 1/4} = r\left(\frac{5}{2} - x\right),$$

porque a parábola $y = x^2 - 5x + 6$ tem vértice em

$$\left(\frac{5}{2}, -4\right)$$

e, por isso, é simétrica em relação à reta vertical $x = 5/2$. Para contornar este problema, precisamos restringir r aos intervalos $(-\infty, 5/2]$ e $[5/2, +\infty)$. No caso desta função, há uma complicação adicional, porque a função não está definida nos pontos 2 e 3. Para evitar estes dois valores, somos obrigados a subdividir cada um dos intervalos $(-\infty, 5/2]$ e $[5/2, +\infty)$ em dois. Teremos, com isso, as duas funções inversíveis,

$$r_- : (-\infty, 2) \cup (2, 5/2) \rightarrow \mathbb{R} \quad \text{e} \quad r_+ : (5/2, 3) \cup (3, +\infty) \rightarrow \mathbb{R}$$

definidas como restrições de r , de modo que

$$r_j(x) = r(x) = \frac{1}{x^2 - 5x + 6} \quad \text{quando} \quad j \in \{-, +\}.$$

Como veremos adiante, a subdivisão dos domínios destas funções levam a uma subdivisão de suas imagens. Antes, porém, precisamos calcular as fórmulas para as inversas de r_- e r_+ . Suponhamos que y esteja no domínio de uma destas funções. Por definição, isto significa que existe um número real x , no domínio da mesma função, tal que

$$\frac{1}{x^2 - 5x + 6} = y.$$

Mas isto equivale a

$$x^2 - 5x + \left(6 - \frac{1}{y}\right) = 0.$$

Resolvendo esta equação quadrática, encontramos

$$x = \frac{1}{2} \left(5 \pm \sqrt{\frac{4+y}{y}} \right)$$

Contudo, para que esta fórmula esteja bem definida é necessário que

$$\frac{4+y}{y} \geq 0,$$

o que pode ocorrer quando y e $4+y$ são ambos positivos ou ambos negativos. No primeiro caso temos que $y \in (0, +\infty)$ e no segundo que $y \in (-\infty, -4]$. O primeiro destes intervalos só contém números positivos e o segundo só contém números negativos. Como

$$x^2 - 5x + 6 \text{ é } \begin{cases} \text{positivo quando} & x < 2 \text{ ou } x > 3 \\ \text{negativo quando} & 2 < x < 3, \end{cases}$$

podemos concluir que

- r_- é positiva em $(-\infty, 2)$ e negativa em $(2, 5/2)$;
- r_+ é negativa em $(5/2, 3)$ e positiva em $(3, +\infty)$.

Para concluir o exemplo, resta-nos explicitar as inversas. Como

$$5 - \sqrt{\frac{4+y}{y}} < 5 + \sqrt{\frac{4+y}{y}},$$

podemos afirmar que

$$r_-^{-1} : (0, +\infty) \cup (-\infty, -4] \longrightarrow (-\infty, 2) \cup (2, 5/2)$$

deve ser definida pela regra

$$r_-^{-1}(x) = \frac{1}{2} \left(5 - \sqrt{\frac{4+y}{y}} \right),$$

ao passo que

$$r_+^{-1} : (0, +\infty) \cup (-\infty, -4] \longrightarrow (5/2, 3) \cup (3, +\infty)$$

deve ser definida por

$$r_+^{-1}(x) = \frac{1}{2} \left(5 + \sqrt{\frac{4+y}{y}} \right),$$

concluindo nossa análise das restrições necessárias para que seja possível inverter r .

Quando o gráfico de uma função inversível $f : X \rightarrow Y$ é conhecido, é possível determinar o gráfico de sua inversa de maneira geométrica, mesmo que a fórmula da inversa seja muito difícil de calcular. O ponto chave da construção é o fato de que

$$(x_0, f(x_0)) \in \Gamma_f \longleftrightarrow (f(x_0), x_0) \in \Gamma_{f^{-1}}.$$

Supondo que $f(x_0) \neq x_0$, a reta que passa por estes pontos tem inclinação

$$\frac{f(x_0) - x_0}{x_0 - f(x_0)} = -1,$$

de modo que sua equação é

$$y - f(x_0) = -(x - x_0).$$

Logo, trata-se de uma reta perpendicular a $y = x$ e que intersecta esta última reta no ponto

$$\left(\frac{x_0 + f(x_0)}{2}, \frac{x_0 + f(x_0)}{2} \right)$$

que fica exatamente no meio do segmento que une $(x_0, f(x_0))$ a $(f(x_0), x_0)$. Portanto, $(x_0, f(x_0))$ e $(f(x_0), x_0)$ são simétricos relativamente à reta $y = x$. Em outras palavras, para obter $\Gamma_{f^{-1}}$ a partir de Γ_f basta espelhar este último gráfico relativamente à reta $y = x$. Por exemplo, a função $p_3 : [0, +\infty) \rightarrow [0, +\infty)$ definida pela regra $p_3(x) = x^2$ tem como gráfico a parábola ilustrada à esquerda da figura 8. Espelhando este gráfico relativamente à reta $y = x$, que também foi desenhada na figura, obtemos o gráfico da função $p_3^{-1} : [0, +\infty) \rightarrow [0, +\infty)$ definida por $p_3^{-1}(x) = \sqrt{x}$ sem a necessidade de calcular explicitamente nenhum de seus pontos, como ilustrado à direita da figura 8.

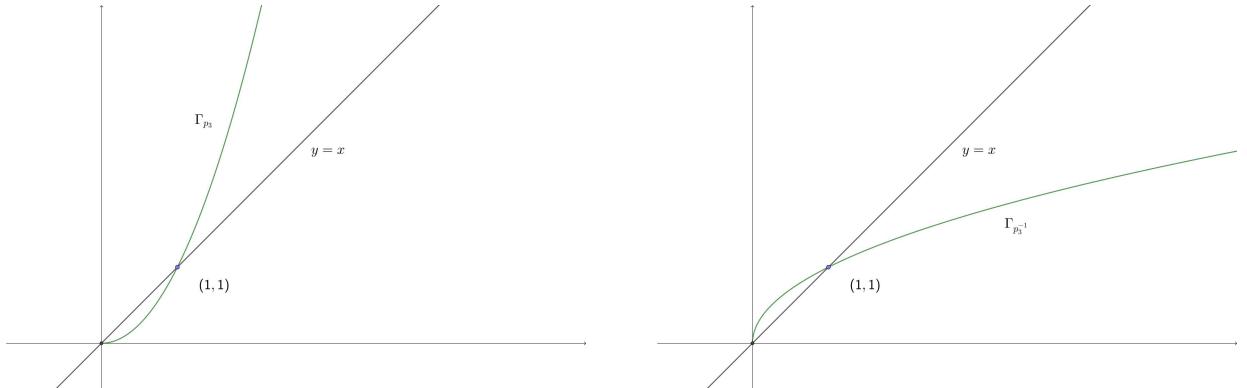


FIGURE 8. Gráficos de p_3 e p_3^{-1} .

Um exemplo mais divertido de uma função e sua inversa é ilustrado na figura 9. A curva verde corresponde ao gráfico de $p_4(x) = x^3$, ao passo que a curva vermelha é o gráfico da inversa $p_4^{-1}(x) = \sqrt[3]{x}$, que foi obtido refletindo a curva verde em torno da reta $y = x$.

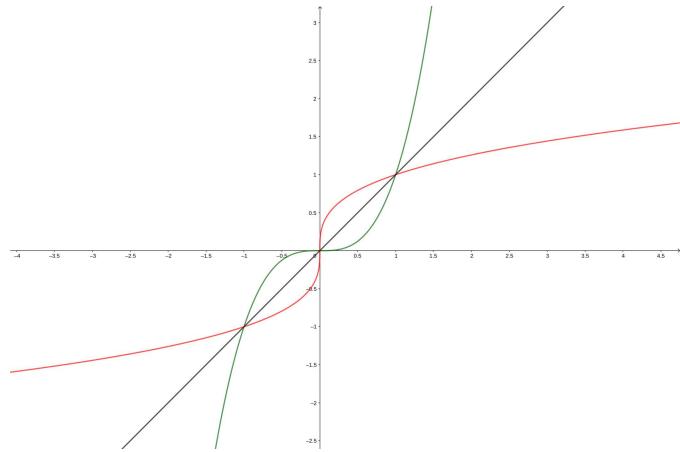


FIGURE 9. Gráficos das funções p_4 e p_4^{-1} .

CHAPTER 5

Funções polinomiais

As funções polinomiais são as únicas funções que conseguimos calcular usando apenas adição e multiplicação. Além disso, muitas funções podem ser aproximadas por polinômios com uma precisão tão grande quanto desejada, como vocês verão no curso de cálculo numérico.

1. Definições e propriedades básicas

Um *função polinomial*, ou simplesmente *polinômio*, é uma função que tem os números reais como seu domínio e contradomínio e cujo valor em $x \in \mathbb{R}$ é dado por

$$(39) \quad f(x) = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0$$

em que $a_n, a_{n-1}, \dots, a_1, a_0 \in \mathbb{R}$. Sempre ordenaremos as parcelas na soma da fórmula de uma função polinomial em ordem decrescente do expoente da variável x . Os números $a_n, a_{n-1}, \dots, a_1, a_0$ são os *coeficientes* do polinômio f . Se $a_n \neq 0$, então n é o *grau* de f , que denotaremos por $\text{grau}(f)$, e a_n é seu *termo líder*, que denotaremos por $\text{ld}(f)$. Um polinômio é *mônico* quando seu termo líder é igual a um. Normalmente, quando escrevermos um polinômio específico, não representaremos os termos cujos coeficientes são nulos; em outras palavras, escreveremos

$$x^5 - 5x^2 - 9 \text{ e não } x^5 + 0 \cdot x^4 + 0 \cdot x^3 - 5x^2 + 0 \cdot x - 9,$$

que é um polinômio mônico de grau 5. De posse desta terminologia, podemos começar a estudar as propriedades básicas dos polinômios

 Em geral, polinômio e função polinomial são conceitos diferentes, contudo, como estamos considerando apenas polinômios definidos sobre os reais podemos usá-los como sinônimos impunemente.

A soma e o produto de duas funções polinomiais f_1 e f_2 também são funções polinomiais. Para entender o porquê disto, suponhamos que

$$f_1(x) = a_n x^n + \cdots + a_1 x + a_0 \text{ e } f_2(x) = b_n x^n + \cdots + b_1 x + b_0$$

em que os a 's e b 's são números reais. Talvez esteja se perguntando se é necessário que os dois polinômios tenham o mesmo grau para que possamos somá-los; a resposta é não.

Para que f_1 e f_2 tivessem o mesmo grau seria necessário que a_n e b_n fossem ambos não nulos, mas não fizemos esta hipótese. Em outras palavras, estamos considerando que, se um dos polinômios tem grau menor que o outro, então acrescentamos algumas parcelas, com coeficientes iguais a zero, ao polinômio de menor grau, de modo que os mesmos expoentes de x apareçam em ambos. Por exemplo, se

$$(40) \quad h_1(x) = 2x^5 + 3x^4 - 6x^3 + 8x + 1 \quad \text{e} \quad h_2(x) = -9x^3 - 5x^2,$$

reescreveríamos ambos na forma

$$h_1(x) = 2x^5 + 3x^4 - 6x^3 + 0 \cdot x^2 + 8x + 1 \quad \text{e} \quad h_2(x) = 0 \cdot x^5 + 0 \cdot x^4 - 9x^3 - 5x^2 + 0 \cdot x + 0.$$

Fazemos isto apenas para evitar uma multiplicidade de casos, dependendo dos graus de cada polinômio.

Voltando à verificação de que a soma das funções polinomiais f_1 e f_2 é uma função polinomial, usamos a definição da soma de funções para escrever

$$(f_1 + f_2)(x) = (a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0) + (b_n x^n + b_{n-1} x^{n-1} + \cdots + b_1 x + b_0).$$

Não esqueça que x denota apenas um número real que não desejamos explicitar, de modo que podemos aplicar todas as propriedades da adição e multiplicação de números reais a estas duas expressões. Levando em conta que a adição de números reais é associativa e comutativa, podemos reescrever a fórmula para $f_1 + f_2$ como

$$(f_1 + f_2)(x) = (a_n + b_n)x^n + (a_{n-1} + b_{n-1})x^{n-1} + \cdots + (a_1 + b_1)x + (a_0 + b_0),$$

que é uma nova função polinomial. Por exemplo, a soma das funções polinomiais h_1 e h_2 definidas em (40) é a função polinomial dada pela regra

$$(h_1 + h_2)(x) = (2x^5 + 3x^4 - 6x^3 + 8x + 1) + (-9x^3 - 5x^2) = 2x^5 + 3x^4 - 15x^3 - 5x^2 + 8x + 1$$

Ao contrário do que fizemos no caso da soma, na análise da multiplicação é mais conveniente supor que

$f_1(x) = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0$ e que $f_2(x) = b_m x^m + b_{m-1} x^{m-1} + \cdots + b_1 x + b_0$ são funções polinomiais para as quais $a_n \neq 0$ e $b_m \neq 0$. Começamos usando a distributividade da multiplicação dos reais sobre a soma, para escrever

$$(41) \quad (f_1 \cdot f_2)(x) = (a_n x^n) \cdot f_2 + (a_{n-1} x^{n-1}) \cdot f_2 + \cdots + (a_1 x) \cdot f_2 + a_0 \cdot f_2.$$

Portanto, basta descobrirmos como calcular um produto da forma $a_k x^k \cdot f_2$, pois já sabemos somar polinômios. Mas, usando a distributividade novamente,

$$a_k x^k \cdot f_2 = a_k x^k \cdot b_m x^m + a_k x^k \cdot b_{m-1} x^{m-1} + \cdots + a_k x^k \cdot b_1 x + a_k x^k \cdot b_0.$$

Como

$$a_k x^k \cdot b_\ell x^\ell = a_k b_\ell x^{k+\ell},$$

obtemos

$$a_k x^k \cdot f_2 = a_k b_m x^{k+m} + a_k b_{m-1} x^{k+m-1} + \cdots + a_k b_1 x^{k+1} + a_k b_0 x^k.$$

substituindo estas expressões em (41), encontramos o polinômio

$$(f_1 \cdot f_2)(x) = c_s x^s + c_{s-1} x^{s-1} + \cdots + c_1 x + c_0$$

em que

$$\begin{aligned} c_0 &= a_0 b_0 \\ c_1 &= a_0 b_1 + a_1 b_0 \\ c_2 &= a_0 b_2 + a_1 b_1 + a_2 b_0 \\ &\vdots \\ c_j &= a_0 b_j + a_1 b_{j-1} + \cdots + a_{j-1} b_1 + a_j b_0 \\ &\vdots \\ c_{n+m} &= a_n b_m. \end{aligned}$$

Como estamos supondo que $a_n \neq 0$ e $b_m \neq 0$, temos que $c_{n+m} = a_n b_m \neq 0$. Levando em conta que a maior potência de x na fórmula do produto $f_1 \cdot f_2$ é obtida multiplicando-se a maior potência de x em f_1 pela maior potência de x em f_2 , concluímos que c_{n+m} é o coeficiente do termo com maior potência de x na fórmula de $f_1 \cdot f_2$. Em outras palavras,

$$\text{grau}(f_1 \cdot f_2) = n + m = \text{grau}(f_1) + \text{grau}(f_2).$$

A fórmula para o grau da soma de dois polinômios é bem menos precisa,

$$\text{grau}(f_1 + f_2) \leq \max\{\text{grau}(f_1), \text{grau}(f_2)\}.$$

Isto ocorre porque vários coeficientes podem se cancelar quando somamos f_1 com f_2 , reduzindo, assim, o grau da soma. Por exemplo, se

$$g_1(x) = x^9 - x^2 + 1 \quad \text{e} \quad g_2(x) = -x^9 - x - 3,$$

então

$$(g_1 + g_2)(x) = -x^{18} + x^{11} - x^{10} - 4x^9 + x^3 + 3x^2 - x - 3$$

tem grau igual a $2 \cdot 9 = 18$; contudo

$$(g_1 + g_2)(x) = -x^2 - x - 2$$

tem grau $2 < 9$. Finalmente, para que as fórmulas para os graus da soma e da multiplicação de polinômios continuem válidas quando um dos polinômios é nulo, definimos o grau deste polinômio como sendo $-\infty$. Assim,

$$\text{grau}(f) = \begin{cases} -\infty & \text{se } f = 0 \\ n & \text{se } f = a_n x^n + \cdots + a_1 x + a_0 \text{ e } a_n \neq 0. \end{cases}$$

Na seção 3 veremos que a quantidade de raízes de um polinômio não pode exceder o seu grau.

