Introdução ao Pensamento Dedutivo

S. C. Coutinho

And oh, how our glory may fade,
Fade...
Well, at least we've learned,
Some things along the way.

Paolo Nutini (Tricks of the trade)

Prefácio

Através dos séculos, e das categorias profissionais, a matemática tem sido encarada de muitas maneiras diferentes. Para muitos ela não parece passar de uma coleção de fórmulas, possivelmente úteis, a serem memorizadas e diretamente aplicadas, quando o "prompt" correto é encontrado. Deste ponto de vista, a matemática não passa de um enorme arquipélago, formado por ilhas com culturas totalmente diferentes, que não se comunicam entre si.

Neste livro adotamos um ponto de vista diferente. Sob este outro enfoque, a matemática é um sistema de ideias, fortemente conectadas entre si, baseadas em um pequeno conjunto de fatos tomados como fundamentais. Não se trata de um ponto de vista novo, já que foi introduzido na Grécia Antiga, em algum momento antes de 300 a.C. por muitos séculos era a maneira padrão de apresentar a matemática.

A beleza deste enfoque, hoje em dia tão impopular, é que as fórmulas cedem lugar as ideias. Sua impopularidade, talvez, tenha origem exatamente no fato de que fórmulas e cálculos cedem lugar as ideias, o que requer mais esforço tanto da parte de estudantes, quando de seus professores. A grande vantagem está em que é muito mais fácil absorver um conjunto de ideais interligadas do que de fórmulas isoladas.

O livro originou-se de um curso, de mesmo nome, oferecido aos estudantes de primeiro período do curso de Ciência da Computação da Universidade Federal do Rio de Janeiro. Por sua vez, o curso foi criado por causa da crescente dificuldade destes alunos em acompanhar disciplinas obrigatórias como cálculo, álgebra linear e matemática discreta. Contudo, é importante frisar que não se trata de um curso tradicional do que se convencionou chamar de pré-cálculo. Afinal, como o nome indica, trata-se de uma introdução ao pensamento dedutivo e não apenas de uma revisão aprofundada de tópicos como números reais, desigualdades e funções. Ainda que este três tópicos sejam abordados, eles o são dentro de uma estrutura dedutiva. Por exemplo nosso estudo de números reais parte das propriedades básicas destes, a partir das quais deduzimos todas as demais de que precisamos. Dito isto, é necessário esclarecer que, tampouco se trata de um curso de introdução a análise. Ainda que os cortes de Dedekind sejam mencionados, isto é feito apenas para indicar claramente o que ficou de fora em nosso estudo dos números reais.

Deixando de lado a introdução, em que o método dedutivo é apresentado em um contexto histórico, o livro consta de seis capítulos. O primeiro introduz lógica proposicional e conjuntos, os dois seguintes tratam de quantificadores, números reais e funções e os três últimos de geometria plana, trigonometria e gráficos de funções. O tratamento é tão rigoroso quanto um curso elementar o permite. Por exemplo, nossa abordagem de geometria é baseada em seis axiomas, dois dos quais tratam da medição de distâncias e ângulos. Portanto, usamos números reais e suas propriedades em nosso estudo de geometria, sem o que nosso tratamento de trigonometria seria inviável. Além disso, os tópicos estão limitados ao que é viável incorporar em uma disciplina semestral. Com isso, nosso estudo de geometria não vai além dos triângulos, e apresentamos quantificadores de maneira extremamente ingênua.

O livro foi usado, sob a forma de apostila distribuída aos alunos, durante vários anos no curso de Ciência da Computação da UFRJ e beneficiou-se das observações, correções e sugestões das dezenas de alunos que fizeram estes cursos e dos professores Juliana Vianna Valério e João Antônio Récio da Paixão que utilizaram o mesmo material ao oferecer a disciplina e, sobretudo ao João Carlos Pereira da Silva, que além de usar o material, leu todo o texto mais de uma vez. Naturalmente todos os erros que, porventura, não tenham sido eliminados são de minha inteira responsabilidade. Também devo agradecimentos especiais à turmas que ingressaram nos primeiros semestres de 2022 e 2023 pelas críticas e sugestões que me ajudaram a melhorar o texto e, sobretudo, a sistematizar, de maneira mais adequada, o capítulo de geometria.

Rio de Janeiro, 10 de novembro de 2026

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, João Carlos Pereira da Silva, 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, Daniel Nocito Falcão Lopes, Júlio Burlamaqui dos Santos, Henrique Lima Cardoso, Carlos Eduardo de Schuller Banjar, Afonso Lustosa Pires Junior, Davi de Souza Gomes Pereira, Raisa Christina Nascimento Gonçalves, Arthur Sobrinho Ferreira da Rocha, Renato Sousa do Nascimento, Gustavo Henrique Lemos de Lima, Marcos Henrique Junqueira Muniz Barbi Silva, Raphael Bandeira de Brito, Julia Vilela da Silva Brito, Pedro Henrique Honório Saito, Vitória Frazão Batista, Rafael Passos Orioli Guimarães, João Vitor Amaral de Oliveira, Beatriz Lacerda Pinto Portella, Lucas Asafe de Souza Carvalho, Miguel Santos de Araujo do Nascimento, Gustavo de Mendonça Freire, Nicole Cardozo, Gabriel Martins de Freire, Giovanna Magalhães Lavouras, Matheus Ávila Abreu de Lima, Erick Cassiano da Silva Passos, Mateus Anastacio de Jesus, João Victor Lopez Pereira, Leonardo Lima Santos, Pedro Kury Kitagawa, Francisco Theodoro Arantes Florêncio, Marcelo Blak, Paulo Roberto Ferreira Godoy Moreira, Yasmim Mirella Paiva de Lima, Kaique Fabrício Eufrásio, Vitor Amaral Bedin, João Victor Borges, Pedro Salazar Pessoa Machado, Nívea Martins, Moises Auad Werneck, Gabriel Alves Trajano, Guilherme Vasconcellos Sobreira de Carvalho, Gabriel Duarte Soares, Igor Coelho de Carvalho, Luiz Vitor Vieira de Mattos, Alexandre Ribeiro de Souza, Luiz Renato Santana Salomão, Gabriel Felix Hilleshein, Igor Peres Barboza de Souza, Gabriel Sarmento Moraes Lomar, Guilherme Aleixo Name, Açucena Santos Martins da Silva, Gabriel Nobre Araújo, Maria de Oliveira Vilas Boas,

Agradecimentos especiais à turma de 2022.1 que ajudou a sistematizar o axioma referente à medida de ângulos.

Contents

CAPÍTULO 1

Introdução

Começaremos com um pequeno apanhado de alguns aspectos da história da matemática que nos ajudarão a esclarecer a importância do conteúdo deste livro no entendimento de muitos tópicos centrais para a matemática estudada por cientistas da computação, matemáticos, físicos e engenheiros.

1. Áreas

Toda criança vai se apropriando, com o passar dos anos, de inúmeras "partículas de informação" que ouve dos adultos, uma das quais é a explicação de que *geometria* é 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 o filósofo neoplatônico Proclo (410-485 d.C.) começa seu famoso *Sumário* de história da geometria,

> 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]; (Thomas, 1991, p. 145).

A tradição grega de que a geometria seria originária do Egito remonta, pelo menos, à *História* de Heródoto (484–425 a.C.). Segundo ele, 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 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 este livro 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*. Séculos mais tarde, 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 da figura 1, 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*, o outro de 2 *khets* e meio por um quarto de *khet*. 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

$$6 + 6 \cdot \frac{1}{4} + 10 \cdot \frac{1}{16} = \frac{65}{8}$$

ou, como os egípcios provavelmente fariam a conta,

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

Até aqui, nossa excursão histórica serviu, apenas, para referendar o fato de que *a área de um retângulo é o produto dos comprimentos dos seus lados*.

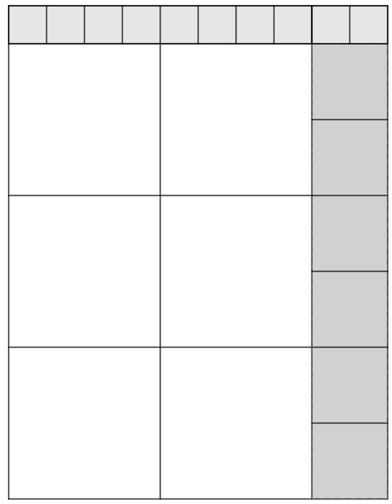


FIGURA 1. Área do retângulo de lados 2 *khets* e meio por 3 *khets* e um quarto.

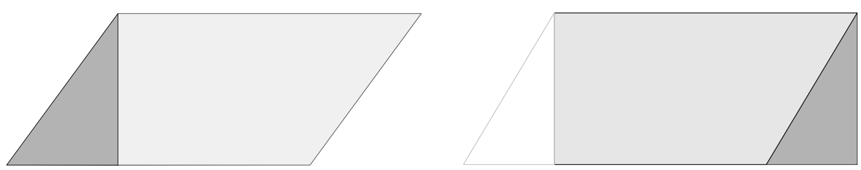


FIGURA 2. Transformando o paralelogramo em um retângulo.

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 2, 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 de sua base por sua altura*.

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 3, na qual, ao triângulo da esquerda, que foi dado, juntamos uma cópia refletida relativamente ao lado maior, para formar um paralelogramo. Portanto, a área do triângulo de baixo 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 de sua base por sua altura*.

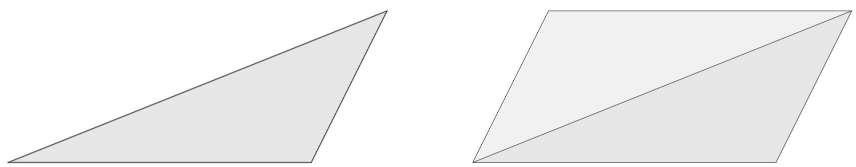


FIGURA 3. Juntando dois triângulos iguais para formar um paralelogramo.

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 todas elas 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

$$\left(2r - \frac{2r}{9}\right)^2 = \frac{256}{81}r^2. \tag{1}$$

Como observamos anteriormente isto corresponde a tomar

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

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 4. Na próxima seção veremos como passar desta figura à fórmula (1).

2. A geometria chega à Grécia

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

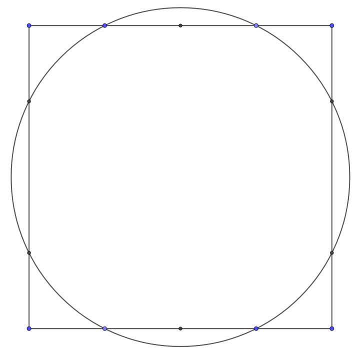


FIGURA 4. Aproximando a área do círculo no Antigo Egito

> 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; (Thomas, 1991, p. 147).

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; (Laertius, 2006, p. 29)

De nosso ponto de vista, o que Tales fez foi usar a semelhança de triângulos, como faremos na seção 5 do capítulo 5. 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.

Voltando ao sumário de Proclo (e deixando de lado "Ameristos, irmão do poeta Estesicorus", mencionado logo adiante na mesma citação e sobre o qual nada sabemos), Pitágoras é o personagem seguinte a ser mencionado. 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; (Thomas, 1991, p. 149).

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".

Em meio a tantas dúvidas, a tabuleta de argila YBC 7289, 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. 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 Fowler, Robson (1998), 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 5.

Apesar de bem-conhecida esta figura não nos ajuda a entender como alguém teria percebido a relação entre os três quadrados. Porém, nada nos impede de especular o que poderia ter acontecido. Para isso, lembre-se que os pisos das casas gregas e romanas eram frequentemente recobertos por mosaicos. 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 mosaicos da casa, ilustrados na figura 6.

Talvez você já tenha tido a experiência de olhar longamente para uma figura, até que, inesperadamente, 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 ocorrido, redesenhamos os mosaicos na figura 7, mantendo apenas aquilo que nosso matemático teria visto em seu momento de epifania.

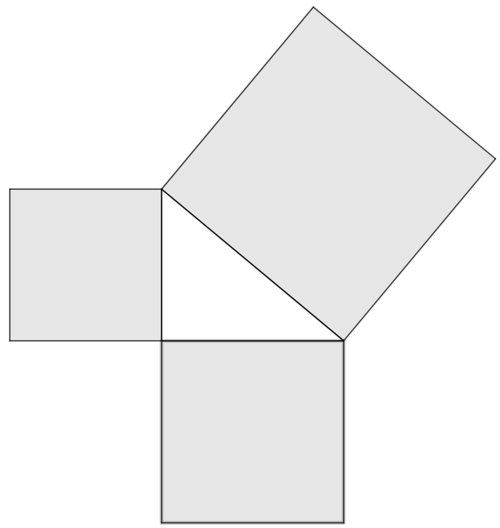


FIGURA 5. O teorema de Pitágoras

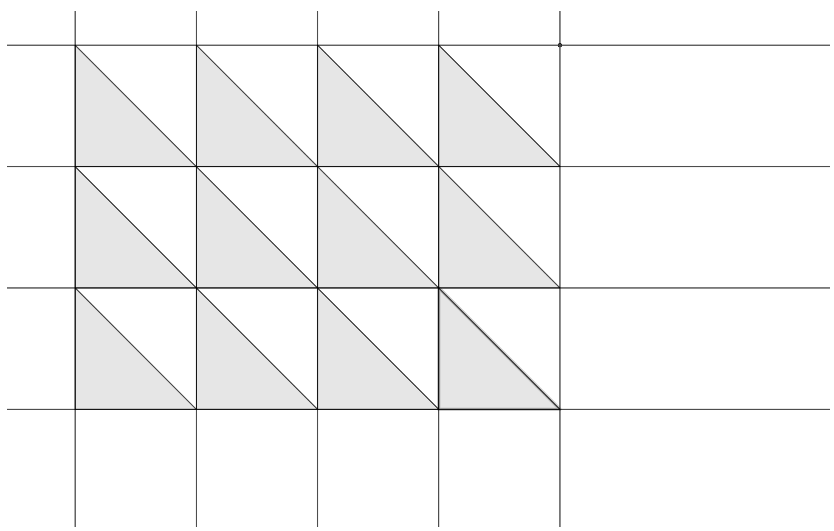


FIGURA 6. O ladrilhamento na casa do banquete

O ponto crucial é que cada um dos quadrados 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. Obtivemos, assim, um caso muito especial da figura 5, porque os catetos do triângulo retângulo que estamos considerando são iguais. A 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

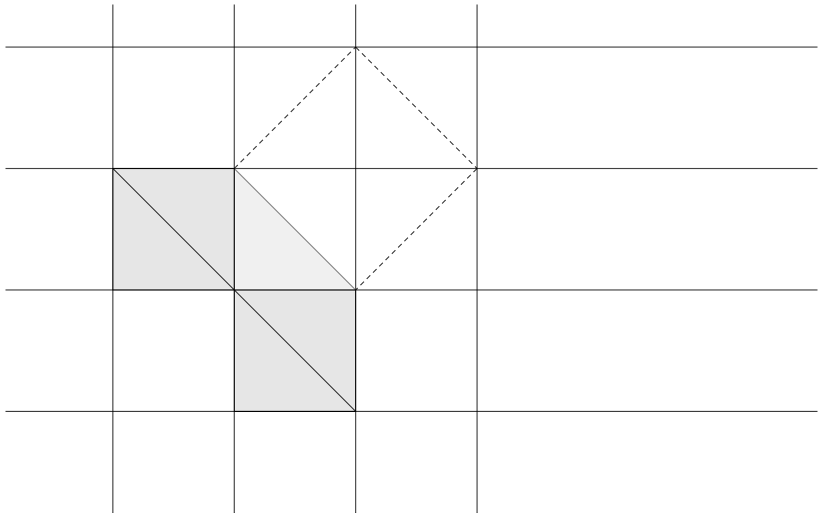


FIGURA 7. O que o matemático viu

antes de Pitágoras. Uma vez que um caso especial tenha sido descoberto é difícil para alguém com propensões matemáticas não se perguntar: vale para outros triângulos? 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 8 construímos um quadrado de lado igual a 8 e dividimos cada um de seus lados em quatro partes iguais. 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 teriam encontrado para o comprimento foi, provavelmente, 9. Mas, tomando 9, em vez de $4\sqrt{5}$ como valor do diâmetro, obtemos

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

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 nessas mudanças.

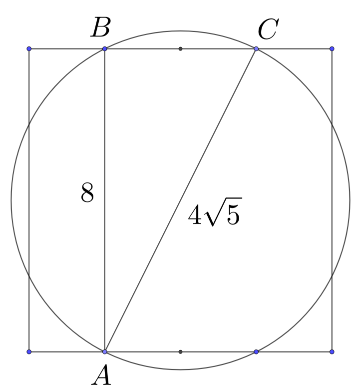


FIGURA 8. Usando o teorema de Pitágoras para aproximar π

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, levando ao aperfeiçoamento da arte da argumentação e à habilidade em justificar, cuidadosamente, a posição pessoal dos participantes.

Estes mesmos fatores influenciaram o desenvolvimento da filosofia, que também fez progressos expressivos nesta mesma época. 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 descreve a si próprio 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; (Aristotle, 1960, p. 37).

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, (Aristotle, 1960, p. 57),

ele observa que

> nem todo conhecimento é obtido através de demonstrações; pois o conhecimento imediato não é demonstrável; (Aristotle, 1960, p. 37)

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 em seus argumentos matemáticos.

À medida que a geometria grega foi se tornando mais teórica e sofisticada, e a quantidade de resultados conhecidos foi aumentando, os matemáticos sentiram a necessidade de sistematizar seu conhecimento. Isto foi feito em tratados, conhecidos como *elementos*, porque estabeleciam os fundamentos da geometria então conhecida. 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, segundo o qual

> no caso das proposições geométricas, chamamos de elementos aquelas cujas demonstrações estão presentes nas demonstrações de todas, ou da maioria, das demais. (Aristotle, 1989, p.117)

Em seus *Comentários*, Proclo dá uma explicação mais pitoresca do mesmo termo:

> assim como chamamos de elementos aos sons, que são os primeiros princípios, mais simples e indivisiveis, da língua, e todas as palavras e discurso a eles estão sujeitas; da mesma maneira, na geometria como um todo, há teoremas que estão em sua base e são a razão dos que vêm depois, alcançando os demais e provendo as demonstrações de suas propriedades, a eles damos o nome de elementos; (Diadochi, 1873, p. 72) (Proclus, 1992, p. 59).

Hipócrates não é o único matemático que Proclo menciona como tendo escrito elementos. Entre os outros matemáticos que ele cita neste contexto estão Leon, estudante de Neoclides, e Teúdios da Magnésia, sobre os quais pouco sabemos. Ainda segundo Proclo, este último "reuniu bons elementos, generalizando várias partes". Um pouco adiante, tendo mencionado os discípulos de Platão, Proclo acrescenta

> não muito mais jovem que estes foi Euclides, aquele que reuniu os elementos e organizou os resultados de Eudoxo e aperfeiçou os de Teeteto; além disso, elevou as demonstrações a níveis mais convincentes que aqueles de seus predecessores; (Thomas, 1991, p. 155).

Na verdade o Resumo de Proclo, que já citamos tantas vezes, aparece na segunda parte do prólogo de seu *Comentário ao primeiro livro dos Elementos de Euclides*, que provavelmente era uma edição de suas lições de geometria para iniciantes da Escola Neo-Platônica de Atenas, da qual era o diretor.

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. Inicia-se, assim, o período histórico conhecido como Helenístico.

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 uma espécie de instituto de pesquisas. 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 muitas 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 referem-se a 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*. De nosso ponto de vista, estas últimas são propriedades básicas dos números reais; por isso vamos considerar em mais detalhes apenas as definições e os postulados.

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; (Thomas, 1991, p. 439).

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. Por exemplo, 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. estender 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 9, corta duas retas r_1 e r_2 e sejam α e β os ângulos formados,

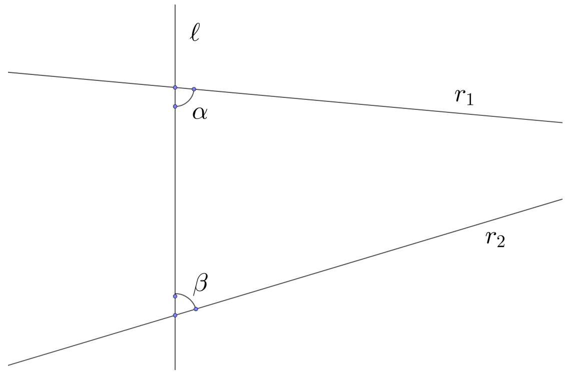


FIGURA 9. Quinto postulado dos *Elementos*

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^o, 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. Múltiplas 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 geometria que respeita os quatro primeiros postulados, mas não o quinto, sem incidir em 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*. Este era o termo usado por Aristóteles e 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 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. Suponhamos que AB e CD são segmentos de retas rotulados de modo que A está à esquerda de B e C à esquerda de D. Para somar AB com CD, um contemporâneo de Euclides executaria as seguintes etapas:

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

Passo 2: com o compasso aberto desta maneira, poria a agulha em B e marcaria, 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 o 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 que temos quatro segmentos A_1B_1, A_2B_2, A_3B_3 e A_4B_4, cujos comprimentos são, respectivamente, c_1, c_2, c_3 e c_4. Para provar a igualdade

$$\frac{c_1}{c_2} = \frac{c_3}{c_4}$$

no âmbito da teoria de Eudoxo, Euclides precisava mostrar que uma das seguintes equivalências é verdadeira para qualquer escolha de inteiros positivos m e n:

$$(mA_1B_1 > nA_2B_2 \text{ se, e somente se, } mA_3B_3 > nA_3B_3);$$
$$(mA_1B_1 = nA_2B_2 \text{ se, e somente se, } mA_3B_3 = nA_3B_3);$$
$$(mA_1B_1 < nA_2B_2 \text{ se, e somente se, } mA_3B_3 < nA_3B_3).$$

Certamente isto vai lhe parecer uma maneira bastante complicada de expressar o que, para nós, é uma mera igualdade de números. A bem da verdade é preciso deixar claro que, ao utilizar a igualdade de números reais, varremos a dificuldade para baixo do tapete. Teremos mais a dizer sobre isto no capítulo 3. O fato é que, complicada ou não, a definição de Eudoxo permitiu aos matemáticos gregos mostrar que não existe nenhuma unidade sob a qual o lado e a diagonal do quadrado tenham ambos, como medidas, números inteiros. Expressamos isto dizendo que o lado e a diagonal do quadrado são *incomensuráveis*.

Esta descoberta fundamental não poderia ter sido feita pelos escribas da Mesopotâmia ou do Antigo Egito, porque estavam interessados em medições práticas. Mas o valor de uma medição produz, apenas, uma quantidade finita de decimais e, portanto, pode ser expresso como uma fração. Por exemplo, medindo o comprimento da diagonal de um quadrado de um metro de lado com a precisão de 50 casas decimais, o que é virtualmente impossível, obteríamos

$$1.41421356237309504880168872420969807856967187537 69$$

que, nada mais é, que a fração

$$\frac{141421356237309504880168872420969807856967187537 69}{100000000000000000000000000000000000000000000000000}.$$

Nas palavras do historiador W. Knorr (Knorr, 2004, p. 87):

> Pode-se, por exemplo, reconhecer as dificuldades computacionais na consideração de quantidades como a raiz quadrada de 2; mas a afirmação e a demonstração de sua *irracionalidade* é uma questão de ordem totalmente diferente. De fato, até mesmo chegar ao nível de *definir* o que a incomensurabilidade significa, para não mencionar a sondagem das várias construções para estabelecer quais delas produzem linhas incomensuráveis, não é possível

> sem um controle preciso dos métodos de demonstração, especialmente do método de demonstração indireta, assim como o reconhecimento de que alguns conceitos muito básicos requerem definições especiais, nada óbvias.

O método de *demonstração indireta* mencionado nesta citação, também chamado de *demonstração por contradição* ou *redução ao absurdo* é, ainda hoje, um dos mais utilizados em matemática. Teremos oportunidade de analisar este método com cuidado na seção 2 do capítulo 3. Mas, para que você tenha, desde já, uma ideia de como a demonstração por contradição funciona, vamos ilustrá-la com dois exemplos.

Imagine, que sua irmã de 10 anos lhe perguntasse porque existem infinitos números. Uma resposta possível seria dizer "qual você acha que seria o maior número?" Se ela respondesse "um quintilhão", você poderia responder "um quintilhão e um também é um número", que é basicamente uma demonstração por contradição da infinidade dos números naturais. Outra demonstração por contradição extremamente simples é a de que não existe um menor número racional positivo. Suponhamos que este número existisse e fosse igual a q; como $q/2 < q$ também é um número racional, temos uma contradição; provando, assim, que não faz sentido supor que há um menor número racional.

A irracionalidade da raiz quadrada de 2, mencionada por Knorr, é o conteúdo do Teorema 2.1 da página 55. A demonstração que apresentamos lá é uma versão modernizada da que aparece na Proposição 117 do Livro X dos *Elementos*; embora o consenso atual seja de que esta demonstração não fazia parte da versão original da obra de Euclides. A mesma demonstração é mencionada por Aristóteles nos *Analíticos anteriores*.[1]

6. O método axiomático

Embora os *Elementos* tenham se tornado o livro de geometria por excelência; mesmo no período Helenístico, matemáticos como Arquimedes e Apolônio não faziam uso sistemático de um esquema tão rígido quanto o de Euclides na apresentação dos seus resultados. Arquimedes chega mesmo a explicar, em um de seus tratados, os métodos heurísticos que havia utilizado para obter alguns dos teoremas, que depois provou usando métodos mais usuais. A história, cheia de reviravoltas, do códice que contém este livro de Arquimedes é contada em Netz, Noel (2008).

Apesar das divergências de apresentação dos *Elementos* com a prática de outros matemáticos, o livro de Euclides tornou-se, com o passar do tempo, o local ao qual recorriam todos que desejavam aprender geometria. Com isso, os *Elementos* foram traduzidos para o árabe, por volta do ano 800 e para o latim no século XII; ao passo que a primeira edição impressa data de 1482 Wardhaugh (2021).

[1]Livro I. 23, 41 a 26–7.

A primazia dos *Elementos* como fonte da geometria só começaria a ser ameaçada no século XVII, com a criação da geometria analítica por Pierre de Fermat e René Descartes. Descartes, em particular, tinha uma opinião não muito lisonjeira da maneira como os geômetras gregos apresentavam seus resultados. Na Regra IV de suas *Regras para a Direção do Espírito*, ele manifesta sua suspeita de que, como os métodos utilizados pelos matemáticos gregos para descobrir seus resultados eram muito "fáceis e simples", eles

> preferiram substituí-los por verdades áridas, demonstradas de maneira dedutiva [...] para que venhamos a nos maravilhar com eles; em vez de nos ensinar seu método, o que teria destruído nossa admiração; (Descartes, 1997, p. 15).

Certamente Descartes não teria tal opinião se tivesse conhecido o tratado de Arquimedes, mencionado acima, que, entretanto, só foi identificado em 1906.

Apesar desta opinião não ter sido partilhada por outros matemáticos, o desenvolvimento do cálculo diferencial e integral esteve de tal modo ligado aos métodos da geometria analítica que a apresentação de novos resultados matemáticos a partir de axiomas, ao estilo de Euclides, foi muito pouco utilizado nos séculos XVII e XVIII. Contudo, isto não significa que os *Elementos* tenham sido esquecidos ao longo destes séculos, já que continou a ser utilizado, como livro didático, em muitos países.

Na verdade, o desenvolvimento acelerado do cálculo diferencial e integral nos séculos XVII e XVIII acabou sendo um dos fatores que contribuíram, ainda que indiretamente, para o retorno dos métodos axiomáticos. Uma boa maneira de ilustrar isto é considerar o comportamento de somas infinitas. Algumas destas somas são iguais a um número, como ocorre com

$$\frac{1}{2}+\frac{1}{4}+\frac{1}{8}+\frac{1}{16}+\cdots, \tag{2}$$

que é igual a 1; ao passo que outras, claramente, produzem valores infinitos, como

$$1+1+1+\cdots. \tag{3}$$

Mas, o que dizer sobre

$$1-1+1-1+1-1+\cdots? \tag{4}$$

À primeira vista, pode parecer que esta soma deveria ser igual a zero mas, se reagruparmos os termos na forma

$$1+(-1+1)+(-1+1)+(-1+1)\cdots;$$

isto não sugere que a soma deveria ser um? Somas que correspondem a um número, como é o caso de (2), são chamadas de *convergentes*, às demais, como (3) e (4), damos o nome de *divergentes*. Por dois séculos as somas infinitas foram utilizadas, sem que houvesse a preocupação em determinar se eram convergentes ou divergentes. Finalmente, no século XIX, vários matemáticos se deram conta do problema, o que deu início a um movimento para tornar mais rigorosos os métodos infinitos que caracterizam o cálculo diferencial e integral.

Influenciados pela onda de rigor da época, alguns matemáticos se voltaram para o tratamento da geometria nos *Elementos*. Como observamos na seção 5, a apresentação de Euclides deixa muito a desejar do ponto de vista da coerência lógica. No exemplo mencionado naquela seção, Euclides menciona um ponto "dentro" do círculo, sem nunca especificar o que isto significa. Dois outros exemplos bem conhecidos, ambos do Livro I, são: a proposição I, onde é usado o fato de que dois círculos cujo raio é igual à distância entre seus centros têm que se intersectar e a proposição IV, em que um triângulo é deslocado de modo a coincidir, com outro triângulo, embora esta operação nunca tenha sido definida.

Para melhor indicar como são entendidas as apresentações axiomáticas modernas, é interessante comparar dois tratamentos da geometria euclidiana do final do século XIX. O primeiro destes é devido a Moritz Pasch (1843-1930). Suas *Lições de geometria moderna*, publicadas em 1882, são o primeiro livro no qual os vários defeitos encontrados nos *Elementos* são finalmente sanados. Pasch propôs oito princípios (Grundsätz) a partir dos quais os resultados da geometria plana poderiam ser obtidos. Segundo Pasch

> espera-se que estes princípios cubram completamente todo o material empírico a ser tratado de maneira matemática; (Pasch, 1903, p. 17).

Mas, ele acrescenta,

> depois que [estes princípios] foram estabelecidos, não podemos mais nos confiar na percepção sensorial.

O segundo livro que desejamos considerar são os *Fundamentos da geometria*, publicado por David Hilbert em 1899. Na versão de Hilbert, a geometria (plana e espacial) se baseia em 20 axiomas, subdividos em cinco grupos: axiomas de ligação, de ordem, de congruência, de paralelismo e de continuidade.

A grande diferença entre as formulações de Hilbert e Pasch está na interpretação dos axiomas. O ponto de vista de Hilbert é de que não importa o que são os objetos (pontos, retas, planos, ...) aos quais os axiomas se referem, 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; (Lan, 2005, p. 714).

Ainda em 1899, Hilbert propôs um enfoque axiomático para o conceito de número e no ano seguinte, em uma conferência que se tornou extremamente famosa, incluiu a axiomatização da física entre os 21 problemas que apresentou aos matemáticos como desafios para o século que se iniciava. Com o passar dos anos, Hilbert transformaria suas ideias em um programa sistemático para axiomatizar toda a matemática. Entretanto, como considerava os axiomas de modo puramente formal, e não como abstrações do mundo natural, Hilbert precisou incluir, como parte de seu programa, a necessidade de provar a *consistência* dos axiomas; isto é, que não conduziriam à demonstração simultânea de um teorema e de sua negativa.

Uma das áreas em que a axiomatização logo fez sentir sua influência foi a teoria de conjuntos. Como veremos no início do capítulo 2, muito cedo esta teoria levou a situações absurdas, o que a tornou um tema promissor para um tratamento axiomático. Em 1908, Ernst Zermelo (1871–1953), que estudou sob Hilbert, publicou um conjunto de axiomas que, na versão aperfeiçoada, independemente, por Abraham Fraenkel (1891–1965) e Thoralf Skolem (1887–1963), é usado até hoje.

Um enfoque ainda mais radical que o de Hilbert foi iniciado por Bertrand Russell (1872–1979), que almejava reduzir toda a matemática à lógica e, com isso, eliminar os problemas aos quais a teoria de conjuntos havia dado lugar. Com este objetivo Russell escreveu, em colaboração com A. N. Whitehead (1861–1947) o *Principia Mathematica*, cujos três volumes foram publicados entre 1910 e 1913. Famosamente, a demonstração de que 1 + 1 = 2, que reproduzimos abaixo, só é feita na página 86 do segundo volume:

$$\begin{aligned}
&\vdash . * 110{\cdot}632 . *101{\cdot}21{\cdot}28{\cdot} \supset \\
&\vdash .1 +_c 1 = \hat{\xi}\{(\exists y).y \in \xi . \xi - \iota^c y \in 1\} \\
&[*54{\cdot}3] = 2. \supset \vdash .\text{Prop}
\end{aligned}$$

A propósito, o primeiro volume do *Principia Mathematica* tem 658 páginas. Imediatamente após esta demonstração, lê-se

> [a] proposição acima é ocasionalmente útil. Ela será usada pelo menos três vezes[.]