2. Divisão com resto e fatoração

Agora que sabemos somar e multiplicar dois polinômios, vejamos como dividi-los. Não se trata, porém, de aplicar a divisão de polinômios como definida na seção 4, mas sim da divisão com resto. Em outras palavras, dados dois polinômios f e $g \neq 0$, desejamos obter dois outros polinômios q e r tais que

$$f = gq + r \quad \text{e} \quad r = 0 \quad \text{ou} \quad \text{grau}(r) < \text{grau}(g).$$

O algoritmo usado para efetuar esta divisão é extremamente simples. Suponhamos que

$f(x) = a_nx^n + a_{n-1}x^{n-1} + \dots + a_1x + a_0$, que $g(x) = b_mx^m + b_{m-1}x^{m-1} + \dots + b_1x + b_0$ e que

$$\text{grau}(f) \geq \text{grau}(g).$$

Calculando

$$f(x) - \frac{a_n}{b_m}x^{n-m}g(x),$$

obtemos

$$\left(a_{n-1} - \frac{a_n b_{m-1}}{b_m}\right)x^{n-1} + \left(a_{n-2} - \frac{a_n b_{m-2}}{b_m}\right)x^{n-2} + \dots + \left(a_0 - \frac{a_n b_0}{b_m}\right)$$

que tem grau menor que n . O algoritmo de divisão consiste, então, em executar passos como este repetidas vezes até obter um polinômio de grau menor que m . Lembrando que $\text{ld}(f)$ é o coeficiente do termo de maior grau de f , o algoritmo pode ser descrito da seguinte maneira.

ALGORITMO DE DIVISÃO DE POLINÔMIOS. *Dados polinômios f e $g \neq 0$, o algoritmo calcula polinômios q e r tais que*

$$f = gq + r \quad \text{e} \quad r = 0 \quad \text{ou} \quad \text{grau}(r) < \text{grau}(g).$$

Etapa 1: Initialize $F = f$ e $Q = 0$.

Etapa 2: Enquanto F tem grau maior ou igual que o grau de g , repita

$$F = F - \frac{\text{ld}(F)}{\text{ld}(g)}x^{n-m}g \quad \text{e} \quad Q = Q + \frac{\text{ld}(F)}{\text{ld}(g)}x^{n-m},$$

em que n é o grau de F e m o grau de g .

Etapa 3: Retorne quociente = Q e resto = F .

Por exemplo, dividindo $6x^6 + 7x^5 + 8x^4 + 1$ por $x^4 + x + 1$ obtemos o quociente $6x^2 + 7x + 8$ e o resto $-6x^3 - 13x^2 - 15x - 7$. Quando a divisão de polinômios é realizada com lápis e papel, é conveniente usar um esquema semelhante ao da divisão longa de inteiros, como ilustrado na tabela 1.

Como no caso dos números inteiros, a divisão pode ser utilizada para provar um teorema sobre a fatoração dos polinômios. Para isto, precisamos de algumas definições. Se f ,

$$\begin{array}{r}
 6x^6 + 7x^5 + 8x^4 + 1 \\
 7x^5 + 8x^4 - 6x^3 - 6x^2 + 1 \\
 \hline
 8x^4 - 6x^3 - 13x^2 - 7x + 1 \\
 -6x^3 - 13x^2 - 15x - 7
 \end{array}
 \quad
 \begin{array}{c}
 |x^4 + x + 1| \\
 \hline
 6x^2 + 7x + 8
 \end{array}$$

TABLE 1. Divisão de $6x^6 + 7x^5 + 8x^4 + 1$ por $x^4 + x + 1$.

g_1 e g_2 são polinômios de modo $f = g_1g_2$, dizemos que g_1 e g_2 são fatores de f . Um fator de f é *próprio* se tem grau maior que zero e menor do que f . Por exemplo, se

$$h(x) = 10 \cdot (x^3 + 1) \cdot (x^2 + 4x + 1)$$

então $(x^3 + 1)$ e $(x^2 + 4x + 1)$ são fatores próprios de h , mas não seu produto. Um polinômio que não é constante e que não admite fatores próprios de grau diferente de zero é chamado de *irreduzível*. Observe que, por esta definição, se p é um polinômio irreduzível e $c \neq 0$ é um número real, então o polinômio cp também é irreduzível. Neste caso, dizemos que p e cp são polinômios irreduzíveis *associados*. Por exemplo,

$$x + 1, 3x + 3 \text{ e } x^2 + 1$$

são todos três polinômios irreduzíveis e os dois primeiros são associados. É possível provar que os únicos polinômios irreduzíveis sobre os números reais têm grau um e dois, mas a demonstração está além dos recursos deste curso.

Temos uma *fatoração* de um polinômio não constante f , quando

$$(42) \quad f = p_1 \cdots p_n.$$

em que p_1, \dots, p_n são polinômios irreduzíveis, não necessariamente distintos. Note que se f é irreduzível, então ele é sua própria fatoração. Suponhamos que

$$f = q_1 \cdots q_m$$

também seja uma fatoração de f , de modo que q_1, \dots, q_m são polinômios irreduzíveis. Dizemos que estas duas fatorações de f são iguais se $m = n$ e se for possível ordenar os q 's de maneira que

$$(p_1, q_1), (p_2, q_2), \dots, (p_n, q_n)$$

sejam todos pares de polinômios associados. Com isto podemos enunciar e provar o seguinte resultado.

TEOREMA DA FATORAÇÃO ÚNICA. *Todo polinômio não constante admite uma única fatoração.*

A demonstração da unicidade da fatoração de polinômios é razoavelmente mais elaborada do que quase todos os resultados que provamos até agora e pode ser uma boa ideia deixá-la de lado em uma primeira leitura deste capítulo.

PROOF. Seja f um polinômio não constante. Precisamos provar que admite uma fatoração e que quaisquer duas de suas fatorações são iguais. A existência da fatoração é muito fácil de provar por indução no grau do polinômio. Se o polinômio tiver grau um, então é necessariamente irreduzível, o que prova a base da indução. Suponhamos que todo polinômio de grau menor que $n > 1$ admita uma fatoração e que $\text{grau}(f) = n$. Já vimos que, se f for irreduzível, então é igual à sua fatoração, de modo que podemos supor que f não seja irreduzível. Mas, neste caso, existem polinômios g_1 e g_2 , ambos de grau positivo e menor que n , tais que $f = g_1g_2$. Pela hipótese de indução, tanto g_1 , quanto g_2 , podem ser fatorados. Para obter a fatoração de f basta multiplicar as fatorações de g_1 e g_2 .

Como no caso de inteiros, a unicidade da fatoração é mais difícil de provar do que sua existência. Daremos uma demonstração por contradição que é um pouco cheia de truques, mas que tem a virtude de ser curta e elementar. Admitindo, por contradição, que existam polinômios que tenham duas fatorações diferentes, podemos supor que

(H) f é um polinômio não constante de menor grau possível entre os polinômios que têm mais de uma fatoração.

Em particular, existem polinômios irreduzíveis p_1, \dots, p_n e q_1, \dots, q_m , tais que

$$(43) \quad f = p_1 \cdots p_n \quad \text{e} \quad f = q_1 \cdots q_m$$

são duas fatorações *distintas* de f . A primeira coisa a observar é que (H) implica que nenhum dos p 's pode ser associado a nenhum dos q 's. Para entender porque, digamos que $q_1 = cp_1$. Substituindo esta igualdade em (43), obtemos

$$p_1p_2 \cdots p_n = c_1p_1q_2 \cdots q_m.$$

Mas, cancelando p_1 ,

$$(44) \quad p_2 \cdots p_n = c_1q_2 \cdots q_m.$$

Como, por definição, um polinômio irreduzível não pode ser constante, temos que $\text{grau}(p_1) \geq 1$, donde

$$\text{grau}(p_2 \cdots p_n) < \text{grau}(f) - \text{grau}(p_1) < \text{grau}(f).$$

Segue-se, portanto, de (H) que $p_2 \cdots p_n$ tem uma única fatoração. Combinando isto com (44), podemos concluir que $m = n$ e que é possível reordenar os q 's de maneira que

$$(q_2, p_2), \dots, (q_n, p_n)$$

são todos pares de polinômios associados. Como estamos supondo que q_1 e p_1 são associados, isto contradiz nossa hipótese de que as fatorações em (43) são distintas. Reduzimos, assim, o problema a mostrar que existe algum p associado a algum q .

Voltando às duas fatorações distintas em (43), suponhamos que os polinômios foram ordenados de modo que p_1 tem o maior grau possível entre os p 's e que q_1 tem o maior

grau possível entre os q 's. Digamos que

$$s = \text{grau}(q_1) \geq r = \text{grau}(p_1) \text{ e que } c = \frac{\text{ld}(q_1)}{\text{ld}(p_1)}.$$

Deste ponto em diante, a demonstração baseia-se nas propriedades do polinômio

$$h = f - cp_1x^{s-r}q_2\cdots q_m.$$

Lembrando que $f = q_1\cdots q_m$, podemos escrever

$$h = q_1\cdots q_m - cp_1x^{s-r}q_2\cdots q_m = (q_1 - cp_1x^{s-r})q_2\cdots q_m.$$

Há, então, dois casos a considerar. No primeiro, $q_1 - cp_1x^{s-r} = 0$, que equivale a

$$q_1 = cp_1x^{s-r},$$

e contradiz a irredutibilidade de q_1 , levando diretamente a uma contradição. No segundo caso, $q_1 - cp_1x^{s-r} \neq 0$. Mas, como vimos em nosso estudo da divisão de polinômios,

$$\text{grau}(q_1 - cp_1x^{s-r}) < \text{grau}(q_1);$$

de modo que

$$(45) \quad h = (q_1 - cp_1x^{s-r})q_2\cdots q_m$$

tem grau menor do que

$$\text{grau}(f) = \text{grau}(q_1q_2\cdots q_m).$$

Nossa escolha de f como o polinômio de menor grau com duas fatorações distintas nos permite concluir que h tem uma única fatoração. Contudo, substituindo $f = p_1\cdots p_n$ na definição de h , encontramos

$$h = f - cp_1x^{s-r}q_2\cdots q_m = p_1(p_2\cdots p_n - cx^{s-r}q_2\cdots q_m).$$

Contudo podemos obter uma fatoração de f multiplicando p_1 por uma fatoração de

$$p_2\cdots p_n - cx^{s-r}q_2\cdots q_m.$$

Portanto, existe uma fatoração de h da qual p_1 é um dos fatores irreduzíveis. Logo, a unicidade da fatoração de h , aplicada a (45), nos permite concluir que p_1 tem que dividir $(q_1 - cp_1x^{s-r})$, já que p_1 não é associado a nenhum dos q 's. Mas isto só é possível se p_1 dividir q_1 , contradizendo, novamente, o fato de que p_1 não é associado a nenhum dos q 's e completando a demonstração do teorema. \square

Ao contrário do que ocorre no caso dos números inteiros, existem algoritmos extremamente eficientes para fatorar polinômios. Na próxima seção estudaremos a relação entre as raízes de um polinômio e seus fatores, o que nos permitirá elaborar um pouco sobre o funcionamento destes algoritmos.

3. Divisão com resto e raízes

Começaremos a seção com duas definições. A primeira nos diz que um número real α é *raiz* de um polinômio f se $f(\alpha) = 0$; a outra, que um polinômio g *divide* um polinômio f se o resto da divisão de f por g for igual a zero. Isto basta para que possamos enunciar e provar o primeiro resultado desta seção.

TEOREMA 3.1. *Seja f um polinômio. Um número real α é raiz de f se, e somente se, o polinômio $x - \alpha$ divide f .*

PROOF. Se $x - \alpha$ divide f , então

$$f = (x - \alpha)q.$$

Substituindo $x = \alpha$ nesta igualdade,

$$f(\alpha) = (\alpha - \alpha)q(\alpha) = 0,$$

logo α é raiz de f . A recíproca é mais interessante. Suponhamos que α é raiz de f . Dividindo f por $x - \alpha$, obtemos

$$f = (x - \alpha)q + r \quad \text{e} \quad r = 0 \quad \text{ou} \quad \text{grau}(r) < 1.$$

Seja como for, $r \in \mathbb{R}$, pois $\text{grau}(r) = 0$ é o mesmo que dizer que r é uma constante. Logo, fazendo $x = \alpha$ na expressão acima, e lembrando que α é raiz de f , encontramos

$$0 = f(\alpha) = g(\alpha)(\alpha - \alpha) + r = r.$$

Lembre-se que r , sendo um número real, não é alterado quando substituímos α no lugar da variável. Mostramos, assim, que $f = (x - \alpha)q$, o que completa nossa demonstração. \square

A seguinte consequência do teorema 3.1 pode ser facilmente provada por indução na quantidade de raízes.

COROLÁRIO 3.2. *Se $\alpha_1, \dots, \alpha_k \in \mathbb{R}$ são raízes distintas de um polinômio f , então o produto $(x - \alpha_1) \cdots (x - \alpha_k)$ divide f .*

O corolário 3.2, por sua vez, nos permite estabelecer a seguinte relação entre a quantidade de raízes de um polinômio e o seu grau.

COROLÁRIO 3.3. *A quantidade de raízes de um polinômio não nulo é sempre menor ou igual do que o seu grau.*

Vimos, na seção 4 do capítulo 1, que duas funções $f_1, f_2 : D \rightarrow C$ são iguais se $f_1(x) = f_2(x)$, para todo $x \in D$. O próximo corolário mostra que, quando f_1 e f_2 são funções polinomiais, há um critério muito simples para determinar se são ou não iguais.

COROLÁRIO 3.4. *Sejam f_1 e f_2 funções polinomiais definidas, respectivamente, por*

$$f_1(x) = a_n x^n + \cdots + a_1 x + a_0 \quad e \quad f_2(x) = b_n x^n + \cdots + b_1 x + b_0.$$

Se $f_1(r) = f_2(r)$, para todo r em um subconjunto infinito $S \subset \mathbb{R}$, então $b_i = a_i$ para todo $1 \leq i \leq n$.

PROOF. Se h é a função polinomial definida por $h(x) = f_1(x) - f_2(x)$, então

$$h(r) = f_1(r) - f_2(r) = 0 \text{ para todo } r \in S.$$

Contudo, pelo corolário 3.3, uma função polinomial não nula não pode ter mais raízes que o seu grau. Como S é infinito, isto implica que

$$h(x) = (a_n - b_n)x^n + (a_{n-1} - b_{n-1})x^{n-1} + \cdots + (a_1 - b_1)x + (a_0 - b_0)$$

tem que ser o polinômio nulo, que é aquele cujos coeficientes são todos iguais a zero. Portanto,

$$a_n - b_n = a_{n-1} - b_{n-1} = \cdots = a_1 - b_1 = a_0 - b_0 = 0,$$

que equivale à afirmação no enunciado do corolário. \square

Encerraremos a seção apresentando um método para calcular as raízes racionais de um polinômio com coeficientes inteiros que, apesar de ser não muito eficiente, tem a virtude de ser fácil de explicar e de aplicar.

TEOREMA 3.5. *Seja f o polinômio definido por*

$$f(x) = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0,$$

em que a_n, \dots, a_1, a_0 são todos números inteiros. Se f tem uma raiz racional α , então o numerador de α divide a_0 e seu denominador divide a_n . Em particular, quando f é mônico, todas as suas raízes racionais são inteiras e dividem a_0 .

PROOF. Suponhamos que $\alpha = r/s$, em que $s \neq 0$ e r e s são inteiros primos entre si. Substituindo $x = r/s$ na expressão que define f , obtemos

$$f\left(\frac{r}{s}\right) = a_n \left(\frac{r}{s}\right)^n + a_{n-1} \left(\frac{r}{s}\right)^{n-1} + \cdots + a_1 \left(\frac{r}{s}\right) + a_0.$$

Eliminando os denominadores e lembrando que $\alpha = r/s$ é raiz de f , obtemos

$$(46) \quad a_n r^n + a_{n-1} s r^{n-1} + \cdots + a_1 s^{n-1} r + a_0 s^n = 0.$$

Mas todas as parcelas da soma acima aparecem multiplicadas por uma potência de r com expoente positivo, menos a última. Passando $a_0 s^n$ para o lado direito e pondo r em evidência do lado esquerdo,

$$r(a_n r^{n-1} + a_{n-1} s r^{n-2} + \cdots + a_1 s^{n-1}) = -a_0 s^n.$$

Logo, r divide $a_0 s^n$. Como estamos supondo que r e s não têm fatores comuns, a equação anterior mostra que r tem que dividir a_0 ; veja, por exemplo, [3, Lema, p. 43].

Para provar que s divide a_n , rearrumamos novamente a equação (46); só que desta vez passamos a primeira parcela para o lado direito, pondo s em evidência do lado esquerdo. Isto nos dá,

$$s(a_{n-1}r^{n-1} + \cdots + a_1s^{n-2}r + a_0s^{n-1}) = -a_nr^n.$$

Portanto, s divide a_nr^n e, como não divide r , tem que dividir a_n . \square

Vejamos como aplicar o teorema 3.5 para determinar as raízes racionais do polinômio

$$f(x) = 2x^4 - 5x^3 + x^2 - x + 6$$

Neste exemplo, o termo constante é 6 e o coeficiente do termo de maior grau é 2. Logo, se houver alguma raiz racional, ela deve ser da forma r/s , em que r divide 6 e s divide 2. Listando os divisores, *positivos e negativos* de 6 e 2, temos

$$r \in \{\pm 1, \pm 2, \pm 3, \pm 6\} \quad s \in \{\pm 1, \pm 2\}.$$

As possibilidades de quocientes entre r e s nos dois conjuntos anteriores são

$$\left\{-6, -3, -2, -\frac{3}{2}, -1, -\frac{1}{2}, \frac{1}{2}, 1, \frac{3}{2}, 2, 3, 6\right\}.$$

Para saber quais destas frações são raízes do polinômio f , calculamos a imagem por f de cada uma delas, o que nos dá

$$\begin{array}{lll} f(-6) = 3720, & f(-3) = 315, & f(-2) = 84 \\ f(-3/2) = 147/4 & f(-1) = 15, & f(-1/2) = 15/2, \\ f(1/2) = 21/4, & f(1) = 3, & f(3/2) = 0 \\ f(2) = 0, & f(3) = 39 & f(6) = 1548. \end{array}$$

Portanto, as únicas raízes racionais de f são $3/2$ e 2 .

4. Fórmulas de Viète

Como consequência do corolário 3.4 podemos provar as relações entre as raízes de um polinômio e seus coeficientes, originalmente obtidas, no caso em que as raízes do polinômio são positivas, por François Viète (1540-1603) e posteriormente generalizadas para raízes positivas e negativas por Albert Girard (1595-1632). Vejamos, primeiramente, como se apresentam no caso em que

$$(47) \quad f(x) = a_3x^3 + a_2x^2 + a_1x + a_0$$

é um polinômio é de grau três com três raízes reais, que denotaremos por α_1 , α_2 e α_3 . Pelo corolário 3.2, temos que

$$f(x) = c(x - \alpha_1)(x - \alpha_2)(x - \alpha_3),$$

para algum número real não nulo c . Multiplicando os três fatores,

$$(48) \quad f(x) = cx^3 - c(\alpha_1 + \alpha_2 + \alpha_3)x^2 + c(\alpha_1\alpha_2 + \alpha_2\alpha_3 + \alpha_1\alpha_3)x - c\alpha_1\alpha_2\alpha_3.$$

Mas, pelo corolário 3.4, dois polinômios são iguais se, e somente se, seus coeficientes coincidem. Comparando os coeficientes do polinômio f em (47) e (48), verificamos que

$$\begin{aligned} a_3 &= c, & a_1 &= c(\alpha_1\alpha_2 + \alpha_2\alpha_3 + \alpha_1\alpha_3) \\ a_2 &= -c(\alpha_1 + \alpha_2 + \alpha_3), & a_0 &= -c\alpha_1\alpha_2\alpha_3; \end{aligned}$$

onde

$$(49) \quad \alpha_1 + \alpha_2 + \alpha_3 = -\frac{a_2}{a_3}, \quad \alpha_1\alpha_2 + \alpha_2\alpha_3 + \alpha_1\alpha_3 = \frac{a_1}{a_3} \quad \text{e} \quad \alpha_1\alpha_2\alpha_3 = -\frac{a_0}{a_3},$$

que são as fórmulas de Viète para o polinômio de grau três.

Como as fórmulas de Viète estabelecem relações entre as raízes e os coeficientes de um polinômio, é quase inevitável que ocorra, a quem se depara com elas, a ideia de usá-las para achar as raízes do polinômio. Infelizmente isto não é possível. Para entender porque, basta considerar o caso em que o polinômio tem grau dois. Digamos que

$$f(x) = a_2x^2 + a_1x + a_0 = a_2(x - \alpha_1)(x - \alpha_2).$$

Neste caso, as fórmulas de Viète nos dão

$$\alpha_1 + \alpha_2 = -\frac{a_1}{a_2} \quad \text{e} \quad \alpha_1\alpha_2 = \frac{a_0}{a_2},$$

o que sugere achar o valor de α_2 na primeira fórmula e substituí-lo na segunda. Fazendo isto, obtemos

$$\alpha_1 \cdot \left(-\alpha_1 - \frac{a_1}{a_2} \right) = \frac{a_0}{a_2}.$$

Simplificando esta equação, resta

$$\alpha_1^2 + \frac{a_1}{a_2}\alpha_1 + \frac{a_0}{a_2} = 0,$$

que equivale a dizer que α_1 é raiz de $f(x)$ e, portanto, em nada nos ajuda a determinar α_1 .

As expressões em termos de α_1 , α_2 e α_3 nas fórmulas (49) são as funções simétricas elementares em três números. Para enunciar as relações de Viète no caso geral, precisamos definir estas funções no caso em que há n números. Dados números reais $\alpha_1, \dots, \alpha_n$, denotaremos por $\sigma_k(\alpha_1, \dots, \alpha_n)$ a soma de todos os produtos possíveis de k destes números; em outras palavras,

$$(50) \quad \sigma_k(\alpha_1, \dots, \alpha_n) = \sum_{1 \leq i_1 < \dots < i_k \leq n} \alpha_{i_1} \cdots \alpha_{i_k}.$$

Assim, quando $n = 3$, σ_1 , σ_2 e σ_3 correspondem ao lado esquerdo das fórmulas de Viète em (49). Com isto podemos enunciar o teorema.

TEOREMA 4.1. *Se o polinômio $f(x) = a_nx^n + \dots + a_1x + a_0$ tem raízes reais $\alpha_1, \dots, \alpha_n$, então,*

$$\sigma_k(\alpha_1 \cdots \alpha_n) = (-1)^k \frac{a_{n-k}}{a_n} \quad \text{para todo } 1 \leq k \leq n.$$

Há várias maneiras de provar este teorema, a que faremos utiliza o princípio de indução finita. Antes, porém, precisamos determinar qual a relação entre os números

$$\sigma_k(\alpha_1, \dots, \alpha_n) \text{ e } \sigma_k(\alpha_1, \dots, \alpha_{n+1});$$

que, de agora em diante, denotaremos por $\sigma_k^{(n)}$ e $\sigma_k^{(n+1)}$, para simplificar a notação. Quando $k = 1$, temos que

$$\sigma_1^{(n+1)} = \alpha_1 + \dots + \alpha_n + \alpha_{n+1} = \sigma_1^{(n)} + \alpha_{n+1}.$$

No outro extremo,

$$\sigma_n^{(n+1)} = \alpha_1 \cdots \alpha_n \cdot \alpha_{n+1} = \sigma_0^{(n)} \cdot \alpha_{n+1}.$$

Resta-nos considerar o caso em que $2 \leq k \leq n - 1$. Para isso, dividimos $\sigma_k^{(n+1)}$ em dois somatórios, um deles com as parcelas em que aparece α_{n+1} e o outro com aquelas em que este número não aparece. Fazendo isto a partir da definição em (50), obtemos

$$(51) \quad \sigma_k^{(n+1)} = \left(\sum_{1 \leq i_1 < \dots < i_k \leq n} \alpha_{i_1} \cdots \alpha_{i_k} \right) + \alpha_{n+1} \left(\sum_{1 \leq i_1 < \dots < i_{k-1} \leq n} \alpha_{i_1} \cdots \alpha_{i_{k-1}} \right).$$

Note que, no segundo somatório à direita, há apenas $k - 1$ índices, porque a k -ésima raiz em cada produto é α_{n+1} , que já pusemos em evidência. Contudo, por definição,

$$\sigma_k^{(n)} = \sum_{1 \leq i_1 < \dots < i_k \leq n} \alpha_{i_1} \cdots \alpha_{i_k} \text{ e } \sigma_{k-1}^{(n)} = \sum_{1 \leq i_1 < \dots < i_{k-1} \leq n} \alpha_{i_1} \cdots \alpha_{i_{k-1}}.$$

Substituindo isto em (51), obtemos

$$\sigma_k^{(n+1)} = \sigma_k^{(n)} + \alpha_{n+1} \sigma_{k-1}^{(n)}.$$

Resumindo,

$$(52) \quad \sigma_k^{(n+1)} = \begin{cases} \sigma_1^{(n)} + \alpha_{n+1} & \text{se } k = 1 \\ \sigma_k^{(n)} + \alpha_{n+1} \sigma_{k-1}^{(n)} & \text{se } 2 \leq k \leq n - 1 \\ \sigma_0^{(n)} \cdot \alpha_{n+1} & \text{se } k = n, \end{cases}$$

que são as fórmulas de que precisamos na demonstração das relações de Viète.

PROOF. Para que

$$f(x) = a_n x^n + \dots + a_1 x + a_0$$

tenha grau n é necessário que $a_n \neq 0$. Dividindo $f(x)$ por a_n , obtemos o polinômio mônico

$$g(x) = \frac{1}{a_n} f(x) = x^n + b_{n-1} x^{n-1} + \dots + b_1 x + b_0,$$

em que

$$b_{n-1} = \frac{a_{n-1}}{a_n}, \dots, b_1 = \frac{a_1}{a_n} \text{ e } b_0 = \frac{a_0}{a_n}.$$

Como $g(\alpha) = 0$, se e somente se, $f(\alpha) = 0$, estes dois polinômios têm as mesmas raízes. Portanto, denotando por $\alpha_1, \dots, \alpha_n$ as raízes destes polinômios, basta mostrarmos que

se $\alpha_1, \dots, \alpha_n$ são as raízes de $g(x) = x^n + \dots + b_1x + b_0$, então

$$\sigma_k^{(n)} = (-1)^k b_{n-k}, \quad \text{para todo } 1 \leq k \leq n.$$

para termos o resultado desejado. Provaremos esta última afirmação por indução no grau do polinômio. A base da indução corresponde ao caso em que o polinômio $g(x)$ é mônico de grau um. Assim, $g(x) = x + b_0$, de modo que

$$\sigma_1 = -b_0,$$

que é a única raiz de $g(x)$. Supondo, agora, que a afirmação acima vale para todo polinômio mônico de grau n , provaremos que também é válida para todo polinômio mônico de grau $n+1$. Mas, polinômio mônico

$$\gamma(x) = (x - \alpha_1) \cdots (x - \alpha_n)(x - \alpha_{n+1})$$

pode ser escrito na forma

$$(53) \quad \gamma(x) = g(x)(x - \alpha_{n+1}) = xg(x) - \alpha_{n+1}g(x)$$

em que

$$g(x) = (x - \alpha_1) \cdots (x - \alpha_n).$$

Contudo, pela hipótese de indução,

$$g(x) = x^n + \sum_{i=1}^n (-1)^i \sigma_i^{(n)} x^{n-i}.$$

Substituindo esta expressão no lado direito de (53), obtemos

$$\gamma(x) = x \left(x^n + \sum_{i=1}^n (-1)^i \sigma_i^{(n)} x^{n-i} \right) + \alpha_{n+1} \left(x^n + \sum_{i=1}^n (-1)^i \sigma_i^{(n)} x^{n-i} \right),$$

onde

$$\gamma(x) = x^{n+1} + \sum_{i=1}^n (-1)^i \sigma_i^{(n)} x^{n-i+1} + \alpha_{n+1} x^n + \sum_{i=1}^n (-1)^i \alpha_{n+1} \sigma_i^{(n)} x^{n-i}.$$

Somando os coeficientes dos termos de mesmo expoente nos dois polinômios, encontramos

$$\gamma(x) = x^{n+1} - (\sigma_1^{(n)} + \alpha_{n+1}) x^n + \sum_{i=2}^n (-1)^i (\alpha_{n+1} \sigma_{i-1}^{(n)} + \sigma_i^{(n)}) x^{n-i} + (-1)^{n+1} \sigma_n^{(n)} \alpha_{n+1}.$$

O resultado desejado é, então, consequência de (52). \square

CHAPTER 6

Trigonometria

Ao contrário da geometria, que já na Grécia Antiga se caracterizava por um tratamento extremamente abstrato e teórico, a trigonometria surgiu na Grécia basicamente como adjunto à astronomia. Contudo, os gregos utilizavam cordas do círculo; o seno e o cosseno que utilizamos hoje em dia foram introduzidos pelos matemáticos indianos antes de 543 E.C. Neste capítulo apresentamos as principais funções trigonométricas de maneira rigorosa, mas usando apenas o que aprendemos sobre geometria no capítulo ??.