Teremos oportunidade de mencionar a obra de Rusell e Whitehead novamente no capítulo 2 quando estudarmos os conectivos lógicos.

As tentativas de provar a consistência de sistemas de axiomas, como o de Hilbert para a geometria e o de Zermelo para conjuntos, naufragaram em 1931. Depois de provar que a lógica de primeira ordem, cujos princípios mais elementares estudaremos nos capítulos 2 e 3, é consistente, Kurt Gödel (1906–1978) mostrou em 1931 que não é possível provar, dentro dos limites propostos pelo próprio Hilbert, a consistência de nenhum sistema de axiomas que inclua os números naturais. O impacto dos teoremas de Gödel na matemática e na filosofia foi enorme, o que os tornou extremamente famosos.

Ainda como parte de seu programa de axiomatização da matemática, Hilbert publicou, em 1928, o livro *Princípios de lógica matemática*, escrito com a colaboração de seu ex-aluno Wilhelm Ackermann. Um dos problemas propostos neste livro é o *Entscheidungsproblem*, ou problema da decibilidade, que pergunta se existe um algoritmo capaz de determinar se uma dada fórmula lógica pode ou não ser provada, em uma quantidade finita de passos, usando as regras da lógica. Este problema foi solucionado, independentemente, por Alan Turing (1912–1954) e Alonzo Church (1903–1995) em 1936. O que eles mostraram é que tal algoritmo não pode existir.

O artigo em que Turing mostra que o problema da decibilidade não tem solução tornou-se especialmente famoso porque é nele que Turing propôs o modelo matemático de computabilidade usado hoje em dia. Uma discussão elementar deste modelo pode ser encontrada em *The Universal Computer* de Martin Davis Davis (2000). O mesmo livro apresenta outros problemas matemáticos que não admitem uma solução algorítmica, entre eles o que trata da existência ou não de um algoritmo para resolver qualquer equação diofantina, e que também foi proposto por Hilbert.

7. Olhando adiante

Tendo chegado a este ponto, talvez você esteja se perguntando se, apesar do título, este não é, simplesmente, mais um livro de geometria. A resposta é não. Há duas razões pela qual a geometria euclidiana plana dominou este capítulo introdutório.

A primeira é histórica. É aos gregos antigos que devemos o enfoque atualmente em voga na matemática e, para eles, um dos principais ramos da matemática era a geometria. Por isso, não é surpreendente que, ao introduzir os princípios da lógica, Aristóteles tenha se inspirado no uso que deles faziam os geômetras de sua época.

A segunda razão é mais prosaica. Como você estudou um pouco de geometria na escola, a terminolgia e os resultados básicos da geometria euclidiana já lhe são familiares. Com isso podíamos mencioná-los sem nenhuma introdução prévia, como seria conveniente em um capítulo inicial como este.

Para você ter uma ideia do que esperar deste livro, convém lhe dar uma visão panorâmica do que estar por vir. Os seis capítulos restantes podem ser agrupados em dois temas principais. Os três primeiros tratam de noções extremamente básicas para o entendimento da matemática apresentada nos cursos de ciência da computação, matemática e engenharia. Já o capítulo cinco contém um tratamento axiomático da geometria plana dos triângulos, que é utilizado, nos dois capítulos finais, para estudar os gráficos de algumas funções elementares e as funções trigonométricas.

Vejamos de que trata cada um destes capítulos, individualmente. O capítulo 2 é uma introdução aos conectivos lógicos e aos conjuntos e suas operações. Nele introduziremos tabelas verdade e veremos como provar as propriedades das operações com conjuntos utilizando o que aprendemos sobre os conectivos lógicos.

No capítulo 3 completamos nosso estudo de lógica discutindo quantificadores e utilizamos tudo o que aprendemos até aí para analisar, com algum detalhe, os números reais. Supondo conhecidas as propriedades algébricas básicas dos números reais, e aquelas referentes à maneira como são ordenados, mostramos que é possível deduzir delas muitas outras propriedades, entre elas as bem conhecidas regras de sinais.

O quarto capítulo tem como tema as funções. Além de definir funções de maneira bastante geral, investigamos com cuidado as funções inversíveis e a operação de composição de funções. Apesar das funções reais serem introduzidas neste capítulo, não mencionaremos seus gráficos porque, para isso, precisamos de conhecimentos de geometria que, oficialmente, só teremos a partir do capítulo seguinte.

O capítulo cinco é dedicado a uma apresentação axiomática da geometria, baseada em números reais. Partindo de cinco axiomas simples, que incluem a possibilidade de medir segmentos e ângulos, provaremos, passo-a-passo, vários dos resultados de geometria plana que aprendemos na escola. Entre estes: os casos de congruência e semelhança de triângulos, o fato de que a soma dos ângulos internos de um triângulos é igual a π radianos e o teorema de Pitágoras.

De posse da geometria elementar de triângulos, podemos investigar os gráficos de funções no capítulo seis. Além de mostrar que o gráfico de uma função linear é uma reta, analisaremos as propriedades do gráfico de uma função quadrática em muito mais detalhe do que é normalmente feito na escola.

O último capítulo é uma introdução às funções trigonométricas, utilizando os teoremas sobre congruência e semelhança de triângulos estudados no capítulo cinco. Na última seção deste capítulo mostraremos como as fórmulas que expressam o seno e o cosseno de um ângulo em função da tangente da metade do ângulo para desenvolver um algoritmo capaz de provar relações trigonométricas.

Como vimos na seção 4, o método científico, como proposto por Aristóteles, toma como ponto de partida afirmações que deveriam ser suficientemente óbvias para que ninguém precisasse ser convencido delas. Deste ponto de vista nossa apresentação da teoria de conjuntos e dos números reais deixa muito a desejar. O problema com os números reais é que não seria muito sensato mostrar, em um livro elementar, como são construídos a partir dos inteiros. Para fazer isto de maneira satisfatória precisaríamos de um livro inteiro; voltaremos a comentar sobre isto, em mais detalhes, na seção 5 do capítulo 3.

A questão dos conjuntos é ainda mais complicada, porque alguns dos axiomas de Zermelo-Frankel, que já foram mencionados na página 19, têm um teor extremamente técnico, porque foram introduzidos para evitar os vários paradoxos dos quais a teoria de conjuntos sofreu em seu início, o primeiro dos quais será mencionado no início do próximo capítulo.

Contudo, nosso enfoque dedutivo será bastante sólido, desde que você aceite as propriedades básicas dos conjuntos e dos números reais, introduzidas, respectivamente, nos capítulos 2 e 3. Em outras palavras, nosso compromisso consiste em utilizar o conteúdo destes dois capítulos para desenvolver uma boa parte da matemática elementar de maneira bastante criteriosa, ainda que não suficientemente formal para satisfazer um matemático.

Exercícios

1. Use o método de recorte e colagem para determinar uma fórmula para a área do trapézio de altura h, cuja base menor mede b e cuja base maior B.

2. Use a fórmula da área de um triângulo para provar cada uma das seguintes afirmações:
   (a) Se ABC é um triângulo e M é o ponto médio do lado BC, então os triângulos ABM e ACM têm a mesma área.
   (b) Sejam ABC e ABD dois triângulos de mesma base AB. Se os vértices C e D destes triângulos estão sobre uma mesma reta paralela à base AB, então os dois triângulos têm a mesma área.
   (c) Sejam ABC e ADE dois triângulos cujas bases BC e DE estão sobre uma mesma reta, então
   $$\frac{\text{área de } ABC}{\text{área de } ADE} = \frac{BC}{DE}.$$

3. Seja ABC um triângulo, P um ponto dentro do triângulo e A' o ponto em que AP corta BC. Então
   $$\frac{\text{área de } ABP}{\text{área de } ACP} = \frac{A'B}{A'C}.$$
   Dica: use o exercício 2(c) três vezes.

4. Prove que a diagonal de um paralelogramo divide-o em dois triângulos de mesma área.

CAPÍTULO 2

Conjuntos

Nosso primeiro capítulo trata da teoria de conjuntos, que estuda uma das noções mais básicas da matemática. Como a axiomatização da teoria de conjuntos suscita questões bastante técnicas, não é conveniente adotá-la em um livro elementar como este. Por isto nosso tratamento inicial de conjuntos será bastante ingênuo. Contudo, uma vez que as noções fundamentais da teoria de conjuntos tenham sido estabelecidas, nosso estudo tomará um caminho mais rigoroso. Em particular, provaremos, cuidadosamente, as propriedades básicas das operações com conjunto. Para isto precisaremos introduzir os conectivos lógicos e investigar seu comportamento.

1. Conjuntos

Conjunto é a noção matemática que sistematiza a ideia de uma coleção de objetos distintos, chamados de *elementos* do conjunto. Quando um dado elemento faz parte do conjunto, dizemos que *pertence* ao conjunto. Assim, as bolas de gude do seu irmão, os livros de uma biblioteca, um cardume de peixes e as moléculas de água em uma garrafa são todos exemplos de conjuntos.

É importante observar que o que dissemos acima sobre conjuntos não constitui uma verdadeira definição porque, a bem da verdade, conjunto e coleção são basicamente sinônimos. Por isso, em matemática, conjunto, elemento e o fato de um elemento pertencer ou não a um conjunto são consideradas noções primitivas; isto é, que não admitem definição. Mas, como usar algo, se não sabemos o que é? A resposta é que, como vimos no capítulo 1, especificamos os *axiomas* que estas noções satisfazem. Neste livro não desenvolveremos a teoria de conjuntos a partir de seus axiomas, porque isto seria inapropriado em um primeiro contato com esta teoria. Por isso vamos nos contentar com a noção intuitiva apresentada no início desta seção.

Antes de prosseguir, precisamos introduzir um pouco da terminologia. Normalmente denotaremos conjuntos por letras maiúsculas e seus elementos por letras minúsculas. Se x é um elemento do conjunto S, escrevemos $x \in S$. O símbolo $\in$ foi introduzido por Giuseppe Peano em 1889 e é uma versão estilizada da letra grega ϵ (épsilon). Diremos que os conjuntos S e S' são iguais se têm exatamente os mesmos elementos; indicamos isto escrevendo $S = S'$.

A maneira mais simples de especificar um conjunto consiste em listar seus elementos entre chaves, por exemplo,

$$C = \{1, 2, 3, 5, 7, 9, 12, 14\}. \tag{5}$$

Como a única relação relevante entre um conjunto e um elemento é a pertinência, para a qual é indiferente a ordem em que os elementos são listados, temos que

$$\{1, 2, 3, 5, 7, 9, 12, 14\} = \{12, 5, 7, 14, 9, 2, 1, 3\}.$$

Da mesma forma, repetir um *mesmo* elemento entre as chaves não altera o fato de pertencer ou não a um dado conjunto; assim,

$$\{1, 2, 3, 5, 7, 9, 12, 14\} = \{1, 1, 1, 2, 3, 5, 5, 5, 7, 9, 9, 9, 9, 9, 9, 12, 14\}.$$

A segunda maneira de especificar um conjunto, já utilizada nos exemplos do início desta seção, consiste em enunciar uma propriedade comum a todos os elementos de um conjunto *e somente a estes elementos*. Logo veremos mais exemplos, antes,porém, precisamos de uma definição.

Quando todos os elementos de um conjunto A também são elementos de um conjunto B, dizemos que A *está contido* em B, ou que A é *subconjunto* de B. Denotamos isto escrevendo $A \subset B$. Em particular, como todos os elementos de B estão contidos em B, temos que $B \subset B$. Quando o subconjunto A de B não é igual a B, dizemos que se trata de um *subconjunto próprio* de B e escrevemos $A \subsetneq B$.

A especificação de um conjunto a partir de uma dada propriedade é especialmente útil quando se trata do subconjunto de um conjunto dado. Por exemplo,

$$S = \{1, 3, 5, 7, 9\} \tag{6}$$

é o subconjunto do conjunto C, definido em (5), cujos elementos são números ímpares. Em geral, especificamos o conjunto Y formado pelos elementos $x \in X$ que satisfazem uma dada propriedade $P(x)$ escrevendo

$$Y = \{x \in X \mid P(x)\}.$$

Por exemplo, no caso do conjunto S definido em (6),

$$P(x) = x \text{ é ímpar};$$

assim,

$$S = \{x \in C \mid x \ \text{ é ímpar}\}.$$

Pode acontecer de um mesmo subconjunto poder ser especificado a partir de mais de uma propriedade diferente; por exemplo,

$$S = \{x \in C \mid 1 \leq x \leq 9\}.$$

Frequentemente é desejável considerar que todos os objetos matemáticos que estamos considerando em um dado momento sejam elementos de um único conjunto. Quando isto ocorre, dizemos que este é nosso *conjunto universo*. Assim, nos exemplos dos últimos parágrafos, o universo era C; na teoria dos números elementar, o universo é o conjunto $\mathbb{Z}$ dos números inteiros; já na geometria plana, o universo é o conjunto dos pontos do plano. De agora em diante, admitiremos sempre que os conjuntos com os quais estamos lidando estão todos contidos em um dado conjunto universo, que denotaremos sistematicamente por $\mathcal{U}$. As únicas exceções a esta prática ocorrerão quando o universo já tem uma notação estabelecida, como é o caso dos conjuntos $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{Q}$ e $\mathbb{R}$.

A especificação de subconjuntos por meio das propriedades que seus elementos satisfazem tem um efeito colateral inesperado. Por exemplo, nada nos impede de escrever

$$V = \{x \in C \mid x > 20\}.$$

Entretanto, nenhum elemento de C é maior que 20, portanto não há elementos com os quais popular V. A mais óbvia é decretar que há casos em que uma propriedade, apesar de fazer sentido, não define um subconjunto. O problema com esta solução é que introduz a necessidade de determinarmos se uma dada propriedade é ou não satisfeita por algum elemento do universo, antes de podermos construir o conjunto. Por isso, a solução efetivamente adotada em matemática consiste em admitir que um conjunto possa não ter nenhum elemento; afinal, uma caixa vazia continua sendo uma caixa. O conjunto sem elementos é conhecido como *conjunto vazio* e denotado por $\varnothing$, de modo que

$$V = \{x \in C \mid x > 20\} = \varnothing.$$

Naturalmente, se fôssemos escrever o conjunto vazio relacionando seus elementos, teríamos que escrever $\varnothing = \{\}$.

Uma propriedade importante do conjunto vazio é que está contido em qualquer conjunto. Para entender o porquê disto, lembre-se que, pela definição de inclusão, para que um conjunto Y *não* esteja contido em outro conjunto X é necessário que haja um elemento em Y que não esteja em X. Mas $\varnothing$ não tem nenhum elemento que não esteja contido em X. Logo, nos vemos obrigados a concluir, por nossa definição, que $\varnothing \subset X$, não importa qual seja o conjunto X.

Como quaisquer objetos podem ser reunidos para formar um conjunto, podemos criar um conjunto cujos elementos também são conjuntos. Por exemplo, dado um conjunto A, podemos formar o conjunto $\mathcal{P}(A)$, cujos elementos são os subconjuntos de A. Quando $A = \{1, 2, 3\}$, este conjunto será

$$\mathcal{P}(A) = \{\varnothing, \{1\}, \{2\}, \{3\}, \{1, 2\}, \{1, 3\}, \{2, 3\}, \{1, 2, 3\}\}.$$

Dizemos que $\mathcal{P}(A)$ é o *conjunto potência*, ou *conjunto das partes* de A. O fato de um conjunto poder ser elemento de outro conjunto significa que

$$\varnothing \subset \{\varnothing\} \quad \text{e} \quad \varnothing \in \{\varnothing\}$$

são ambas verdadeiras: a primeira porque, como já vimos, o conjunto vazio está contido em qualquer conjunto; a segunda, porque o único elemento do conjunto $\{\varnothing\}$ é o conjunto vazio. Assim, $\{\varnothing\}$ é como uma caixa que contém apenas uma caixa vazia.

No início de século XX os matemáticos descobriram que não faz sentido falar de um conjunto universo absoluto, que contenha todos os conjuntos possíveis como subconjuntos. Para entender o porquê disto, digamos que um tal conjunto existisse. Denotando-o por Ω, poderíamos definir o subconjunto

$$\mathcal{R} = \{x \in \Omega \mid x \notin x\}.$$

Uma vez tendo definido este conjunto, podemos nos perguntar se $\mathcal{R}$ é, ou não um elemento dele próprio. Infelizmente isto nos leva a um beco sem saída. Lembre-se que o que determina se um dado $x \in \Omega$ pertence ou não a $\mathcal{R}$ é a condição $x \notin x$. Logo, $\mathcal{R} \in \mathcal{R}$ só poderia ocorrer quando $\mathcal{R} \notin \mathcal{R}$. Por outro lado, se $\mathcal{R} \notin \mathcal{R}$, então $\mathcal{R}$ satisfaz a condição que define a si próprio, donde $\mathcal{R} \in \mathcal{R}$. Em ambos os casos obtivemos afirmações contraditórias, de modo que temos um paradoxo. Portanto, para garantir a saúde da teoria de conjuntos é necessário que o conjunto Ω não possa ser construído. Detalhes de como isto é feito podem ser encontrados (Devlin, 1993, pp. 46–50).

O paradoxo do parágrafo anterior, conhecido como *paradoxo de Russell*, é parte de uma história um tanto dramática. Em 1893, o filósofo Gottlob Frege (1848–1925) havia publicado o primeiro volume do seu livro *Leis Básicas da Aritmética*, no qual propunha uma maneira de derivar completamente a aritmética a partir de princípios lógicos. Em 1903, quando o manuscrito do segundo volume já estava pronto para ser enviado à gráfica, Frege recebeu uma carta de Bertrand Russell, na qual Russell explicava que o argumento acima podia ser derivado de uma das leis básicas estabelecidas no primeiro volume da obra de Frege. Isto levou Frege a incluir um apêndice no volume 2 das *Leis Básicas da Aritmética* onde tenta sanar o problema detectado por Russell. O apêndice começa, famosamente, com a seguinte declaração

> [d]ificilmente algo pior pode ocorrer ao escritor de um trabalho científico do que ter as fundações do seu edifício abaladas quando o trabalho já foi concluído. Uma carta do senhor Bertrand Russell me pôs exatamente nesta posição quando a impressão deste volume estava sendo concluída.

2. Operações com conjuntos

Nesta seção introduzimos as operações elementares de conjuntos e investigamos, de maneira heurística, suas propriedades básicas. A primeira destas operações é a união. Se A e B são subconjuntos de $\mathcal{U}$, definimos sua *união* $A \cup B$ por

$$A \cup B = \{x \in \mathcal{U} \mid (x \in A) \text{ ou } (x \in B)\}.$$

Em outras palavras, $A \cup B$ é o conjunto que resulta quando juntamos os elementos de A e de B em um mesmo conjunto. Por exemplo, no caso em que o universo é o conjunto

$$\mathbb{N} = \{1, 2, 3, \dots\}$$

dos números naturais, a união dos subconjuntos

$$S = \{1, 2, 3, 4, 5, 7\} \ \text{ e } \ S' = \{5, 6, 7, 8, 9\}$$

será o conjunto

$$S \cup S' = \{1, 2, 3, 4, 5, 6, 7, 8, 9\}.$$

Voltando ao caso geral, em que A e B são subconjuntos de $\mathcal{U}$, note que, juntar os elementos de A aos elementos de B, ou os elementos de B aos de A, produz o mesmo conjunto; isto é,

$$A \cup B = B \cup A,$$

de modo que a união de conjuntos é *comutativa*. Além disso, como $\varnothing$ não tem elementos, um conjunto não se altera quando é unido ao vazio,

$$A \cup \varnothing = A.$$

Portanto, o conjunto vazio é o *elemento neutro*, às vezes também chamado *elemento identidade*, da união de conjuntos. Uma propriedade semelhante afirma que todo conjunto é *idempotente* relativamente à união; isto é, que

$$A \cup A = A,$$

qualquer que seja o conjunto $A \subset \mathcal{U}$. No extremo oposto de unir um conjunto com o vazio, está uni-lo ao universo. Como todos os elementos de A pertencem a $\mathcal{U}$, obtemos

$$A \cup \mathcal{U} = \mathcal{U};$$

propriedade que é conhecida como *absorção*.

Para entender a importância da próxima propriedade da união é necessário lembrar que só definimos a união de dois conjuntos de cada vez. Para unir três conjuntos, precisamos começar unindo dois deles e, então, juntar ao resultado, o terceiro conjunto. A *associatividade* nos garante que, não importa como fizermos isto, sempre obteremos o mesmo resultado. Mais precisamente, se C também é subconjunto de $\mathcal{U}$, então

$$A \cup (B \cup C) = (A \cup B) \cup C.$$

A segunda operação de conjuntos que desejamos considerar é a *interseção* de dois subconjuntos A e B de $\mathcal{U}$, denotada por $A \cap B$, e definida por

$$A \cap B = \{x \in \mathcal{U} \mid (x \in A) \ \text{ e } \ (x \in B)\}.$$

Note que a única diferença desta definição para a da união é que a conjunção *ou* da união foi trocado por um *e*. Assim, $A \cap B$ é o conjunto dos elementos que pertencem, simultaneamente, a A e B. Voltando ao exemplo usado para ilustrar a união, se

$$S = \{1, 2, 3, 4, 5, 7\} \ \text{ e } \ S' = \{5, 6, 7, 8, 9\} \ \text{ então } \ S \cap S' = \{5, 7\}.$$

Mesmo que dois conjuntos não tenham elementos em comum, sua interseção é um conjunto que, neste caso, será vazio. Dizemos que dois conjuntos cuja interseção é vazia são *disjuntos*. As propriedades da interseção são semelhantes às da união, mas $\varnothing$ e $\mathcal{U}$ aparecem com papéis trocados relativamente ao elemento neutro e absorção. Supondo que A, B e C sejam subconjuntos de $\mathcal{U}$, temos que

$$A \cap B = B \cap A, \quad A \cap \mathcal{U} = A \text{ e } A \cap \varnothing = \varnothing$$

e também que

$$A \cap (B \cap C) = (A \cap B) \cap C.$$

Todas as propriedades que vimos até agora dizem respeito à combinação de conjuntos usando uma única operação, mas nada nos impede de misturar a união com a interseção de conjuntos em uma mesma expressão. Neste caso, as propriedades mais importantes são aquelas que nos permitem distribuir, tanto a união sobre a interseção,

$$A \cup (B \cap C) = (A \cup B) \cap (A \cup C),$$

quanto a interseção sobre a união,

$$A \cap (B \cup C) = (A \cap B) \cup (A \cap C).$$

A tabela 1 resume todas as propriedades da união e da interseção que vimos acima, com seus respectivos nomes.

Note que a expressão $A \cup B \cap C$ é ambígua porque ser interpretada como $A \cup (B \cap C)$ ou como $(A \cup B) \cap C$ e os resultados destas expressões podem ser totalmente diferentes. Por exemplo, quando

$$A = \{1\}, \quad B = \{2\} \text{ e } C = \{1, 2\}$$

temos que

$$A \cup (B \cap C) = \{1\} \text{ ao passo que } (A \cup B) \cap C = \{1, 2\}.$$

Por isso, é necessário posicionar os parêntesis corretamente sempre que haja algum risco de ambiguidade e você precisa manter-se alerta para isso ao resolver os exercícios.

Associativa	$A \cup (B \cup C) = (A \cup B) \cup C$	$A \cap (B \cap C) = (A \cap B) \cap C$
Comutativa	$A \cup B = B \cup A$	$A \cap B = B \cap A$
Idempotente	$A \cup A = A$	$A \cap A = A$
Elemento neutro	$A \cup \varnothing = A$	$A \cap \mathcal{U} = A$
Absorção	$A \cup \mathcal{U} = \mathcal{U}$	$A \cap \varnothing = \varnothing$
Distributiva	$A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$	$A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$

TABELA 1. Propriedades da união e da interseção de conjuntos.

A última operação com conjuntos que consideraremos neste capítulo é o complementar. Se A é um subconjunto do universo $\mathcal{U}$, então seu *complementar*, denotado por A^c, é o conjunto constituído pelos elementos de $\mathcal{U}$ que não pertencem a A; isto é

$$A^c = \{x \in \mathcal{U} \mid x \notin A\}.$$

Por exemplo, quando o universo é

$$\mathcal{U} = \{1,2,3,4,5,6,7,8,9\} \text{ e } S = \{5,6,7,8,9\} \text{ então } S^c = \{1,2,3,4\}.$$

Caso o universo seja o conjunto dos inteiros, o complementar do conjunto dos números pares é o conjunto dos números ímpares e vice-versa; já o complementar do conjunto dos números primos é formado pela união do conjunto dos números compostos com $\{-1,0,1\}$, porque estes três números não são, nem primos, nem compostos.

Supondo que $\mathcal{U}$ é o universo e A e B são subconjuntos de $\mathcal{U}$, é fácil ver que

$$(A^c)^c = A, \quad \varnothing^c = \mathcal{U} \text{ e } \mathcal{U}^c = \varnothing,$$

mas é bem mais difícil se convencer de que

$$(A \cup B)^c = A^c \cap B^c \text{ e } (A \cap B)^c = A^c \cup B^c.$$

Estas duas últimas propriedades, conhecidas como *leis de De Morgan*, nos dizem que o complementar de uma união não é a união, mas sim a interseção dos complementares dos dois conjuntos, ao passo que o complementar da interseção é a união dos complementares.

Ainda que tenhamos discutido, com algum detalhe, as três operações básicas de conjuntos, as definições que oferecemos não são suficientes para que sejamos capazes, por exemplo, de provar suas propriedades. Para isso, precisamos explicitar, de maneira precisa, o comportamento das conjunções utilizadas para definir a união e a interseção, assim como a negação usada na definição do complementar.

3. Conectivos lógicos

Começaremos com algumas definições básicas. Uma *proposição* é uma afirmação que podemos determinar, pelo menos em princípio, se é verdadeira ou falsa. Assim, verdadeiro e falso, abreviados como V e F, são os *valores verdade* que uma proposição pode assumir. Por exemplo, *está chovendo* é uma proposição, independentemente do fato de estar chovendo agora ou não. Se estiver chovendo, a proposição é verdadeira, caso contrário, é falsa; mas, em ambos os casos é uma proposição. Por outro lado, *está chovendo?* e *feche a porta!* não são proposições, porque não faz sentido tentar determinar se são verdadeiras ou falsas. Contudo, para que tenhamos uma proposição não é necessário que saibamos atribuir um valor a uma afirmação neste momento; *existe vida em Marte* é uma proposição, embora ainda não saibamos se é verdadeira ou falsa.

Assim como sentenças curtas podem ser combinadas, usando conjunções, para formar outras mais longas, proposições podem ser combinadas usando *conectivos*, para formar proposições mais complexas. Começaremos nosso estudo dos conectivos por $\vee$ (lê-se “ou”) e $\wedge$ (lê-se "e"), porque é a partir deles que definimos a união e a interseção de conjuntos na seção 1. Assim, se A e B são subconjuntos de um conjunto universo $\mathcal{U}$, então

$$A \cup B = \{x \in \mathcal{U} \mid (x \in A) \vee (x \in B)\} \text{ e } A \cap B = \{x \in \mathcal{U} \mid (x \in A) \wedge (x \in B)\}. \tag{7}$$

A primeira coisa a notar é que, para um dado $x \in \mathcal{U}$, podemos, ao menos em princípio, determinar, tanto se $x \in A$, quanto se $x \in B$, de modo que estas duas sentenças matemáticas são proposições, o que nos permite conectá-las usando $\vee$ e $\wedge$.

Como mencionamos ao final da seção anterior, para que possamos provar, com segurança, as propriedades da união e da interseção de conjuntos é necessário tornar preciso o uso dos conectivos $\vee$ e $\wedge$, porque as conjunções correspondentes em português são usadas de maneira muito imprecisa. Isto é mais flagrante no caso do "ou", que pode ser tanto *inclusivo*, como quando dizemos "Pedro ou Paulo virão à festa", o que não exclui que ambos possam vir; quanto *exclusivo*, como na frase "você tem que escolher entre virar a página ou fechar o livro", que pressupõe que não podemos executar estas duas ações ao mesmo tempo.

Antes de podermos especificar o uso de $\vee$ e $\wedge$ precisamos de algumas definições. Diremos que uma proposição é *composta* se é construída combinando outras proposições através de conectivos; uma proposição que não é composta, é chamada de *atômica*. Por exemplo, *cheguei em casa e preparei o jantar* é a proposição composta, porque pode ser representada na forma $P \wedge Q$, em que P e Q são as proposição atômicas *cheguei em casa* e *preparei o jantar*, respectivamente.

A maneira como os valores verdade de uma proposição composta dependem das proposições atômicas que a constituem é explicitada através de uma *tabela verdade*. Se a proposição composta é formada por n proposições atômicas, sua tabela verdade começa com n colunas, indexadas pelas proposições atômicas, e acaba com uma coluna para a composta. Neste caso, as primeiras n posições de cada linha serão preenchidas com V ou F, de modo que todas as possíveis combinações destes valores apareça na tabela. A coluna final de uma dada linha contém o valor verdade que a proposição composta assume para os valores verdade dados nas primeiras n posições daquela linha.

Por exemplo, a tabela verdade relativa ao conectivo $\vee$ terá três colunas, indexadas, por P, Q e $P \vee Q$, em que os símbolos P e Q representam as proposições atômicas. A terceira coluna conterá o valor verdade de $P \vee Q$ que resulta da escolha de valores verdades para P e Q feitas nas duas primeiras posições daquela linha.

Na tabela 2 incluímos colunas especificando tanto os valores de $\vee$, quanto os de $\wedge$. Como você pode constatar, $P \vee Q$ só é falso quando ambos, P e Q, são falsos, o que corresponde a afirmar que $\vee$ é o "ou inclusivo". É importante ressaltar que, em matemática, o "ou" é sempre considerado como inclusivo. Já em computação, o "ou" exclusivo ou XOR é bastante utilizado. No exercício 18 você será chamado a descrever o ou exclusivo a partir

P	Q	$P \vee Q$	$P \wedge Q$
V	V	V	V
V	F	V	F
F	V	V	F
F	F	F	F

TABELA 2. Tabelas verdade dos conectivos $\vee$ e $\wedge$.

dos demais conectivos. Já o $\wedge$ não apresenta nenhuma surpresa, uma vez que normalmente esperamos que $P \wedge Q$ só seja verdade quando P e Q forem ambas verdadeiras.

Assim como a união e a interseção são operações com conjuntos, os conectivos $\vee$ e $\wedge$ são operações entre proposições, já que nos permitem combiná-las para formar novas proposições. Por exemplo, se P, Q e R são proposições, é possível formar

$$(P \vee Q) \vee R, \quad (P \vee Q) \wedge R \text{ e } (P \wedge Q) \vee R,$$

entre muitas outras combinações possíveis. Podemos escrever tabelas verdade para especificar o comportamento destas proposições a partir dos valores verdade de P, Q e R, utilizando o que já sabemos sobre os conectivos $\vee$ e $\wedge$. Por exemplo, no caso de $(P \vee Q) \wedge R$, precisaremos de três colunas, uma para cada uma das proposições P, Q e R. Em seguida, usamos a tabela referente ao conectivo $\vee$ para construir a coluna referente a $P \vee Q$. Finalmente, usando os valores verdade de $P \vee Q$ e R, construímos a coluna correspondente a $(P \vee Q) \wedge R$ usando a tabela verdade para $\wedge$. O resultado da aplicação destas etapas é ilustrado na tabela 3.

P	Q	R	$P \vee Q$	$(P \vee Q) \wedge R$
V	V	V	V	V
V	V	F	V	F
V	F	V	V	V
V	F	F	V	F
F	V	V	V	V
F	V	F	V	F
F	F	V	F	F
F	F	F	F	F

TABELA 3. Tabela verdade de $(P \vee Q) \wedge R$.

Uma vez que estamos considerando $\vee$ e $\wedge$ como operações, podemos nos perguntar que propriedades satisfazem. Por exemplo, será que $\wedge$ distribui sobre $\vee$, de modo que

$$(P \vee Q) \wedge R \text{ e } (P \wedge R) \vee (Q \wedge R)$$

têm a mesma tabela verdade? A propósito, duas proposições P_1 e P_2, que têm os mesmos valores verdade são *equivalentes*. Indicamos isto escrevendo $P_1 \equiv P_2$. Portanto, o que estamos perguntando é se

$$(P \vee Q) \wedge R \equiv (P \wedge R) \vee (Q \wedge R). \tag{8}$$

No caso de proposições construídas usando conectivos, como é o caso de (8), a equivalência corresponde a dizer que as duas proposições precisam ter os mesmos valores verdade *quaisquer que sejam as escolhas de valores verdade das proposições atômicas que as compõem*. Uma maneira de verificar isto consiste em construir as colunas referentes às duas proposições em uma mesma tabela e determinar se são ou não iguais. Fizemos isto, no caso da distributividade de $\wedge$ sobre $\vee$, na tabela 4.

P	Q	R	$P \vee Q$	$(P \vee Q) \wedge R$	$P \wedge R$	$Q \wedge R$	$(P \wedge R) \vee (Q \wedge R)$
V	V	V	V	V	V	V	V
V	V	F	V	F	F	F	F
V	F	V	V	V	V	F	V
V	F	F	V	F	F	F	F
F	V	V	V	V	F	V	V
F	V	F	V	F	F	F	F
F	F	V	F	F	F	F	F
F	F	F	F	F	F	F	F

TABELA 4. Distributividadede $\wedge$ sobre $\vee$.

A distributividade de $\vee$ sobre $\wedge$, assim como a associatividade e a comutatividade destes dois conectivos, pode ser facilmente provada de maneira análoga. Antes de listar todas as propriedades dos conectivos que vimos até agora, precisamos introduzir duas proposições especiais: $\top$ corresponde à proposição que é sempre verdadeira e $\perp$ àquela que é sempre falsa. As propriedades

$$\perp \vee P \equiv P, \quad \top \vee P \equiv \top, \quad \perp \wedge P \equiv \perp, \quad \text{e} \quad \top \wedge P \equiv P,$$

são provadas na tabela verdade 5.

O que fizemos até aqui basta para que possamos provar as propriedades da união e da interseção de conjuntos. Mas, para definir formalmente o complementar e provar suas

P	$\top$	$\bot$	$\bot \vee P$	$\top \vee P$	$\bot \wedge P$	$\top \wedge P$
V	V	F	V	V	F	V
F	V	F	F	V	F	F

TABELA 5. Elemento neutro e absorção para $\vee$ e $\wedge$.

propriedades precisamos introduzir a *negação*, denotada por $\neg$ ou ~, que tem como efeito trocar o valor verdade de uma proposição de verdadeiro para falso e vice-versa. Uma consequência imediata desta definição é que a *dupla negativa* retorna uma proposição a seu valor verdade original; isto é $\neg(\neg P) \equiv P$.

O comportamento da disjunção e da conjunção sob a negação é determinado pelas leis de De Morgan:

$$\neg(P \vee Q) \equiv \neg P \wedge \neg Q \ \text{ e } \ \neg(P \wedge Q) \equiv \neg P \vee \neg Q.$$

Provamos, na tabela 6 a primeira destas leis como exemplo de uma tabela verdade que envolve a negação.

P	Q	$P \vee Q$	$\neg(P \vee Q)$	$\neg P$	$\neg Q$	$\neg P \wedge \neg Q$
V	V	V	F	F	F	F
V	F	V	F	F	V	F
F	V	V	F	V	F	F
F	F	F	V	V	V	V

TABELA 6. Demonstração da lei de De Morgan.

Outras duas propriedades da negação,

$$P \vee \neg P \equiv \top \quad \text{e} \quad P \wedge \neg P \equiv \bot .$$

serão extremamente importantes em demonstrações de capítulos posteriores. A primeira destas propriedades é conhecida como *princípio do meio excluído* porque afirma que qualquer proposição é verdadeira ou falsa, sem que haja uma terceira possibilidade, supostamente um meio caminho entre as duas. Note que estas propriedades são consequências imediatas das tabelas verdade de $\vee$ e $\wedge$, porque, se uma de duas proposições ligadas por $\vee$ é verdadeira, a proposição resultante é verdadeira; ao passo que, se uma das proposições ligadas pelo $\wedge$ é falsa, a resultante é falsa. Um sumário geral das propriedades dos conectivos que estudamos neste capítulo pode ser encontrada na tabela 7. Note que, nesta tabela, P, Q e R denotam três proposições quaisquer, *sejam elas atômicas ou não*.

Propriedade	$\vee$	$\wedge$
Associativa	$P \vee (Q \vee R) \equiv (P \vee Q) \vee R$	$P \wedge (Q \wedge R) \equiv (P \wedge Q) \wedge R$
Comutativa	$P \vee Q \equiv Q \vee P$	$P \wedge Q \equiv Q \wedge P$
Idempotente	$P \vee P \equiv P$	$P \wedge P \equiv P$
Elemento neutro	$P \vee \bot \equiv P$	$P \wedge \top \equiv P$
Absorção	$P \vee \top \equiv \top$	$P \wedge \bot \equiv \bot$
Meio excluído	$P \vee \neg P \equiv \top$	$P \wedge \neg P \equiv \bot$
De Morgan	$\neg(P \vee Q) \equiv \neg P \wedge \neg Q$	$\neg(P \wedge Q) \equiv \neg P \vee \neg Q$
Distributiva	$P \vee (Q \wedge R) \equiv (P \vee Q) \wedge (P \vee R)$	$P \wedge (Q \vee R) \equiv (P \wedge Q) \vee (P \wedge R)$

TABELA 7. Propriedades dos conectivos $\vee$, $\wedge$ e $\neg$.

4. Aplicações das propriedades dos conectivos

Nesta seção aplicaremos o que aprendemos na seção anterior a dois problemas: a demonstração das propriedades da união e da interseção enunciadas na seção 2 e a verificação das equivalências entre proposições formadas a partir dos conectivos $\vee$, $\wedge$ e $\neg$.

Começaremos ilustrando o método a ser utilizado provando a associatividade da união de conjuntos, segundo a qual, se A, B e C são subconjuntos de um conjunto universo $\mathcal{U}$, então

$$A \cup (B \cup C) = (A \cup B) \cup C.$$

Usando a definição de união como

$$A \cup B = \{x \in \mathcal{U} \mid (x \in A) \vee (x \in B)\}, \tag{9}$$

temos que

$$A \cup (B \cup C) = \{x \in \mathcal{U} \mid (x \in A) \vee (x \in (B \cup C))\}.$$

Por sua vez,

$$B \cup C = \{x \in \mathcal{U} \mid (x \in B) \vee (x \in C)\};$$

de modo que

$$A \cup (B \cup C) = \{x \in \mathcal{U} \mid (x \in A) \vee ((x \in B) \vee (x \in C))\}. \tag{10}$$

Definindo

$$P(x) = (x \in A), \quad Q(x) = (x \in B) \text{ e } R(x) = (x \in C),$$

podemos reescrever (10) na forma

(11) $$A \cup (B \cup C) = \{x \in \mathcal{U} \mid P(x) \vee (Q(x) \vee R(x))\}.$$

Note, porém, que $P(x)$, $Q(x)$ e $R(x)$ só são proposições para um dado valor de x; caso contrário, não temos como decidir se são afirmações verdadeiras ou falsas. Um raciocínio análogo nos permite escrever

$$(A \cup B) \cup C = \{x \in \mathcal{U} \mid (P(x) \vee Q(x)) \vee R(x)\}.$$

Mas, por definição, dois conjuntos são iguais quando têm os mesmos elementos. Contudo, $x \in \mathcal{U}$ só será elemento de $A \cup (B \cup C)$ quando $P(x) \vee (Q(x) \vee R(x))$ for verdadeira e só será elemento de $(A \cup B) \cup C$ quando $(P(x) \vee Q(x)) \vee R(x)$ for verdadeira. Portanto, para que os dois conjuntos sejam iguais é necessário que a equivalência

$$P(x) \vee (Q(x) \vee R(x)) \equiv (P(x) \vee Q(x)) \vee R(x)$$

seja válida, não importa qual seja o valor escolhido para $x \in \mathcal{U}$. Mas isto já sabemos que é verdade, porque a associatividade do $\vee$ garante que a igualdade $P \vee (Q \vee R) \equiv (P \vee Q) \vee R$ vale, quaisquer que sejam as proposições P, Q e R.

Todas as outras propriedades das operações de conjuntos podem ser tratadas da mesma maneira. Por exemplo, para provar que

$$(A \cup B)^c = A^c \cap B^c$$

usamos a definição do complementar para escrever o lado esquerdo como

$$(A \cup B)^c = \{x \in \mathcal{U} \mid x \notin (A \cup B)\} = \{x \in \mathcal{U} \mid \neg(x \in (A \cup B))\}$$

e o lado direito como

$$A^c \cap B^c = \{x \in \mathcal{U} \mid x \notin A\} \cap \{x \in \mathcal{U} \mid x \notin B\}.$$

Recorrendo, agora, à definição da união de conjuntos, obtemos

$$(A \cup B)^c = \{x \in \mathcal{U} \mid \neg((x \in A) \vee (x \in B))\}.$$

Por outro lado,

$$A^c \cap B^c = \{x \in \mathcal{U} \mid (x \notin A) \wedge (x \notin B)\} = \{x \in \mathcal{U} \mid \neg(x \in A) \wedge \neg(x \in B)\}.$$

Logo, se

$$P(x) = (x \in A) \text{ e } Q(x) = (x \in B),$$

então,

$$(A \cup B)^c = \{x \in \mathcal{U} \mid \neg(P(x) \vee Q(x))\} \text{ e } A^c \cap B^c = \{x \in \mathcal{U} \mid \neg P(x) \wedge \neg Q(x)\}.$$

Portanto, para que $(A \cup B)^c = A^c \cap B^c$, basta que

$$\neg(P(x) \vee Q(x)) \equiv \neg P(x) \wedge \neg Q(x)$$

qualquer que seja o x escolhido em $\mathcal{U}$. Como esta igualdade já foi provada na tabela 6, podemos concluir que $(A \cup B)^c = A^c \cap B^c$, como desejávamos mostrar.

Antes de continuar, convém observar que, embora tenhamos definido, nos argumentos acima, as proposições que definem a união e a interseção de conjuntos em termos da pertinência aos conjuntos dados, nada nos impede de fazer exatamente o oposto. Suponhamos que $P(x)$ e $Q(x)$ são proposições acerca dos elementos de um conjunto universo $\mathcal{U}$. Se

$$A = \{x \in \mathcal{U} \mid P(x)\} \text{ e } B = \{x \in \mathcal{U} \mid Q(x)\}$$

então

$$A \cup B = \{x \in \mathcal{U} \mid P(x) \vee Q(x)\} \text{ e } A \cap B = \{x \in \mathcal{U} \mid P(x) \wedge Q(x)\}.$$

Portanto, a união e a interseção de dois conjuntos são definidas a partir da aplicação dos conectivos $\vee$ e $\wedge$ às propriedades que os definem.

Na seção 3 vimos que duas proposições são equivalentes quando admitem os mesmos valores verdade e que isto pode ser provado usando uma tabela verdade. Contudo, agora que conhecemos as propriedades básicas dos conectivos $\vee$, $\wedge$ e $\neg$, podemos usá-las para provar a equivalência de proposições de maneira mais direta. Por exemplo, como já provamos, na tabela 6, que

$$\neg(P \vee Q) \equiv \neg P \wedge \neg Q \tag{12}$$

podemos usar esta equivalência para provar a outra lei de De Morgan; isto é,

$$\neg(P \wedge Q) \equiv \neg P \vee \neg Q.$$

A primeira coisa a lembrar é que mostramos que (12) é valida *quaisquer que sejam as proposições P e Q*. Logo, a equivalência tem que continuar valendo se substituírmos P e Q por suas negativas. Fazendo isto, obtemos

$$\neg(\neg P \vee \neg Q) \equiv \neg(\neg P) \wedge \neg(\neg Q).$$

Aplicando, agora, a dupla negativa, encontramos

$$\neg(\neg P \vee \neg Q) \equiv P \wedge Q.$$

Negando os dois lados desta equivalência,

$$\neg(\neg(\neg P \vee \neg Q)) \equiv \neg(P \wedge Q)$$

e aplicando novamente a dupla negativa

$$\neg P \vee \neg Q \equiv \neg(P \wedge Q)$$

que é a equivalência que queríamos mostrar.

Para um outro exemplo, um pouco mais elaborado, provaremos que

$$(P \wedge \neg R) \vee ((P \vee Q) \wedge (\neg Q \vee P)) \equiv P$$

Começamos usando a comutatividade do $\vee$, que nos dá

$$(P \wedge \neg R) \vee ((P \vee Q) \wedge (\neg Q \vee P)) \equiv (P \wedge \neg R) \vee ((P \vee Q) \wedge (P \vee \neg Q)).$$

Aplicando, em seguida, a distributividade de $\vee$ sobre $\wedge$ para pôr P em evidência, temos que

$$(P \wedge \neg R) \vee ((P \vee Q) \wedge (P \vee \neg Q)) \equiv (P \wedge \neg R) \vee (P \vee (Q \wedge \neg Q)).$$

Mas, $Q \wedge \neg Q \equiv \perp$, que é o elemento neutro do $\vee$, donde

$$(P \wedge \neg R) \vee (P \vee (Q \wedge \neg Q)) \equiv (P \wedge \neg R) \vee P.$$

Usando a distributiva de $\vee$ sobre $\wedge$ e depois a idempotência do $\vee$, verificamos que

$$(P \wedge \neg R) \vee P \equiv (P \vee P) \wedge (\neg R \vee P) \equiv P \wedge (\neg R \vee P).$$

Neste ponto procedemos de maneira um pouco menos intuitiva. Como $\perp$ é o elemento neutro do $\vee$, podemos escrever

$$P \wedge (\neg R \vee P) \equiv (\perp \vee P) \wedge (\neg R \vee P);$$

que, pela comutativa e distributiva, nos dá

$$(\perp \vee P) \wedge (\neg R \vee P) \equiv P \vee (\perp \wedge \neg R).$$

Finalmente, por absorção,

$$(\perp \wedge \neg R) \vee P \equiv \perp \vee P$$

que é equivalente a P, pois $\perp$ é o elemento neutro do $\vee$.

5. Implicação

Nosso próximo conectivo é muito importante, mas mais problemático que todos os anteriores. Importante porque a construção *se P então Q* está por trás de quase todas as deduções em matemática; problemático, porque seu comportamento foi, na Grécia Antiga, fonte de discussão e divergência e, ainda hoje, causa estranheza a quem o encontra pela primeira vez.

Escrevendo no segundo século d.C., o filósofo cético Sextus Empiricus relata em seu tratado *Contra os lógicos* que

> todos os dialéticos declaram uma implicação [se P então Q] como sendo correta quando seu consequente [Q] segue de seu antecedente [P], mas sobre, em que circunstâncias, e como seguem um do outro eles se desentendem abertamente. Assim, segundo Fílon, a implicação não é verdadeira quando começa com algo verdadeiro e se encerra com algo falso. De modo que, segundo ele, a implicação é verdadeira em três casos e falsa em apenas um.[1]

[1] Contra os lógicos II.112-114.

Em outras palavras, segundo Fílon, a implicação *se P então Q* só é falsa quando P é verdadeira e Q é falsa. Contudo Diodoro Crônos, que fora professor de Fílon, discordava desta afirmação. Segundo ele uma implicação só poderia ser verdadeira quando o antecedente *nunca* possa levar a uma conclusão falsa.

A escolha de Fílon para os valores verdade da implicação aparece, de maneira implícita, nos *Anteriores analíticos* de Aristóteles. Um dos exemplos que Aristóteles propõe é o seguinte [2]:

> se um ser humano é uma pedra e uma pedra é um animal então um ser humano é um animal.

Para entender melhor o que está por trás do exemplo, digamos que H é o conjunto dos seres humanos, P é o conjunto das pedras e A é o conjunto dos animais. Então o exemplo de Aristóteles pode ser reformulado como

$$(H \subset P) \wedge (P \subset A) \Longrightarrow (H \subset A).$$

O ponto central é que, a partir de $H \subset P$ e $P \subset A$ podemos deduzir corretamente que $H \subset A$, que é uma afirmação verdadeira, muito embora $H \subset P$ e $P \subset A$ sejam ambas falsas. Isto sugere que é razoável admitir que, de uma afirmação falsa, seja possível deduzir uma afirmação verdadeira.

Ainda que a posição de Fílon tenha acabado por prevalecer, ela só se tornou padrão a partir dos trabalhos de G. Frege, no final do século XIX, e do *Principia Mathematica* de B. Russell e A. N. Whitehead, cujo primeiro volume foi publicado em 1910. Segundo Russel e Whitehead a implicação (Russell, Whitehead, 1910, p. 7)

> é uma função proposicional com dois argumentos p e q, e é a proposição de que não-p ou q é verdadeira, isto é, é a proposição $\sim p \vee q$. Assim, se p é verdadeira, então $\sim p$ é falsa e, consequentemente, a única alternativa deixada pela proposição $\sim p \vee q$ é que q seja verdadeira. Em outras palavras, se p e $\sim p \vee q$ são ambas verdadeiras, então q é verdadeira. Neste sentido nos referiremos à proposição $\sim p \vee q$ como p implica q. A ideia contida nesta função proposicional é tão importante que requer um simbolismo que, com simplicidade direta, represente a proposição como contendo p e q, sem a intervenção direta de $\sim p$. Mas "implica" como usado aqui não expressa nada além da conexão entre p e q também expressa pela disjunção "não-p ou q".

Denotando a implicação por $\Longrightarrow$, podemos dizer que, segundo Fílon, Frege, Russell e Whitehead,

$$(P \Longrightarrow Q) \equiv \neg P \vee Q, \tag{13}$$

[2] *Analíticos anteriores* 53b27

P	Q	$P \Longrightarrow Q$
V	V	V
V	F	F
F	V	V
F	F	V

TABELA 8. Tabela verdade de $P \Longrightarrow Q$.

cuja tabela verdade é dada na tabela 8.

Antes de prosseguir, faremos um exemplo em que provamos a equivalência de duas proposições que envolvem a implicação. Uma *tautologia* é uma proposição que é sempre verdade, quaisquer que sejam os valores verdade de suas proposições atômicas. Em outras palavras, uma tautologia é um proposição equivalente a $\top$. Nosso objetivo é provar

$$((P \wedge Q) \Longrightarrow R) \Longrightarrow (P \Longrightarrow (Q \Longrightarrow R))$$

é uma tautologia, usando as propriedades dos conectivos. Começamos substituindo a implicação por sua definição em termos de $\neg$ e $\vee$. Fazendo isto verificamos que a proposição dada é equivalente a

$$\neg(\neg(P \wedge Q) \vee R) \vee (\neg P \vee (\neg Q \vee R)).$$

Aplicando a lei de De Morgan à parcela da esquerda, obtemos

$$\neg((\neg P \vee \neg Q) \vee R) \vee (\neg P \vee (\neg Q \vee R)).$$

Mas, pela propriedade associativa, a parcela da direita é equivalente a

$$\neg((\neg P \vee \neg Q) \vee R) \vee ((\neg P \vee \neg Q) \vee R).$$

Finalmente, o resultado segue pelo princípio do meio excluído.

A definição de $P \Longrightarrow Q$ como $\neg P \vee Q$ nos permite determinar facilmente sua negativa, pois, pelas leis de De Morgan e a dupla negativa,

$$\neg(P \Longrightarrow Q) \equiv \neg(\neg P \vee Q) \equiv P \wedge \neg Q;$$

donde

$$\neg(P \Longrightarrow Q) \equiv P \wedge \neg Q.$$

Portanto, ao contrário do que poderia lhe parecer à primeira vista, *a negação de uma implicação não é obtida invertendo o sentido da seta*. Isto pode ser constatado na tabela verdade, comparando a terceira e quarta colunas na tabela 9 da página 42. A propósito, a implicação $Q \Longrightarrow P$ é a *recíproca* de $P \Longrightarrow Q$; ao passo que $\neg Q \Longrightarrow \neg P$ é sua contrapositiva. Esta última tem uma propriedade extremamente importante. Para obtê-la aplicamos a definição (13), que nos dá

$$(\neg Q \Longrightarrow \neg P) \equiv \neg(\neg Q) \vee \neg P;$$

donde, pela dupla negativa e comutatividade do $\vee$,

$$(\neg Q \Longrightarrow \neg P) \equiv Q \vee \neg P \equiv \neg P \vee Q.$$

Portanto,

$$(\neg Q \Longrightarrow \neg P) \equiv (P \Longrightarrow Q).$$

Do ponto de vista prático isto significa que *tanto faz provar uma implicação ou sua contrapositiva*, porque ambas têm exatamente os mesmos valores verdade. Teremos muitas oportunidades de aplicar a contrapositiva quando começarmos nosso estudo de funções, no próximo capítulo; mas é conveniente analisar um exemplo simples aqui, até porque ilustra várias das noções mencionadas nesta seção.

Seja n um número inteiro par. Por definição, isto significa que podemos escrevê-lo na forma $n = 2q$, em que q é o quociente da divisão de n por 2. Elevando ao quadrado os dois lados desta última equação, obtemos

$$n^2 = (2q)^2 = 4q^2.$$

Podemos, assim, concluir que

$$(n \text{ é par}) \Longrightarrow (n^2 \text{ é par}). \tag{14}$$

Será que a recíproca desta afirmação também é verdadeira? Isto é, será que

$$(n^2 \text{ é par}) \Longrightarrow (n \text{ é par}) \tag{15}$$

também vale? A primeira coisa a observar é que inverter o argumento usado para provar (14) não vai funcionar. Se tentássemos fazer isto, teríamos que começar supondo que n^2 é par. Em outras palavras, dividindo n^2 por 2 teríamos resto nulo, o que nos permitiria escrever $n^2 = 2s$, para algum inteiro positivo s. Até este ponto o argumento funciona: mas como sair de $n^2 = 2s$ para uma equação que envolve apenas n elevado à primeira potência? A resposta, naturalemente, é que basta extrair a raiz quadrada dos dois lados de $n^2 = 2s$. O problema é que, ao fazer isto, obtemos

$$n = \sqrt{2} \cdot \sqrt{s}.$$

O argumento empaca neste ponto, porque não sabemos nem se $\sqrt{2}$ é inteiro (não é!), nem se $\sqrt{s}$ é inteiro; o que não nos permite concluir nada sobre a paridade de n. A solução é lançar mão da contrapositiva de (15), que é

$$\neg(n \text{ é par}) \Longrightarrow \neg(n^2 \text{ é par}).$$

Como um número inteiro que não é par tem que ser ímpar, podemos reformular isto na forma

$$(n \text{ é ímpar}) \Longrightarrow (n^2 \text{ é ímpar}). \tag{16}$$

Mas, se n é ímpar, ao dividi-lo por 2 obtemos resto 1. Denotando por q o quociente desta divisão, isto nos permite escrever

$$n = 2q + 1.$$

Elevando os dois lados ao quadrado e arrumando a equação resultante,

$$n^2 = (2q + 1)^2 = 4q^2 + 4q + 1 = 2(2q^2 + 2q) + 1.$$

Portanto, quando n é ímpar o resto da divisão de n^2 por 2 é igual a 1, o que mostra que n^2 também é ímpar. Com isto, provamos que (16) é verdadeira. Como uma proposição tem o mesmo valor que sua contrapositiva, isto também mostra que se o quadrado de um número inteiro é par então o número tem que ser par. Formularemos isto como um teorema para que possamos usá-lo no futuro.

TEOREMA 5.1. *Seja n um número inteiro. Então, n é par se, e somente se, n^2 é par.*

Matemáticos e filósofos costumam identificar os papéis de P e Q na implicação $P \Longrightarrow Q$ em termos de condições suficientes e condições necessárias. Como basta que P seja verdadeira, para que Q também seja, dizemos que P é uma *condição suficiente* para que Q seja verdadeira. Por outro lado, Q é uma *condição necessária* para que P ocorra, porque Q sempre está sempre presente quando P é verdadeira. Infelizmente, na linguagem do dia-a-dia a diferença entre o que é suficiente e o que é necessário é muito pouco clara, o que tende a criar uma certa confusão. Os seguintes exemplos talvez ajudem um pouco a esclarecer a diferença entre os dois termos:

> é *necessário* que haja nuvens para que possa chover, porque sempre que chove as nuvens estão presentes, mas não é *suficiente*, porque nem sempre chove em um dia nublado.

Por outro lado,

> é *suficiente* que eu tenha uma caneta para que possa escrever, mas não é *necessário*, porque posso escrever com um lápis.

Assim como $\vee$ e $\wedge$, a implicação pode ser interpretada em termos de conjuntos. Se A e B são conjuntos, então,

$$(x \in A) \Longrightarrow (x \in B)$$

nos diz que todo elemento que está em A, também está em B, o que, por definição, equivale a afirmar que $A \subset B$. Como $A = B$ ocorre exatamente quando $A \subset B$ e $B \subset A$, podemos definir a igualdade de conjuntos pela proposição

$$((x \in A) \Longrightarrow (x \in B)) \wedge ((x \in B) \Longrightarrow (x \in A)).$$

Na verdade,

$$(P \Longrightarrow Q) \wedge (Q \Longrightarrow P)$$

ocorre com tanta frequência que tem um símbolo próprio, sendo denotado por $\iff$, que representa a fusão de uma seta apontando para a direita com uma seta apontando para a esquerda. Quando $P \iff Q$ dizemos que P *vale se, e somente se,* Q *vale*. A tabela verdade para $\iff$ pode ser obtida, facilmente, a partir das tabelas para $\implies$ e $\wedge$, como ilustrado na tabela 9. Como seria de esperar $P \iff Q$ só é verdadeira quando ambos P e Q têm os mesmos valores verdade.

P	Q	$P \implies Q$	$Q \implies P$	$(P \implies Q) \wedge (Q \implies P)$
V	V	V	V	V
V	F	F	V	F
F	V	V	F	F
F	F	V	V	V

TABELA 9. Tabela verdade de $P \iff Q \equiv (P \implies Q) \wedge (Q \implies P)$.

À primeira vista pode parecer que $P \implies Q$ é um conectivo secundário, porque pode ser expresso a partir de $\neg$ e $\vee$. Contudo, as leis de De Morgan também nos permitem expressar $\wedge$ em termos de $\neg$ e $\vee$, pois

$$P \wedge Q \equiv \neg(\neg P \vee \neg Q)$$

e $\vee$ em função de $\neg$ e $\wedge$, já que

$$P \vee Q \equiv \neg(\neg P \wedge \neg Q).$$

Assim, qualquer proposição pode ser escrita usando apenas $\vee$ e $\neg$ ou apenas $\wedge$ e $\neg$. Por exemplo, usando a equivalência (13), podemos escrever

$$(P \implies Q) \vee (P \wedge (\neg Q \implies R))$$

na forma

$$(\neg P \vee Q) \vee (P \wedge (\neg Q \vee R))$$

que, pelas leis de De Morgan equivale a

$$\neg((P \wedge \neg Q) \wedge \neg(P \wedge \neg(Q \wedge \neg R))).$$

Antes que você comece a achar que, embora $\vee$ e $\wedge$ sejam intercambiáveis, a negação é, de alguma forma, um conectivo fundamental. Note que, segundo a tabela 10,

$$\neg P \equiv (P \implies \perp).$$

Portanto, quais conectivos você deseja tomar como fundamentais é, em grande parte, uma questão de conveniência.

P	$\perp$	$P \Longrightarrow \perp$
V	F	F
F	F	V

TABELA 10. Tabela verdade de $P \Longrightarrow \perp$.

Exercícios

1. Quais das seguintes sentenças descrevem um conjunto?
   (a) Números dos CPFs de cada brasileiro;
   (b) palavras do dicionário, enumeradas em ordem alfabética;
   (c) números inteiros que são simultaneamente pares e ímpares.

2. Quais das afirmações abaixo são verdadeiras e quais são falsas?
   (a) 234 é um elemento do conjunto das placas de carros do Brasil;
   (b) $[a, e, i, o, u]$ é o conjunto das vogais;
   (c) a pertence ao conjunto das vogais;
   (d) $\{a, a, e, c, b, b\} = \{a, e, c, b\}$;
   (e) $\{c\}$ pertence ao conjunto das consoantes;
   (f) $\varnothing$ é subconjunto do conjunto das vogais;
   (g) $\varnothing \in \{\}$;
   (h) $\varnothing \in \{\varnothing\}$;
   (i) $\varnothing \subset \{\varnothing\}$.

3. Quais dos seguintes conjuntos são diferentes: $\varnothing, \{0\}, \{\varnothing\}$? Justifique sua resposta.

4. Dados dois subconjuntos A e B de um conjunto universo $\mathcal{U}$, se $a \in A$ e $b \in B$, definimos
$$(a, b) = \{\{a\}, \{a, b\}\}.$$
   Prove, usando a definição acima e a definição de igualdade de conjuntos, que $(a, b) = (c, d)$ só pode ocorrer quando $a = c$ e $b = d$.

5. Sejam A e B dois subconjuntos de um conjunto $\mathcal{U}$, ambos diferentes de $\varnothing$ e $\mathcal{U}$. Prove ou dê um exemplo para o qual a proposição é falsa:
   (a) quaisquer que sejam os subconjuntos X e Y de $\mathcal{U}$, se $A \cup X = A \cup Y$, então $X = Y$;
   (b) qualquer que seja o subconjunto $X \subset \mathcal{U}$, se $X \cap A = X \cap B$, então $A = B$.

6. Sejam P_1 e P_2 as proposições *dirigi acima de* 100 *Km/h* e *fui multado*, respectivamente. Escreva as seguintes afirmações como expressões nas proposições P_1 e P_2, usando os conectivos $\vee$, $\wedge$ e $\neg$:
   (a) não fui multado;
   (b) apesar de dirigir acima de 100 Km/h, não fui multado;

(c) não dirigi acima de 100 Km/h mas acabei sendo multado.

7. Quais das seguintes sentenças são proposições?
   (a) $2 > 3$;
   (b) Espera por mim.
   (c) Isaac Newton nasceu em Grantham.
   (d) Uma curva só tem tangente em um ponto p se seu gradiente em p não for nulo.
   (e) O enunciado desta questão é uma proposição.

8. Sejam P e Q proposições. A equivalência
$$(P \vee \bot) \wedge (Q \vee \top) \equiv P$$
é válida?

9. Nas sentenças abaixo, identifique as proposições atômicas, numere-as na forma $P_1, P_2, \ldots$ e escrever a sentença dada como combinação de suas proposições atômicas com conectivos:
   (a) Quando duas retas se cortam, os ângulos opostos pelo vértice são iguais.
   (b) Um triângulo tem seus três lados iguais se, e somente se, tem seus três ângulos internos iguais.
   (c) Dois triângulos de mesma base e altura, têm as mesmas áreas.

10. Use o método de tabela verdade para provar as seguintes equivalências:

$(\mathrm{a})\quad ((P_1 \Longrightarrow P_2) \wedge P_1) \Longrightarrow P_2 \equiv \top \qquad (\mathrm{b})(P_1 \wedge P_2) \Longrightarrow P_3 \equiv P_1 \Longrightarrow (P_2 \Longrightarrow P_3).$

11. Usando tabela-verdade, mostre que
   (a) se A é um subconjunto do conjunto vazio então $A = \varnothing$.
   (b) $(A \setminus B) \subset (A \cup B)$.

12. Usando as propriedades dos conectivos, mostre que:
$$((R \vee P) \wedge (\neg P \vee R)) \vee (R \wedge \neg Q) \equiv R$$
Indique em cada passo qual propriedade foi usada.

13. Use as propriedades dos conectivos para provar as equivalências abaixo. Você deve indicar qual a propriedade utilizada em cada etapa de sua demonstração.
   (a) $P \Longleftrightarrow Q \equiv (P \wedge Q) \vee (\neg P \wedge \neg Q)$;
   (b) $(P \wedge \neg R) \vee ((P \vee Q) \wedge (\neg Q \vee P)) \equiv P$.

14. Considere a proposição $P \Longrightarrow (Q \Longrightarrow R)$.
   (a) Reescreva esta proposição usando apenas os conectivos $\neg$, $\wedge$ e $\vee$.
   (b) Expresse a negativa da proposição dada usando apenas P, Q, R e $\neg P$, $\neg Q$ e $\neg R$.
   (c) Expresse a contrapositiva da proposição dada usando apenas de P, Q, R e $\neg P$, $\neg Q$ e $\neg R$.

15. Lembre-se que uma *tautologia* é um proposição equivalente a $\top$. Prove que as expressões

$$\neg(\neg P \vee (P \Longrightarrow Q)) \Longrightarrow \neg Q \text{ e } (P \vee \neg(Q \wedge R)) \vee (Q \Longrightarrow R)$$

são ambas tautologias, usando:
(a) uma tabela verdade;
(b) as propriedades dos conectivos.