1. Seno e cosseno

Se α for um dos ângulos internos de um triângulo retângulo, diferente do ângulo reto, podemos facilmente definir seu seno e cosseno. Denotando os vértices do triângulo por O , A e B , de modo que α seja o ângulo do vértice O , como ilustrado na figura 1, definimos

$$(54) \quad \text{sen}(\alpha) = \frac{\overline{AB}}{\overline{OB}} \quad \text{e} \quad \cos(\alpha) = \frac{\overline{OA}}{\overline{OB}}.$$

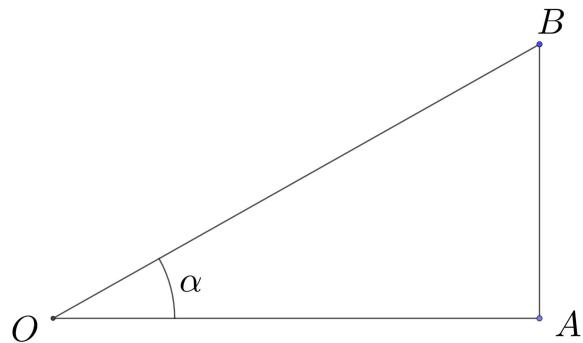


FIGURE 1. Triângulo retângulo com ângulo α

Como,

$$\sin^2(\alpha) + \cos^2(\alpha) = \left(\frac{\overline{AB}}{\overline{OB}}\right)^2 + \left(\frac{\overline{OA}}{\overline{OB}}\right)^2 = \frac{\overline{AB}^2 + \overline{OA}^2}{\overline{OB}^2},$$

temos, pelo teorema de Pitágoras, que

$$(55) \quad \sin^2(\alpha) + \cos^2(\alpha) = 1.$$

Em particular, se construirmos o triângulo retângulo de modo que sua hipotenusa tenha comprimento igual a 1, então o vértice B cai sobre o círculo de raio 1. Isto nos permite redesenhar a figura 1 na forma abaixo. A propósito: neste contexto frequentemente nos referimos ao círculo de raio como *o círculo trigonométrico*.

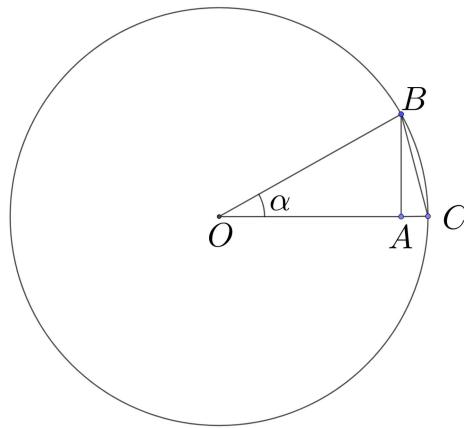


FIGURE 2. Triângulo retângulo inscrito no círculo

Sob a hipótese de que a hipotenusa OB é igual a 1, as fórmulas em (54) resumem-se a

$$(56) \quad \sin(\alpha) = \overline{AB} \text{ e } \cos(\alpha) = \overline{OA}.$$

Como observado na introdução deste capítulo, o seno e o cosseno foram introduzidos pelos matemáticos indianos; os matemáticos gregos utilizavam a corda \overline{BC} . Podemos calculá-la, facilmente, usando o seno, o cosseno e o teorema de Pitágoras, pois \overline{BC} é a hipotenusa do triângulo retângulo ABC cujos catetos são $\sin(\alpha)$ e $1 - \cos(\alpha)$. Assim,

$$\overline{BC}^2 = \sin^2(\alpha) + (1 - \cos(\alpha))^2,$$

de modo que, por (55),

$$\overline{BC}^2 = 2 - 2\cos(\alpha).$$

Uma das vantagens do seno e cosseno sobre as cordas, é que os dois primeiros são mais fáceis de aproximar.

Para que senos e cossenos possam ser efetivamente utilizados, precisamos de uma maneira de medir ângulos. Faremos isto indicando que fração do comprimento 2π da circunferência de raio 1 é subtendida pelo ângulo. Isto nos dá a medida de ângulos em *radianos*. Como não provamos a fórmula para o comprimento de uma circunferência, a atribuição de 2π ao comprimento da circunferência de raio 1 vai soar um pouco arbitrário. Mas isto não é um problema. Na verdade, estamos em muito boa companhia, porque na matemática dos povos da Mesopotâmia a circunferência toda valia 360 e, durante a Revolução Francesa, foi introduzida uma medição decimal de ângulos com o ângulo reto valendo 100. No primeiro caso temos a medição em graus, que você conhece, no segundo a unidade é chamada de *grado*, que é uma unidade oficial na União Européia e na Suíça.

Como, para cada triângulo retângulo de hipotenusa igual a um, existe apenas um cateto oposto ao ângulo $\alpha = \widehat{COB}$, as fórmulas em (56) defininem o seno e o cosseno como funções de α . Contudo, como os ângulos opostos a um cateto de um triângulo retângulo têm que ter soma igual a $\pi/2$, as fórmulas em (56) só nos permitem calcular o seno e o cosseno de α quando $0 < \alpha < \pi/2$. Contornaremos este problema generalizando as definições do seno e do cosseno, para que possamos tratá-los como funções cujo domínio é todo o conjunto \mathbb{R} .

Como primeiro passo na direção desta generalização, posicionaremos o centro do círculo de raio 1 na origem, como na figura 3. Com isso, o cosseno \overline{OA} e o seno \overline{AB} do ângulo α correspondem à abscissa e ordenada do vértice B do triângulo OAB . Podemos, então, redefinir, o seno e o cosseno de $\alpha = \widehat{COB}$ como sendo a ordenada e a abscissa do ponto B . Em outras palavras, se $B = (b_1, b_2)$, então

$$\text{sen}(\alpha) = b_2 \quad \text{e} \quad \cos(\alpha) = b_1.$$

Como $\alpha = 0$ se, e somente se, $B = (1, 0)$, segue-se desta nova definição que

$$\text{sen}(0) = 0 \quad \text{e} \quad \cos(0) = 1.$$

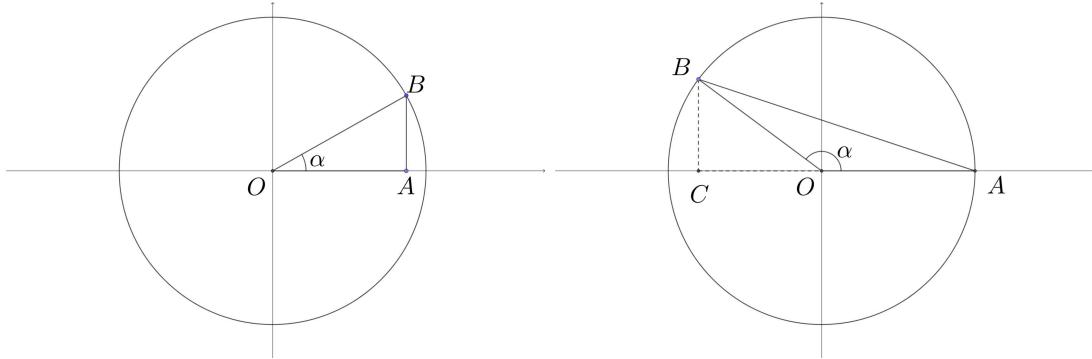
Analogamente, $\alpha = \pi/2$ quando $B = (0, 1)$, de modo que

$$\text{sen}(\pi/2) = 1 \quad \text{e} \quad \cos(\pi) = 0.$$

Se continuarmos a aumentar o ângulo α , o ponto B vai sendo movido de quadrante em quadrante, e o sinal do seno e do cosseno vão sendo modificados como na tabela 1 da página 91. Por exemplo, se $\pi/2 < \alpha < \pi$, como na figura 4, o seno \overline{BC} continuará sendo positivo, ao passo que o cosseno \overline{OC} será negativo.

Com isso podemos determinar o seno e o cosseno dos ângulos que são múltiplos de $\pi/2$. Por exemplo, quando o ângulo α é igual a π , o ponto correspondente ao vértice B tem coordenadas $(-1, 0)$, de modo que

$$\text{sen}(\pi) = 0 \quad \text{e} \quad \cos(\pi) = -1.$$

FIGURE 3. $0 < \alpha < \pi/2$ FIGURE 4. $\pi/2 < \alpha < \pi$

Por outro lado, quando $\alpha = 3\pi/2$, temos que $B = (0, -1)$, o que nos dá

$$\sin\left(\frac{3\pi}{2}\right) = -1 \quad \text{e} \quad \cos\left(\frac{3\pi}{2}\right) = 0.$$

Tomando $\alpha = 2\pi$, teremos dado uma volta completa na circunferência, de modo que o seno e o cosseno de 2π serão os mesmos do ângulo zero. De maneira mais geral, qualquer que seja o ângulo α com que começamos, ao completarmos uma volta na circunferência a partir de α , estaremos de volta ao mesmo ponto do círculo. Assim,

$$\sin(\alpha + 2\pi) = \sin(\alpha) \quad \text{e} \quad \cos(\alpha + 2\pi) = \cos(\alpha).$$

Outra relação muito útil é aquela entre os senos e cossenos de ângulos que diferem por $\pi/2$ radianos. Suponhamos que, como na figura 5, $\beta = \alpha + \pi/2$. Como

$$\beta + \widehat{DOC} = \pi,$$

temos que

$$(57) \quad \alpha + \widehat{DOC} = \frac{\pi}{2}.$$

Mas os ângulos em \widehat{OAB} e \widehat{ODC} medem $\pi/2$ e a soma dos ângulos em um triângulo é igual a π , de modo que, por (57),

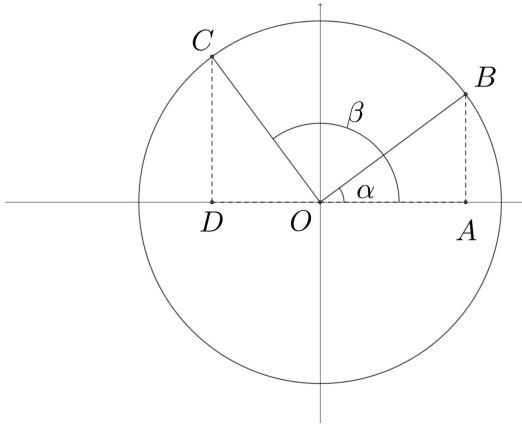
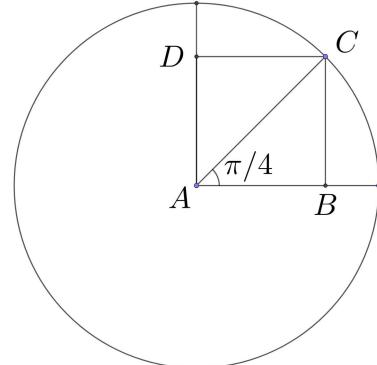
$$\widehat{OBA} \equiv \widehat{DOC} \quad \text{e} \quad \widehat{AOB} \equiv \widehat{OCD}.$$

Combinando isto com

$$\overline{OC} \equiv \overline{OB},$$

pois ambos são raios do mesmo círculo, temos pelo caso ALA que os triângulos OBA e ODC são congruentes. Logo,

$$\overline{OD} \equiv \overline{AB} \quad \text{e} \quad \overline{DC} \equiv \overline{OA},$$

FIGURE 5. $\beta - \alpha = \pi/2$ FIGURE 6. Seno e cosseno de $\pi/4$

já que os vértices O , A e B de um dos triângulos correspondem, respectivamente, a C , D e O do outro. Portanto,

$$\sin(\alpha) = \overline{AB} = -\overline{OD} = -\cos(\beta) \quad \text{e} \quad \cos(\alpha) = \overline{OA} = \overline{CD} = \sin(\beta).$$

Em outras palavras,

$$(58) \quad \sin\left(\alpha + \frac{\pi}{2}\right) = \cos(\alpha) \quad \text{e} \quad \cos\left(\alpha + \frac{\pi}{2}\right) = -\sin(\alpha).$$

Mais interessante é o cálculo do seno e do cosseno do ângulo de $\pi/4$ radianos, correspondente a um oitavo de circunferência. Quando $\alpha = \pi/4$, o vértice B cai, exatamente, a meio caminho entre zero e $\pi/2$, como mostra a figura 6. Assim,

$$\widehat{BAC} = \widehat{CAD} = \frac{\pi}{4}.$$

Mas,

$$\widehat{ADC} \equiv \widehat{ABC}$$

são ambos ângulos retos, o que nos permite concluir, pelo teorema 6.8, que

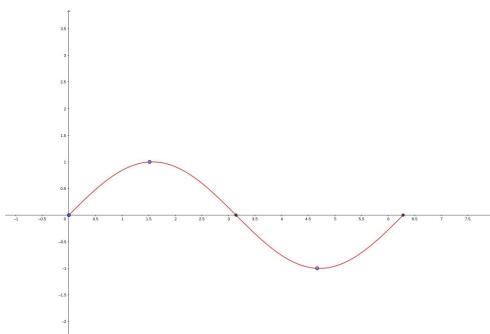
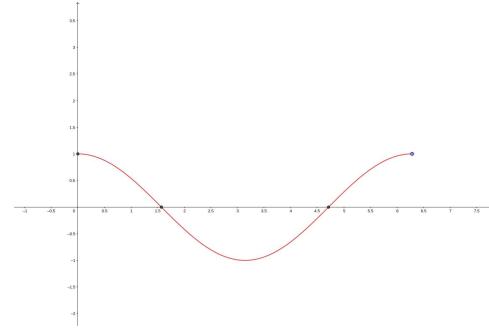
$$\widehat{ACB} \equiv \widehat{ACD}.$$

Logo, os triângulos ABC e ADC são congruentes pelo caso ALA, porque \overline{AC} é um lado comum a ambos. Em particular,

$$\cos\left(\frac{\pi}{4}\right) = \overline{AB} = \overline{AD} = \sin\left(\frac{\pi}{4}\right).$$

Levando isto em conta, obtemos por (55) que

$$2 \cos^2\left(\frac{\pi}{4}\right) = 1.$$

FIGURE 7. O seno de 0 a 2π .FIGURE 8. O cosseno de 0 a 2π .

Como $\pi/4 < \pi/2$, o cosseno e o seno serão positivos, donde

$$\sin\left(\frac{\pi}{4}\right) = \cos\left(\frac{\pi}{4}\right) = \frac{\sqrt{2}}{2}.$$

Dos ângulos mais conhecidos, falta-nos, apenas, calcular o seno e o cosseno de $\pi/3$ e $\pi/6$. Faremos isto, na próxima seção, como aplicação das fórmulas do seno e do cosseno de uma soma de dois ângulos. Antes, porém, usaremos os valores que já temos para esboçar os gráficos do seno e do cosseno, como ilustram as figuras 7 e 8.