16. Sejam P_1 e P_2 proposições. Use o método de tabela verdade para determinar quais das seguintes equivalências são verdadeiras.

$$(\text{a})\ (\neg P_1 \vee P_2) \wedge P_2 \equiv P_1; \qquad (\text{b})\ (\neg P_1 \vee P_2) \wedge P_1 \equiv P_2.$$

17. Sejam P_1, P_2 e P_3 proposições. Use as propriedades dos conectivos para provar as equivalências abaixo. Você deve indicar qual a propriedade utilizada em cada etapa de sua demonstração.
(a) $\neg(P_1 \vee (\neg P_1 \wedge P_2)) \equiv \neg P_1 \wedge \neg P_2$;
(b) $\neg(P_1 \vee P_2) \vee P_3 \equiv (\neg P_1 \vee P_3) \wedge (\neg P_2 \vee P_3)$;

18. Descreva a tabela verdade do "ou exclusivo", também conhecido como XOR, e denotado por $\oplus$, e use-a para provar que se P, Q e R são proposições, então
(a) $P \oplus Q \equiv Q \oplus P$;
(b) $P \oplus (Q \oplus R) \equiv (P \oplus Q) \oplus R$;
(c) $P \oplus \perp \equiv P$;
(d) $P \oplus P \equiv \perp$.

19. Sejam A e B subconjuntos de $\mathcal{U}$. Mostre que se $A \subset B$, então $B^c \subset A^c$.

20. Se A e B são subconjuntos de um universo $\mathcal{U}$, definimos sua *diferença*,

$$A \setminus B = \{x \in \mathcal{U} \mid (x \in A) \wedge (x \notin B)\}.$$

Sabendo-se que C também é um subconjunto de $\mathcal{U}$, prove as seguintes propriedades da diferença de conjuntos:
(a) $A \setminus B = A \cap B^c$;
(b) $A \setminus (B \cup C) = (A \setminus B) \cap (A \setminus C)$;
(c) $A \setminus (B \cap C) = (A \setminus B) \cup (A \setminus C)$;
(d) $A \setminus \varnothing = A$;
(e) $\varnothing \setminus A = \varnothing$;
(f) $A \setminus (B \setminus C) = (A \cap C) \cup (A \setminus B)$.

21. A *diferença simétrica* dos subconjuntos A e B de um conjunto $\mathcal{U}$ é o conjunto

$$A \Delta B = (A \setminus B) \cup (B \setminus A)$$

(a) Mostre que

$$A \Delta B = \{x \in \mathcal{U} \mid (x \in A) \oplus (x \in B)\}$$

(b) Mostre que a diferença simétrica é associativa e comutativa.
(c) Determine o elemento neutro da diferença simétrica.
(d) Mostre que $A \Delta A = \varnothing$.
(e) Determine o subconjunto X de $\mathcal{U}$ que satisfaz a igualdade $A \Delta X = B$.

Dica: use a associatividade do $\oplus$ ao provar que a diferença simétrica é associativa.

CAPÍTULO 3

Números reais

Neste capítulo começamos a aplicar o que aprendemos sobre lógica à demonstração de resultados matemáticos. Nas duas primeiras seções trataremos do processo de dedução de maneira mais ou menos geral; nas seções seguintes aplicaremos o que aprendemos para provar alguns resultados sobre números reais que serão utilizados nos três capítulos finais.

1. Quantificadores

Os conceitos de lógica que aprendemos até aqui nos permitem expressar simbolicamente muitas proposições matemáticas ou não, mas não é claro como deveríamos utilizá-los para expressar muitas outras. Suponhamos, por exemplo, que nosso universo $\mathcal{U}$ seja um suconjunto de $\mathbb{N}$. À primeira vista, não é claro que seja possível usar $\neg$, $\vee$ e $\wedge$ para expressar as seguintes proposições como:

A1: todos os números em $\mathcal{U}$ são menores que 10;
A2: há um número par no conjunto $\mathcal{U}$.

Naturalmente, o problema está nas palavras como *há* e *todos*. Embora nunca tenham aparecido em nenhuma das proposições que modelamos até aqui, tais palavras ocorrem frequentemente em matemática.

Como primeiro passo para expressar estas proposições, introduzimos duas as seguintes propriedades, referentes aos elementos do conjunto $\mathcal{U}$:

$$P_1(x) = x < 10;$$
$$P_2(x) = x \text{ é par.}$$

Com isso, podemos reformular as proposições acima como

A1: todos os $x \in \mathcal{U}$ satisfazem $P_1(x)$;
A2: existem $x \in \mathcal{U}$ que satisfazem $P_2(x)$.

Tivemos o cuidado de nos referir a $P_1(x)$ e $P_2(x)$ como *propriedades*, porque nenhuma das duas é uma proposição. De fato, a não ser que escolhamos x como sendo um elemento específico de $\mathcal{U}$, não temos como determinar se estas propriedades são verdadeiras ou falsas..

Começaremos tratando do caso, radicalmente simples, no qual

$$\mathcal{U} = \{1, 2\}.$$

Neste caso, **A1** será verdadeira exatamente quando $P_1(1)$ e $P_1(2)$ forem ambas verdadeiras. A solução, naturalmente, está em escrever

$$P_1(1) \wedge P_1(2).$$

À primeira vista, **A2** representa um problema mais complicado. Contudo, analisando-a mais de perto, verificamos que só pode haver um número par em $\mathcal{U}$ se 1 for par ou se 2 for par; o que equivale a escrever

$$P_2(1) \vee P_2(2).$$

No final das contas, a solução acabou sendo simples, mas pode lhe parecer extremamente suspeita, porque escolhemos um universo com apenas dois elementos. O que faríamos para escrever as mesmas proposições se

$$\mathcal{U} = \{1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12\}?$$

Como $\wedge$ só é verdadeira se cada uma das proposições que conecta são verdadeiras, teríamos que expressar **A1** na forma

$$P_1(1) \wedge P_1(2) \wedge P_1(3) \wedge P_1(4) \wedge P_1(5) \wedge P_1(6) \wedge P_1(7) \wedge P_1(8) \wedge P_1(9) \wedge P_1(10) \wedge P_1(12).$$

Analogamente, **A2** pode ser escrita na forma

$$P_2(1) \vee P_2(2) \vee P_2(3) \vee P_2(4) \vee P_2(5) \vee P_2(6) \vee P_2(7) \vee P_2(8) \vee P_2(9) \vee P_2(10) \vee P_2(12),$$

pois basta que uma das proposições ligadas por $\vee$ seja verdadeira para que a proposição resultante também o seja.

Embora as expressões de **A1** e **A2** tenham-se tornado inconvenientes quando $\mathcal{U}$ tem doze elementos, não há nenhum obstáculo conceitual que nos impeça de escrevê-las mesmo que o universo tenha mil ou, até mesmo, infinitos elementos. Contudo, neste casos, seria necessário recorrer a abreviações. Por exemplo, no caso em que $\mathcal{U} = \mathbb{N}$, as mesmas proposições seriam escritas como

$$\bigwedge_{x \in \mathbb{N}} P_1(x) \quad \text{e} \quad \bigvee_{x \in \mathbb{N}} P_2(x).$$

Esta é notação é adotada em alguns livros, como (Grzegorczyk, 1974, pp. 118-119), mas tem alguns incovenientes.

Além de ser utilizada em poucos livros, ela esconde o significado real das proposições. Por exemplo, para ler

$$\bigvee_{x \in \mathbb{N}} P_2(x)$$

como *existe um natural que satisfaz* $P_2(x)$, você precisa decodificar um *ou* como significando *existe*, que não é algo que fazemos naturalmente.

Por causa disto, foram inventados símbolos especiais para denotar *existe* e *para todo*. O primeiro a aparecer foi $\exists$, introduzido pelo matemático em 1897 pelo matemático G. Peano. O E maiúsculo invertido foi escolhido por lembrar a palavra italiana *esiste* (português *existe*). Por sua vez, o A maiúsculo invertido, usado para denotar *para todo*, foi inventado por G. K. E. Gentzen em 1935 por lembrar o alemão *alle* (inglês *all*). Usando esta notação, escreveremos

$$\forall x \in \mathbb{N}, (P_1(x)) \quad \text{em vez de} \quad \bigwedge_{x \in \mathbb{N}} P_1(x)$$

e também

$$\exists x \in \mathbb{N}, (P_2(x)) \quad \text{em vez de} \quad \bigvee_{x \in \mathbb{N}}, (P_2(x)).$$

Apesar da mudança na notação ser muito pequena, a facilidade em interpretar o significado das proposições é considerável

Os operadores $\exists$ e $\forall$ são conhecidos como *quantificadores* porque porque especificam a quantidade de elementos de um conjunto $\mathcal{U}$ que têm uma certa propriedade. Contudo, a maneira como um quantificador "especifica a quantidade de elementos" é bastante vaga: ele apenas indica se há *algum* elemento do conjunto que satisfaz uma propriedade, ou se isto vale para *todos* os elementos do conjunto; o primeiro é chamado de *quantificador existencial* e o segundo de *quantificador universal*.

É hora de resumir o que fizemos até aqui, em termos da notação que acabamos de introduzir. Sejam $\mathcal{U}$ o conjunto que fará o papel de universo e $P(x)$ é uma propriedade que se aplica aos elementos $x \in \mathcal{U}$. Escreveremos

$$\forall x \in \mathcal{U}, (P(x)), \tag{17}$$

para indicar que $P(x)$ é verdadeiro para *todos* os elementos $x \in \mathcal{U}$ e

$$\exists x \in X, (P(x)), \tag{18}$$

para indicar que existe, pelo menos, um elemento $x \in \mathcal{U}$ para o qual $P(x)$ é verdadeira. Ainda precisamos investigar como determinar a negativa de uma proposição que envolve um quantificador. Antes, porém, ilustraremos o uso de quantificadores através de algumas proposições menos ingênuas que as que investigamos até aqui.

Como nosso primeiro exemplo, considere a proposição *a união de conjuntos admite um elemento neutro*. Neste enunciado, *admite* faz o papel de existe, de modo que podemos reformular a proposição na forma *existe um elemento neutro para a união de conjuntos*. Tendo identificado o quantificador a ser usado, sabemos que a proposição deverá seguir o modelo em (18). O próximo passo consiste em identificar o conjunto universo. Como se trata de uma operação entre conjuntos, os elementos de $\mathcal{U}$ terão que ser conjuntos. Por isso, escolheremos nosso universo como sendo $\mathcal{P}(X)$, o conjunto das partes de um conjunto X.

Com isso, a proposição que queremos modelar afirma que *existe um subconjunto especial de* X que funciona como elemento neutro da união de conjuntos.

Em seguida, precisamos determinar a proposição $P(X)$ que está sob o quantificador. Denotando por V o conjunto que fará o papel de elemento neutro, $P(X)$ afirmará que a igualdade

$$A \cup V = A,$$

é válida, qualquer que seja o subconjunto A de X. Com isso, nos vemos às voltas com a necessidade de usar um segundo quantificador, uma vez que a maneira como formulamos $P(X)$ faz surgir a palavra *qualquer*, indicando a necessidade de usar o quantificador universal. Seguindo o modelo em (17), e levando em conta que A é subconjunto de X, a proposição $P(X)$ terá a forma

$$\forall A \in \mathcal{P}(X), (A \cup V = A).$$

Juntando isto ao que havíamos feito no parágrafo anterior, nossa tradução de *a união de conjuntos admite um elemento neutro* será

$$\exists\, V \in \mathcal{P}(X), (\forall A \in \mathcal{P}(X), (A \cup V = A)). \tag{19}$$

Naturalmente sabemos que $V = \varnothing$, mas não seria conveniente adotar esta notação diretamente na formulação acima. Afinal, o conjunto vazio é um conjunto específico de $\mathcal{P}(X)$ e já sabemos que ele *existe*; teríamos, apenas, que verificar que $A \cup \varnothing = A$, para sabar que é o elemento neutro da união. Mas isto não requer nenhum quantificador existencial, apenas um quantificador universal.

Um detalhe importante no uso de quantificadores é que é fundamental considerar a ordem em que eles aparecem na proposição. Assim, (19) afirma que existe um elemento $V \in \mathcal{P}(X)$ com a propriedade de que $A \cup V = A$ é válido para qualquer $A \in \mathcal{P}(X)$. Porém, se invertermos a ordem dos quantificadores, obtemos

$$\forall A \in \mathcal{P}(X), (\exists V \in \mathcal{P}(X), (A \cup V = A)). \tag{20}$$

A diferença é que, em princípio esta segunda formulação não especifica que *o mesmo* V tem que satisfazer $A \cup V = A$ para qualquer $A \in \mathcal{P}(X)$. Assim, escolher $V = A$ é compatível com (20), mas não com (19).

Outro exemplo interessante diz respeito à afirmação de que *a união de dois subconjuntos de* $\mathcal{P}(X)$ *também é um subconjunto de* $\mathcal{P}(X)$. Da maneira como enunciamos esta proposição, não há nenhum quantificador à vista. Mas, naturalmente, o que está sendo dito é que a afirmação é válida *quaisquer que sejam os subconjuntos de* X *que escolhermos*. Assim, o que esta proposição afirma deve ser traduzido como

$$\forall A \in \mathcal{P}(X), (\forall B \in \mathcal{P}(X), (A \cup B \in \mathcal{P}(X))). \tag{21}$$

Para facilitar a interpretação das proposições em que um mesmo quantificador é aplicado a mais de uma variável de um mesmo conjunto, vamos escrevê-lo apenas uma vez, de modo que

$$(\forall A, B \in \mathcal{P}(X)), (A \cup B \in \mathcal{P}(X))$$

é uma abreviação para (21). Fica por sua conta escrever as demais propriedades das operações de conjuntos, seguindo o mesmo modelo.

Vejamos um exemplo mais elaborado. O quinto postulado de Euclides, que já encontramos no capítulo 1, afirma que

> para toda reta r e todo ponto p, fora de r, existe uma reta r', paralela a r, que contém p.

Para escrever esta proposição usando quantificadores é precisamos introduzir alguns símbolos adicionais. Chamaremos de $\mathbb{P}$ e $\mathbb{L}$ os conjuntos dos pontos e das retas do plano. Denotanto por $r \parallel r'$ o fato de r ser paralela a r', o quinto postulado pode ser escrito na forma

$$\forall r \in \mathbb{L}, (\forall p \in \mathbb{P}, (p \notin r \Longrightarrow (\exists r' \in \mathbb{L}, ((p \in r') \wedge (r' \parallel r))))). \tag{22}$$

Veremos muitos outros exemplos deste gênero quando estudarmos números reais nas seções 3 e 4 deste capítulo. Como toda técnica, formular uma proposição usando quantificadores requer prática. Você encontrará vários exercícios ao final deste capítulo que lhe ajudarão a obtê-la.

Encerraremos o capítulo investigando como proceder para negar uma proposição em que aparecem quantificadores. Lembre-se que o conectivo $\neg$ se caracteriza por inverter o valor verdade de uma proposição. Assim, nosso ponto de partida deve ser a pergunta: sob que condições $\forall x \in X, (P(x))$ é falsa? Por exemplo, se $\mathcal{B}$ é um conjunto de bolas coloridas e $A(x)$ representa a sentença x *é azul*, então $\forall x \in \mathcal{B}, (A(x))$ equivale a dizer que *todas as bolas do conjunto* $\mathcal{B}$ *são azuis*. Portanto, basta que haja, em $\mathcal{B}$, uma bola que não seja azul, para que $\forall x \in \mathcal{B}, A(x)$ seja falsa. Logo,

$$\neg \forall x \in \mathcal{B}, A(x) \equiv \exists x \in \mathcal{B}, (\neg A(x)).$$

Reciprocamente, se *não existe* uma bola azul em $\mathcal{B}$ isto ocorre porque todas as bolas de $\mathcal{B}$ não são azuis ou, como seria mais natural dizer em português, nenhuma bola de $\mathcal{B}$ é azul. Em símbolos,

$$\neg \exists x \in \mathcal{B}, A(x) \equiv \forall x \in \mathcal{B}, (\neg A(x)).$$

Em geral, se $\mathcal{U}$ é um conjunto e $P(x)$ é uma proposição relativa aos elementos de $\mathcal{U}$, teremos que

$$\neg(\forall x \in \mathcal{U}, P(x)) \equiv \exists x \in \mathcal{U}, (\neg P(x)). \tag{23}$$

para a negativa do quantificador universal e

$$\neg(\exists x \in \mathcal{U}, P(x)) \equiv \forall x \in \mathcal{U}, (\neg P(x)), \tag{24}$$

para a negativa do quantificador existencial.

Optamos por identificar as negativas dos quantificadores partindo de exemplos concretos, mas poderíamos ter feito o mesmo usando as definições dos quantificadores $\forall$ e $\exists$ em termos dos conectivos $\wedge$ e $\vee$. Assim, como

$$\forall x \in \mathcal{U}, (P(x)) \equiv \bigwedge_{x \in \mathcal{U}} (P(x)),$$

podemos usar a lei de Morgan, obtendo

$$\neg(\forall x \in \mathcal{U}, (P(x))) \equiv \bigvee_{x \in \mathcal{U}} \neg P(x).$$

Mas, esta última expressão também pode ser traduzida por

$$\exists x \in \mathcal{U}, (P(x)),$$

o que nos leva de volta a (23). Analogamente,

$$\neg(\exists x \in \mathcal{U}, (P(x))) \equiv \neg \bigvee_{x \in \mathcal{U}} P(x).$$

que, pela lei de de Morgan, é o mesmo que

$$\bigwedge_{x \in \mathcal{U}} \neg P(x) \equiv \forall x \in \mathcal{U}, (P(x)),$$

de modo que chegamos a (24) por outro caminho.

Aplicando a negativa de $\exists$ ao exemplo do início desta seção, vemos que

$$\neg(\exists\, V \in \mathcal{P}(X), (\forall A \in \mathcal{P}(X), (A \cup V = A)))$$

pode ser escrita como

$$\forall\, V \in \mathcal{P}(X), \neg(\forall A \in \mathcal{P}(X), (A \cup V = A)).$$

Usando, agora, a negativa do $\forall$, obtemos

$$\forall\, V \in \mathcal{P}(X), (\exists A \in \mathcal{P}(X), \neg(A \cup V = A)).$$

Esta proposição, que também podemos escrever como

$$\forall\, V \in \mathcal{P}(X), (\exists A \in \mathcal{P}(X), (A \cup V \neq A)),$$

é a negativa desejada. Outro exemplo simples é

$$\neg(\forall A \in \mathcal{P}(X)(\forall B \in \mathcal{P}(X), (A \cup B \in \mathcal{P}(X)))).$$

Negando o primeiro $\forall$, encontramos

$$\exists A \in \mathcal{P}(X), \neg(\forall B \in \mathcal{P}(X), (A \cup B \in \mathcal{P}(X)))$$

e, negando o segundo,

$$\exists A \in \mathcal{P}(X), (\exists B \in \mathcal{P}(X), \neg(A \cup B \in \mathcal{P}(X))),$$

que é equivalente a

$$\exists A \in \mathcal{P}(X), (\exists B \in \mathcal{P}(X), (A \cup B \notin \mathcal{P}(X))).$$

O quinto postulado de Euclides oferece um exemplo mais elaborado e mais interessante.

Como vimos na seção 5 do capítulo 1, existem geometrias que não obedecem, não à proposição que modelamos em (22), mas à sua negativa. Para identificar exatamente o que isto significa, aplicaremos o que aprendemos sobre a negativa dos quantificadores para determinar

$$\neg(\forall r \in \mathbb{L}, (\forall p \in \mathbb{P}, (p \notin r \implies (\exists r' \in \mathbb{L}, ((p \in r') \wedge (r' \parallel r)))))).$$

Usando (23) uma vez,

$$\exists r \in \mathbb{L}, \neg(\forall p \in \mathbb{P}, (p \notin r \implies (\exists r' \in \mathbb{L}, ((p \in r') \wedge (r' \parallel r)))))$$

e mais uma

$$\exists r \in \mathbb{L}, (\exists p \in \mathbb{P}, \neg(p \notin r \implies (\exists r' \in \mathbb{L}, ((p \in r') \wedge (r' \parallel r))))).$$

Traduzindo a implicação em termos de $\vee$ e $\neg$,

$$\exists r \in \mathbb{L}, (\exists p \in \mathbb{P}, \neg(\neg(p \notin r) \vee (\exists r' \in \mathbb{L}, ((p \in r') \wedge (r' \parallel r))))).$$

Usando as leis de De Morgan e a dupla negação,

$$\exists r \in \mathbb{L}, (\exists p \in \mathbb{P}, ((p \notin r) \wedge \neg(\exists r' \in \mathbb{L}, ((p \in r') \wedge (r' \parallel r))))).$$

Finalmente, por (24) e De Morgan,

$$\exists r \in \mathbb{L}, (\exists p \in \mathbb{P}, ((p \notin r) \wedge (\forall r' \in \mathbb{L}, (\neg(p \in r') \vee \neg(r' \parallel r)))));$$

que podemos reescrever na forma

$$\exists r \in \mathbb{L}, (\exists p \in \mathbb{P}, ((p \notin r) \wedge (\forall r' \in \mathbb{L}, ((p \in r') \implies \neg(r' \parallel r)))));$$

Portanto, em geometria não euclidiana,

> existem uma reta r e um ponto $p \notin r$, de maneira que qualquer reta r' que contém p não é paralela a r.

Em nosso estudo de geometria, no capítulo 5, encontraremos uma versão do quinto postulado, mais forte do que a que modelamos (22).

2. O jogo dedutivo

Como vimos na seção 4 do capítulo 1, a primeira tentativa de sistematizar as regras do raciocínio humano foi feita por Aristóteles (384–322 a.C.) nos *Analíticos anteriores*, obra que se inicia com o seguinte parágrafo,

> [c]omeçamos dizendo de que trata nossa investigação, que é sobre a demonstração e o conhecimento demonstrativo. Em seguida, determinaremos o que são uma premissa, um termo e um silogismo.

Um pouco mais adiante, Aristóteles define a *premissa* como uma "sentença em que se afirma ou nega algo sobre alguma coisa", o *termo* como "aquilo em que uma premissa pode ser analisada, como sujeito e predicado" e o *silogismo* como "uma forma de discurso em que, algo sendo afirmado, algo diferente do que foi dito segue-se necessariamente da afirmação feita anteriormente". O exemplo de silogismo que se tornou padrão é:

todos os homens são mortais,

Sócrates é um homem,

logo Sócrates é mortal.

Denotando por $\mathcal{U}$ o conjunto de todos os seres, reais ou imaginários, por H o conjunto dos homens e por M o conjunto daqueles seres que são mortais, podemos reescrever o silogismo na forma

$$((\forall x \in \mathcal{U}, (x \in H \Longrightarrow x \in M)) \wedge (\text{Sócrates} \in H)) \Longrightarrow \text{Sócrates} \in M.$$

Para Aristóteles, o silogismo é o instrumento por excelência do pensamento dedutivo e teremos muitas oportunidades de utilizá-lo ao longo deste livro. Apesar de sua inegável importância, os silogismos não são suficientes, nem mesmo para expressar completamente os detalhes de uma demonstração dos *Elementos* de Euclides.

Um dos métodos de demonstração discutidos por Aristóteles é o que ele chama de *demonstração indireta*, mais conhecido hoje em dia como *redução ao absurdo* ou *demonstração por contradição*. De acordo com Aristóteles, [1]

> todos que fazem uma demonstração por absurdo provam uma conclusão falsa usando silogismos, mostrando a afirmação original a partir das hipóteses, quando algo impossível resulta da afirmação oposta.

Para entender o que Aristóteles quis dizer, suponhamos que $H_1, \ldots, H_k$ são certas hipóteses, a partir das quais queremos provar uma proposição P. Em outras palavras, queremos mostrar que

$$(H_1 \wedge \cdots \wedge H_k) \Longrightarrow P$$

é uma proposição verdadeira. O método de demonstração por contradição consiste em provar que

$$(H_1 \wedge \cdots \wedge H_k \wedge \neg P) \Longrightarrow \perp;$$

isto é, adicionando a negativa $\neg P$ da *conclusão* às hipóteses, almejamos obter uma contradição. Como consequência da tabela verdade da implicação, isto só pode ocorrer quando

$$(H_1 \wedge \cdots \wedge H_k \wedge \neg P) \equiv \perp .$$

[1] Analíticos nteriores 41a 24–27

Contudo, como estamos supondo que as hipóteses $H_1, \ldots, H_k$ são verdadeiras, a tabela verdade do $\wedge$, nos permite, então, concluir que $\neg P$ é falsa. Mas, pelo princípio do meio excluído,

$$P \vee \neg P \equiv \top;$$

o que significa que P tem que ser verdadeira.

As demonstrações por contradição são mais naturais do que a análise acima pode lhe fazer crer. Um exemplo mais satisfatório, e muito mais importante, é a demonstração, já mencionada na seção 5 do capítulo 1, de que o lado e a diagonal de um quadrado são incomensuráveis. Em outras palavras, não existe nenhuma unidade de medida relativamente à qual o lado e a diagonal do quadrado tem comprimentos inteiros. Supondo que o quadrado tem lado igual a 1, temos pelo Teorema de Pitágoras, que sua diagonal terá comprimento igual a $\sqrt{2}$. Portanto, o que precisamos mostrar se reduz ao conteúdo do seguinte teorema.

TEOREMA 2.1. *O número $\sqrt{2}$ não é uma fração.*

DEMONSTRAÇÃO. Suponhamos, por contradição, que $\sqrt{2}$ seja um número racional. Por definição, isto significa que existem inteiros m e $n \neq 0$, tais que

$$\sqrt{2} = \frac{m}{n}.$$

Eliminando os termos comuns entre o numerador e o denominador, podemos supor que m e n *não* são números ambos pares. Elevando os dois lados da equação anterior ao quadrado e multiplicando a equação resultante por n^2, obtemos

$$2n^2 = m^2. \tag{25}$$

Mas, pelo teorema 5.1 da página 41, isto só pode ocorrer se m for par. Logo, existe algum inteiro k de modo que $m = 2k$. Substituindo em (25),

$$2n^2 = (2k)^2 = 4k^2.$$

Dividindo os dois lados por 2, obtemos

$$n^2 = 2k^2.$$

Logo, usando novamente o teorema 5.1, podemos afirmar que n tem que ser par, contrariando nossa afirmação de que m e n não podem ser ambos pares. Portanto, $\sqrt{2}$ não pode ser uma fração, logo é um número irracional. □

Esta demonstração é mencionada, pela primeira vez, nos *Analíticos anteriores* de Aristóteles, onde ele afirma que *o lado e a diagonal de um quadrado não são comensuráveis porque um número natural não pode ser, simultaneamente, par e ímpar*. Tendo lido a demonstração, você reconhecerá que o comentário de Aristóteles, apesar de extremamente enigmático, vai direto ao ponto crucial da demonstração. Ao longo deste livro teremos oportunidades de estudar muitas demonstrações por contradição, a começar pela seção 3.

Nosso ponto de partida serão, mais uma vez, as obras de Aristóteles sobre lógica. Nos *Analíticos posteriores*[2] Aristóteles define o que poderíamos chamar de conhecimento científico da seguinte maneira

> consideramos que conhecemos de maneira absoluta [ou científica], e não à maneira acidental dos sofistas, quando consideramos que conhecemos a causa do fato, que aquela é de fato sua causa, e que não é possível que fosse de outra maneira.

Os sofistas, mencionados neste texto, eram professores itinerantes que se propunham a ensinar um pouco de tudo, preparando os filhos das famílias ricas para a vida pública nas cidades gregas. O compromisso de muitos sofistas com o treinamento para o sucesso público, em detrimento da verdade, deu à toda a categoria uma fama ruim, refletida em português na palavra sofisma, que significa um "argumento concebido a fim de produzir uma ilusão de verdade".

O ponto central da definição de Aristóteles é que o conhecimento científico de um fato requer que conheçamos sua causa. Para os gregos, o exemplo por excelência de conhecimento científico era a matemática. Não é a toa que a própria palavra matemática seja derivada do verbo grego que significa aprender, estudar. A sistematização da geometria, que já havia começado no tempo de Aristóteles, transformou-a em uma ciência dedutiva, de modo que um resultado matemático só poderia ser aceito quando provado a partir de outros, suas premissas, que desempenham, em matemática, o mesmo papel que as causas nas ciências físicas. Nos *Posteriores analíticos*[3] Aristóteles esclarece que essas premissas devem ser verdadeiras, primárias e imediatas, além de melhor conhecidas do que a conclusão à que queremos chegar. Em outras palavras, para que o conhecimento de um fato possa ser considerado científico, é necessário que possa ser *provado a partir de premissas mais bem conhecidas que o fato que queremos afirmar*.

Tomando por base o que aprendemos com Aristóteles, podemos formular o método axiomático como um jogo. As regras deste jogo são dadas pelo comportamento dos conectivos lógicos e pelo uso correto dos quantificadores. Quando você começa o jogo, os únicos fatos que conhece são os axiomas e tudo o que provar tem que tomá-los como ponto de partida. Mas cada novo teorema que prova provê um novo fato, a partir do qual pode provar novos resultados. Forma-se, assim, uma rede interligada que, partindo dos axiomas, avança cada vez mais longe.

Parte do problema de pôr este programa em prática reside na dificuldade em se certificar que apenas os axiomas, e resultados já provados a partir deles, estão sendo usados nas demonstrações. Em outras palavras, em ter certeza que não estamos violando nenhuma das regras do jogo. Naturalmente o risco é tanto maior, quanto mais familiar é o assunto, como

[2] II.71b09

[3] I.II 71b.20

é o caso da geometria e das propriedades básicas dos números reais, que consideraremos nas próximas seções.

Quando você lê sobre um resultado, mas não sabe prová-lo a partir dos axiomas, está na mesma situação de uma pessoa que foi levada, de olhos vendados, de um ponto conhecido da cidade, a outro onde nunca esteve. Se for abandonado lá, estará perdido, sem saber o caminho de volta. Já alguém que lê a demonstração de um dado resultado, e sabe como relacioná-lo a outros, que já lhe eram conhecidos (e, portanto, em última análise, aos axiomas) é como alguém que chega, de olhos abertos, a um determinado local da cidade que ainda não conhecia, de modo que sabe retornar com segurança ao ponto de partida.

Resumindo, as regras do jogo dedutivo são as seguintes:

1. o ponto de partida são os axiomas, ou outras propriedades tomadas como fundamentais;
2. as deduções devem utilizar os conectivos lógicos e os quantificadores, e respeitar suas propriedades;
3. os únicos fatos admissíveis em um argumento são os axiomas (ou outras propriedades fundamentais) mencionados acima, ou fatos que já foram corretamente provados a partir deles.

Nas próximas seções utilizaremos o jogo dedutivo para analisar as propriedades dos números reais. Você verá, então, que muitas das propriedades que aprendeu sobre estes números, introduzidas como fatos isolados, são, na verdade, parte de uma densa rede de fatos interligados que têm sua origem em alguns poucos axiomas.

3. Os números reais: álgebra

Nas seções anteriores apresentamos, sucintamente, o jogo dedutivo, que é, por excelência, o método de exposição empregado na matemática desde o início do século XX. Nesta seção começaremos a usá-lo para investigar os números reais de maneira mais ou menos sistemática. Tomaremos como axiomas algumas das propriedades fundamentais destes números. Contudo, nossa apresentação deixa bastante a desejar no que diz respeito aos critérios estabelecidos por Aristóteles para identificar bons axiomas. Discutiremos esta questão em mais detalhe na seção 5.

AXIOMAS ALGÉBRICOS. *O conjunto $\mathbb{R}$, dos número reais, admite duas operações, denotadas por + (adição) e $\cdot$ (multiplicação), que satisfazem os seguintes axiomas, que valem quaisquer que sejam os números reais a, b e c:*

Para simplificar a notação, frequentemente escreveremos ab em vez de $a \cdot b$, para denotar a multiplicação entre a e b.

Axioma		
Associatividade	$\forall a, b \in \mathbb{R}, (a + (b + c) = (a + b) + c)$	$\forall a, b \in \mathbb{R}, (a \cdot (b \cdot c) = (a \cdot b) \cdot c)$
Comutatividade	$\forall a, b \in \mathbb{R}, (a + b = b + a)$	$\forall a, b \in \mathbb{R}, (a \cdot b = b \cdot a)$
Elemento neutro	$\exists 0 \in \mathbb{R}, (\forall a \in \mathbb{R}, (a + 0 = a))$	$\exists 1 \in \mathbb{R}, (\forall a \in \mathbb{R}, (a \cdot 1 = a))$
Simétrico	$\forall a \in \mathbb{R}, (\exists -a \in \mathbb{R}, (a + (-a) = 0))$	
Inverso	$\forall a \in \mathbb{R} \setminus \{0\}, (\exists a^{-1} \in \mathbb{R}, (a \cdot a^{-1} = 1))$	
Distributiva	$\forall a, b \in \mathbb{R}, (a \cdot (b + c) = a \cdot b + a \cdot c)$	

Vamos começar mostrando que o conjunto dos números reais não pode ter mais do que um elemento neutro para a adição. Para isso, imagine que duas pessoas diferentes encontraram números reais 0 e $0'$ como elementos neutros da adição em $\mathbb{R}$. Como estamos supondo que 0 é um elemento neutro para a adição, teremos, necessariamente, que

$$0' + 0 = 0'.$$

Porém, como a adição é comutativa,

$$0' + 0 = 0 + 0'.$$

Contudo, estamos supondo que $0'$ também é um elemento neutro para a adição em $\mathbb{R}$, de modo que

$$0 + 0' = 0.$$

Logo, $0' = 0$, confirmando nossa afirmação original de que só pode haver um elemento neutro para a adição de números reais. Um argumento análogo mostra que a multiplicação de números reais só pode ter um elemento neutro.

Tendo provado a unicidade dos elementos neutros da adição e multiplicação de números reais, precisamos, em seguida, mostrar que cada número real tem um único simétrico. Para provar isto, digamos que duas pessoas diferentes encontraram a_1 e a_2 como simétricos de um dado número real a. Pela definição de simétrico,

$$a + a_1 = 0 \text{ e } a + a_2 = 0.$$

Assim,

$$a_1 = a_1 + 0 = a_1 + (a + a_2).$$

Logo, pela comutatividade e associatividade da adição,

$$a_1 = (a + a_1) + a_2 = 0 + a_2 = a_2$$

o que mostra que ambos encontramos os mesmos simétricos. Em particular, como a e $-(-a)$ são ambos simétricos de $-a$, podemos afirmar que se $a \in \mathbb{R}$, então

$$-(-a) = a \qquad (26)$$

Assim como o simétrico, o inverso de cada número real não nulo também é único e satisfaz $(a^{-1})^{-1} = a$.

Vejamos, para começar, o que acontece quando multiplicamos um número real por zero, porque esta é uma das propriedades mais simples que não foi listada acima. Seja $a \in \mathbb{R}$. Como as propriedades fundamentais não nos dizem qual deveria ser o valor de $a \cdot 0$, vamos denotá-lo por z. Por outro lado, como 0 é o elemento neutro da soma, temos que

$$0 + 0 = 0.$$

Multiplicando os dois lados desta igualdade por a, obtemos

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

Contudo, pela propriedade distributiva,

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

Combinando estas duas últimas equações, encontramos

$$z + z = z.$$

Como z é um número real, ele tem um simétrico $-z$. Somando $-z$ aos dois lados desta equação, obtemos

$$-z + (z + z) = -z + z;$$

donde, pela associatividade da adição,

$$(-z + z) + z = -z + z;$$

Mas, pela comutatividade e definição de $-z$, sabemos que $-z + z = 0$, donde

$$0 + z = 0.$$

Finalmente, a comutatividade e o fato de 0 ser o elemento neutro da adição nos dão $z = 0$. Como $z = a \cdot 0$, mostramos que

$$a \cdot 0 = 0. \tag{27}$$

Note que a recíproca desta propriedade também é verdadeira; isto é, se $ab = 0$, então $a = 0$ ou $b = 0$. De fato, se $b = 0$, então $ab = 0$ por (27); por outro lado, se $b \neq 0$, então, multiplicando $ab = 0$ por b^{-1} encontramos

$$(ab)b^{-1} = 0 \cdot b^{-1}.$$

Aplicando a comutatividade da multiplicação e (27) ao lado direito desta igualdade, temos que

$$(ab)b^{-1} = 0;$$

de modo que, pela associatividade da multiplicação nos dão

$$a(bb^{-1}) = 0.$$

Como $bb^{-1} = 1$, que é o elemento neutro da multiplicação, obtemos $a = 0$. Podemos resumir tudo o que fizemos neste parágrafo na seguinte equivalência

$$\forall a, b \in \mathbb{R}, ((a \cdot b = 0) \iff ((a = 0) \vee (b = 0))). \tag{28}$$

A propriedade que acabamos de provar, pode ser usada para provar outras. Por exemplo, como $1 + (-1) = 0$, a equação (28) nos dá

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

Porém, somando $-a$ aos dois lados desta equação, encontramos

$$-a = a \cdot (-1). \tag{29}$$

Mostramos, com isso, que multiplicar um número real qualquer pelo simétrico de 1 produz o simétrico deste número. Em particular, quando $a = -1$, obtemos

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

Logo, pela equação (26),

$$(-1) \cdot (-1) = 1, \tag{30}$$

explicando, assim, o porquê da regra de sinais usual. Outra consequência de (29), que precisamos mencionar, é que

$$\forall a \in \mathbb{R}, (a \neq 0 \implies -a \neq 0). \tag{31}$$

Provaremos isto apelando para a contrapositiva da afirmação acima, segundo a qual

$$\forall a \in \mathbb{R}, ((-a = 0) \implies (a = 0)).$$

Por (29) e (26),

$$a = (-1) \cdot (-a),$$

de modo que, por (28),

$$a = (-1) \cdot (-a) = (-1) \cdot 0 = 0,$$

provando, assim, (31) através de sua contrapositiva.

Do ponto de vista da aritmética que aprendemos na escola, duas operações ficaram de fora de nossa investigação dos números reais: a subtração e a divisão. Isto ocorre porque, no enfoque desenvolvido a partir do final do século XIX, nenhuma das duas é uma operação fundamental. Assim, a subtração $a - b$ nada mais é que a soma de a com o simétrico de b; ao passo que, quando $b \neq 0$, a divisão a/b é igual ao produto de a pelo inverso de b. Apesar disto, há uma propriedade da divisão que vale à pena explicitar. Para isso, digamos que a_1, a_2, b_1, b_2 são números reais e que b_1 e b_2 são ambos não nulos. A propriedade que queremos provar afirma que

$$a_1/b_1 = a_2/b_2 \tag{32}$$

se, e somente se,

$$a_1 b_2 = a_2 b_1.$$

Suponhamos, para começar, que (32) seja verdadeira. Pela definição de divisão dada acima, isto equivale a

$$a_1 b_1^{-1} = a_2 b_2^{-1}.$$

Mas, multiplicando os dois lados por $b_1 b_2$, obtemos

$$(a_1 b_1^{-1})(b_1 b_2) = (a_2 b_2^{-1})(b_1 b_2).$$

Pela propriedade associativa, isto nos dá

$$a_1 (b_1^{-1} b_1) b_2 = a_2 (b_2^{-1} (b_1 b_2)).$$

Usando, agora associatividade e comutatividade, encontramos

$$a_1 (b_1 b_1^{-1}) b_2 = a_2 (b_1 (b_2 b_2^{-1}));$$

donde,

$$(a_1 \cdot 1) b_2 = a_2 (b_1 \cdot 1).$$

Finalmente, como 1 é o elemento neutro da multiplicação,

$$a_1 b_2 = a_2 b_1.$$

Como enunciamos o resultado a ser provado como uma equivalência entre duas fórmulas, ainda precisamos provar a recíproca da direção que acabamos de fazer. Não faremos isto simplesmente porque, neste caso, para provar a recíproca basta ler a demonstração acima de trás-para-frente, já que cada um dos passos é facilmente revertido.

Passando, agora, aos produtos notáveis, vejamos como obtê-los a partir dos axiomas algébricos. Distribuindo o número real $a + b$ sobre a soma de a e b, obtemos

$$(a + b)^2 = (a + b) \cdot (a + b) = (a + b)a + (a + b)b.$$

Usando, então, a comutatividade da multiplicação e distribuindo, em seguida, a e b em cada parcela sobre a soma $a + b$,

$$(a + b)^2 = a^2 + ba + ab + b^2.$$

A comutatividade da multiplicação nos dá, então,

$$(a + b)^2 = a^2 + 2ab + b^2, \tag{33}$$

que é a fórmula usual. Para obter

$$(a - b)^2 = a^2 - 2ab + b^2, \tag{34}$$

basta substituir b por $-b$ na fórmula anterior. Outra fórmula muito útil resulta quando subtraímos (34) de (33), obtendo

$$4ab = (a + b)^2 - (a - b)^2. \tag{35}$$

Do ponto de vista geométrico, a equação (35) pode ser interpretada como expressando a área de um retângulo de lados a e b como a diferença de dois quadrados. Como se trata de uma afirmação sobre áreas de quadrados e retângulos, podemos ilustrá-la como na figura 1.

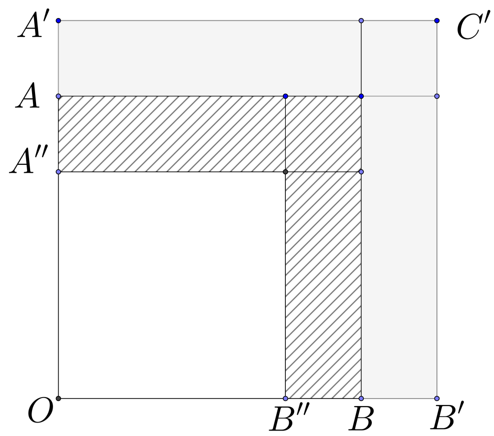


FIGURA 1. Um retângulo como diferença de quadrados

Para facilitar a discussão da figura, chamaremos de *gnômon* (não confundir com gnômo!) a figura em forma de L maiúsculo, formada pela justaposição de dois retângulos perpendiculares entre si. A terminologia remonta à Grécia Antiga, onde palavra originalmente designava o esquadro de marceneiro, que tem a mesma forma.

Voltando à figura 1, digamos que os segmentos OA e OB têm comprimento a, ao passo que

$$AA', AA'', BB' \text{ e } BB''$$

têm comprimento igual a b. Portanto, o quadrado $OA'C'B'$ tem área igual a

$$(a+b)^2$$

e o quadrado branco tem área

$$(a-b)^2.$$

A diferença entre estas duas áreas corresponde à dois gnômons. O gnômon menor, que aparece hachurado na figura, é formado por um retângulo horizontal, cujos lados têm comprimento a e b, e um retângulo menor, vertical, cujos lados têm comprimentos $a-b$ e b. Portanto, a área deste gnômon é igual a

$$ab + (a-b)b = 2ab - b^2.$$

Já os retângulos que formam o gnômon maior, pintado de cinza na figura, têm lados iguais a $a+b$ e b, no caso do retângulo horizontal, e a e b, no caso do triângulo vertical. Logo, a área do gnômon maior será

$$(a+b)b + ab = 2ab + b^2.$$

Somando as áreas dos dois gnômons, obtemos

$$(2ab - b^2) + (2ab + b^2) = 4ab,$$

como afirma a equação 35.

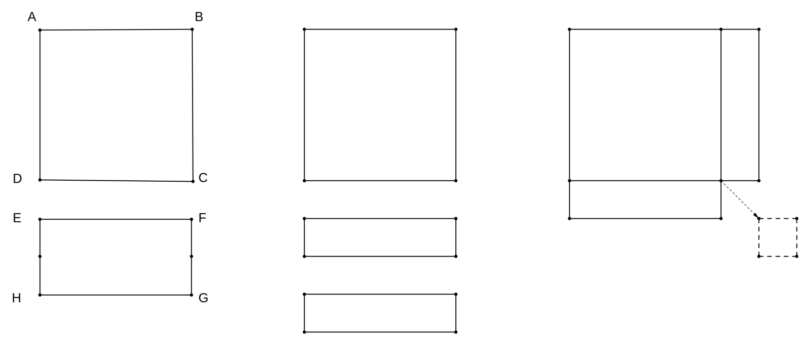


FIGURA 2. Completando quadrados

O produto notável (33) pode ser usado para provar a fórmula usual para a solução de uma equação do segundo grau. Para isto, suponhamos que a equação é

$$ax^2 + bx + c = 0,$$

em que $a \neq 0$, b e c são números reais. Admitindo que x exista e seja um número real, podemos aplicar à equação acima as propriedades algébricas que aprendemos. A ideia é rearrumar os termos de modo que seja possível resolver a equação apenas extraindo uma raiz quadrada.

Começamos reescrevendo esta equação na forma

$$x^2 + \frac{b}{a}x + \frac{c}{a} = 0. \tag{36}$$

Em seguida, utilizaremos o método de *completar quadrados*, cujo passo-a-passo é ilustrado na figura2. Nela o quadrado $ABCD$ tem lado x e o retângulo $EFGH$ tem base igual a x e altura igual a b/a. Como as áreas destes dois quadriláteros são iguais a x^2 e bx/a, respectivamente, a área total dos dois é

$$x^2 + bx/a$$

Em seguida dividimos $EFGH$ em dois retângulos de altura $b/2a$ e colamos um deles abaixo do quadrado e o outro do seu lado direito. Podemos indicar isto algebricamente escrevendo

$$x^2 + bx/a = x^2 + 2\frac{b}{2a}x. \tag{37}$$

A figura obtida nesta etapa só não é um quadrado de lado $x + b/2a$, por falta do pequeno quadrado de lado $b/2a$ que aparece tracejado na figura. Portanto,

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

Substituindo isto em (36), obtemos

$$\left(x + \frac{b}{2a}\right)^2 - \left(\frac{b}{2a}\right)^2 + \frac{c}{a} = 0, \tag{38}$$

que equivale a

$$\left(x + \frac{b}{2a}\right)^2 = \left(\frac{b}{2a}\right)^2 - \frac{c}{a}.$$

Reescrevendo o lado direito sobre um mesmo denominador

$$\left(x + \frac{b}{2a}\right)^2 = \frac{b^2 - 4ac}{4a^2}.$$

Extraindo, agora, a raiz quadrada dos dois lados,

$$x + \frac{b}{2a} = \pm\sqrt{\frac{b^2 - 4ac}{4a^2}};$$

donde

$$x = -\frac{b}{2a} \pm \sqrt{\frac{b^2 - 4ac}{4a^2}}. \tag{39}$$

Simplificando, obtemos

$$x = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}, \tag{40}$$

que é a fórmula usual para a solução da equação quadrática.

4. Os números reais: ordem

As duas últimas propriedades dos números reais de que precisamos dizem respeito à existência de um subconjunto $\mathbb{P}$ de $\mathbb{R}$, *que não contém o* 0 e cujos elementos satisfazem as propriedades listadas na tabela 1.

Axioma	
Tricotomia	$\forall a \in \mathbb{R}, ((a \in \mathbb{P}) \vee (a = 0) \vee (-a \in \mathbb{P}));$
Fechamento	$\forall a, b \in \mathbb{R}, ((a \in \mathbb{P}) \wedge (b \in \mathbb{P}) \Longrightarrow (a + b \in \mathbb{P}) \wedge (ab \in \mathbb{P})).$

TABELA 1. Propriedades da ordenação dos números reais

Outra propriedade relevante, mas que podemos deduzir destas duas, nos diz que um número e seu simétrico não podem, ambos, pertencer a $\mathbb{P}$. Provaremos isto usando o método de redução ao absurdo, introduzido na seção 2. Suponhamos, por contradição, que a e $-a$ pertençam a $\mathbb{P}$. Pela positividade da adição, teríamos que

$$a + (-a) \in \mathbb{P}.$$

Contudo, $a + (-a) = 0$, que não é um elemento de $\mathbb{P}$. A contradição a que chegamos mostra que a suposição de que a e $-a$ são ambos elementos de $\mathbb{P}$ não é viável, como havíamos afirmado. Como, por hipótese, $a \in \mathbb{P}$, podemos concluir que $-a \notin \mathbb{P}$ e vice-versa.

O que fizemos já basta para que sejamos capazes de mostrar que o quadrado de qualquer número real não nulo pertence a $\mathbb{P}$. Se $a \in \mathbb{P}$, isto é consequência do fechamento. Por outro lado, se $a \notin \mathbb{P}$ não for igual a zero, então $-a \in \mathbb{P}$ pela tricotomia; donde

$$(-a)^2 \in \mathbb{P}. \tag{41}$$

Contudo, por (29),

$$(-a)^2 = (-a) \cdot (-a) = ((-1) \cdot a) \cdot ((-1) \cdot a).$$

Mas, utilizando a comutatividade e a associatividade da multiplicação, podemos reescrever isto na forma

$$((-1) \cdot a) \cdot ((-1) \cdot a) = ((-1) \cdot (-1)) \cdot (a \cdot a);$$

que, por (30), nos dá

$$(-a)^2 = (-a) \cdot (-a) = 1 \cdot (a \cdot a) = a^2.$$

Portanto, por (41),

$$a^2 = (-a)^2 \in \mathbb{P},$$

como queríamos mostrar. Em particular, isto garante que

$$1 = 1^2 \in \mathbb{P},$$

o que nos permite, também, afirmar que $\mathbb{P}$ não pode ser vazio.

Provavelmente você já se deu conta de que $\mathbb{P}$ corresponde ao conjunto dos números reais positivos e deve estar se perguntando o porquê de considerar o conjunto $\mathbb{P}$, em vez da relação $>$, com a qual já está habituado. A resposta é que, como nossa meta consiste em provar as propriedades da desigualdade a partir da tricotomia e do fechamento, é preferível usar uma notação diferente da usual para evitarmos cair no problema, apontado na seção 2, de introduzir, sem nos darmos conta, propriedades das desigualdades que nos são muito familiares, mas que ainda não foram provadas, violando assim as regras do jogo dedutivo.

Contudo, como as próximas propriedades são mais fáceis de enunciar usando a relação $>$, passaremos a utilizá-la a partir deste ponto. Assim, de agora em diante, se $a, b \in \mathbb{R}$, escreveremos $a > b$ para indicar que $a - b \in \mathbb{P}$. Em particular, como 0 é o elemento neutro da soma,

$$\forall a \in \mathbb{R}, (a > 0 \iff a \in \mathbb{P}).$$

Suponhamos que $a, b, c \in \mathbb{R}$. Se $b > a$ e $c > b$, temos, por definição, que

$$b - a \in \mathbb{P} \text{ e } c - b \in \mathbb{P};$$

donde, por fechamento,

$$(b - a) + (c - b) \in \mathbb{P}.$$

Como

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

podemos concluir que $c - a \in \mathbb{P}$; isto é, que $c > a$. Provamos, assim, a transitividade da relação $>$, segundo a qual

$$\forall a, b, c \in \mathbb{R}, ((b > a) \wedge (c > b) \Longrightarrow (c > a)).$$

Em particular, se $b > 0$ e $a < 0$, isto nos dá $a < b$, mostrando que todo número negativo é menor que qualquer número positivo. Por outro lado, da equação (29), obtemos

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

o que nos permite concluir que

$$\forall a, b, c \in \mathbb{R}, (a > b \Longrightarrow (a + c) > (b + c)).$$

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

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

donde, por (29) e pela distributiva,

$$ac - bc > 0,$$

que equivale a dizer que

$$ac > bc.$$

Em particular, quando $-c > 0$, esta desigualdade nos dá,

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

Logo, por (29),

$$-ac > -bc;$$

de modo que

$$bc - ac > 0.$$

Denotando $-\alpha > 0$ por $\alpha < 0$, como usual, podemos resumir o que dizendo que, quaisquer que sejam os números reais a, b e c,

$$a > b \Longrightarrow \begin{cases} ac > bc & \text{quando } c > 0; \\ ac < bc & \text{quando } c < 0. \end{cases} \tag{42}$$

Antes de prosseguir, precisamos de mais uma definição. Dados $a, b \in \mathbb{R}$, escreveremos

$$a \geq b \Longleftrightarrow (a = b) \vee (a > b)$$

e uma definição análoga se aplica a $a \leq b$. Com isto, podemos deduzir de (42) que

$$\forall a, b \in \mathbb{R}, (((a \leq 0) \wedge (b \leq 0)) \vee ((a \geq 0) \wedge (b \geq 0)) \implies ab \geq 0), \tag{43}$$

ao passo que

$$\forall a, b \in \mathbb{R}, (((a \geq 0) \wedge (b \leq 0)) \vee ((a \leq 0) \wedge (b \geq 0)) \implies ab \leq 0). \tag{44}$$

As contrapositivas destas duas propriedades são muito úteis. Mostraremos como obter a contrapositiva de (43) porque precisaremos dela abaixo, a outra ficará como exercício para você. Por definição, temos que a contrapositiva de (43) é

$$\forall a, b \in \mathbb{R}, (ab < 0 \implies \neg(((a \leq 0) \wedge (b \leq 0)) \vee ((a \geq 0) \wedge (b \geq 0)))).$$

Aplicando a lei de De Morgan, à conclusão, vemos que é igual a

$$((a > 0) \vee (b > 0)) \wedge ((a < 0) \vee (b < 0));$$

donde, pela distributividade dos conectivos $\vee$ e $\wedge$,

$$((a > 0) \wedge (a < 0)) \vee ((a > 0) \wedge (b < 0)) \vee ((b > 0) \wedge (a < 0)) \vee ((b > 0) \wedge (b < 0)).$$

Contudo,

$$(a > 0) \wedge (a < 0) = (b > 0) \wedge (b < 0) = \perp,$$

que é o elemento neutro do conectivo $\vee$. Levando isto em conta, a contrapositiva desejada é

$$\forall a, b \in \mathbb{R}, (ab < 0 \implies ((a < 0) \wedge (b > 0)) \vee ((a > 0) \wedge (b < 0))). \tag{45}$$

A tabela 2 resume as principais propriedades que provamos nesta seção e na anterior.

Propriedade	
Absorção	$0 \cdot a = 0$
Multiplicação por -1	$(-1) \cdot a = -a$
Produtos notáveis	$(a \pm b)^2 = a^2 + b^2 \pm 2ab$ e $a^2 - b^2 = (a + b)(a - b)$
Positividade do 1	$1 > 0$
Sinal do simétrico	$(a > 0) \implies (-a < 0)$
Transitividade	$(b > a) \wedge (c > b) \implies (c > a)$
Produto por número positivo	$(a > b) \wedge (c > 0) \implies ac > bc$
Produto por número negativo	$(a > b) \wedge (c < 0) \implies ac < bc$
Multiplicação cruzada	$((bd \neq 0) \wedge (a/b = c/d)) \implies ad = cb$

TABELA 2. Sumário das propriedades dos números reais

As propriedades de ordem nos ajudam a entender de que forma as raízes da equação do segundo grau

$$ax^2 + bx + c = 0,$$

dependem dos valores de a, b e c. O número

$$\Delta = b^2 - 4ac$$

é o *discriminante* da equação e desempenhará um papel muito importante em seções posteriores. Como $a, b, c \in \mathbb{R}$, o mesmo vale para Δ. Por tricotomia há três casos a analisar. Se $\Delta < 0$, a raiz quadrada não é um número real, de modo que $ax^2 + bx + c = 0$ não tem raízes reais. Se $\Delta = 0$, a equação tem apenas $-b/2a$ como raiz, ao passo que se $\Delta > 0$, a equação tem duas raízes reais, uma para cada escolha de sinal na fórmula (40).

Também podemos usar desigualdades para construir os subconjuntos de $\mathbb{R}$ que mais aparecem em aplicações, os *intervalos*. Se $a < b$ são números reais, escrevemos

$$[a, b] = \{x \in \mathbb{R} \mid a \leq x \leq b\} \quad \text{e} \quad (a, b) = \{x \in \mathbb{R} \mid a < x < b\};$$

note que são usados parênteses quando a desigualdade é estrita (<); caso contrário usamos colchetes. Colchetes e parênteses podem ser combinados em um mesmo intervalo; assim,

$$[a, b) = \{x \in \mathbb{R} \mid a \leq x < b\} \quad \text{e} \quad (a, b] = \{x \in \mathbb{R} \mid a < x \leq b\}.$$

Quando uma das extremidades que delimitam o intervalo está contida no intervalo, dizemos que ele é *fechado* nesta extremidade, caso contrário dizemos que é *aberto* naquela extremidade. Por exemplo, $(1, 2]$ é aberto em 1 e fechado em 2, ao passo que $[2, 3]$ é fechado em 2 e fechado em 3. O intervalo é aberto quando suas duas extremidades são abertas e fechado se ambas são fechadas; assim $(1, 2)$ é aberto e $[2, 3]$ é fechado. Quando ilustramos um intervalo, a extremidade aberta é identificada por um ponto vazado e a extremidade fechada por um ponto cheio, como exemplificado na figura 3.

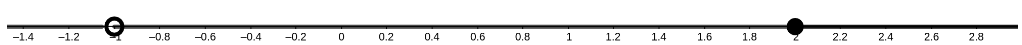


FIGURA 3. O intervalo $(-1, 2]$.

O último resultado desta seção é uma desigualdade extremamente útil mas, para enunciá-la precisamos de uma definição: o *valor absoluto*, também chamado de *módulo*, de um número real a é o número real

$$\operatorname{abs}(a) = |a| = \begin{cases} +a & \text{se } a \geq 0 \\ -a & \text{se } a < 0; \end{cases}$$

Nossa meta é provar a *desigualdade triangular*, segundo a qual, se $a, b \in \mathbb{R}$, então

$$|a + b| \leq |a| + |b|. \tag{46}$$

Antes, porém, precisamos mostrar que

$$|ab| \geq ab. \tag{47}$$

Quando $ab \geq 0$ temos, por definição do valor absoluto, que $|ab| = ab$; de modo que (47) é válida neste caso. Por outro lado, se $ab < 0$, então por (45) temos que

$$((a < 0) \wedge (b > 0)) \vee ((a > 0) \wedge (b < 0)).$$

Se $a < 0$ e $b > 0$, então, multiplicando $|a| \geq a$ por b e usando (42), obtemos

$$|a|b \geq ab.$$

Mas, $b > 0$ implica que $|b| = b$. Combinando isto à desigualdade anterior, encontramos

$$|a||b| = |a|b \geq ab.$$

O caso em que $a > 0$ e $b < 0$ pode ser tratado da mesma maneira; os detalhes ficarão por sua conta.

Para provar a desigualdade triangular a partir de (47), basta lembrar que

$$(a + b)^2 = a^2 + b^2 + 2ab$$

nos permite afirmar que

$$(a + b)^2 \leq a^2 + b^2 + 2|a| \cdot |b|.$$

Como

$$a^2 + b^2 + 2|a| \cdot |b| = (|a| + |b|)^2,$$

a desigualdade anterior pode ser escrita na forma

$$|a + b|^2 \leq (|a| + |b|)^2.$$

Logo,

$$0 \geq |a + b|^2 - (|a| + |b|)^2 = (|a + b| - (|a| + |b|))(|a + b| + (|a| + |b|)).$$

Mas, por (45), isto só pode ocorrer quando

$$|a + b| - (|a| + |b|) \geq 0 \quad \text{e} \quad |a + b| + (|a| + |b|) \leq 0$$

ou quando

$$|a + b| - (|a| + |b|) \leq 0 \quad \text{e} \quad |a + b| + (|a| + |b|) \geq 0$$

No primeiro caso,

$$|a + b| \leq -|a| - |b|$$

que só pode ocorrer quando $a = b = 0$. Portanto, caso a ou b não sejam nulos, teremos necessariamente que

$$|a + b| - (|a| + |b|) \leq 0 \quad \text{e} \quad |a + b| + (|a| + |b|) \geq 0$$

A primeira destas duas últimas desigualdade é a que queremos provar, já a segunda é inteiramente óbvia, já que todo número positivo é maior que qualquer número negativo.

5. O axioma de continuidade

Todas as propriedades dos números reais que consideramos nas seções anteriores são também satisfeitas pelos números inteiros. Entretanto, como vimos na seção 4 do capítulo 1, uma das coisas que distingue a matemática da Grécia Antiga daquela praticada no Egito e na Mesopotâmia, é, justamente, que os primeiros entenderam que nem toda medida de segmento corresponde a uma fração. O quê, então, faz com que os números reais sejam tão diferentes dos números racionais?

Surpreendentemente, um axioma basta para distinguir entre números racionais e números reais. Conhecido como *axioma de continuidade* ou *axioma de completude*, ele pode ser formulado de muitas maneiras diferentes. A que adotaremos aqui é uma das mais fáceis de usar na prática, mas, para poder enunciá-la, precisamos de um par de definições.

Seja S um subconjunto de $\mathbb{R}$. Dizemos que S tem uma *cota superior* se existe um número real c tal que

$$\forall x \in S, (x \leq c).$$

Note que se c é uma cota superior para S, então todo número real maior que c também é uma cota superior para S. Por isso sempre usaremos o artigo indefinido quando tratarmos de cotas superiores. A menor cota superior de S é o seu *supremo*, que denotaremos por $\sup(S)$. Por exemplo, 100 é uma cota superior do conjunto

$$\{1, 2, 8/3, 11/2\}$$

cujo supremo é $11/2$. Neste caso, o supremo é parte do conjunto, mas isto nem sempre acontece; assim, o supremo do intervalo $(1, 2)$ é 2. O axioma que nos falta é o seguinte.

Axioma de continuidade. *Se um subconjunto de números reais admite uma cota superior então tem um supremo.*

Diante dos dois exemplos acima, o axioma não parece estar dizendo grande coisa. Precisamos de um exemplo mais robusto para que você entenda quão poderoso é este axioma. Considere o subconjunto

$$C = \{q \in \mathbb{Q} \mid q^2 \leq 2\}. \tag{48}$$

A primeira coisa a notar é que não se trata de um intervalo, porque estamos considerando apenas números racionais neste conjunto. Em segundo lugar, C tem 2 como uma cota superior. Para mostrar isto, observe que, de

$$x^2 \leq 2 < 4$$

podemos deduzir

$$(x - 2)(x + 2) < 0.$$

Logo, um dos termos é negativo e o outro positivo. Contudo, como

$$x - 2 < x + 2,$$

temos,forçosamente, que

$$x < 2,$$

o que confirma nossa afirmação de que 2 é uma cota para C. Logo, pelo axioma de continuidade, S tem um supremo. Diremos, por definição, que

$$\sup(C) = \sqrt{2}.$$

Assim, de nosso ponto de vista, o que fizemos foi provar que *existe* um número real que satisfaz a equação $x^2 = 2$.

Neste ponto é importante chamar sua atenção para uma propriedade bastante simples do supremo. Digamos que $S \subset \mathbb{R}$ e que $\sigma = \sup(S)$. Se r for um número real menor que σ, então, pela minimalidade do supremo, tem que existir algum $x \in S$ que é maior do que r. Aplicando isto ao conjunto C definido em (48), vemos que, por maior que seja o inteiro $n > 0$, existe um número $q \in C$, tal que

$$\sqrt{2} - \frac{1}{10^n} \leq q < \sqrt{2}.$$

Em outras palavras, podemos obter um número racional que é uma aproximação de $\sqrt{2}$ com um erro menor que 10^{-n}, qualquer que seja o inteiro positivo n. Um argumento semelhante pode ser usado para provar que *qualquer número real pode ser aproximado por um número racional com um erro tão pequeno quanto desejado*.

De todos os axiomas que definem os números reais, o axioma de continuidade foi o que mais tempo levou para ser corretamente especificado. Isto só veio a ocorrer do meio para o final do século XIX, quando vários problemas com os fundamentos do cálculo diferencial e integral levaram a um aumento significativo no rigor exigido para que uma demonstração fosse considerada correta.

Um dos primeiros matemáticos a confrontar esta questão com sucesso foi Richard Dedekind (1831–1916). Em 1858, quando foi contratado pelo Instituto Politécnico de Zurique (atual ETHzürich), Dedekind viu-se responsável, pela primeira vez, pelo curso introdutório de cáculo diferencial e integral. O primeiro dos tópicos na ementa que preparou para este curso é *números reais: continuidade* (Dugac, 1976, p. 151). No início de seu ensaio *Continuidade e os números irracionais* Dedekind (1963), Dedekind menciona que, ao preparar este curso, sentiu "mais intensamente do que em qualquer momento anterior a falta de um fundamento realmente científico para a aritmética". Isto o conduziu à ideia que deveria ser possível construir os números reais diretamente a partir dos números racionais. Para isto, Dedekind inventou a noção de corte, publicada pela primeira vez em 1872 no ensaio mencionado acima.

Um subconjunto C dos números racionais é um *corte* se

1. C não é vazio nem igual a $\mathbb{Q}$;

2. se $x \in C$ e $y \notin C$, então $x < y$;

3. C não contém um número racional maior que todos os elementos de C.

É provável que a semelhança da noção de corte com o axioma de continuidade não lhe tenha passado despercebida; não se trata de mera coincidência: nossa formulação do axioma é baseada na noção de corte. Assim, o conjunto C definido em (48) é um corte no sentido de Dedekind.

A esta altura é importante explicitar a diferença entre nosso tratamento dos números reais e o que foi exposto por Dedekind. O que fizemos foi explicitar as propriedades que especificam o conjunto dos números reais, entre vários outros conjuntos possíveis. Diante disto, podemos adotar uma, entre duas atitudes possíveis. A primeira é que o conjunto dos números reais é pré-existente; apenas listamos suas propriedades básicas. A segunda é que as propriedades que identificamos servem para definir o quê realmente é o conjunto dos números reais. Ambas as atitudes são problemáticas. De um lado, ninguém realmente pode dizer que já viu este conjunto; de outro, nada garante que uma dada lista de propriedades é compatível, no sentido de não gerar nenhuma contradição. Com seus cortes, Dedekind resolve simultaneamente estes dois problemas. Sua estratégia assume como pré-existente apenas o conjunto dos números racionais e, a partir deles, *constrói* um conjunto que, como Dedekind mostra em *Continuidade e os números irracionais*, satisfaz todas as propriedades que enunciamos acima. Em particular, isto mostra que um conjunto com as propriedades desejadas *de fato existe*. A dificuldade em fazer este esquema funcionar não deve ser subestimada. Se aos cortes desejamos fazer corresponder os números irracionais, é necessário, por exemplo, explicar como dois cortes podem ser somados, subtraídos, multiplicados e divididos.