Estas duas figuras partem do princípio de que o domínio de ambas as funções corresponde aos valores que α toma quando damos uma única volta no círculo trigonométrico. Contudo, nada nos impede de dar mais de uma volta no círculo, tanto no sentido anti-horário, quanto no sentido horário. Naturalmente, a cada volta, o ângulo aumenta de 2π . Por outro lado, dois ângulos que diferem de 2π determinam exatamente o mesmo ponto no círculo trigonométrico, de modo que têm exatamente os mesmos valores do seno e do cosseno. Assim, admitindo que o ângulo tome qualquer valor real, os gráficos destas funções são obtidos copiando e colando os gráficos das figuras 7 e 8 nos intervalos,

$$\dots, [-4\pi, -2\pi], [-2\pi, 0], [0, 2\pi], [2\pi, 4\pi], \dots,$$

como ilustrado para o caso do seno na figura 9. Quando os valores de uma função se repetem a intervalos fixos de comprimento p , dizemos que a função é *periódica de período p* . Portanto, o seno e o cosseno são ambos funções periódicas de período 2π . Não é coincidência que estas funções tenham sido introduzidas originalmente por matemáticos interessados em astronomia que, na antiguidade, era o domínio por excelência dos fenômenos periódicos.

2. Seno e cosseno de soma de ângulos

Nesta seção provaremos as fórmulas para o seno e o cosseno da soma de dois ângulos. Todo o nosso argumento será baseado na figura 10, na qual o ângulo \widehat{OCQ} é reto. Note

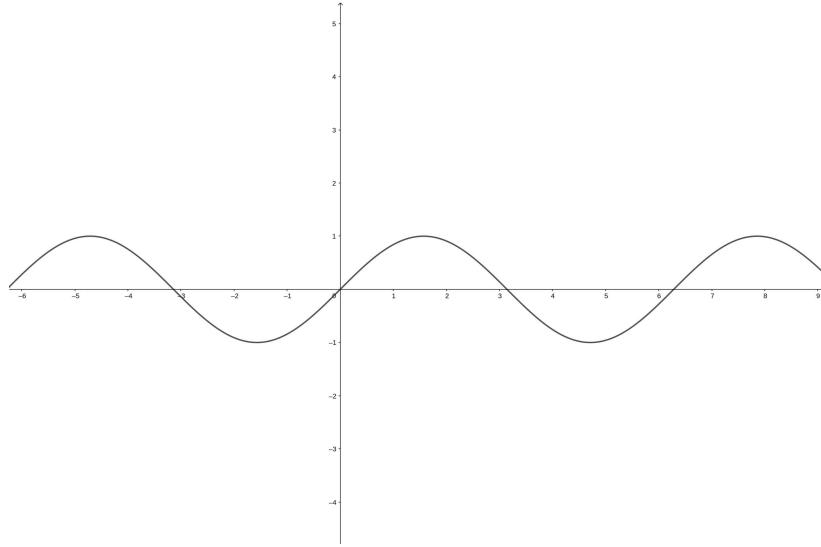


FIGURE 9. O seno como função periódica

que

$$\widehat{OEA} \equiv \widehat{CED},$$

porque são opostos pelo vértice. Como os triângulos OBC e CEQ são ambos retângulos, segue-se que

$$\widehat{CQE} = \widehat{BOC} = \alpha.$$

Isto nos permite calcular o seno e o cosseno de α a partir dos triângulos OBC , QDC e QCE . Faremos uso disto no cálculo de ambas as fórmulas para soma de ângulos.

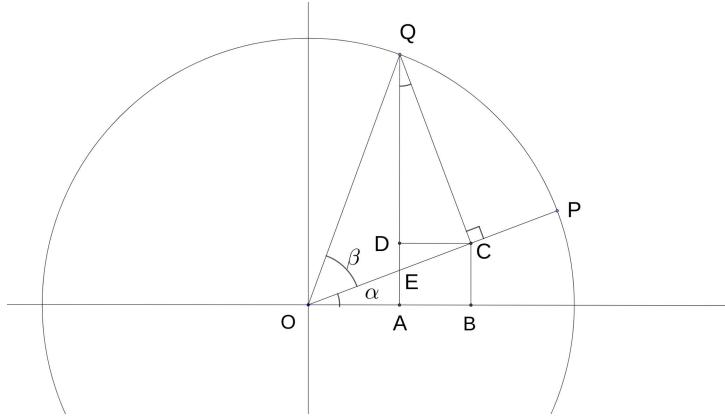


FIGURE 10. Seno e cosseno de uma soma de ângulos.

Começaremos calculando

$$(59) \quad \sin(\alpha + \beta) = \overline{QA} = \overline{QD} + \overline{DA}.$$

Mas, do triângulo retângulo QDC ,

$$\cos(\alpha) = \frac{\overline{QD}}{\overline{QC}}; \quad \text{isto é,} \quad \overline{QD} = \overline{QC} \cos(\alpha).$$

Por outro lado, do triângulo retângulo OQC ,

$$\sin(\beta) = \frac{\overline{QC}}{\overline{OQ}},$$

onde

$$(60) \quad \overline{QC} = \sin(\beta),$$

pois $\overline{OQ} = 1$ é o raio do círculo. Logo,

$$(61) \quad \overline{QD} = \cos(\alpha) \sin(\beta).$$

Para calcular \overline{DA} , note, primeiramente, que $\overline{DA} = \overline{CB}$, de modo que

$$\sin(\alpha) = \frac{\overline{BC}}{\overline{OC}}; \quad \text{isto é,} \quad \overline{BC} = \overline{OC} \sin(\alpha).$$

Porém, pelo triângulo OQC ,

$$(62) \quad \cos(\beta) = \frac{\overline{OC}}{\overline{OQ}} = \overline{OC}.$$

Assim,

$$(63) \quad \overline{BC} = \cos(\beta) \sin(\alpha).$$

Combinando (59) com (61) e (63), obtemos

$$(64) \quad \sin(\alpha + \beta) = \cos(\alpha) \sin(\beta) + \cos(\beta) \sin(\alpha),$$

que é a primeira das duas fórmulas que almejamos obter. Um caso particular muito útil é obtido fazendo $\beta = \alpha$ em (64), que nos dá,

$$(65) \quad \sin(2\alpha) = 2 \sin(\alpha) \cos(\alpha).$$

O argumento usado para determinar a fórmula do cosseno da soma de dois ângulos é muito semelhante ao anterior. Como a hipotenusa \overline{OQ} do triângulo OQA é igual a 1,

$$(66) \quad \cos(\alpha + \beta) = \overline{OA} = \overline{OB} - \overline{AB}.$$

Mas, calculando o cosseno de α relativamente ao triângulo OCB ,

$$(67) \quad \cos(\alpha) = \frac{\overline{OB}}{\overline{OC}}, \quad \text{isto é,} \quad \overline{OB} = \cos(\alpha) \overline{OC}.$$

Calculando, agora, o seno de α relativamente ao triângulo DCQ ,

$$(68) \quad \sin(\alpha) = \frac{\overline{DC}}{\overline{QC}}, \quad \text{onde,} \quad \overline{DC} = \sin(\alpha) \overline{QC}.$$

Lembrando que $\overline{AB} = \overline{DC}$, basta combinar as equações (66), (67) e (68) para obter

$$\cos(\alpha + \beta) = \cos(\alpha)\overline{OC} - \sin(\alpha)\overline{QC}.$$

Portanto, por (60) e (62),

$$(69) \quad \cos(\alpha + \beta) = \cos(\alpha)\cos(\beta) - \sin(\alpha)\sin(\beta).$$

Substituindo $\beta = \alpha$ em (69),

$$(70) \quad \cos(2\alpha) = \cos(\alpha)^2 - \sin(\alpha)^2 = 2\cos(\alpha)^2 - 1,$$

que é a fórmula para o cosseno do dobro de um ângulo.

As fórmulas que obtivemos nesta seção nos permitem determinar o seno e o cosseno de $\pi/3$ e de $2\pi/3$. Para isto substituímos $\alpha = 2\pi/3$ e $\beta = \pi/3$ em (64), obtendo

$$\sin\left(\frac{2\pi}{3} + \frac{\pi}{3}\right) = \sin\left(\frac{\pi}{3}\right)\cos\left(\frac{2\pi}{3}\right) + \cos\left(\frac{\pi}{3}\right)\sin\left(\frac{2\pi}{3}\right)$$

Mas,

$$\sin\left(\frac{2\pi}{3} + \frac{\pi}{3}\right) = \sin(\pi) = 0;$$

onde

$$\sin\left(\frac{\pi}{3}\right)\cos\left(\frac{2\pi}{3}\right) = -\cos\left(\frac{\pi}{3}\right)\sin\left(\frac{2\pi}{3}\right)$$

Usando, agora, (65) e (70),

$$\sin\left(\frac{\pi}{3}\right) \cdot \left(2\cos^2\left(\frac{\pi}{3}\right) - 1\right) = -\cos\left(\frac{\pi}{3}\right) \cdot 2\sin\left(\frac{\pi}{3}\right)\cos\left(\frac{\pi}{3}\right).$$

Como $\pi/3$ não é múltiplo de π , seu seno é não nulo. Cancelando-o na equação anterior, encontramos

$$2\cos^2\left(\frac{\pi}{3}\right) - 1 = -\cos\left(\frac{\pi}{3}\right) \cdot 2\cos\left(\frac{\pi}{3}\right);$$

que equivale a

$$4\cos^2\left(\frac{\pi}{3}\right) = 1.$$

Extraindo a raiz quadrada,

$$\cos\left(\frac{\pi}{3}\right) = \pm\frac{1}{2}.$$

Porém, como $\pi/3$ está no primeiro quadrante, seu cosseno é positivo. Combinando isto com

$$\sin^2\left(\frac{\pi}{3}\right) + \cos^2\left(\frac{\pi}{3}\right) = 1,$$

e levando em conta que o seno de $\pi/3$ também é positivo,

$$\sin\left(\frac{\pi}{3}\right) = \frac{\sqrt{3}}{2}.$$

Por outro lado, pela fórmula do cosseno do dobro de um ângulo,

$$\cos\left(\frac{\pi}{3}\right) = \cos\left(\frac{2\pi}{6}\right) = 2\cos^2\left(\frac{\pi}{6}\right) - 1;$$

onde

$$\frac{1}{2} = 2\cos^2\left(\frac{2\pi}{6}\right) - 1.$$

Assim

$$\cos^2\left(\frac{2\pi}{6}\right) = \frac{3}{4}.$$

Como $\pi/6$ também está no primeiro quadrante, seu cosseno é positivo. Logo,

$$\cos\left(\frac{2\pi}{6}\right) = \frac{\sqrt{3}}{2}.$$

Mas, pela fórmula do seno do dobro de um ângulo,

$$\sin\left(\frac{\pi}{3}\right) = \sin\left(\frac{2\pi}{6}\right) = 2\sin\left(\frac{\pi}{6}\right)\cos\left(\frac{\pi}{6}\right),$$

da qual obtemos

$$\sin\left(\frac{\pi}{6}\right) = \frac{\sin\left(\frac{\pi}{3}\right)}{2\cos\left(\frac{\pi}{6}\right)} = \frac{\sqrt{3}/2}{\sqrt{3}/2} = \frac{1}{2}.$$

Na tabela 1 listamos os valores dos senos e cosenos dos ângulos do primeiro quadrante que calculamos até aqui.

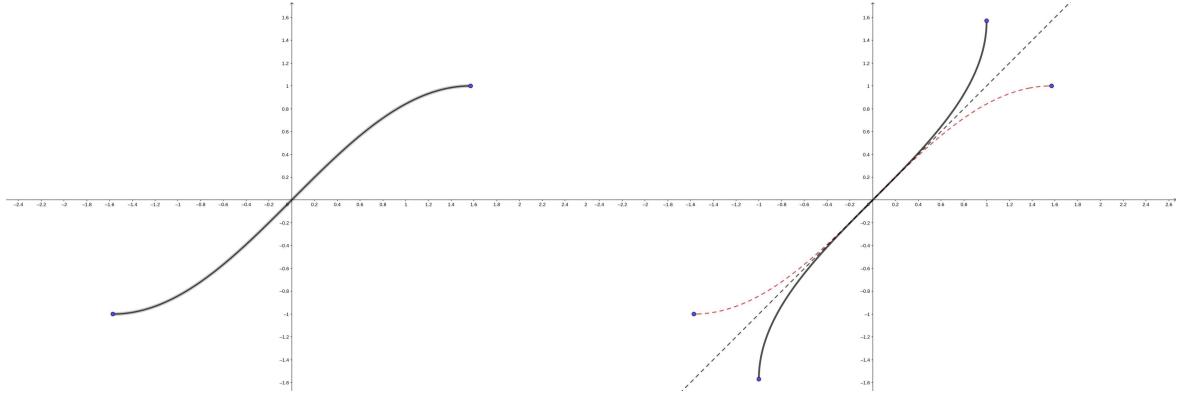
Ângulo	0	$\pi/6$	$\pi/4$	$\pi/3$
Seno	0	$1/2$	$\sqrt{2}/2$	$\sqrt{3}/2$
Cosseno	1	$\sqrt{3}/2$	$\sqrt{2}/2$	$1/2$

TABLE 1. Senos e cosenos de ângulos do primeiro quadrante

Encerraremos esboçando o gráfico das funções inversas do seno e do cosseno. Lembre-se que, para que uma função tenha inversa é necessário que seja injetora e sobrejetora. Como as imagens do seno e do cosseno correspondem ao intervalo $[-1, 1]$, este será também o domínio de suas inversas. Contudo, nem o seno, nem o cosseno são injetivas no intervalo $[0, 2\pi]$. Contornamos isto limitando seus domínios. Para o domínio do seno, escolheremos o intervalo $[-\pi/2, \pi/2]$ por duas razões: o seno é uma função injetora neste intervalo e seu gráfico é simétrico em relação aos eixos. Com isso, podemos construir

$$\text{sen}^{-1} : [-1, 1] \longrightarrow [-\pi/2, \pi/2],$$

que é conhecida como *arco-seno*, e usualmente denotada por arcsen . Voltando à escolha de domínio para o seno, optamos por um intervalo em que a curva é simétrica relativamente aos eixos porque isto facilita a construção do gráfico de sua inversa a partir da reta $y = x$; ver a seção 4. Na figura 11 temos, à direita o gráfico do arco-seno, sobreposto aos gráficos

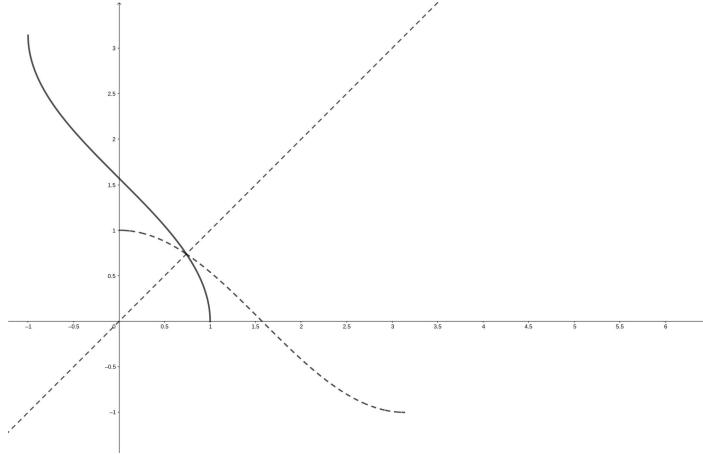
FIGURE 11. Gráficos do seno em $[-\pi/2, \pi/2]$ e do arco-seno.

tracejados do seno, restrito ao intervalo $[0, \pi]$, e da reta $y = x$. Como os esboços das duas curvas se sobrepõe, o gráfico do seno foi desenhado separadamente à esquerda para facilitar a comparação.