Para encerrar, vejamos até que ponto nossa análise dos números reais se conforma aos critérios de pensamento científico estabelecidos por Aristóteles, que mencionamos na seção 2. A primeira coisa a notar é que, sem dúvida alguma, mostramos que os axiomas enunciados nas seções 3 e 4 são as "causas" das propriedades enunciadas na tabela 2 e que "não é possível que fosse de outra maneira". Contudo, nosso enfoque não se sai tão bem, quando consideramos as propriedades que tomamos como ponto de partida. O problema é que a análise de Aristóteles requer que as premissas das quais parte uma demonstração devam ser, em última análise, mais simples do que o resultado que estamos provando. Entretanto, ainda que a afirmação de que o zero é o elemento neutro da adição seja bastante elementar, é discutível se, por exemplo, a propriedade distributiva e a existência do simétrico são mais conformes à nossa intuição do que o fato da multiplicação por zero dar sempre zero.

Em defesa do que fizemos, estas propriedades foram escolhidas porque são comuns a uma grande quantidade de sistemas matemáticos, que vão dos números naturais às matrizes e dos inteiros modulares às funções reais. Portanto, ainda que não sejam, todas elas, facilmente acessíveis à intuição, estas propriedades são fundamentais no sentido de estarem presentes em inúmeros sistemas extremamente diversos.

Finalmente, é importante frisar que o tratamento matemático rigoroso dos números reais nada deixa a desejar aos critérios estabelecidos nos *Posteriores analíticos*. As deficiências apontadas acima têm sua origem no fato dos números reais não serem objetos simples, mas sim construídos a partir dos números naturais. Assim, um tratamento mais sofisticado dos reais partiria dos cinco axiomas de Peano, que estabelecem as propriedades fundamentais dos números naturais. Os quatro primeiros axiomas são:

Axioma 1: 1 é um número natural;
Axioma 2: todo número natural n tem um sucessor $S(n)$;
Axioma 3: $S(x) \neq x$;
Axioma 4: se $S(x) = S(x')$, então $x = x'$;

ao passo que o quinto é o *princípio de indução finita*. Tomando estes cinco axiomas como ponto de partida construímos primeiramente os inteiros; de posse destes, podemos construir os números racionais como pares de inteiros sujeitos a uma certa relação. Finalmente, os número reais são construídos como cortes de Dedekind. Infelizmente esta maneira de proceder é longa e envolve a demonstração de uma grande quantidade de resultados intermediários, o que a torna inviável em um livro elementar. Uma versão extremamente detalhada desta construção pode ser encontrada nos *Fundamentos da análise* de Edmund Landau (1877–1938), que tem uma tradução para o inglês Landau (1951). Landau foi um mestre no uso do método dedutivo, que aplica sistematicamente ao longo do *Grundlagen*. Em vista do que fizemos, vale à pena reproduzir o segundo e terceiro itens no *prefácio para o estudante* deste livro:

2. peço apenas que tenha a habilidade de ler alemão e pensar logicamente–nada da matemática do ensino médio e, certamente, nenhuma matemática avançada;
3. por favor esqueça tudo o que aprendeu na escola; você não aprendeu.

Caso esteja curioso, o primeiro item é: *por favor não leia o prefácio para o professor*.

Exercícios

1. Traduza a proposição
$$\forall a \in \mathbb{R}, (\forall b \in \mathbb{R}, ((a < b) \Longrightarrow (\exists c \in \mathbb{R}, ((a < c) \wedge (c < b)))))$$
em português e determine sua negativa. A negativa deve ser construída passo-a-passo e expressa em termos da negativa de suas proposições atômicas, indicando a propriedade dos conectivos usadas em cada passo.

2. Uma variação do quantificador existencial é $\exists!$, que significa, *existe um e apenas um*.
   (a) Modele a sentença $\exists! x, P(x)$, em que $P(x)$ é uma proposição, usando apenas $\exists$ e os conectivos lógicos.
   (b) Determine a negativa de $\exists! x, P(x)$.

Use as informações a seguir para fazer as questões 3 e 6. Seja Ω o conjunto universo que consiste no quadrado da figura 3 em que algumas casas estão preenchidas com figuras geométricas de várias cores. Considere as seguintes proposições relativas às figuras geométricas x e y na figura .

$T(x)$ = x é um triângulo;
$C(x)$ = x é um círculo;
$Q(x)$ = x é um quadrado;
$B(x)$ = x é branco;
$P(x)$ = x é preto;
$N(x,y)$ = x está ao norte de y;
$O(x,y)$ = x está a oeste de y;
$M(x,y)$ = x tem a mesma cor que y.

N

				○		
	○		○	▲		
O	■	△			□	L
				△		
			□		□	

S

TABELA 3. O quadrado

3. Traduza cada uma das seguintes afirmações em português e determine quais são verdadeiras e quais são falsas:
   (a) $\forall y, (C(y) \Longrightarrow \exists x, N(y,x))$;
   (b) $\forall x, (T(x) \Longrightarrow \exists y, (Q(y) \wedge M(x,y)))$;
   (c) $\exists x, (Q(x) \wedge \forall y, (C(y) \Longrightarrow O(x,y)))$;
   (d) $\forall x, (Q(x) \Longrightarrow \exists y, (C(y) \wedge M(x,y)))$;

4. Determine a negativa de cada uma das proposições do exercício 3.

5. Ache, passo-a-passo a negação da seguinte fórmula:

$$(((\exists x)(\forall y)(p(x,y))) \Rightarrow ((\forall y)(\exists x)(p(x,y))))$$

6. Usando a notação do exercício anterior, traduza cada uma das seguintes afirmações como uma fórmula lógica a partir de suas proposições atômicas, determine sua negativa e traduza esta última em português
   (a) para cada círculo, há um quadrado da mesma cor;
   (b) existe um quadrado a oeste de todos os triângulos;
   (c) todo quadrado vermelho está ao norte de algum triângulo;
   (d) todos os quadrados têm a mesma cor que algum triângulo.

7. Modele as proposições abaixo usando $\neg$, $\vee$, $\wedge$, $\to$, $\leftrightarrow$ e os quantificadores $\forall$ e $\exists$.
   (a) Todo número racional não nulo tem um inverso multiplicativo.
   (b) Existem números primos.
   (c) Se a e n são inteiros tais que $\mathrm{mdc}(a,n) = 1$, então $ab - 1$ é divisível por n, para algum inteiro b.
   (d) Dados três segmentos de retas de comprimentos a, b e c, se $a + b \geq c$ existe um triângulo cujos lados são estes segmentos.
   (e) Toda equação do segundo grau cujo discriminante é maior ou igual a 0 tem, pelo menos, uma raiz real.

8. Prove que as seguintes afirmações são verdadeiras:
   (a) $\forall a,b,c,d \in \mathbb{R}, (((a < b) \wedge (c < d)) \Longrightarrow a + c < b + d)$;
   (b) $\forall a,b \in \mathbb{R}, ((a < b) \Longrightarrow (-b < -a))$;
   (c) $\forall a,b,c,d \in \mathbb{R}, (((a < b) \wedge (c > d)) \Longrightarrow a - c < b - d)$;
   (d) $\forall a,b \in \mathbb{R}, (((a < 0) \wedge (b > 0)) \Longrightarrow ab < 0)$;
   (e) $\forall a \in \mathbb{R}, ((a > 1) \Longrightarrow (a^2 > a))$;
   (f) $\forall a \in \mathbb{R}, ((0 < a < 1) \Longrightarrow (a^2 < a))$;
   (g) $\forall a,b,c,d \in \mathbb{R}, (((0 \leq a < b) \wedge (0 \leq c < d)) \Longrightarrow ac < bd)$;
   (h) $\forall a,b \in \mathbb{R}, ((0 \leq a < b) \Longrightarrow a^2 < b^2)$;
   (i) $\forall a,b \in \mathbb{R}, (ab \geq 0 \Longrightarrow ((a \geq 0 \wedge b \geq 0) \vee (a \leq 0 \wedge b \leq 0)))$.

9. Considere a seguinte proposição

$$\forall a,b \in \mathbb{R}, (((a > 0) \wedge (b > 0) \wedge (a < b)) \Longrightarrow b^{-1} \leq a^{-1}).$$

   (a) Determine a contrapositiva desta proposição.
   (b) Prove a proposição original por contradição, justificando cada um dos passos do argumento a partir das propriedades fundamentais dos números reais.

10. Mostre que

$$\forall a, b \in \mathbb{R}, \Big((0 < a < b) \Longrightarrow \Big(a < \sqrt{ab} < \frac{a+b}{2} < b\Big)\Big).$$

Dica: use a equação (35).

11. Prove que, quaisquer que sejam os números reais a e b, temos que
    (a) $|ab| = |a| \cdot |b|$;
    (b) $|a - b| \leq |a| + |b|$;
    (c) $|a| - |b| \leq |a - b|$;
    (d) $||a| - |b|| \leq |a - b|$;.

12. Ache todos os números reais x para os quais:
    (a) $|x - 3| = 8$;
    (b) $|x - 3| < 8$;
    (c) $|x - 3| > 8$;
    (d) $|x - 1| + |x + 1| < 1$.

13. Sejam a e b números reais. Ache todos os números reais x para os quais
    (a) $(x - a)(x - b) > 0$;
    (b) $(x - a)(x - b) < 0$.

14. Ache todos os valores de x para os quais as seguintes desigualdades são verdadeiras, justificando cada passo de seus cálculos usando os resultados da apostila e do exercício anterior. A resposta deve ser explicitada na forma de um intervalo ou união de intervalos.
    (a) $7 - 3x < 5 - 2x$;
    (b) $x^2 + 35 > 12x$;
    (c) $x^2 - 2 < x$;
    (d) $(x - 2)/(x - 3) > 0$.

15. 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 $2 = 1$.

CAPÍTULO 4

Funções

Neste capítulo iniciamos nosso estudo das funções, que constituem um dos conceitos mais importantes da matemática, como é atualmente praticada. Como veremos na seção 1, ainda que a formulação atual date do século XIX, as funções vêm sendo utilizadas desde a pré-história.

1. Uma longa história

Imagine um pastor da pré-história, vivendo antes mesmo que os números houvessem sido inventados. Como poderia se certificar de que todas as ovelhas voltaram ao redil, onde passarão a noite? Se o rebanho é pequeno, poderia dar um nome a cada ovelha. Mas, isto seria difícil se o rebanho fosse grande. Neste caso, poderia fazer corresponder a cada ovelha que saísse para o pasto uma pequena indentação em um osso ou madeira. À noite, passaria o dedo sobre os entalhes, um de cada vez, à medida que as ovelhas fossem entrando no redil. No primeiro caso, o pastor associou a cada ovelha seu nome, no segundo, um entalhe. É este tipo de associação que está na origem de nosso conceito atual de função. Ossos com entalhes, como os imaginados acima, foram encontrados na África e na Europa e datados de até 43.000 anos atrás.

Naturalmente o processo de contagem não passa de uma versão mais refinada, e mais prática, do processo imaginado no parágrafo anterior. Em vez de pedras, ou de nomes específicos para um rebanho, a cada ovelha são associadas palavras construídas de maneira padronizada. As vantagens são óbvias: além das mesmas palavras poderem ser utilizadas mesmo que haja mudanças no rebanho, elas permitem que pastores diferentes comparem facilmente as quantidades de ovelhas em seus respectivos rebanhos.

Ao passo que os vestígios da pré-histórico são poucos e abertos a outras interpretações, o mesmo não ocorre com os tabletes de argila encontrados no território da Antiga Mesopotâmia. Por exemplo, o tablete AO 6481, encontrado em Uruk, apresenta uma tabela, em várias colunas, relativa à posição de Marte no período que vai de 188 a.C. a 109 a.C.. A primeira coluna lista 79 anos da Era selêucida, cujo começo ocorre mais ou menos em 311 a.C.. Já a quarta coluna especifica a posição do primeiro ponto estacionário de Marte em termos de sua longitude celeste, que é sua distância do equinócio de verão, medida ao longo da eclíptica. Um ponto estacionário de um planeta ocorre quando sua trajetória na abóbada celeste parece parar e então retroceder. Portanto, estas duas colunas da tabela representam uma função que a cada ano associa a longitude de Marte em seus primeiros pontos estacionários daquele ano. Mais detalhes sobre este tablete podem ser encontrados em (Neu, 1983, p. 335).

Desde então astrônomos e matemáticos produziram uma enorme quantidade de tabelas, cada uma das quais hoje interpretamos como uma função. Entre as mais conhecidas estão as funções trigonométricas e os logaritmos. Por exemplo, o *Pañca-siddhāndikā*, que data do século sexto d.C., contém uma lista de valores para a função seno para arcos no primeiro quadrante, baseada em uma circunferência de raio igual a 120. Esta lista e sua construção são discutidas em (Plofker, 2009, p. 51). Já os logaritmos foram originalmente definidos por John Napier em seu livro *Descrição do Maravilhoso Cânon dos Logaritmos*, publicada em 1614. Napier incluiu, neste livro, noventa páginas em que tabela valores do seno e seus respectivos logaritmos, com a finalidade de simplificar cálculos em geometria esférica.

Ainda que as incontáveis listas e tabelas de números produzidas desde a Antiga Mesopotâmia representem, de nosso ponto de vista, uma grande variedade de funções, o conceito matemático de função só foi formulado no século XVII, a partir do desenvolvimento da mecânica e do cálculo diferencial. A variação da posição e da velocidade de um objeto com o passar do tempo aparece claramente na obra de Galileu. Newton introduz o termo *fluent* para designar uma quantidade que varia com o tempo. Contudo, o uso da palavra *função* neste contexto parece ser devida a Leibniz. Em uma carta de 1698, endereçada a Leibniz, Johan Bernoulli concorda em utilizar esta palavra para designar “uma quantidade formada, de alguma maneira, começando por x e as constantes”. É também a Leibniz que devemos a introdução em matemática dos termos constante, variável e parâmetro. Mais detalhes e referências podem ser encontrados em (Bourbaki, 2004, p. 154).

Entretanto, a utilização do conceito de função como fundamento do cálculo infinitesimal remonta a *Introductio in analysin infinitorum* publicada por L. Euler em 1748 (Kline, 1990, p. 405). Ainda assim, 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, para cada valor atribuído à variável, seja possível deduzir um valor correspondente da função; (Dieudonné, 1986, 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 criação de funções cada vez mais exóticas, causando uma cascata de problemas que só foram completamente esclarecidos no século XX.

2. Função, domínio e imagem

Nossa primeira definição de função é uma elaboração da definição de Cauchy, citada na seção anterior. Tomaremos como ponto de partida dois conjuntos, que chamaremos de X e Y. A definição é a seguinte: uma *função* f de X em Y é uma regra que, a cada elemento de X associa um, e um único, elemento de Y. Assim, para que uma regra defina uma função,

(i) não pode haver elementos em X aos quais não é associado um elemento de Y;

(ii) a um elemento de X não pode ser associado mais de um elemento de Y.

O conjunto X é o *domínio* da função f e Y é seu contra-domínio. Abreviamos isto escrevendo

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

Frequentemente vamos nos referir aos elementos de X como sendo os *argumentos* ou *variáveis independentes* da função f. Se $x \in X$, denotamos por $f(x)$ o elemento associado a x pela regra que define a função f e diremos que $f(x)$ é a *imagem* de x por f.

No caso do pastor que conta suas ovelhas, o domínio da função é o rebanho e seu contra-domínio é o conjunto dos números naturais. Como nenhum rebanho é infinito, à maioria dos números naturais não estará associada a nenhuma ovelha do rebanho. Este exemplo também mostra que a "regra" que define uma função não precisa ser uma fórmula. Outro exemplo da mesma natureza é a função

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

que a cada letra do alfabeto latino $\mathcal{L}$ associa sua posição no alfabeto. Assim, $p(A) = 1$ 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 das funções mais simples, é frequentemente 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 $d(n) = 2n$. 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 $q(n) = n^2$. Certamente você conhece muitas outras funções cujo domínio são os números naturais, ainda que sob um nome diferente; porque uma sequência (também chamada de progressão)

$$a_1, a_2, a_3, \ldots,$$

formada por elementos de um conjunto Y, nada mais é que a função $\mathbb{N} \longrightarrow Y$ que ao número natural n associa o elemento $a_n \in Y$. Assim, as progressões aritmética e geométrica são funções $\mathbb{N} \longrightarrow \mathbb{Z}$: a cada número natural n a primeira destas funções associa o inteiro $a + n \cdot r$ e a segunda aq^n, em que $a, r, q \in \mathbb{Z}$.

Nem toda função pode ser descrita por regras tão simples; provavelmente o exemplo mais conhecido é a função *valor absoluto*, também conhecida como *função módulo*,

$$\mathrm{abs} : \mathbb{Z} \longrightarrow \mathbb{N} \cup \{0\}$$

que é definida por duas regras, dependendo se o número inteiro é positivo ou negativo,

$$\mathrm{abs}(n) = \begin{cases} -n & \text{quando } \ n < 0; \\ +n & \text{quando } \ n \geq 0. \end{cases}$$

Outra função muito interessante, e que requer duas regras distintas, dependendo do valor do argumento, é a *função de Collatz*, $C : \mathbb{N} \longrightarrow \mathbb{N}$, definida por

$$C(n) = \begin{cases} n/2 & \text{quando } n \text{ é par} \\ 3n+1 & \text{quando } n \text{ é ímpar.} \end{cases}$$

Esta função tem um comportamento extremamente enigmático, como veremos na seção 6. Naturalmente, nada nos impede de inventar funções definidas por quantas fórmulas diferentes desejarmos.

Especificar a regra que define a função é conveniente, mas nem sempre é possível. Na prática, muitas funções são definidas por suas propriedades. Um caso bem conhecido é o da *sequência de Fibonacci*, que é a função $F : \mathbb{N} \longrightarrow \mathbb{N}$ que satisfaz as propriedades

$$F(1) = F(2) = 1 \ \text{ e } \ F(n+1) = F(n) + F(n-1), \quad \text{quando } \ n \geq 3.$$

Surpreendentemente, a mesma função pode ser definida pela fórmula

$$F(n) = \frac{\phi^n - \psi^n}{\sqrt{5}},$$

em que

$$\phi = \frac{1+\sqrt{5}}{2} \ \text{ e } \ \psi = \frac{1-\sqrt{5}}{2}.$$

Embora às vezes seja chamada de fórmula de Binet, esta maneira de expressar $F(n)$ foi descoberta por Abraham de Moivre (1667–1754). E esta nem mesmo é a única fórmula que define $F(n)$, pois

$$F(n) = \frac{1}{\sqrt{5}} \lfloor \phi \rceil,$$

na qual $\lfloor r \rceil$ denota o inteiro mais próximo do número real r.

Este último exemplo chama a atenção para uma questão extremamente importante: quando dizer que duas funções são iguais? Como mostra nossa discussão da Função de Fibonacci, isto não pode depender da regra usada para definir a função, porque uma mesma associação de um elemento do domínio a um elemento do contradomínio pode ser efetuada usando regras totalmente distintas. Isto sugere a seguinte definição: as funções $f, g : X \longrightarrow Y$ são *iguais* se

$$\forall x \in X, (f(x) = g(x)).$$

Portanto, ao menos em princípio, podemos determinar se duas funções são iguais mesmo que sejam calculadas a partir de caixas pretas cujo funcionamento ignoramos.

É importante observar que, o fato de $f : X \longrightarrow Y$ ser uma função não significa que todo elemento de Y é imagem de um elemento de X. Por exemplo, no caso da função valor absoluto, só os números inteiros não negativos são imagem de algum inteiro por esta função. Isto sugere a seguinte definição. A imagem de $f : X \longrightarrow Y$ é o conjunto

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

Portanto,

$$\operatorname{im}(\mathrm{abs}) = \mathbb{N} \cup \{0\},$$

ao passo que

$$\operatorname{im}(p) = \{1, 2, \dots, 26\}$$

porque o alfabeto só tem 26 letras. Já para a função de Collatz temos que

$$\operatorname{im}(C) = \mathbb{N}$$

porque para todo $n \in \mathbb{N}$ temos que

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

uma vez que $2n$ é par.

A esta altura é necessário chamar sua atenção para dois pontos importantes. O primeiro é que uma função só está completamente definida quando explicitamos seu domínio, seu contradomínio e a regra que a define. O segundo é que, mesmo os matemáticos, raramente são tão cuidadosos e você encontrará, frequentemente, afirmações como

$$\text{considere a função } f(x) = \frac{x-2}{x-3},$$

sem nenhuma menção de quem seriam seu domínio e contradomínio. Isto ocorre normalmente em situações em que é possível concluir, a partir do contexto, quem deveriam ser estes conjuntos. Por exemplo, em um curso de cálculo, o argumento de uma função sempre é um número real e o mesmo acontece com suas imagens. Entretanto, pode ocorrer que seja necessário excluir do domínio alguns número reais nos quais a função não está definida. Assim, no exemplo acima, o domínio da função f seria escolhido como $\mathbb{R} \setminus \{3\}$, para evitar uma divisão por zero no denominador da fração que define a função.

3. Funções reais

Dada sua origem na mecânica e na geometria, onde imperam as variáveis contínuas, não é surpreendente que as funções estudadas no cálculo diferencial e integral tenham como domínio intervalos de $\mathbb{R}$. Como se trata de uma primeira introdução às funções reais, suporemos também que o contradomínio de nossas funções é sempre $\mathbb{R}$, ainda que, em mecânica, o plano e o espaço sejam contradomínios mais frequentes. Começaremos nossa investigação destas funções com alguns exemplos bem elementares.

A *função constante* $\mathbf{k}_c : \mathbb{R} \to \mathbb{R}$ é definida pela regra $\mathbf{k}_c(x) = c$, em que $c \in \mathbb{R}$. Assim, a imagem de qualquer número real por $\mathbf{k}_c$ é igual a c; donde

$$\mathrm{im}(\mathbf{k}_c) = \{c\}$$

Uma função constante que aparece com muita frequência 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* $\lambda : \mathbb{R} \to \mathbb{R}$ por

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

Para determinar a imagem de λ basta lembrar que um número real y pertence à imagem de λ se, e somente se, existe $x \in \mathbb{R}$ tal que $y = \lambda(x)$. Usando a regra que define λ isto equivale a

$$y = ax + b.$$

Resolvendo esta equação, obtemos

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

Logo todo elemento de $\mathbb{R}$ é imagem de algum número real por λ; donde

$$\mathrm{im}(\lambda) = \mathbb{R}.$$

A terceira família de exemplos que desejamos investigar é a das funções quadráticas. Uma *função quadrática* $q : \mathbb{R} \to \mathbb{R}$ é definida por

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

em que $a \neq 0$, b e c são números reais. Ao contrário das funções lineares, a imagem das funções quadráticas não coincide com todo o conjunto $\mathbb{R}$. Por exemplo, a imagem de $q(x) = x^2$ não contém nenhum número negativo, porque, como vimos na seção 4 do capítulo 3, o quadrado de um número real é sempre positivo. Não é difícil calcular a imagem de $q(x)$. Para isto, basta lembrar que $y \in \mathrm{im}(q)$ equivale a afirmar que existe $x \in \mathbb{R}$ tal que $q(x) = y$; isto é,

$$ax^2 + bx + c = y.$$

Considerando y como um número fixado, precisaríamos resolver a equação

$$ax^2 + bx + (c - y) = 0, \tag{49}$$

para encontrar os valores de x que são levados em y. Mas, para que um tal x exista é necessário que o discriminante desta equação seja positivo. No caso da equação (49), isto significa que

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

donde,

$$4ay \geq -b^2 + 4ac.$$

Denotando $b^2 - 4ac$ por Δ como usual, verificamos que

$$y \geq -\frac{\Delta}{4a}$$

quando $a > 0$. Por outro lado, quando $a < 0$,

$$y \leq -\frac{\Delta}{4a};$$

porque, neste caso, o sinal de desigualdade inverte quando dividimos por $4a$. Mostramos, assim, que

$$\mathrm{im}(q) = \begin{cases} [-\Delta/4a, +\infty) & \text{se } a > 0 \\ (-\infty, -\Delta/4a] & \text{se } a < 0. \end{cases} \tag{50}$$

Faremos uma análise mais detalhada desta função na seção 3 do capítulo 7.

As três famílias de funções reais que definimos até aqui são exemplos de funções polinomiais. Em geral, uma *função polinomial* é definida por uma regra da forma

$$p(x) = a_n x^n + \cdots + a_1 x + a_0,$$

em que $n \geq 0$ é um número inteiro e $a_0, \ldots, a_n$ são números reais. Como o cálculo da imagem de um número real por uma função polinomial requer apenas a adição, subtração e multiplicação, seu domínio é sempre todo o conjunto $\mathbb{R}$. A determinação da imagem de uma função polinomial geral requer um conhecimento básico do cálculo diferencial e, por isso, está além do que é possível fazer em um livro elementar como este.

Quando a divisão aparece entre as operações utilizadas para definir uma função real, torna-se necessário excluir do domínio aqueles números para os quais o denominador se anula. Por exemplo, os domínios das funções r_1, r_2 e r_3 definidas por

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

são, respectivamente, $\mathbb{R} \setminus \{0\}$, $\mathbb{R} \setminus \{2\}$ e $\mathbb{R} \setminus \{-\sqrt{2}, \sqrt{2}\}$. Contudo uma função pode ser definida através de uma divisão mas, mesmo assim, ter $\mathbb{R}$ como domínio, como é o caso de

$$r_4(x) = \frac{1}{x^2+1},$$

já que o polinômio $x^2 + 1$ não tem nenhuma raiz real. Estas funções são chamadas de *racionais* porque, assim como os números racionais são quocientes de números inteiros, elas são definidas por quocientes de polinômios.

Neste ponto, é conveniente chamar sua atenção para a diferença, nem sempre respeitada quando se tratam de funções reais, entre o nome de uma função e seu valor no argumento. 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 os 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", costume ao qual, por ser prático, vamos aderir ao longo deste livro. Contudo, é útil ter em mente esta diferença, por duas razões. A primeira, é que muitas vezes é necessário diferenciar entre o nome da função e seu valor em x; a segunda é que, ao identificar uma função com sua fórmula, omitimos completamente a especificação do seu domínio. 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 $\sqrt{} : [0, +\infty] \to \mathbb{R}$ 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 reais que consideramos foram definidas por uma única regra em todo o seu domínio. Como na seção 2, a função mais conhecida, à qual isto não se aplica, é $\text{abs} : \mathbb{R} \to \mathbb{R}$, definida por

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

chamada *função valor absoluto* ou *função módulo*. Um detalhe importante: não confunda *função módulo* com *função modular*. Estas últimas são funções de variável complexa, as mais conhecidas das quais são as funções elípticas, que têm este nome porque medem o comprimento de um arco de elipse.

Uma função bem mais exótica, mas que também requer duas regras, é a *função de Dirichlet* $\mathcal{D} : \mathbb{R} \longrightarrow \mathbb{R}$ 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,

$$f(x) = \begin{cases} 0 & \text{se } x \leq -1; \\ -x+1 & \text{se } -1 \leq x \leq 0; \\ 1-2x & \text{se } 0 < x \leq 1; \\ -1 & \text{se } x \geq 1; \end{cases}$$

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

Como os números reais podem ser somados, subtraídos, multiplicados e divididos, podemos definir operações correspondentes para as funções reais. Sejam, então, f_1 e f_2 duas funções cujo domínio é um conjunto $X \subset \mathbb{R}$ e cujo contradomínio é $\mathbb{R}$. Definimos 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

$$(f_1 + f_2)(x) = f_1(x) + f_2(x) \quad \text{e} \quad (f_1 \cdot f_2)(x) = f_1(x) \cdot f_2(x).$$

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 $g_1, g_2 : \mathbb{R} \longrightarrow \mathbb{R}$ são as funções

$$g_1(x) = x^2 + 2 \quad \text{e} \quad g_2(x) = -x^2 + 2x - 1 \tag{51}$$

então $g_1 + g_2 : \mathbb{R} \longrightarrow \mathbb{R}$ é definida por

$$(g_1 + g_2)(x) = g_1(x) + g_2(x) = (x^2 + 2) + (-x^2 + 2x - 1) = 2x + 1,$$

e $g_1 \cdot g_2 : \mathbb{R} \to \mathbb{R}$ por

$$(g_1 \cdot g_2)(x) = g_1(x) \cdot g_2(x) = (x^2 + 2) \cdot (-x^2 + 2x - 1) = -x^4 + 2x^3 - 3x^2 + 4x - 2.$$

Note que não há necessidade de definir separadamente a multiplicação de uma função por um número real, nem a subtração de funções, porque ambas se reduzem às duas operações definidas acima. Assim, supondo que $a \in \mathbb{R}$ e que $\mathbf{k}_a : X \to \mathbb{R}$ é a função constante $\mathbf{k}_a(x) = a$, para todo $x \in \mathbb{R}$, denotaremos por af_1 a função

$$af_1 = \mathbf{k}_a \cdot f_1$$

e por $f_1 - f_2$ a função

$$f_1 + (\mathbf{k}_{(-1)} \cdot f_2);$$

de modo que

$$(af_1)(x) = af_1(x) \quad \text{e} \quad (f_1 - f_2)(x) = f_1(x) - f_2(x).$$

A adição e a multiplicação de funções satisfazem as mesmas propriedades que as operações de mesmo nome em $\mathbb{R}$. Não que isto seja surpreendente, porque, afinal, as primeiras são definidas diretamente a partir das segundas. Por exemplo, a adição e a multiplicação são comutativas porque, como $f_1(x)$ e $f_2(x)$ são números reais, para todo $x \in X$, a comutatividade dos números reais nos permite concluir que

$$f_1(x) + f_2(x) = f_2(x) + f_1(x) \quad \text{e} \quad f_1(x) \cdot f_2(x) = f_2(x) \cdot f_1(x)$$

para todo $x \in X$. Logo, pelas definições da adição e da multiplicação de funções, que

$$f_1 + f_2 = f_2 + f_1 \quad \text{e} \quad f_1 \cdot f_2 = f_2 \cdot f_1,$$

mostrando que, de fato, estas operações com funções são comutativas. As funções polinomiais podem ser caracterizadas como aquelas que podem ser construídas combinando as funções constantes e a função identidade $\mathrm{id}(x) = x$, usando apenas a adição e multiplicação de funções. Em particular, isto nos permite concluir que a soma e o produto de funções polinomiais tem que ser uma função polinomial.

Supondo, ainda, que f_1 e f_2 são funções de $X \subset \mathbb{R}$ em $\mathbb{R}$, podemos definir a divisão de funções por

$$\left(\frac{f_1}{f_2}\right)(x) = \frac{f_1(x)}{f_2(x)}.$$

Ao contrário da adição e multiplicação de funções, a divisão nem sempre produz uma função de X em $\mathbb{R}$, porque precisamos excluir de seu domínio os elementos de X em que f_2 se anula. Portanto, domínio de f_1/f_2 é o conjunto

$$X \smallsetminus \{x \in X \mid f_2(x) = 0\}.$$

A importância das operações com funções torna-se mais clara nos cursos de cálculo, nos quais reinam as chamadas *funções elementares*, que são obtidas combinando a função identidade e as funções constantes e trigonométricas, assim como exponenciais e logaritmos, usando as operações de adição, multiplicação, divisão e composição de funções.

4. Funções com vários argumentos

Como vimos na seção 1, uma das primeiras manifestações da ideia de função foi sob a forma de tabelas que relacionavam entre si duas quantidades. Uma outra maneira pela qual as funções apareceram ao longo do tempo, e que hoje em dia usamos intensamente, são as tabelas de duas entradas, também conhecidas como planilhas. Por exemplo na planilha 1 temos as notas, em duas provas, de três pessoas de uma turma.

Nome	Prova 1	Prova 2
Andrea	8,8	9,5
Daniel	10,0	9,8
Severino	5,4	6,2

TABELA 1. Uma planilha de notas

A pergunta que imediatamente se põe é: como expressar como uma função algo que associa um número a dois argumentos diferentes? A resposta é simples: basta produzir um conjunto cujos elementos são compostos por estes dois argumentos. Faremos isto introduzindo uma nova operação entre conjuntos, o produto cartesiano.

Sejam X e Y dois conjuntos. O *produto cartesiano* $X \times Y$ é o conjunto formado pelos pares (x, y) em que $x \in X$ e $y \in Y$ que, vamos supor, satisfazem a propriedade

$$\forall x, x' \in X, (\forall y, y' \in Y, ((x', y') = (x, y) \Longrightarrow ((x = x') \wedge (y = y')))). \tag{52}$$

Em outras palavras, trata-se de *pares ordenados*, uma vez que a ordem em que os elementos são escritos é essencial para determinar a igualdade entre dois de tais pares. Vamos nos referir aos elementos x e y como a primeira e a segunda *coordenadas* do par (x, y). Por exemplo, se

$$X = \{1, 2, 3, 4\} \quad \text{e} \quad Y = \{a, b, c\}$$

então $X \times Y$ é o conjunto

$$\{(1,a),(1,b),(1,c),(2,a),(2,b),(2,c),(3,a),(3,b),(3,c),(4,a),(4,b),(4,c)\}.$$

Com isto estamos prontos para representar planilhas como funções.

Sejam X e Y os conjuntos cujos elementos indexam, respectivamente, as linhas e as colunas de uma dada planilha. Cada casa da planilha pode ser, então, identificada com um par do produto cartesiano $X \times Y$. Como cada é preenchida com um único elemento de um conjunto Z, podemos usar esta associação para definir uma função

$$f : X \times Y \longrightarrow Z.$$

Por exemplo, no caso da planilha 1, temos que

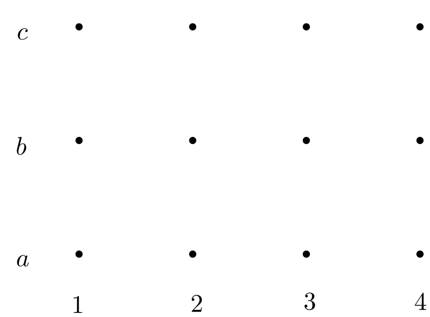
$$X = \{\text{Andrea, Daniel, Severino}\} \ \text{ e } \ Y = \{\text{Prova 1, Prova 2}\}$$

e podemos tomar $Z = \mathbb{R}$.

Tudo isto sugere que o produto cartesiano pode ser representado como uma tabela de duas entradas, com as colunas identificadas pelos elementos de X e as linhas pelos elementos de Y. Aplicando esta interpretação ao conjunto acima obtemos a seguinte tabela.

c	$(1,c)$	$(2,c)$	$(3,c)$	$(4,c)$
b	$(1,b)$	$(2,b)$	$(3,b)$	$(4,b)$
a	$(1,a)$	$(2,a)$	$(3,a)$	$(4,a)$
	1	2	3	4

Uma variação desta interpretação consiste em associar a cada par ordenado um ponto do plano. No exemplo acima, obtemos uma malha de 12 pontos, um para cada par ordenado, como ilustrado na figura abaixo.



Ao usar esta interpretação, vamos nos referir às duas entradas dos pares ordenados como as *coordenadas* do ponto que corresponde a este par. A primeira coordenada será sempre representada ao longo da horizontal e a segunda ao longo da vertical.

Se X e Y são conjuntos, um subconjunto do produto cartesiano $X \times Y$ é conhecido como uma *relação*. É possível que você esteja se perguntando: o que isto tem a ver com o que costumamos chamar de relação como $=$, $\neq$ e $\leq$? À primeira vista, pode parecer que se trata de mais um caso do fenômeno descrito pelo poeta alemão J. W. Goethe; segundo ele,

> o que quer que você diga [aos matemáticos], eles traduzem em sua própria linguagem e, consequentemente, aquilo se torna algo totalmente diferente.

Neste caso, contudo, o que costumamos chamar de relação entre dois objetos é razoavelmente bem traduzido pelo conceito matemático e vice-versa. Considere, por exemplo, a relação entre pessoas e bairros, estabelecida pelo fato de que uma dada pessoa mora em um dado bairro. Se Ana, João e Marcos moram em Botafogo, Maria mora na Penha, ao passo que Paulo mora na Tijuca, podemos definir

$$X = \{\text{Ana}, \text{João}, \text{Marcos}, \text{Paulo}, \text{Maria}\} \ \text{ e } \ Y = \{\text{Botafogo}, \text{Penha}, \text{Tijuca}\}.$$

e representar a relação *mora em* pelo subconjunto

$$\{(\text{Ana}, \text{Botafogo}), (\text{João}, \text{Botafogo}), (\text{Marcos}, \text{Botafogo}), (\text{Paulo}, \text{Tijuca}), (\text{Maria}, \text{Penha})\}$$

Note que podemos, facilmente, descobrir "quem mora onde" a partir deste conjunto e, portanto, reconstruir a relação como originalmente dada.

Vejamos alguns exemplos de natureza mais matemática. Quando

$$X = \{1, 2, 3, 4, 5\}$$

o subconjunto correspondente à relação $=$ definida em X é

$$\mathcal{R}_1 = \{(1,1), (2,2), (3,3), (4,4), (5,5)\},$$

ao passo que o subconjunto correspondente à relação $<$ no mesmo conjunto é

$$\mathcal{R}_2 = \{(1,2), (1,3), (1,4), (1,5), (2,3), (2,4), (2,5), (3,4), (3,5), (4,5)\};$$

já o subconjunto correspondente à relação $\leq$ em X é

$$\mathcal{R}_1 \cup \mathcal{R}_2.$$

Para um exemplo mais substancial, considere o subconjunto

$$\mathcal{P} = \{(r, r') \in \mathbb{L} \times \mathbb{L} \mid r \cap r' = \varnothing\}, \tag{53}$$

em que $\mathbb{L}$ denota o conjunto das retas do plano. Como duas retas sem ponto em comum são *paralelas*, $\mathcal{P}$ define a relação de paralelismo entre retas. Escrevendo

$$r \parallel r' \iff (r, r') \in \mathcal{P},$$

obtemos a representação usual da noção de paralelismo.

A maneira mais geral e abstrata de definir funções é como um tipo especial de relação. Nosso ponto de partida será o conceito de *gráfico de uma função* $f : X \longrightarrow Y$, que é o conjunto

$$\Gamma_f = \{(x, y) \in X \times Y \mid y = f(x)\}.$$

Como se trata de um subconjunto do produto cartesiano $X \times Y$, o gráfico Γ_f é um exemplo de relação. Neste ponto convém fazer uma pausa, porque você talvez tenha sido pego de surpresa por nossa definição de gráfico de uma função. Afinal, onde, provavelmente, esperava uma figura, acabou encontrando um subconjunto de um produto cartesiano. Na verdade, o que costumamos chamar de "gráfico" de uma função é a representação geométrica do gráfico definido como um subconjunto de produto cartesiano. Como vimos acima, um produto cartesiano pode ser representado como uma malha de pontos do plano; desenhamos o gráfico da função f indicando apenas aqueles pontos que pertencem a Γ_f. Por exemplo, a representação geométrica do gráfico da função

$$d : \{1, 2, 3, 4, 5, 6, 7\} \longrightarrow \{1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14\}$$

que a cada $n \in \{1, 2, 3, 4, 5, 6, 7\}$ associa seu dobro é ilustrado na figura 1; do lado esquerdo da qual está ilustrado parte do gráfico da função de Fibonacci. O "L" foi acrescentado à esquerda de cada figura à guisa de "sistema de eixos" e não faz parte do gráfico; estão lá apenas para facilitar a identificação da escala das figuras.

Voltando à questão da definição de uma função como uma relação, a ideia é que, a partir do gráfico de uma função podemos, facilmente reconstruir a função. Para entender como fazer isto, digamos que $\Gamma_f \subset X \times Y$ é o gráfico de uma função. A primeira coisa a observar é que, como Γ_f é um subconjunto de $X \times Y$, então f deve ser uma função cujo domínio é X e cujo contradomínio é Y. Para reconstruir a regra que define a função, procedemos da seguinte maneira. Dado um $x \in X$, procuramos o par em Γ_f cujo primeiro elemento é x. Se y é a segunda coordenada deste par, então $f(x) = y$. Por exemplo, se

$$X = \{1, 2, 3, 4, 5\} \quad \text{e} \quad \Gamma_f = \{(1, 1), (2, 4), (3, 9), (4, 16), (5, 25)\} \subset X \times \mathbb{N},$$

então $f : X \longrightarrow \mathbb{N}$ é a função cuja regra é dada por

$$f(1) = 1, \; f(2) = 4, \; f(3) = 9, \; f(4) = 16 \;\text{ e }\; f(5) = 25;$$

isto é, f leva cada elemento de X em seu quadrado.

Uma questão, relacionada ao gráfico de uma função, que é bem mais interessante consiste em determinar que relações correspondem a gráficos de funções. Para isso precisamos relembrar que, para que uma regra defina uma função duas coisas devem ocorrer:

(1) a cada elemento do seu domínio deve ser associado um elemento do contradomínio;

(2) a um mesmo elemento do domínio não podem ser associados dois elementos do contradomínio.

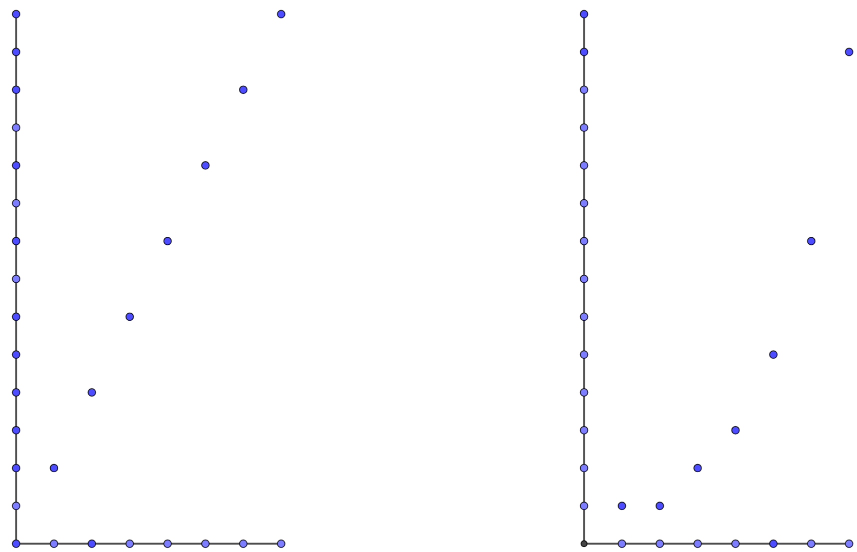


FIGURA 1. Gráficos das funções dobro e Fibonacci.

Para descobrirmos que relações são gráficos de funções, precisamos identificar como estas duas propriedades se refletem no gráfico da função. Seja $f : X \longrightarrow Y$ uma função. A propriedade (1) equivale a dizer que se $x \in X$, então existe um par $(x, y) \in \Gamma_f$; já a propriedade (2) garante que há um único par com esta propriedade. Com isto podemos dar uma definição totalmente abstrata de função, sem nenhuma menção de qualquer regra. Uma relação $\mathcal{F} \subset X \times Y$ é uma *função* se

$$\forall x \in X, (\exists y \in Y, ((x, y) \in \mathcal{F})) \tag{54}$$

mas, também que

$$\forall x \in X, (\forall y, y' \in Y, ((((x, y) \in \mathcal{F}) \wedge ((x, y') \in \mathcal{F})) \Longrightarrow (y = y'))) \tag{55}$$

Definir uma função a partir do seu gráfico tem a vantagem de, ao menos em princípio, ser independente de qualquer regra. O *ao menos em princípio* é necessário porque, a não ser que o domínio seja finito, não é possível apresentar todos os pares que constituem o gráfico da função. A saída é descrever o gráfico a partir de alguma propriedade comum aos seus pontos; isto é, a partir de uma regra.

Chegados a este ponto você pode estar pensando, talvez já sem paciência: 'por que ele não ilustra todas as funções como gráficos? Ia ser muito mais fácil". Sei disso e concordo;

então, por que não há o desenho de nenhum gráfico de função real à vista? A resposta é que o compromisso com o enfoque (mais ou menos) axiomático não me permite. Para desenhar gráficos de funções precisamos conhecer os princípios básicos da geometria; o que só faremos no capítulo 5. Por isso, só no capítulo 7 estaremos aptos a desenhar os gráficos que desejamos, sem infringir o acordo tácito sob o qual estamos operando.

Par encerrar, voltaremos aos pares ordenados com os quais começamos a seção. A definição desta noção que apresentamos é puramente operacional. Em outras palavras não definimos *o que é* um par ordenado, mas sim como os pares ordenados *se comportam*. Podemos contornar este problema, definindo o par ordenado (x,y) como sendo o conjunto $\{\{x\},\{x,y\}\}$. Naturalmente, para confirmar que esta definição é satisfatória precisamos provar que a propriedade (52) vale para estes conjuntos. Digamos que

$$x, x' \in X \quad \text{e} \quad y, y' \in Y,$$

mas também que

$$\{\{x\},\{x,y\}\} = \{\{x'\},\{x',y'\}\}.$$

Suponhamos, primeiramente, que $x = x'$. Neste caso, $\{x\} = \{x'\}$ e podemos ter que

$$\{x,y\} = \{x'\} \quad \text{ou} \quad \{x,y\} = \{x',y'\}.$$

Pela definição de igualdade de conjuntos, o primeiro caso só é possível se $x = y = x'$; donde

$$\{\{x\},\{x,x\}\} = \{\{x\},\{x,y'\}\}. \tag{56}$$

Como

$$\{\{x\},\{x,x\}\} = \{\{x\}\},$$

podemos concluir de (56) que $y' = x$. Portanto, no primeiro caso, $x = x' = y = y'$, o que prova (52). Por outro lado,

$$\{x,y\} = \{x',y'\}$$

equivale a

$$\{x,y\} = \{x,y'\},$$

pois estamos supondo que $x' = x$. Logo, $y' = y$ ou $y' = x$. Se $y' = y$, provamos (52) ao passo que, $y' = x$ nos dá

$$\{x,y\} = \{x\};$$

donde $y = x$; isto é $y = x = y' = x'$, que é um caso particular de (52). Finalmente, $x \neq x'$ implica que

$$\{x\} \neq \{x'\},$$

de modo que

$$\{x\} = \{x',y'\}.$$

Mas isto implica que $x' \in \{x\}$, o que só seria possível se $x = x'$; o que mostra que este caso não pode ocorrer e completa a prova de que a propriedade (52) vale se usarmos $\{\{x\},\{x,y\}\}$ como definindo (x,y).

5. Funções inversas

Nesta seção almejamos responder à pergunta: *que condições uma função $f : X \longrightarrow Y$ deve satisfazer para que a regra*

$$g(y) = x \iff f(x) = y, \tag{57}$$

defina uma função $g : Y \longrightarrow X$? Em outras palavras, queremos saber para que funções é possível inverter o sentido da seta de $X \longrightarrow Y$ para $Y \longrightarrow X$ e continuar obtendo uma função.

Para responder a esta pergunta precisamos determinar quais devem ser as propriedades de f para que g respeite as condições (i) e (ii) da página 79, que podemos reformular na forma:

(i) não há elementos em Y aos quais g não associa um elemento de X;
(ii) um elemento de Y não é associado por g a mais de um elemento de X.

Começaremos analisando (i), que equivale a

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

Usando a equivalência em (57), isto é o mesmo que

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

Em outras palavras, todo elemento de Y é imagem de algum elemento de X. Logo,

$$\operatorname{im}(f) = \{y \in Y \mid f(x) = y\} = Y.$$

Quando isto acontece, dizemos que f é *sobrejetiva* ou *sobrejetora*. Como vimos na seção 2, a função de Collatz é sobrejetiva, mas não a função valor absoluto.

Passando a condição (ii), podemos reformulá-la como

$$\forall y \in Y, (((x = g(y)) \wedge (x' = g(y))) \Longrightarrow (x = x')).$$

Procedendo como na análise de (i), usaremos (57), para reescrever esta sentença como

$$\forall y \in Y, (((f(x) = y) \wedge (f(x') = y)) \Longrightarrow (x = x'));$$

segundo a qual dois elementos de X só podem ter a mesma imagem por f se forem iguais. Quando isto ocorre dizemos que f é *injetiva* ou *injetora*.

Nem a função valor absoluto, nem a função de Collatz são injetivas, porque $\operatorname{abs}(-1) = \operatorname{abs}(1)$, ao passo que

$$C(3) = 3 \cdot 3 + 1 = 10 = \frac{20}{2} = C(20).$$

Uma função injetiva bem conhecida é a função linear

$$\lambda : \mathbb{R} \longrightarrow \mathbb{R}$$

definida por $\lambda(x) = ax + b$, com $a, b \in \mathbb{R}$ e $a \neq 0$. Para mostrar que esta função é injetiva, suponhamos que $\lambda(x) = \lambda(x')$. Usando a definição, obtemos

$$ax + b = ax' + b.$$

Subtraindo b dos dois lados e multiplicando o resultado por $1/a$, encontramos $x = x'$, confirmando nossa afirmação de que esta função é realmente injetiva. Contudo, como vimos na seção 3, $\operatorname{im}(\lambda) = \mathbb{R}$ de modo que esta função também é sobrejetiva. Quando uma função é, simultaneamente, sobrejetiva e injetiva, dizemos que é *bijetiva*.

Passando à função quadrática; já sabemos de (50), que a imagem desta função não é todo o conjunto dos números reais. Portanto, quando $a \neq 0$, a função

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

não é sobrejetiva se considerarmos seu contradomininío como sendo todo o conjunto dos números reais. Para determinar se $q(x)$ é injetiva, suponhamos que $q(x) = q(x')$; isto é, que

$$ax^2 + bx + c = a(x')^2 + bx' + c.$$

Subtraindo o segundo lado do primeiro, obtemos

$$a(x^2 - (x')^2) + b(x - x') = 0. \tag{58}$$

Usando produtos notáveis, podemos escrever

$$(x^2 - (x')^2) = (x - x')(x + x').$$

Aplicando isto em (58), encontramos

$$(x - x')(a(x + x') + b) = 0.$$

Logo, $q(x) = q(x')$ pode ocorrer quando $x \neq x'$, desde que

$$a(x + x') + b = 0$$

que equivale a dizer que

$$x' = \frac{-ax - b}{a}.$$

Portanto, a função quadrática nunca é injetiva se estamos supomos que seu domínio é todo o conjunto $\mathbb{R}$.

A análise do início desta seção pode ser resumida no seguinte teorema; cuja recíproca será provada na próxima seção.

TEOREMA 5.1. *Se uma função $f : X \longrightarrow Y$ é bijetiva então a regra $g(y) = x$, sempre que $f(x) = y$, define uma função $g : Y \longrightarrow X$.*

A função g do teorema acima é a *inversa* de f e é normalmente denotada por f^{-1}. Por exemplo, para determinar a inversa da função linear $\lambda(x) = ax + b$ basta lembrar que

$$y = ax + b \Longleftrightarrow x = \frac{y - b}{a},$$

o que nos permite afirmar que

$$\lambda^{-1} : \mathbb{R} \longrightarrow \mathbb{R}$$

é definida pela regra

$$\lambda^{-1}(y) = \frac{y-b}{a}.$$

Embora seja sugestivo usar y para o argumento da função inversa, é claro que a mesma regra pode ser definida usando qualquer nome para sua variável independente.

Cuidado para não confundir a função inversa f^{-1} com a função definida pela regra $1/f(x)$. Por exemplo,

$$\lambda^{-1}(x) = \frac{x-b}{a} \quad \text{mas} \quad \frac{1}{\lambda(x)} = \frac{1}{ax+b}.$$

Neste caso, o domínio de λ^{-1} é $\mathbb{R}$; ao passo que o domínio de $1/\lambda(x)$ é $\mathbb{R} \setminus \{-b/a\}$.

Outra função bijetiva cuja inversa podemos calcular facilmente é

$$g : \mathbb{Q} \setminus \{0\} \longrightarrow \mathbb{Q} \setminus \{0\},$$

definida por $g(x) = 1/x$. Se y é uma fração não nula, então $y = 1/x$ equivale a $x = 1/y$. Logo, $y = g(1/y)$, de modo que todo elemento do contradomínio é imagem de um único elemento do domínio. Portanto, g é bijetiva e sua inversa é a função

$$g^{-1} : \mathbb{Q} \setminus \{0\} \longrightarrow \mathbb{Q} \setminus \{0\}$$

definida por $g^{-1}(y) = 1/y$. Em particular, neste caso, g é sua própria inversa. Um exemplo semelhante a este, mas um pouco mais elaborado, é dado pela função

$$h : \mathbb{Q} \setminus \{3\} \longrightarrow \mathbb{Q} \setminus \{1\}$$

cuja regra é

$$h(x) = \frac{x+2}{x-3}.$$

Supondo que $y \neq 1$ é um elemento do contradomínio de h, temos que y estará na imagem de h se, e somente se,

$$y = \frac{x+2}{x-3}.$$

Resolvendo esta equação relativamente a x, encontramos

$$x = \frac{3y+2}{y-1}.$$

Logo, como no exemplo anterior, para cada $y \neq 1$ existe um único x tal que $h(x) = y$. Portanto, h é bijetiva e

$$h^{-1} : \mathbb{Q} \setminus \{1\} \longrightarrow \mathbb{Q} \setminus \{3\}$$

tem por regra

$$h^{-1}(y) = \frac{3y+2}{y-1}.$$

Ao longo deste livro encontraremos muitos outros exemplos de funções inversíveis, especialmente nos dois capítulos finais.

Para encerrar, faremos uma aplicação do que vimos nesta seção. Se existe uma função injetiva $f : X \longrightarrow Y$, entre dois conjuntos *finitos* X e Y, então à quantidade de elementos em X é menor ou igual a quantidade de elementos de Y. Para entender o porquê disto, basta lembrar que, sendo f injetiva, as imagens de elementos distintos de X são necessariamente distintas. Assim,

$$\#X = \#\operatorname{im}(f) \le \#Y;$$

pois $\operatorname{im}(f) \subset Y$. No caso particular em que f é bijetiva, sua inversa existe e também é bijetiva. Com isso, temos também que

$$\#Y = \#\operatorname{im}(f^{-1}) \le \#X,$$

o que prova o seguinte resultado.

TEOREMA 5.2. *Dois conjuntos finitos têm a mesma quantidade de elementos se existir uma função bijetiva que tem um deles como domínio e o outro como contradomínio.*

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 à 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 ser 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

$$\phi_n : B^n \longrightarrow \mathcal{P}(S)$$

definida pela regra

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

Por exemplo,

$$\phi_5(1, 0, 1, 1, 0) = \{1, 3, 4\}$$

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

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

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

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

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,

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

e, no segundo,

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

Como, em ambos os casos,

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

podemos concluir que ϕ_n 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.

6. Composição de funções

Sejam $f : X \longrightarrow Y$ e $g : Z \longrightarrow W$ duas funções. Quando a imagem de f está contida no domínio de g, podemos calcular

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

e usar isto para definir uma nova função, a *composta*

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

pela regra

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

Podemos ilustrar a definição usando o seguinte diagrama

$$\begin{tikzcd}
X \arrow[r, "f"] \arrow[ddr, "g \circ f"'] & \mathrm{im}(f) \subset Y \arrow[d, phantom, "\cap"] \\
& Z \arrow[d, "g"] \\
& W
\end{tikzcd}$$

Quando $g \circ f$ existe, seu domínio é sempre igual ao domínio de f e seu contradomínio ao contradomínio de g.

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*.

Alguns exemplos ajudarão a entender melhor a definição. Na seção 2, definimos as funções

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

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

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

é a função definida pela regra

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

para toda letra maiúscula ℓ. Assim,

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

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. Para um exemplo em que ambas as compostas estão definidas, considere a função $q : \mathbb{Z} \longrightarrow \mathbb{Z}$, definida por $q(n) = n^2$. Como d e q têm $\mathbb{Z}$ como contradomínio, temos que

$$q \circ d : \mathbb{Z} \longrightarrow \mathbb{Z} \text{ e } d \circ q : \mathbb{Z} \longrightarrow \mathbb{Z}$$

são definidas pelas regras

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

Note que $q \circ d \neq d \circ q$, porque, para que duas funções sejam iguais, é necessário que cada elemento do domínio tenha a mesma imagem por cada uma delas, 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

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

três funções. Como

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

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

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

Quando $f : X \longrightarrow X$ tem contradomínio igual a seu domínio, podemos 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 a definição recursiva:

$$f^n(x) = \begin{cases} f(x) & \text{quando } n = 1 \\ f^{n-1}(f(x)) & \text{quando } n > 1, \end{cases} \tag{59}$$

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.

Vejamos o que ocorre quando calculamos potências da função de Collatz, definida na seção 2. Por exemplo,

$$C(3) = 10, \quad C(10) = 5 \ \text{ e } \ C(5) = 16,$$

donde, pela fórmula recursiva (59),

$$C^3(3) = C(C^2(3)).$$

Usando a mesma forma novamente, desta vez para C^2, verificamos que

$$C^2(3) = C(C(10)) = C(5) = 16,$$

logo

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

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

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

Fazendo a mesma coisa com as imagens de 7, obtemos

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

e estamos de volta em 1. Até onde foi possível estender os cálculos, verifica-se que todas as potências consecutivas de C, calculadas em um determinado argumento, acabam por atingir 1, ainda que, como nestes dois casos, o expoente da potência que produz 1 dependa do valor escolhido para o argumento. Outros exemplos:

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

Não existe uma demonstração de que isto sempre acontece. O matemático Paul Erdös (1913-1996) sugeriu que "talvez a matemática ainda não esteja madura o suficiente para abordar problemas como este".

Assim como a adição e multiplicação de números inteiros admitem um elemento neutro, o mesmo ocorre para a composição de funções. Dado um conjunto X, definimos a *função identidade* $\mathrm{id}_X : X \longrightarrow X$ pela regra

$$\mathrm{id}_X(x) = x.$$

Se Y é um segundo conjunto e $f : X \longrightarrow Y$ é uma função, então, pela definição de composição de funções,

$$(f \circ \mathrm{id}_X)(x) = f(\mathrm{id}_X(x)) = f(x)$$

vale para todo x. Logo, pela definição de igualdade de funções

$$f \circ \mathrm{id}_X = f.$$

Note que não podemos calcular $\mathrm{id}_X \circ f$, porque id_X tem X como domínio, ao passo que o contradomínio de f é Y. Contudo, nada nos impede de construir a função $\mathrm{id}_Y : Y \longrightarrow Y$, definida pela regra $\mathrm{id}_Y(y) = y$. Com isso,

$$(\mathrm{id}_Y \circ f)(x) = \mathrm{id}_Y(f(x)) = f(x),$$

que é perfeitamente válido, pois $f(x) \in Y$. Desta última equação e da definição de igualdade de funções podemos concluir que

$$\mathrm{id}_Y \circ f = f.$$

Portanto há duas identidades que podemos compor com uma função de X em Y: id_X, que pode ser composta à direita, e id_Y, que pode ser composta à esquerda.

Na seção 5, vimos que se $f : X \longrightarrow Y$ é uma função bijetiva então sua inversa $f^{-1} : Y \longrightarrow X$ satisfaz

$$y = f(x) \Longleftrightarrow x = f^{-1}(y).$$

Podemos expressar isto em termos da composição de funções, observando que, pela equivalência acima,

$$f^{-1}(f(x)) = f^{-1}(y) = x \ \text{ e } \ f(f^{-1}(y)) = f(x) = y,$$

que podemos reescrever na forma

$$f^{-1} \circ f = \mathrm{id}_X \ \text{ e } \ f \circ f^{-1} = \mathrm{id}_Y.$$

Por exemplo, quando $f : \mathbb{Z} \longrightarrow \mathbb{Z}$ é a função definida por $f(x) = x + 1$, temos que $f^{-1}(y) = y - 1$, donde

$$(f \circ f^{-1})(y) = f(f^{-1}(y)) = f(y - 1) = (y - 1) + 1 = y,$$

ao passo que

$$(f^{-1} \circ f)(x) = f^{-1}(f(x)) = f^{-1}(x + 1) = (x + 1) - 1 = x,$$

como esperávamos. Talvez lhe surpreenda descobrir que uma função pode ter uma inversa à esquerda, mas não à direita, e vice-versa. Por exemplo, a função

$$g : \{0, 1, 2, 3, 4\} \longrightarrow \mathbb{N},$$

definida por $g(n) = n + 5$ é uma inversa à direita da função

$$f : \mathbb{N} \longrightarrow \{0, 1, 2, 3, 4\}$$

que leva cada número natural n em seu resto na divisão por 5, porque, se $0 \leq r \leq 4$, então

$$(f \circ g)(r) = f(r + 5) = r,$$

pois, como $0 \leq r \leq 4$, ele terá que ser o resto da divisão de $r + 5$ por 5. Contudo, embora

$$g \circ f : \mathbb{N} \longrightarrow \mathbb{N}$$

esteja bem definida, a regra

$$(g \circ f)(n) = g(r) = r + 5,$$

em que r é o resto da divisão de n por 5, não é igual a n para nenhum $n > 10$. Portanto,

$$g \circ f \neq \mathrm{id}_{\mathbb{N}}.$$

Logo, f tem uma inversa à direita, que não é inversa à esquerda. Mais detalhes sobre inversas à direita e à esquerda podem ser encontradas nos exercícios 20 e 21. Encerraremos a seção provando a recíproca do Teorema 5.1 da seção anterior.

TEOREMA 6.1. *Toda função inversível (à direita e à esquerda) é bijetiva.*

DEMONSTRAÇÃO. Seja $f : X \longrightarrow Y$ uma função inversível e seja

$$f^{-1} : Y \longrightarrow X$$

sua inversa. Como

$$f \circ f^{-1} = \mathrm{id},$$

temos que se $y \in Y$, então

$$f(f^{-1}(y)) = \mathrm{id}(y) = y.$$

Logo, f leva $f^{-1}(y) \in X$ em y, o que mostra que f é sobrejetiva. Por outro lado se $x, x' \in X$ são tais que

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

então, compondo à esquerda com f^{-1} dos dois lados da equação, encontramos

$$f^{-1}(f(x)) = f^{-1}(f(x')).$$

Como $f^{-1} \circ f = \mathrm{id}$, isto equivale a dizer que $x = x'$. Portanto, f também é injetiva. □

Combinando os Teoremas 5.1 e 6.1, obtemos o seguinte resultado, que resume o que fizemos de mais importante nesta seção.

TEOREMA 6.2. *Uma função é inversível se, e somente se, é bijetiva.*

Este teorema nos deixa livres para, de agora em diante, usar as noções de função bijetiva e de função inversível praticamente como sinônimos. Por isso, de agora em diante, passaremos de uma noção à outra sem nem mesmo comentar o que estamos fazendo.

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;
   (b) seu contradomínio;
   (c) sua imagem;
   (d) se é injetiva;
   (e) se é sobrejetiva;
   (f) se é bijetiva.

3. Determine quais das regras dadas abaixo definem uma função. Caso a regra não seja uma função, você deve explicar o porquê disto.
   (a) $f_1 : \mathbb{Z} \to \mathbb{N}$ tal que $f_1(x) = x^2 - 2$;
   (b) $f_2 : \mathbb{N} \to \mathbb{Z}$ tal que $f_2(x) = x^2 - 2$;
   (c) $f_3 : \mathbb{N} \to \mathbb{Q}$ tal que $f_3(x) = x/(x+1)$;
   (d) $f_4 : \mathbb{Z} \to \mathbb{Q}$ tal que $f_4(x) = x/(x+1)$;
   (e) $f_5 : \mathbb{R} \to \mathbb{R}$ tal que $f_5(x) = x - \lfloor x \rfloor$;
   (f) $f_6 : \mathbb{R} \to \mathbb{N}$ tal que $f_6(x) = \lfloor x \rfloor$;
   em que $\lfloor x \rfloor$ denota a parte inteira do número real x.

4. Determine quais das regras abaixo definem funções.

$$h_1 : \mathbb{Z} \longrightarrow \mathbb{Z} \qquad x \longmapsto \begin{cases} 2x & \text{se} \quad x < 0 \\ -x & \text{se} \quad x \geq 0 \end{cases}$$

$$h_2 : \mathbb{Z} \longrightarrow \mathbb{Z} \qquad x \longmapsto \begin{cases} 2x & \text{se} \quad x < 0 \\ -x & \text{se} \quad x \geq -1 \end{cases}$$

5. Para cada uma das regras abaixo, determine um subconjunto $D \subset \mathbb{R}$ sobre o qual a regra define uma função e ache a imagem da função obtida desta maneira.
   (a) $f_1(x) = 2x$;
   (b) $f_2(x) = |x-1|$;
   (c) $f_3(x) = \sqrt{2-x}$;
   (d) $f_4(x) = \sqrt{x^2}$;
   (e) $f_5(x) = 2^x$;
   (f) $f_6(x) = 1/x$;
   (g) $f_7(x) = \sqrt{1+x^2}$;
   (h) $f_8(x) = \sqrt{1-x^2}$;
   (i) $f_9(x) = 4/(|x|-3)$;
   (j) $f_{10}(x) = \sqrt{|x|-3}$;

(k) $f_{11}(x) = (x-1)/(\sqrt{x}-7)$;
(l) $f_{12}(x) = x - \lfloor x \rfloor$;
(m) $f_{13}(x) = 1/(x - \lfloor x \rfloor)$;
(n) $f_{14}(x) = (3x-1)/(3|x|-1)$;
(o) $f_{15}(x) = -3 - |2x|$;
(p) $f_{16}(x) = \sqrt{x^2-1} + \sqrt{1-x^2}$;
(q) $f_{17}(x) = \sqrt{1-x} + \sqrt{x-2}$;
(r) $f_{18}(x) = \sqrt{1-\sqrt{1-x^2}}$.