Como $\cos(-a) = \cos(a)$, para todo número real a , o cosseno não é uma função injetora em nenhum intervalo da forma $[-a, a]$. Por isso, restringiremos o domínio do cosseno ao intervalo $[0, \pi]$ para poder invertê-lo. Com isso, a função inversa cosseno, chamada de *arco-cosseno*, é dada por

$$\arccos = \cos^{-1} : [-1, 1] \longrightarrow [-\pi, \pi].$$

Na figura 12 a curva cheia representa o gráfico do arco-cosseno, ao passo que os gráficos do cosseno, restrito ao intervalo $[0, \pi]$, e da reta $y = x$ foram desenhados tracejados.

FIGURE 12. Gráficos do cosseno em $[0, \pi]$ e do arco-cosseno.

3. Tangente

Tendo investigado o seno e o cosseno, passamos a analisar a *tangente* de um ângulo α que é definida por

$$\tan(\alpha) = \frac{\sin(\alpha)}{\cos(\alpha)}.$$

Note que este quociente *não* está definido quando

$$\alpha \in \left\{ (2k+1)\frac{\pi}{2} \mid k \in \mathbb{Z} \right\},$$

porque o cosseno se anula nos ângulos deste conjunto. Portanto, a tangente é uma função cujo domínio é o intervalo $(-\pi/2, \pi/2)$. Usando os valores dos senos e cossenos calculados nas seções anteriores é fácil verificar que

$$\tan(0) = 0, \quad \tan\left(\frac{\pi}{6}\right) = \frac{\sqrt{3}}{3}, \quad \tan\left(\frac{\pi}{4}\right) = \frac{\sqrt{2}}{2} \quad \text{e} \quad \tan\left(\frac{\pi}{3}\right) = \sqrt{3}.$$

Como no caso do seno e do cosseno, a tangente pode ser identificada com um dos lados de um triângulo associado ao círculo trigonométrico. Para isto precisamos analisar os triângulos OBA e OCD da figura 13, que foi construída de modo que as retas DC e OC sejam perpendiculares em C . Como AB também é perpendicular a OA e como o ângulo \widehat{AOB} é comum aos triângulos OBA e OCD , podemos concluir, pelo teorema 8.2, que estes triângulos são semelhantes. Logo,

$$\frac{\overline{CD}}{1} = \frac{\overline{AB}}{\overline{OA}} = \frac{\sin(\alpha)}{\cos(\alpha)};$$

onde,

$$\overline{CD} = \tan(\alpha).$$

Isto também explica porque a tangente tem este nome, já que a reta DC é tangente à circunferência em C . Para entender o porque esta última afirmação é válida, considere um ponto qualquer da reta DC , que seja diferente de C . Digamos que este ponto seja D . Por Pitágoras,

$$\overline{OD} = \sqrt{1 + \overline{DC}^2}.$$

Como estamos supondo que $D \neq C$, temos que $\overline{DC} > 0$, de modo que, pela fórmula anterior, $\overline{OD} > 1$. Como o círculo é formado por todos os pontos que estão à distância 1 da origem, podemos afirmar que D não está sobre o círculo. Portanto, CD só intersecta o círculo em C , coincidindo, assim, com a tangente neste ponto.

A definição da tangente de α em termos do segmento \overline{CD} mostra que, ao contrário do seno e do cosseno, o valor da tangente não tem nenhum limite. Mais precisamente, à medida que α aumenta em direção a $\pi/2$, o valor da tangente aumenta de maneira ilimitada, ao passo que $\tan(\alpha)$ diminui de maneira ilimitada quando α se aproxima de $-\pi/2$. Isto nos

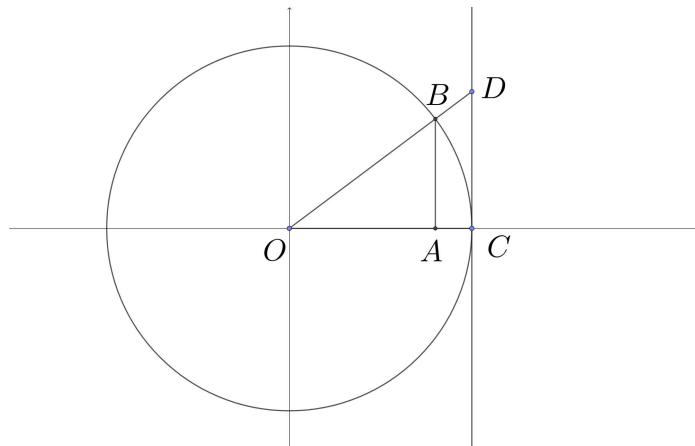
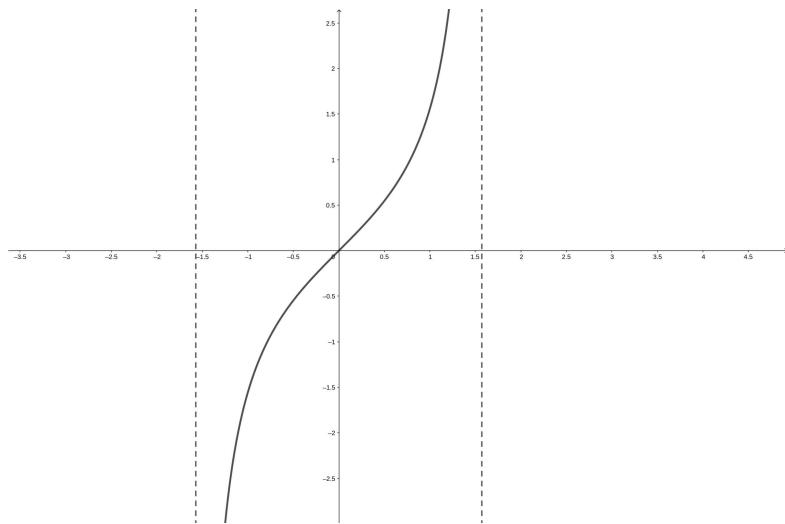


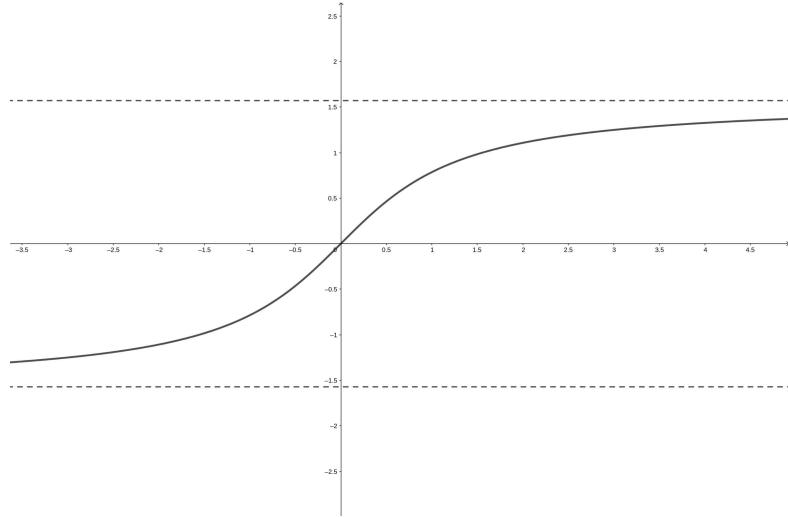
FIGURE 13. Interpretação geométrica da tangente

permite esboçar o gráfico da tangente no intervalo $(-\pi/2, \pi/2)$, como ilustrado na figura 14.

FIGURE 14. Gráfico da tangente no intervalo $(-\pi/2, \pi/2)$.

A função inversa da tangente, conhecida como *arco-tangente* e denotada por arctan, tem \mathbb{R} como domínio e o intervalo $(-\pi/2, \pi/2)$ como imagem. Como no caso do seno e do cosseno, seu gráfico pode ser facilmente determinado simetrizando o gráfico da figura 14 relativamente à reta $y = x$, como ilustrado na figura 15, na qual as retas horizontais têm equações $y = -\pi/2$ e $y = \pi/2$.

Agora que sabemos como definir a tangente de um ângulo α em termos do seu seno e cosseno de α , vejamos como inverter os papéis e determinar o seno e o cosseno deste

FIGURE 15. Gráfico do arco-tangente no intervalo $(-\pi/2, \pi/2)$.

ângulo a partir da tangente, não de α , mas de $\alpha/2$. As fórmulas resultantes são muito úteis, como teremos oportunidade de ver quando estudarmos identidades entre funções trigonométricas. Nossa demonstração das fórmulas para o seno e o cosseno de α relativamente à tangente de $\alpha/2$ será baseada na figura 16.

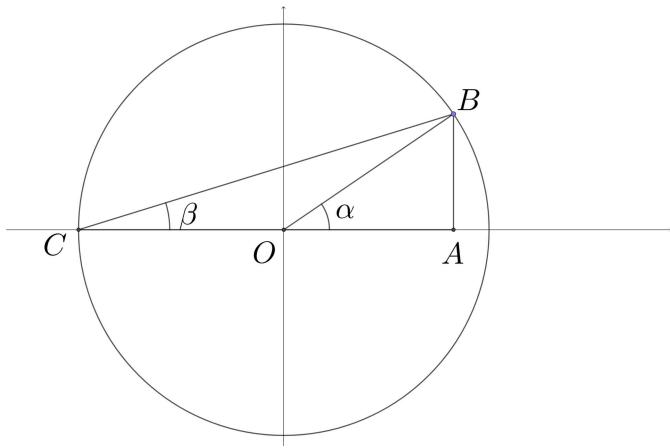


FIGURE 16. Seno e cosseno relativamente à tangente do meio ângulo

A primeira coisa a observar é que o triângulo COB é isósceles, porque \overline{OC} e \overline{OB} são raios do círculo unitário. Logo $\angle CBO = \beta$ pelo teorema 5.6. Podemos, então, concluir do teorema 6.8 que

$$2\beta + (\pi - \alpha) = \pi;$$

onde

$$\beta = \frac{\alpha}{2}.$$

Calculando a tangente de β relativamente ao triângulo ABC , obtemos

$$(71) \quad \tan(\beta) = \frac{\overline{AB}}{\overline{AC}}.$$

Por outro lado, como o círculo tem raio igual a um, o triângulo OBA nos dá

$$\overline{OA} = \cos(\alpha) \text{ e } \overline{AB} = \sin(\alpha).$$

Logo,

$$\overline{AC} = 1 + \overline{OA} = 1 + \cos(\alpha).$$

Como precisaremos que $\overline{AC} \neq 0$, suporemos, de agora em diante, que

$$(72) \quad \alpha \notin \{(2k+1)\pi \mid k \in \mathbb{Z}\}.$$

Substituindo as expressões para \overline{AC} e \overline{AB} em (71),

$$\tan(\beta) = \frac{\sin(\alpha)}{1 + \cos(\alpha)}.$$

Para simplificar a notação, escreveremos

$$s = \sin(\alpha), \quad c = \cos(\alpha) \text{ e } t = \tan(\beta);$$

de modo que a fórmula anterior toma a forma

$$(73) \quad t = \frac{s}{1 + c}.$$

Elevando ao quadrado,

$$t^2 = \frac{s^2}{(1+c)^2} = \frac{1-c^2}{(1+c)^2}.$$

Lembrando que $1-c^2 = (1-c)(1+c)$ e que estamos supondo que $c \neq -1$, podemos cancelar $1+c$ entre o numerador e o denominador, obtendo

$$t^2 = \frac{1-c}{1+c}.$$

Resolvendo esta última equação relativamente a c , encontramos

$$c = \frac{1-t^2}{1+t^2}.$$

Substituindo isto em (73),

$$s = t(1+c) = t \left(1 + \frac{1-t^2}{1+t^2}\right) = \frac{2t}{1+t^2}.$$

Retornando, finalmente, à notação usual,

$$\cos(\alpha) = \frac{1-t^2}{1+t^2} \text{ e } \sin(\alpha) = \frac{2t}{1+t^2}, \text{ em que } t = \tan\left(\frac{\alpha}{2}\right);$$

desde que α satisfaça a condição (72).

4. Identidades trigonométricas

Por ser escrita...

CHAPTER 7

Exponenciais e logaritmos

O que segue é apenas um esboço inicial deste capítulo, porque não vai dar tempo de escrever todos os detalhes antes do final deste período.

1. Potências

Sejam $n > 0$ um número inteiro e b um número real não nulo. A maneira mais ingênua de definir a *potência* b^n é como o produto

$$\underbrace{b \cdots b}_{n \text{ vezes}} .$$

Contudo, é preferível usar a recursão

$$(74) \quad b^n = \begin{cases} b & \text{se } n = 1 \\ b \cdot b^{n-1} & \text{se } n \geq 2, \end{cases}$$

para defini-la porque é mais adequada à demonstração de nosso primeiro resultado, que utilizará o princípio de indução finita.

TEOREMA 1.1. *Se m e n são inteiros positivos e b um número real positivo, então:*

$$b^m \cdot b^n = b^{m+n} \quad \text{e} \quad (b^m)^n = b^{nm}.$$

PROOF. Começaremos provando a primeira fórmula por indução em n . Se $n = 1$, então pela definição da recursão,

$$b^{m+1} = b^m \cdot b,$$

o que prova a base da indução. Suponhamos que

$$(75) \quad b^{m+k} = b^m \cdot b^k.$$

Mas, por (74),

$$b^{m+k+1} = b^{m+k} \cdot b.$$

Combinando isto com (75), obtemos

$$b^{m+k+1} = (b^m \cdot b^k) \cdot b = b^m \cdot (b^k \cdot b).$$

Apelando mais uma vez para a recursão, encontramos

$$b^{m+k+1} = b^m \cdot (b^k \cdot b) = b^m \cdot b^{k+1}.$$

Portanto, pelo princípio de indução finita, a fórmula $b^m \cdot b^n = b^{m+n}$ vale para qualquer valor de n .

Passando, agora, à segunda fórmula do enunciado do teorema, temos que

$$(b^n)^1 = b^n$$

por (74), o que prova a base da indução. Suponhamos, então, que

$$(76) \quad (b^m)^k = b^{km}.$$

Pela recursão,

$$(b^m)^{k+1} = (b^m)^k \cdot b^m.$$

Logo, por (76),

$$(b^m)^{k+1} = b^{mk} \cdot b^m.$$

Finalmente, utilizando a fórmula provada anteriormente,

$$(b^m)^{k+1} = b^{mk+m} = b^{m(k+1)}.$$

Portanto, pelo princípio de indução finita, $(b^m)^n = b^{nm}$ vale para todo inteiro positivo n . \square

 Note que na demonstração das fórmulas do teorema acima, começamos com um valor qualquer de m e aplicamos indução em n , de modo que ambas as fórmulas são válidas qualquer que sejam os valores dos inteiros positivos m e n .

As definições de potências que vimos até aqui aplicam-se apenas ao caso em que o expoente é um inteiro positivo. Nosso próximo passo consistirá em estender a definição ao caso em que o expoente é igual a zero. O ponto crucial é que b^0 deveria ser definido de modo que as fórmulas do teorema 1.1 continuassem válidas. Mas, para que isto aconteça é necessário que

$$b^{m+0} = b^m \cdot b^0;$$

onde

$$b^m = b^m \cdot b^0.$$

Como $b \neq 0$, podemos cancelar b^m dos dois membros desta equação, o que nos dá

$$1 = b^0.$$

Portanto, para que a primeira das duas fórmulas do teorema seja válida, b^0 precisa ser definido como 1. A verificação de que a segunda fórmula também vale sob essa definição fica por sua conta.

Podemos usar uma argumentação semelhante para tratar o caso em que o expoente é negativo. Tomando como ponto de partida o fato de que a primeira fórmula do teorema deveria continuar sendo válida, temos que se n for um inteiro positivo, então

$$b^n \cdot b^{-n} = b^{n-n} = b^0 = 1.$$

Lembrando que $b \neq 0$, podemos dividir os dois lados da equação por b^n , obtendo

$$b^{-n} = \frac{1}{b^n}.$$

Resumindo, se n é um inteiro positivo e $b \neq 0$ for um número real, então

$$b^0 = 1, \quad b^1 = b, \quad b^{n+1} = b^n \cdot b, \quad \text{e} \quad b^{-n} = \frac{1}{b^n}.$$

Finalmente, vale lembrar que se $b < 0$ a potência b^k será positiva, quando k for par, e negativa, quando k for ímpar, independentemente de k ser positivo ou negativo.

2. A função exponencial

O que fizemos na seção anterior nos permite definir potências sempre que o expoente for inteiro, mas isto não basta para definirmos a função exponencial, cujo domínio são os números reais. Para isso ainda vamos estender a definição de potência ao caso em que o expoente é um número real. Faremos isto em duas etapas, estendendo a definição primeiramente aos expoentes fracionários e utilizando-os para o caso em que os expoentes são números irracionais. Ao longo de toda esta seção suporemos sempre que a base das potências são números reais *positivos*.

O caso em que o expoente é uma fração sucumbe a um argumento semelhante ao usado na seção anterior. Mais uma vez, o ponto crucial é que desejamos que as fórmulas do teorema 1.1 continuem valendo quando os expoentes são frações. Levando isto em conta e supondo que m e n são inteiros positivos e que b é um número real positivo, a segunda fórmula do teorema nos dá

$$(b^{1/n})^n = b^{n \cdot 1/n} = b^1 = b.$$

Extraindo a raiz n -ésima dos dois lados da equação, obtemos

$$b^{1/n} = \pm \sqrt[n]{b}.$$

A convenção usual prescreve que escolhamos a raiz positiva, de modo que definimos

$$b^{1/n} = + \sqrt[n]{b}.$$

 Note que há uma enorme diferença entre as fórmulas que obtivemos na seção anterior e a fórmula acima. Enquanto as fórmulas para expoentes inteiros nos permitem calcular facilmente o valor da exponencial, usando apenas as quatro operações básicas, a fórmula para $b^{1/n}$ apenas reescreve este número em termos de raízes n -ésimas que, possivelmente, soam mais familiares. A verdade é que ambos os problemas são igualmente difíceis de resolver sem a ajuda de uma calculadora ou um computador e, provavelmente, você não sabe como nenhum dos dois efetua este cálculo.

Fazendo de conta que sabemos calcular potências com expoentes fracionários, vejamos como usá-las para calcular potências cujo expoente é um número real r qualquer. Para isso utilizaremos a expansão decimal de r , que escreveremos na forma

$$r = a_0, a_1 a_2 a_3 \dots$$

em que $0 \leq a_0, a_1, \dots \leq 9$ são algarismos decimais. Truncando esta expansão na n -ésima casa decimal obtemos a fração

$$f_n = a_0, a_1 a_2 \dots a_n = \frac{c_n}{10^n},$$

em que

$$c_n = a_0 \cdot 10^n + \dots + a_1 \cdot 10 + a_n \in \mathbb{Z}.$$

Portanto, a sequência de frações

$$f_0, f_1, f_2, \dots$$

produz números racionais que se aproximam progressivamente de r . A ideia é, então, que os números

$$(77) \quad b^{f_0}, b^{f_1}, b^{f_2}, \dots,$$

produzem, por sua vez, aproximações cada vez melhores de um número, que definimos como sendo b^r . Infelizmente não temos, neste curso, uma especificação suficientemente precisa dos números reais que nos permita provar isto, mas podemos ilustrar o resultado com um exemplo. Digamos, por exemplo, que tivéssemos uma calculadora capaz de extrair raízes décimas, além de efetuar as operações aritméticas básicas. Como,

$$b^{1/10^n} = (b^{1/10^{n-1}})^{1/10}$$

isto basta para calcularmos aproximações de b^r usando o método descrito acima. Por exemplo,

$$\pi \approx 3.141592 \dots$$

de modo que a sequência nos dá aproximações sucessivas de 2^π .

$$\begin{array}{ccccccc} 2^3 & 2^{3,14} & 2^{3,141} & 2^{3,1415} & 2^{3,14159} & 2^{3,141592} & \dots \\ \| & \| & \| & \| & \| & \| & \\ 8 & 8.815240 & 8.821353 & 8.824411 & 8.824961 & 8.824973 & \dots \end{array}$$

Uma vez que saibamos calcular potências com quaisquer expoentes reais, podemos definir a *exponencial de base b* como sendo a função

$$\exp_b : \mathbb{R} \longrightarrow \mathbb{R}$$

definida pela regra

$$\exp_b(x) = b^x.$$

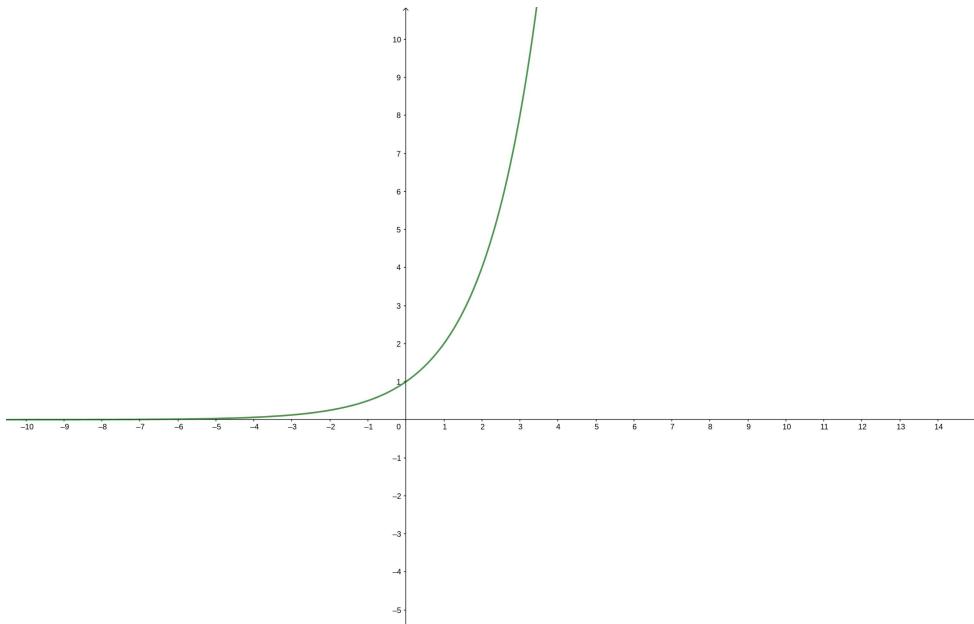


FIGURE 1. Gráfico da exponencial de base 2

Como estamos supondo que $b > 0$, a imagem de \exp_b está contida no intervalo $[0, +\infty)$. Uma análise mais refinada do que podemos fazer aqui, mostra que a imagem de fato coincide com o conjunto dos números reais positivos. O gráfico da exponencial de base 2 é ilustrado na figura 1.

3. O logaritmo

Como na seção anterior, suporemos, de agora em diante, que as bases de todas as nossas potências são número reais *positivos*. Seja, pois, $b > 0$ um número real. A função *logaritmo na base b* é a inversa da função exponencial na base b . Assim,

$$\log_b(\exp_b(x)) = \log_b(b^x) = x$$

qualquer que seja o número real x e

$$\exp_b(\log_b(z)) = b^{\log_b(z)} = z$$

qualquer que seja o número real *positivo* z . Como

$$\text{im}(\exp_b) = [0, \infty),$$

temos que

$$\log_b : [0, \infty) \longrightarrow \mathbb{R}.$$

Na verdade, como \mathbb{R} é o domínio de \exp_b , então

$$\text{im}(\log_b) = \mathbb{R}.$$

O gráfico da função logaritmo de base 2 é ilustrado na figura 2. Como nos vários exemplos de função inversa com os quais nos deparamos até aqui, o gráfico do logaritmo foi construído simetrizando o gráfico de \exp_b relativamente à reta $y = x$, que aparecem tracejados na figura.

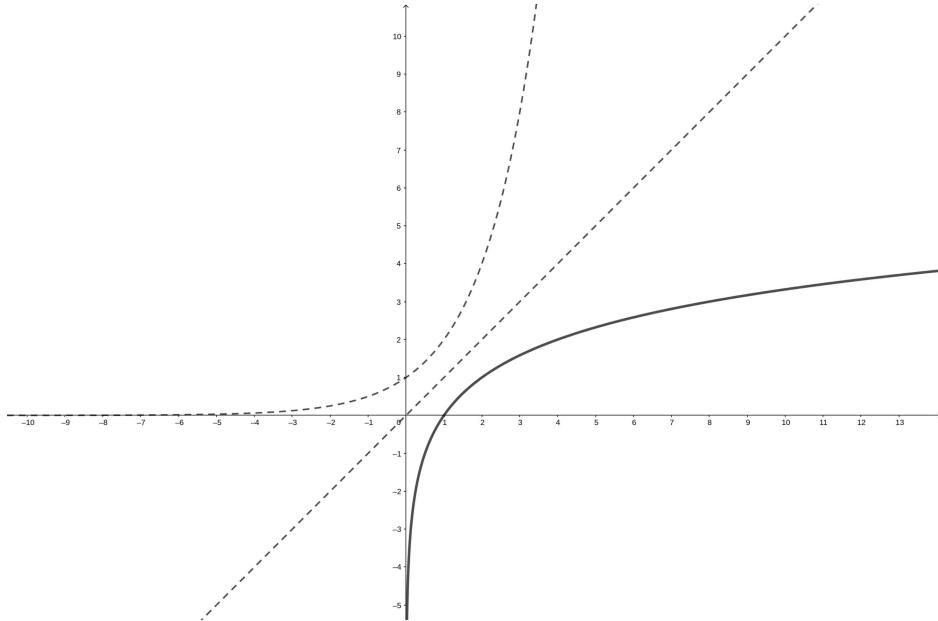


FIGURE 2. Gráfico do logaritmo de base 2

As propriedades do logaritmo são obtidas, a partir das propriedades da exponencial, usando o fato de que \exp_b e \log_b são inversas uma da outra. Por exemplo, aplicando \log_b a

$$\exp_b(x_1 + x_2) = \exp_b(x_1) \cdot \exp_b(x_2),$$

obtemos

$$(78) \quad x_1 + x_2 = \log_b(\exp_b(x_1 + x_2)) = \log_b(\exp_b(x_1) \cdot \exp_b(x_2)).$$

Porém como,

$$y_1 = \exp_b(x_1) \text{ e } y_2 = \exp_b(x_2)$$

correspondem, respectivamente, a

$$x_1 = \log_b(y_1) \text{ e } x_2 = \log_b(y_2)$$

podemos reescrever (78) como

$$\log_b(y_1) + \log_b(y_2) = \log_b(y_1 \cdot y_2);$$

isto é, o logaritmo de um produto é a soma dos logaritmos dos termos deste produto. Para obter a outra propriedade de que precisamos, digamos que

$$(79) \quad a = b^z,$$

de modo que, pela definição de logaritmo,

$$z = \log_b(a).$$

Substituindo (79) em a^x , obtemos

$$a^x = b^{xz} = \exp_b(xz).$$

Aplicando logaritmo na base b aos dois lados desta igualdade, encontramos

$$\log_b(a^x) = xz = x \log_b(a),$$

Resumindo, mostramos que

$$(80) \quad \log_b(y_1y_2) = \log_b(y_1) + \log_b(y_2) \text{ e } \log_b(a^x) = x \log_b(a).$$

Até a popularização das calculadoras eletrônicas na década de 1970, os logaritmos eram utilizados para simplificar cálculos de potências e produtos, já que as fórmulas (80) nos permitem transformar o cálculo de uma produto em uma soma e o de uma potência em uma multiplicação. Por exemplo, para calcular

$$2^{1,98}$$

aplicaríamos logaritmo na base 10, obtendo

$$\log_{10}(2^{1,98}) = \log_{10}(2^{1,98}) = 1,98 \cdot \log_{10}(2).$$

Consultando uma tábua de logaritmos, como a ilustrada na figura 3 da página 142, vemos que

$$\log_{10}(2) = 0,3010300$$

onde

$$\log_{10}(2^{1,98}) = \log_{10}(2^{1,98}) = 1,98 \cdot 0,3010300 = 0,5960394.$$

Apelando, agora, para a seção da mesma tábua de logaritmos ilustrada na figura ??, procuramos, na quarta coluna a partir da esquerda, os quatro primeiros algarismos do logaritmo cujo número desejamos identificar. Em nosso exemplo, os quatro primeiros algarismos são 5960. Procuramos, então, nas colunas imediatamente à direita de deste número, o número cujos algarismos seguintes mais se aproximam de 394 que, nesta tabela, são 3881. Uma aproximação do número cujo logaritmo conhecemos aparece, então, na terceira coluna a partir da direita da mesma tabela. Neste caso o número que encontramos é 3948. Finalmente, como

$$\log_{10}(1) = 0 < 0,5960394 < 1 = \log_{10}(10)$$

temos que o número desejado está entre 0 e 10, de modo que a vírgula deve ser posicionada entre o primeiro e o segundo algarismos de 3948, o que nos dá que

$$2^{1,98} \approx 3,948.$$

Com um pouco mais de trabalho é possível extrair desta tábua de logaritmos uma aproximação melhor para este número, mas nosso propósito aqui é, apenas, dar-lhe uma ideia de como o cálculo seria feito, por isso não entraremos em mais detalhes.

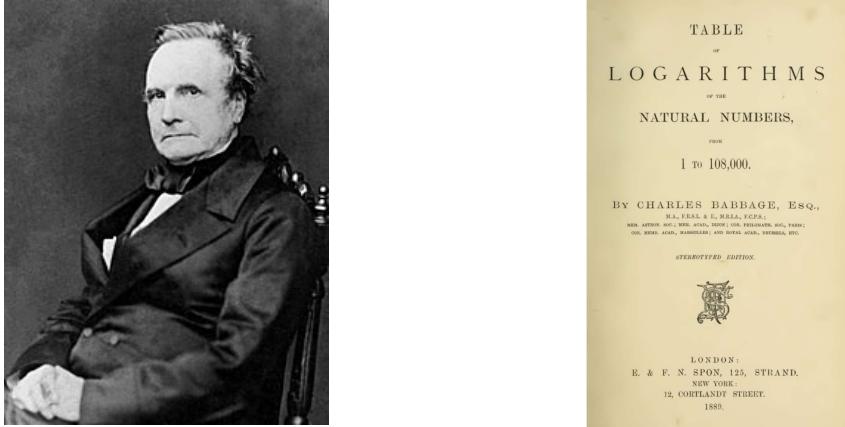


FIGURE 5. Charles Babbage e sua tábua de logaritmos

Um problema com que frequentemente nos deparamos, e que ainda não abordamos, é o de converter logaritmos em uma base para logaritmos em outra base. Para isto, escrevemos um dado número y , simultaneamente, como potência de números reais positivos a e b . Digamos que

$$y = a^x \quad \text{e} \quad y = b^z.$$

Pela definição de logaritmo, temos que

$$(81) \quad x = \log_a(y) \quad \text{e} \quad z = \log_b(y).$$

Aplicando logaritmo na base b a ambos os lados de $a^x = b^z$ e usando a segunda propriedade em (80), encontramos

$$x \log_b(a) = z \log_b(b) = 1.$$

Combinando isto com (81), chegamos a

$$(82) \quad \log_a(y) = \frac{\log_b(y)}{\log_b(a)}$$

que é conhecida como *fórmula de mudança de base* para logaritmos. Por exemplo, se $1 \leq n, m \leq N$ são números inteiros, então a quantidade de divisões que precisam ser executadas para calcular o máximo divisor comum entre m e n é, no máximo, igual a

$$(83) \quad [\log_\phi(\sqrt{5}N)] - 2,$$

em que $\lceil x \rceil$ denota o menor inteiro maior que x e

$$\phi = \frac{1 + \sqrt{5}}{2}$$

é a razão áurea, que já encontramos na seção 7 do capítulo 3. Para a demonstração da fórmula (83) consulte [8, Corollary L, p. 343]. Contudo, mesmo um computador, não tem uma função que calcula diretamente logaritmos cuja base é ϕ . Portanto, se quisermos saber quantas divisões seriam necessárias para calcular o máximo divisor comum de números

com, por exemplo, menos de 200 algarismos, teremos que apelar para mudança de base. Mas, usando (80),

$$\log_\phi(\sqrt{5} \cdot 10^{200}) = \frac{1}{2} \log_\phi(5) + 200 \log_\phi(10),$$

precisamos, apenas, conhecer os logaritmos de 5 e 10 na base ϕ . Contudo, pela fórmula de mudança de base (82),

$$\log_\phi(5) = \frac{\log_{10}(5)}{\log_{10}(\phi)} \quad \text{e} \quad \log_\phi(10) = \frac{\log_{10}(10)}{\log_{10}(\phi)} = \frac{1}{\log_{10}(\phi)}.$$

Como

$$\log_{10}(\sqrt{5}) \approx 0.6989700043 \quad \text{e} \quad \log_{10}(\phi) \approx 0.2089876402,$$

obtemos

$$\log_\phi(5) \approx 3.3445518763 \quad \text{e} \quad \log_\phi(10) \approx 4.7849719667$$

onde

$$\log_\phi(\sqrt{5} \cdot 10^{200}) \approx 958.6666692945.$$

Portanto, para achar o máximo divisor comum entre números menores ou iguais a 10^{200} são necessárias, no máximo,

$$\lceil 958.6666692945 \rceil - 2 = 959 - 2 = 957$$

divisões.

4. Aplicações

As funções exponencial e logaritmo têm uma enorme gama de aplicações, porque vários fenômenos naturais são regidos por *leis exponenciais* que são aquelas que podem ser expressas na forma

$$f(t) = ab^{ct},$$

em que $a \neq 0$, $b > 0$ e $c \neq 0$ são números reais e t representa, tipicamente, o tempo. Exemplos de leis exponenciais incluem o crescimento de bactérias em uma cultura, a reação nuclear em cadeia, o esquema de pirâmide em economia e o espalhamento de memes na rede. Em ciência da computação, o custo de muitos algoritmos, medido pela quantidade de passos que executam, é exponencial. Este é o caso, por exemplo, dos algoritmos para fatoração de números inteiros.

Antes de considerarmos alguns exemplos concreto, precisamos fazer algumas considerações sobre as bases mais utilizadas nas aplicações da exponencial e do logaritmo. No passado a base 10 era, sem dúvida, a mais utilizada, porque as tábuas de logaritmo usadas nos cálculos de engenharia listavam, como vimos na seção 3, os logaritmos dos números nesta base. A escolha da base 10 para estas tábuas era mera consequência de representarmos os números em notação posicional em base 10. Depois que as tábuas de logaritmos e

Isótopo	Meia vida em anos
Urânio 238	$4,51 \cdot 10^9$
Plutônio 239	$2,44 \cdot 10^4$
Carbono 14	5568
Rádio 226	1600

TABLE 1. Meia vida de alguns isótopos

régulas de cálculo foram suplantadas pelas calculadoras eletrônicas, as bases mais comuns passaram a ser

$$e \approx 2,7182818284590452354$$

em ciências e engenharia, e 2, em computação. O problema com a base e é que o conhecimento que temos neste curso é insuficiente para justificarmos porque alguém haveria de usar como base um número tão estranho e que, além de tudo, é irracional. A verdade é que este número surge, de maneira natural, na solução de muitos problemas de engenharia e física.

Analisaremos, para começar, o caso do decaimento radioativo, em que a quantidade $N(t)$, de núcleos de um dado isótopo é dada, depois de um tempo t , por

$$(84) \quad N(t) = N(0) \exp(-\lambda t),$$

em que $N(0)$ é a quantidade de núcleos deste isótopo no instante inicial e λ é a constante de decaimento radioativo. Para calcularmos λ , precisamos saber qual é a *meia vida* do isótopo em questão, que é definida como o tempo necessário para que metade dos átomos deste isótopo sofram decaimento radioativo. A tabela 1 lista a meia vida de alguns isótopos bem conhecidos.

Por definição, se $t_{1/2}$ é a meia vida de um dado isótopo, então segue-se da equação (84) que

$$\frac{1}{2} = \frac{N(t_{1/2})}{N(0)} = \exp(-\lambda t_{1/2}).$$

Tomando logaritmos naturais dos dois lados,

$$\ln\left(\frac{1}{2}\right) = -\lambda t_{1/2};$$

onde

$$\lambda = \frac{\ln(2)}{t_{1/2}}.$$

Por exemplo, a constante de decaimento radioativo do carbono 14, que utilizaremos no próximo exemplo, é igual a

$$\frac{\ln(2)}{5568} \approx 0,00012448.$$

Uma aplicação interessante da lei de decomposição radioativa diz respeito ao cálculo da idade de uma amostra de madeira em um sítio arqueológico, baseada no método de datação por carbono 14. Ao contrário do carbono 12, que é estável, o carbono 14 é radioativo, sendo produzido quando a atmosfera é atingida por prótons e núcleos atômicos que se movem através do espaço com velocidades próximas à da luz. O carbono 14 da atmosfera é absorvido pelas plantas durante o processo de fotossíntese. Como a absorção cessa com a morte da planta, a comparação entre a quantidade de carbono 14 em uma amostra e a quantidade do mesmo isótopo na atmosfera nos permite estimar a idade da amostra de madeira. Para que o procedimento seja viável é necessário conhecermos que fração do carbono presente na atmosfera corresponde ao isótopo ^{14}C . Trabalharemos sob a hipótese de que esta fração permanece constante ao longo do tempo, embora hoje saibamos, que isto não é verdade, de modo que idades medidas pelo método de carbono-14 precisam ser devidamente calibradas usando outros métodos de datação; mais detalhes podem ser encontrados em

Para os propósitos de nosso exemplo, digamos que uma amostra de madeira em um sítio arqueológico foi testada em um laboratório e que a razão entre a quantidade de ^{14}C e de ^{12}C foi determinada como sendo igual a 25% da quantidade deste isótopo presente na atmosfera atual. Em outras palavras,

$$\frac{25}{100} = \frac{N(t)}{N(0)} = \exp(-0,00012448 \cdot t),$$

em que t é a idade da amostra de madeira. Aplicando logaritmo natural aos dois lados da igualdade, obtemos

$$\ln\left(\frac{1}{4}\right) = -0,00012448 \cdot t;$$

de modo que

$$t = \frac{-\ln(4)}{-0,00012448} \approx 11136,68\text{anos}.$$

Outra aplicação interessante dos logaritmos, está relacionada à medida da magnitude de um terremoto. A intensidade de um tremor, que denotaremos por I , corresponde à amplitude da curva registrada por um sismógrafo localizado a 100 Km do epicentro do terremoto. Por definição, a intensidade S de um “terremoto padrão” é igual à amplitude de 1 micron = 10^{-4} cm. A *magnitude* de um terremoto, medida na escala Richter, é, então, dada por

$$M = \log_{10}\left(\frac{I}{S}\right).$$

Elevando 10 aos dois lados da equação, obtemos

$$10^M = \frac{I}{S},$$

Local	Data	Magnitude
San Francisco	1906	8,3
Loma Prieta	1989	6,9
Bay Area	2014	6,0

TABLE 2. Magnitudes na escala Richter de alguns terremotos

de modo que um terremoto de magnitude M tem intensidade

$$I = S \cdot 10^M.$$

Para nos ajudar a esclarecer o comportamento da escala Richter, compararemos as intensidades de dois terremotos cujas magnitudes diferem de apenas uma unidade. Se I_0 é a intensidade de um terremoto de magnitude M_0 , então a intensidade do terremoto de magnitude $M_0 + 1$ será, pela fórmula acima,

$$I = S \cdot e^{M_0+1} = S \cdot 10^{M_0} \cdot 10 = 10 \cdot I_0.$$

Portanto, se a magnitude de um terremoto é um ponto a mais na escala Richter que o de outro terremoto, sua amplitude será dez vezes maior.

A tabela 2 contém a magnitude, na escala Richter, de alguns terremotos ocorridos na Califórnia. O terremoto de Loma Prieta levou ao colapso da ponte que liga San Francisco a Oakland, do outro lado da baía, e teve intensidade igual a

$$I_{LM} = S \cdot 10^{6,9}$$

ao passo que o terremoto de 1906 teve intensidade igual a

$$I_{SF} = S \cdot 10^{8,3}.$$

Como

$$\frac{I_{SF}}{I_{LM}} = \frac{10^{8,3}}{10^{6,9}} \approx 25,12$$

o tremor de 1906 foi vinte e cinco vezes mais intenso que o de Loma Prieta. Não é de surpreender, portanto, que este terremoto, e o incêndio que o sucedeu, tenham causado a destruição de mais de 80% da área de San Francisco. Já o tremor de 2014, que causou uma destruição bem menor, teve intensidade

$$\frac{I_{BA}}{I_{LM}} = \frac{10^{6,0}}{10^{6,9}} \approx 0.126$$

vezes menor que o de Loma Prieta.

Para encerrar, vejamos um exemplo na interseção da lógica com a ciência da computação. O problema que vamos abordar é conhecido como *eliminação de quantificadores* e pode ser formulado da seguinte maneira:

dada uma fórmula construída combinando com \wedge , \vee e \rightarrow várias igualdades e desigualdades de polinômios com coeficientes inteiros, na qual algumas variáveis aparecem quantificadas, produza uma fórmula equivalente na qual *nenhuma* variável é quantificada.

Um exemplo elementar é dado pela fórmula

$$\exists x \in \mathbb{R}, ((a \neq 0) \wedge (ax^2 + bx + c > 0));$$

para a qual temos a versão sem quantificadores

$$(a \neq 0) \wedge (b^2 + 4ac > 0),$$

respondendo, assim, à questão de quando um polinômio quadrático com coeficientes inteiros tem uma raiz real. Outro problema que pode ser abordado de maneira semelhante pergunta *quais são as condições necessárias e suficientes para que uma elipse cujos eixos são paralelos aos eixos coordenados esteja contida em um círculo de raio um*. Podemos descrever este problema usando a fórmula

$$(a > 0) \wedge (b > 0) \wedge \left(\forall x \in \mathbb{R}, \forall y \in \mathbb{R} \left(\frac{(x - c)^2}{a^2} + \frac{(y - d)^2}{b^2} = 1 \implies x^2 + y^2 = 1 \right) \right).$$

Em 1988 Daniel Lazard mostrou que toda solução para este problema tem que conter um certo polinômio de grau 12 em a , b , c e d que tem 104 termos; para mais detalhes veja [2]. O tamanho da solução obtida para este segundo problema sugere que qualquer algoritmo usado para resolvê-lo terá um custo bastante alto. De fato, J. H. Davenport e J. Heintz provaram em 1988 que qualquer algoritmo que resolva o problema de eliminação de quantificadores é duplamente exponencial no número de variáveis dos polinômios; veja [4]. Por exemplo, se os polinômios têm n variáveis, o custo vai ser da ordem de 2^{2^n} , o que é uma péssima notícia se precisamos resolver um problema através deste método.

Bibliography

1. Nicolas Bourbaki, *Elementos de historia de las matemáticas*, Alianza Editorial, Madrid, 1976.
2. B. F. Caviness and J. R. Johnson (eds.), *Quantifier elimination and cylindrical algebraic decomposition*, Texts and Monographs in Symbolic Computation, Springer-Verlag, Vienna, 1998.
3. S. C. Coutinho, *Números inteiros e criptografia RSA*, segunda ed., Coleção Matemática e Aplicações, Instituto Nacional de Matemática Pura e Aplicada (IMPA), Rio de Janeiro, 2014.
4. James H. Davenport and Joos Heintz, *Real quantifier elimination is doubly exponential*, J. Symbolic Comput. **5** (1988), no. 1-2, 29–35.
5. Jean Dieudonné, *Abrégé d'histoire des mathématiques 1700-1900. Nouv. version. (A short history of mathematics 1700-1900)*, Hermann, Paris, 1986.
6. David Fowler and Eleanor Robson, *Square root approximations in old Babylonian mathematics: YBC 7289 in context*, Historia Math. **25** (1998), no. 4, 366–378.
7. Morris Kline, *Mathematical thought from ancient to modern times. Vol. 2*, second ed., The Clarendon Press, Oxford University Press, New York, 1990. MR 1058202
8. Donald E. Knuth, *The art of computer programming. Vol. 2*, second ed., Addison-Wesley Series in Computer Science and Information Processing, Addison-Wesley Publishing Co., Reading, Mass., 1981, Seminumerical algorithms. MR 633878