A propósito: a função $f_4(x)$ coincide com a função identidade $\mathrm{id}(x) = x$?

6. Determine a imagem das funções abaixo e indique se a função é ou não sobrejetiva. Quais destas regras definem funções injetivas?
(a) $f : \mathbb{N} \to \mathbb{Q}$ tal que $f(x) = x/(x+1)$;
(b) $g : \mathbb{R} \to \mathbb{R}$ tal que $g(x) = x - \lfloor x \rfloor$.

7. Determine a imagem de cada uma das funções de $\mathbb{N}$ em $\mathbb{N}$ definidas abaixo:

(a) $f_1(x) = x + 1$;

(b) $f_2(x) = 2x$;

(c) $f_3(x) = x^2$;

(d) $f_4(x) = \begin{cases} 2x & \text{se } x \text{ é par} \\ x+1 & \text{se } x \text{ é ímpar.} \end{cases}$;

(e) $f_5(x) = |x| = \begin{cases} x & \text{se } x \geq 0 \\ -x & \text{se } x < 0. \end{cases}$

8. Quais das funções do exercício anterior são: sobrejetivas? injetivas? bijetivas?

9. 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 $g(x) = 7x$;
(b) $f(x) = 2x$ e $g(x) = 9x^2 + 1$;
(c) $f(x) = 2x$ e $g(x) = x - 3$;
(d) $f(x) = x^2$ e $g(x) = -7x$.

10. Considere as funções

$$f : \mathbb{R} \longrightarrow \mathbb{R}, \quad x \longmapsto x^3 \qquad g : [0, +\infty) \longrightarrow \mathbb{R}, \quad x \longmapsto \sqrt{x} \qquad h : \mathbb{R} \longrightarrow \mathbb{R}, \quad x \longmapsto 2x.$$

Calcule as compostas abaixos que estiverem bem definidas:

$$f \circ g, \quad f \circ h, \quad f \circ f, \quad g \circ g, \quad h \circ g, \quad h \circ f, \quad h \circ h, \quad g \circ f, \quad g \circ h.$$

11. Considere as funções

$$f : \mathbb{R} \longrightarrow \mathbb{R}, \quad x \longmapsto 3x+1 \qquad g : [0,1] \longrightarrow \mathbb{R}, \quad x \longmapsto \sqrt{-x^2+2x+6} \qquad h : \mathbb{R} \setminus \{1/2\} \longrightarrow \mathbb{R}, \quad x \longmapsto 9x/(2x-1).$$

Calcule as compostas abaixos que estiverem bem definidas:

$$f \circ g, \quad f \circ h, \quad f \circ f, \quad g \circ g, \quad h \circ g, \quad h \circ f, \quad h \circ h, \quad g \circ f, \quad g \circ h.$$

12. Considere as funções

$$f : \mathbb{R} \longrightarrow \mathbb{R} \qquad g : [0, +\infty) \longrightarrow \mathbb{R} \qquad h : (0,1) \longrightarrow \mathbb{R}$$
$$x \longmapsto x^3 \qquad x \longmapsto x^2 \qquad x \longmapsto x/2.$$

Calcule as compostas abaixos que estiverem bem definidas:

$$f \circ g, \quad f \circ h, \quad f \circ f, \quad g \circ g, \quad h \circ g, \quad h \circ f, \quad h \circ h, \quad g \circ f, \quad g \circ h.$$

13. Considere as funções

$$g_1 : \mathbb{Z} \to \mathbb{Z} \qquad\qquad g_2 : \mathbb{Z} \to \mathbb{N}$$
$$x \mapsto \begin{cases} x & \text{se } x \text{ é par} \\ -x & \text{se } x \text{ é ímpar} \end{cases} \qquad\qquad x \mapsto \begin{cases} -2x & \text{se } \ x < 0 \\ 2x + 1 & \text{se } \ x \geq 0 \end{cases}$$

(a) Calcule, as funções compostas $g_1 \circ g_2$ e $g_2 \circ g_1$.
(b) Determine quais das duas funções do item (a) são bijetivas e calcule suas inversas indicando seus domínios e contradomínios. Caso a função não seja bijetiva você deve explicar porque.

14. Escreva cada uma das funções abaixo na forma $f \circ g$, em que f e g são duas funções de $\mathbb{R}$ em $\mathbb{R}$. Por exemplo, a função $h(x) = \sqrt{x-2}$; pode ser escrita na forma desejada tomando-se $f(x) = \sqrt{x}$ e $g(x) = x - 2$.

(a) $h_1(x) = \sqrt{x^2 + 1}$;
(b) $h_2(x) = \sqrt{x^3}$;
(c) $h_3(x) = \sqrt{|x|}$;
(d) $h_4(x) = \sqrt{x^3 - 7x^2 - 9}$;

15. Dizemos que uma função $f : \mathbb{R} \longrightarrow \mathbb{R}$ é
par : se $f(-x) = f(x)$, para todo $x \in \mathbb{R}$;
ímpar: se $f(-x) = -f(x)$, para todo $x \in \mathbb{R}$.
Por exemplo, função $h_1(x) = x$ é ímpar, ao passo que $h_2(x) = x^2$ é par. Prove que:
(a) a composta de duas funções ímpares é uma função ímpar;
(b) a composta de duas funções pares é uma função par;
(c) a composta de uma função par com uma função ímpar, em qualquer ordem, é uma função par.

16. Calcule uma inversa para cada uma das funções abaixo, restringindo seu domínio e imagem, quando necessário:

(a) $f_1(x) = -x^2$;
(b) $f_2(x) = x^2 - 5x + 6$;
(c) $f_3(x) = |x|$;
(d) $f_4(x) = 1/x$;
(e) $f_5(x) = (x-1)/(x-2)$;
(f) $f_6(x) = (x-2)/(x-1)$.

17. Para cada uma das regras abaixo: determine o maior domínio que torna a função injetiva, a imagem correspondente e a inversa da função bijetiva construída a partir do domínio e imagem que você encontrou.

(a) $f_1(x) = (x-1)/(x-2)$;
(b) $f_2(x) = x^2 - 5x + 6$;
(c) $f_3(x) = (x-1)/(x^2+3)$;
(d) $f_4(x) = (x^2+2)/(x^2+1)$;
(e) $f_5(x) = \sqrt{x^2-1}$;
(f) $f_6(x) = \operatorname{abs}(x)$;
(g) $f_7(x) = |x^2-1|$;
(h) $f_8(x) = \sqrt{-x^2+2x+6}$;
(i) $f_9(x) = x/(x^2+1)$;
(j) $f_{10}(x) = x/|x-1|$;
(k) $f_{11}(x) = x^3+1$;
(l) $f_{12}(x) = x/(1-x^2)$;
(m) $f_{13}(x) = (x+1)/(x^2+1)$;
(n) $f_{14}(x) = x^3$.

18. Sejam $a < b$ números reais. Uma função $f : (a,b) \longleftarrow \mathbb{R}$ é crescente se
$$\forall x, y \in (a,b), ((x < y) \Longrightarrow f(x) < f(y)).$$
Mostre que função crescente é injetiva.

19. Se f e g são funções crescentes de $\mathbb{R}$ em $\mathbb{R}$, quais das seguintes funções são crescentes?
(a) $f + g$;
(b) $f \cdot g$;
(c) $f \circ g$.

20. A função $g : Y \longrightarrow X$ é uma *inversa à direita* da função $f : X \longrightarrow Y$ se $f \circ g = \mathrm{id}_Y$. Mostre que uma função é sobrejetiva se, e somente se, admite uma inversa à direita.

21. A função $f : X \longrightarrow Y$ é uma *inversa à direita* da função $g : Y \longrightarrow X$ se $f \circ g = \mathrm{id}_X$. Mostre que uma função é injetiva se, e somente se, admite uma inversa à esquerda.

22. Sejam X, Y e Z três conjuntos e $\mathcal{F}(Y,Z)$ o conjunto cujos elementos são as funções com domínio em Y e contradomínio em Z. Uma função
$$f : X \times Y \longrightarrow Z$$
pode ser interpretada como a função
$$F : X \longrightarrow \mathcal{F}(Y,Z)$$
que leva $x \in X$ na função
$$f_x : Y \longrightarrow Z$$
definida por $f_x(y) = f(x,y)$. O processo que converte f em F é chamado de *currying*, em homenagem ao lógico Haskell Curry (1900–1982). Aplique o processo de currying a cada uma das funções de $\mathbb{Z} \times \mathbb{Z} \longrightarrow \mathbb{Z}$ dadas abaixo:
(a) $f(m,n) = |n-m|$;
(b) $g(n,m) = 2mn - 1$;
(c) $h(n,m) = n^2 + 2m + 7$;

(d) $p(n,m) = n^3 - 3m^2 - 2nm + 1$.

23. Sejam X, Y e Z três conjuntos e $\mathcal{F}(Y,Z)$ o conjunto cujos elementos são as funções com domínio em Y e contradomínio em Z. O processo que converte uma função $G : X \longrightarrow \mathcal{F}(Y,Z)$ em uma função $g : X \times Y \longrightarrow Z$ é chamado de *uncurrying*. Defina como g deve ser construída a partir de G de modo que aplicando uncurrying ao currying de $f : X \times Y \longrightarrow Z$ retorne a função f.

CAPÍTULO 5

Geometria

Neste capítulo investigaremos a geometria dos triângulos de maneira semelhante ao que fizemos com os números reais. Assim, depois de estabelecer alguns axiomas básicos, provaremos várias propriedades dos triângulos, que serão usadas nos capítulos que tratam de trigonometria e dos gráficos de funções.

1. Introdução

Como vimos na seção 5 do capítulo 1, ainda que todas as proposições do Livro I dos *Elementos* devessem ser provadas usando, como ponto de partida, apenas cinco postulados e oito noções comuns, já na antiguidade, vários matemáticos chamaram a atenção de que isto não ocorre. No caso da Proposição 4, por exemplo, Euclides prova o caso de congruência de triângulos conhecido como LAL (Lado-Ângulo-Lado) movendo um triângulo sobre o outro; veja p. 125. Contudo, além de nenhum postulado tratar do movimento de figuras, nenhuma das definições explicita que duas figuras são congruentes quando uma pode ser perfeitamente sobreposta à outra. Como argumentos de sobreposição são raramente usados nos *Elementos*, é possível que o próprio Euclides tivesse consciência que sua definição era problemática.

Há duas maneiras de contornar este problema. A primeira consiste em definir exatamente o que deve ser entendido pelo movimento de uma figura e o quê significa que uma figura pode ser perfeitamente sobreposta à outra. Pôr isto em prática é mais difícil do que possa parecer, porque as propriedades requeridas pelos movimentos não são fáceis de aplicar. Por isso não adotaremos esta solução neste livro, apesar de ser a mais satisfatória do ponto de vista de nossa intuição geométrica.

A segunda maneira de contornar o problema é a que foi a adotada por David Hilbert em seu livro *Fundamentos da Geometria* Hilbert (1910). A ideia é introduzir, como axioma, o caso de congruência de triângulos provado por Euclides na Proposição 4, mencionada acima, e usá-lo para demonstrar os demais casos de congruência. Desde então, esta tem sido a maneira mais comumente adotada para introduzir de maneira precisa a congruência de triângulos.

Os enfoques da geometria adotados por Hilbert e Euclides têm em comum a ausência de qualquer tipo de medição. Ainda que isto possa ser justificado em termos do que almejavam fazer, tornar esta escolha precisa requer uma grande quantidade de axiomas, o que a torna inadequada em um livro elementar. Para evitar estes problemas, adotaremos, neste livro, um terceiro enfoque, introduzido por George Birkhoff (1884–1944) em 1932, com o objetivo de simplificar a exposição axiomática da geometria elementar. Este enfoque é apresentado por Birkhoff no artigo Birkhoff (1932) e também no livro Birkhoff, Beatley (1959), escrito em colaboração com Ralph Beatley.

A principal diferença entre o sistema de Birkhoff, e o que fora proposto por Hilbert, está na introdução de axiomas que explicitam como devemos proceder para medir comprimentos e ângulos. Além de nos permitir tratar os problemas de geometria de maneira mais natural, a escolha de Birkhoff também permite que trabalhemos com um número menor de axiomas. A grande desvantagem deste enfoque é a necessidade de usar números reais para efetuar estas medidas porque, como vimos no capítulo 3, descrever estes números de maneira rigorosa é bastante difícil. Portanto, em nosso tratamento da geometria usaremos, sistematicamente, as propriedades dos números reais enunciadas nas seções 3, 4 e 5 do capítulo 3.

Dito isto, é necessário fazer uma ressalva. Se você comparar os axiomas estabelecidos por Birkhoff com os que apresentaremos neste capítulo, verá que não seguimos sua apresentação à risca. Para evitar algumas demonstrações, bastante delicadas, de resultados que parecem óbvios, optamos por introduzir alguns axiomas que, no tratamento dado por Birkhoff, são consequência de outros axiomas. Em outras palavras, nossos axiomas não são independentes uns dos outros. Além disso, embora sigamos Birkhoff ao tratar congruência como um caso especial da semelhança; preferimos incluir como axioma uma generalização do caso ALA (Ângulo-Lado-Ângulo) de congruência, ao passo que Birkhoff adota uma generalização do caso LAL (Lado-Ângulo-Lado).

Uma questão que se põe quando tratamos, matematicamente, um assunto de caráter tão visual quanto a geometria plana é o papel das figuras nas demonstrações. Na geometria grega, elas desempenhavam um papel absolutamente fundamental. Apesar disto, algumas das figuras nos manuscritos que chegaram até nós parecem ter sido deliberadamente desenhadas com distorções, seguindo, talvez, o adágio segundo o qual *a geometria é a arte de argumentar corretamente a partir de figuras erradas*. Embora todos os nossos argumentos tenham que ser respaldados nos axiomas, ou 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 não nos deixar levar pela figura e, com isso, fazer afirmações que não podem ser justificadas com base nos axiomas.

Finalmente, uma observação sobre terminologia. É costume rotular resultados de lemas, teoremas, ou corolários, dependendo de sua importância e do papel que desempenham. Normalmente, os resultados mais importantes são chamados de *teoremas*, do verbo grego que significa “contemplar”. Já um resultado cujo papel é apenas o de auxiliar na demonstração de um teorema é 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. Retas e pontos

Por um *plano* entenderemos um conjunto infinito $\mathbb{P}$, cujos elementos chamaremos de pontos. Como fizemos com os números reais, nossa análise deste conjunto será feita a partir dos axiomas que estabeleceremos para alguns de seus subconjuntos. O mais fundametal destes subconjuntos são as *retas*, que sempre suporemos serem subconjuntos próprios e não vazios. Portanto, nenhuma reta pode conter todos os pontos do plano. Os primeiros axiomas tratam, justamente, da relação entre retas e pontos.

Como estamos estudando geometria plana, todos os pontos que considerarmos pertencerão ao plano $\mathbb{P}$, que estamos supondo ser único. Com isso, podemos dizer simplesmente “P é um ponto”, sem precisar acrescentar “do plano”.

AXIOMA DE INCIDÊNCIA. *Dois pontos distintos definem uma e somente uma reta.*

Denotaremos por $\overleftrightarrow{AB}$ a reta determinada pelos pontos $A \neq B$ do plano. Esta notação só é viável porque, pelo axioma de incidência, dois pontos do plano sempre definem uma, e não mais que uma, reta. Antes de provar nosso primeiro teorema, precisamos de um pouco de terminologia. Quando três pontos pertencem a uma mesma reta, diremos que são *colineares*; já duas retas que não têm ponto em comum são chamadas de *paralelas*. Note que, segundo esta definição, uma reta *não* é paralela a ela própria.

TEOREMA 2.1. *Se duas retas distintas não são paralelas, então têm exatamente um ponto comum.*

DEMONSTRAÇÃO. Sejam ℓ_1 e ℓ_2 duas retas distintas que não são paralelas. Queremos provar que $\ell_1 \cap \ell_2$ consiste em apenas um ponto. Como, por hipótese, as retas não são paralelas então, por definição, têm que ter *pelo menos* um ponto P em comum. Suponhamos, por contradição, que $Q \neq P$ seja um segundo ponto de interseção das mesmas retas. Neste caso, pelo Axioma de Incidência, ℓ_1 e ℓ_2 têm que coincidir, 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. $\square$

Nosso primeiro axioma trata, apenas, das posições relativas de pontos e retas, mas não nos permite fazer nenhuma medição no plano. Para isso precisamos do próximo axioma.

AXIOMAS DE DISTÂNCIA.

Primeiro: *Existe uma função* $d : \mathbb{P} \times \mathbb{P} \longrightarrow [0, +\infty)$*, chamada de* distância*, que a cada par de pontos* (A, B) *associa um número real não negativo* $d(A, B)$*, de modo que* $d(A, B) = d(B, A)$ *e* $d(A, A) = 0$*, quaisquer que sejam os ponto* A *e* B *do plano.*

Segundo: *Dada uma reta* ℓ *do plano e um ponto* $O \in \ell$ *existe uma bijeção* $x : \ell \longrightarrow \mathbb{R}$ *tal que*

$$x(O) = 0 \ \textit{ e } \ |x(P) - x(Q)| = d(P, Q),$$

quaisquer que sejam $P, Q \in \ell$*.*

A função x tem o papel de dotar a reta sobre a qual é definida de coordenadas. Por isso diremos que x é uma *função coordenada* para ℓ, com origem em O. Sempre que conveniente, escreveremos x_P, em vez de $x(P)$ para a imagem de P por x. Note que a função d desempenha um papel essencial no Segundo Axioma. Sem ela, não teríamos como comparar as distâncias entre pontos estabelecidas pelas funções coordenadas de diferentes retas.

Antes de passar às consequências dos Axiomas de Distância, precisamos nos perguntar sobre a unicidade das funções d e x, estabelecidas nestes axiomas. A primeira coisa a notar é que, se α é um número real positivo, então a função d' definida por $d'(A, B) = \alpha d(A, B)$, quaisquer que sejam os pontos A e B, também satisfaz as propriedades do Primeiro Axioma. Isto não é surpreendente, porque, ao multiplicar d por α, apenas alteramos a escala com que as distâncias estão sendo medidas. Passando ao Segundo Axioma de Distância, o próximo teorema mostra que, supondo fixadas a função d e a reta ℓ, as escolhas de coordenadas para ℓ são bastante limitadas.

TEOREMA 2.2. *Se* ℓ *é uma reta e* x *e* x' *são funções associadas a* ℓ *e aos pontos* O *e* O' *de* ℓ*, então*

$$x'(P) = \begin{cases} x(P) + x'(O), & \textit{para todo } \ P \in \ell; \\ \qquad \textit{ou} & \\ -x(P) + x'(O), & \textit{para todo } \ P \in \ell. \end{cases}$$

DEMONSTRAÇÃO. Definindo

$$x''(P) = x'(P) - x'(O)$$

para todo $P \in \ell$, obtemos uma nova função bijetiva $x'' : \ell \longrightarrow \mathbb{R}$, que satisfaz

$$x''(O) = x'(O) - x'(O) = 0$$

e também

$$|x''(P) - x''(Q)| = |(x'(P) - x'(O)) - (x'(Q) - x'(O))| = |x'(P) - x'(Q)| = d(P, Q),$$

sempre que $P, Q \in \ell$. Começaremos investigando as possíveis relações entre x'' e x. O Segundo Axioma de Distância nos permite afirmar que

$$|x(A) - x(B)| = |x''(A) - x''(B)|, \tag{60}$$

quaisquer que sejam os pontos $A, B \in \ell$, porque estes dois números são iguais a $d(A, B)$. Suponhamos, por contradição, que existam pontos Q e Q' em $\ell \setminus \{O\}$ tais que

$$x''(Q) = x(Q) \text{ mas } x''(Q') = -x(Q').$$

Somando estas duas equações, obtemos

$$x''(Q) + x''(Q') = x(Q) - x(Q').$$

Combinando isto com (60),

$$|x''(Q) + x''(Q')| = |x''(Q) - x''(Q')|,$$

que só é possível se $x''(Q) = 0$ ou $x''(Q') = 0$. Como x é uma função bijetiva, isto implica que $Q = O$ ou $Q' = 0$, levando à contradição desejada. Portanto,

$$x''(P) = x(P) \text{ para todo } P \in \ell,$$

ou,então,

$$x''(P) = -x(P) \text{ para todo } P \in \ell.$$

Para obter o resultado desejado, basta lembrar que, por definição, $x''(P) = x'(P) - x'(O)$. □

René Descartes foi o primeiro a enunciar de forma precisa a relação entre números e pontos de uma reta codificada no Segundo Axioma de Distância. No início de sua *Geometria* (Descartes, 1997, p. 3), ele afirma que

> [t]odos os problemas de geometria podem ser facilmente reduzidos a termos tais que só há necessidade de conhecer certas linhas retas para sua construção. E, como toda a aritmética não é composta senão de quatro ou cinco operações, que são a adição, a subtração, a multiplicação, a divisão e a extração de raízes, que podemos tomar por uma espécie de divisão, assim, não há outra coisa a fazer em geometria, no que diz respeito às linhas que procuramos, que lhes adicionar ou subtrair outras [linhas]. Ou, tendo um segmento, que chamarei de unidade para melhor relacioná-la aos números, e que pode, em geral, ser escolhida de maneira arbitrária, e, sendo dadas duas outras linhas, podemos encontrar uma quarta que está para uma delas, como a outra está para a unidade, o que é o mesmo que a multiplicação.

Seguindo o costume corrente na sua época, Descartes usava palavra *linha* para descrever, tanto uma reta, quanto um segmento de reta. Na seção 5, voltaremos a examinar este texto de Descartes para esclarecer a menção que ele faz à multiplicação como proporção entre segmentos.

O axioma que trata da medição de distâncias tem várias consequências simples que enunciamos a seguir.

COROLÁRIO 2.3. *Toda reta tem uma infinidade de pontos.*

DEMONSTRAÇÃO. Sejam ℓ uma reta e O um ponto de ℓ. O Segundo Axioma de Distância estabelece a existência de uma função bijetiva que a cada ponto de ℓ associa um número real. Como o conjunto dos números reais é infinito, a bijetividade da função garante que ℓ é um conjunto infinito de pontos, provando o corolário. □

Como consequência dos axiomas de incidência e distância podemos provar que o plano contém uma infinidade de retas.

COROLÁRIO 2.4. *Por um ponto qualquer, fora de uma reta, passam infinitas retas distintas que intersectam a reta dada.*

DEMONSTRAÇÃO. Seja ℓ uma reta do plano. Como, por definição, nenhuma reta contém todos os pontos do plano, existe um ponto O fora de ℓ. Mas, pelo Axioma de Incidência, para cada ponto $P \in \ell$ existe uma reta r_P que contém a ambos. Suponhamos que $Q \neq P$ é outro ponto de ℓ e que r_Q é a reta por O e Q. Então, pelo teorema 2.1,

$$r_P \cap \ell = \{P\} \neq \{Q\} = r_Q \cap \ell,$$

de modo $r_Q \neq r_P$. Logo, retas que passam por O e por pontos distintos de ℓ não podem ser iguais. Para finalizar, basta lembrar que, pelo corolário 2.3, ℓ tem uma infinidade de pontos. □

A medição de distâncias, nos permite definir, de maneira rigorosa, segmentos e semirretas. Sejam A e B dois pontos do plano e ℓ a reta que contém a ambos. O segmento AB é definido a partir da função x que atribui coordenadas a ℓ e se anula em um ponto $O \in \ell$. Quando $x_A < x_B$, definimos

$$AB = \{P \in \ell \,|\, x_A \leq x_P \leq x_B\};$$

ao passo que, quando $x_A > x_B$, o segmento será

$$AB = \{P \in \ell \,|\, x_A \geq x_P \geq x_B\}.$$

Assim, usamos a ordenação dos números reais, e o fato da função coordenada ser bijetiva, para ordenar os pontos de ℓ e com isso definir o que significa um de seus pontos estar entre outros dois. Note que, por estas definições, AB e BA sempre representam o mesmo segmento, cujo *comprimento* é a distância $d(A,B)$ entre os pontos que o definem.

Para definir uma semirreta, precisamos de uma reta ℓ e um ponto $O \in \ell$. Escolhemos, então, uma função coordenada x associada a ℓ e O. O ponto O divide a reta em duas *semirretas*, definidas por

$$\overrightarrow{s_+} = \{P \in \ell \,|\, x_P \geq 0\} \text{ e } \overrightarrow{s_-} = \{P \in \ell \,|\, x_P \leq 0\}.$$

Se $P \neq O$ também escrevemos $\overrightarrow{OP}$ para designar a semirreta que contém P e começa em O.

Antes de encerrar a seção, 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, que definiremos na próxima seção.

3. Ângulos e triângulos

Como definiremos o interior de um ângulo como uma interseção de semiplanos, precisamos de uma axioma que especifique de que maneira uma reta subdivide um plano.

AXIOMAS DE SEPARAÇÃO.

Primeiro: *Os pontos do plano fora de uma reta podem ser agrupados em dois subconjuntos disjuntos, não vazios, chamados de semiplanos, que contêm todos os pontos do plano que estão fora da reta.*

Segundo: *Se S é um dos semiplanos em que uma reta ℓ divide o plano, então $P, Q \in S$ se, e somente se, o segmento PQ não intersecta a reta ℓ.*

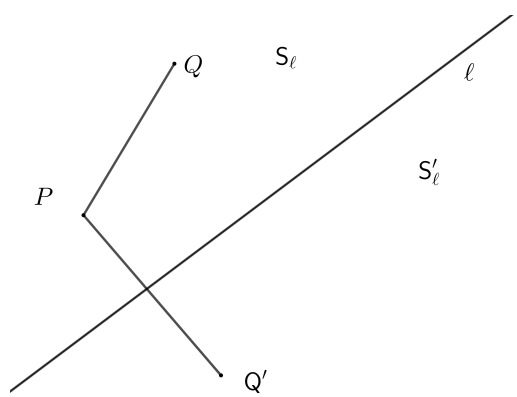


FIGURA 1. A reta ℓ e seus dois semiplanos.

Denotaremos por S_ℓ e S'_ℓ os dois semiplanos em que a reta ℓ divide o plano. Pelo Primeiro Axioma de Separação, temos que

$$\mathbb{P} = S_\ell \cup \ell \cup S'_\ell.$$

Quando $\ell = \overleftrightarrow{OA}$ escreveremos simplesmente S_{OA} e S'_{OA}. A figura 1 ilustra os dois casos que podem ocorrer quando dois pontos fora de uma reta são unidos por um segmento, conforme o Segundo Axioma de Separação. Quando P e Q estão em um mesmo semiplano, denotado por S_ℓ na figura, todo o segmento PQ está contido em S_ℓ. Por outro lado, se $P \in S_\ell$ e $Q' \in S'_\ell$, então o segmento PQ' necessariamente intersecta a rela ℓ.

Frequentemente nos referiremos aos semiplanos em que a reta ℓ divide o plano como os *lados* de uma reta. Assim, o Segundo Axioma de Separação também pode ser enunciado dizendo que *dois pontos P e Q estão de um mesmo lado de uma reta se todo o segmento PQ está contido deste mesmo lado da reta.*

Com isso podemos passar à definição de ângulo. Sejam O, A e B três pontos não colineares. Se $B \in S_{OA}$ e $A \in S_{OB}$, então o *ângulo* $\sphericalangle AOB$ é a região do plano entre as semirretas $\overrightarrow{OA}$ e $\overrightarrow{OB}$. Isto sugere que o ângulo $\sphericalangle AOB$ deveria ser definido como a região $S_{OA} \cap S_{OB}$, que é a região cinza na figura 2. Entretanto, como é conveniente incluir os pontos das semirretas $\overrightarrow{OA}$ e $\overrightarrow{OB}$ entre os pontos do ângulo $\sphericalangle AOB$, diremos que

$$\sphericalangle AOB = (S_{OA} \cap S_{OB}) \cup \overrightarrow{OA} \cup \overrightarrow{OB}.$$

Por isso, vamos nos referir à região $S_{OA} \cap S_{OB}$ como sendo o *interior* do ângulo.

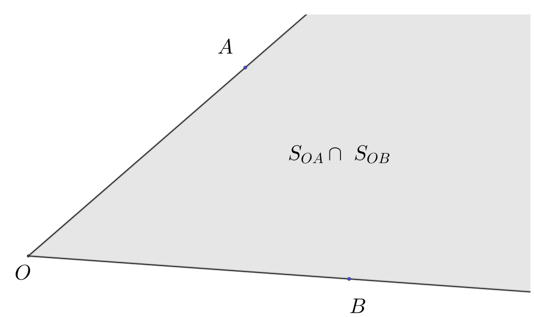


FIGURA 2. Ângulo $\sphericalangle AOB$.

Quando a interseção entre dois ângulos coincide com uma das semirretas que os definem, dizemos que são *adjacentes*. Na figura 3 os ângulos $\sphericalangle AOA'$ e $\sphericalangle A'OB$ são adjacentes, e o mesmo vale $\sphericalangle A'OB$ e $\sphericalangle B'OB$; contudo, $\sphericalangle AOA'$ e $\sphericalangle BOB'$ não são adjacentes.

Nosso próximo conjunto de axiomas especifica como mediremos ângulos. Faremos isto em duas etapas, de maneira semelhante ao que foi feito quando definimos medidas de segmentos.

AXIOMAS DE ÂNGULOS. *Seja $\mathcal{A}$ o conjunto dos ângulos do plano.*

Primeiro: *Existe uma função* $\operatorname{rad} : \mathcal{A} \longrightarrow [0, \pi]$, *que a cada ângulo do plano associa um número real* $\operatorname{rad}(AOB) \in [0, \pi]$, *de modo que, se $C \in \sphericalangle AOB$, então*

$$\operatorname{rad}(AOB) = \operatorname{rad}(AOC) + \operatorname{rad}(COB).$$

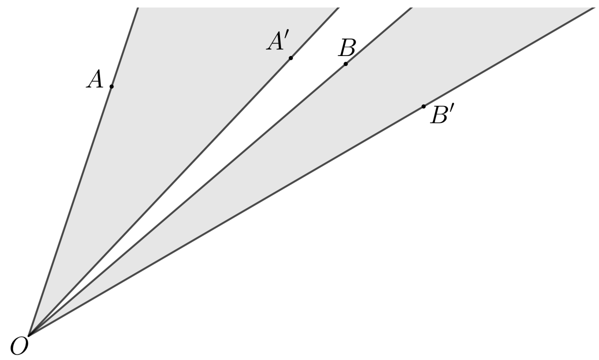


FIGURA 3. $\sphericalangle AOA'$ e $\sphericalangle BOB'$ são adjacentes.

Segundo: *Se ℓ é uma reta, O e P são pontos distintos de ℓ e S_ℓ é um dos semiplano em que ℓ divide o plano, então a função que a cada semirreta $\overrightarrow{OA} \subset S_\ell \cup \ell$ associa $\mathrm{rad}(POA)$ define uma bijeção entre as semirretas de $S_\ell \cup \ell$ e o intervalo $[0,\pi]$.*

Usaremos a expressão *aditividade da medida de ângulos* quando quisermos nos referir à propriedade referente à soma de ângulos adjacentes enunciada no Primeiro Axioma de Ângulos.

Nossa escolha de unidades para medir ângulos requer uma justificativa. Há três unidades diferentes para medir ângulos: graus, grados e radianos. A medição em graus é a mais antiga. Ela foi inventada na Mesopotâmia, cujo sistema de numeração era sexagesimal, o que explica porque um ângulo de meia volta seja igual a $180^o = 3 \cdot 60^o$. A medida em grados é decimal. Ela foi inventada na época da Revolução Francesa e é uma das medidas oficiais de ângulo na União Europeia. A medida em *radianos*, adotada neste livro, é, de certa maneira, a mais natural, porque π corresponde ao comprimento da metade de uma circunferência. Começamos nossas aplicações dos axiomas de medida de ângulos determinando quais ângulos que medem 0 e quais medem π.

TEOREMA 3.1. *Sejam O, $A \neq O$ e B três pontos.*

(a) $\mathrm{rad}(AOB) = 0$ *se, e somente se,* $B \in \overrightarrow{OA}$.
(b) $\mathrm{rad}(AOB) = \pi$ *se, e somente se,* B *pertence à reta* $\overleftrightarrow{OA}$*, mas não à semirreta* $\overrightarrow{OA}$.

DEMONSTRAÇÃO. Como

$$A \in \sphericalangle AOA = \overrightarrow{OA}$$

temos, pelo Primeiro Axioma de Ângulos, que

$$\mathrm{rad}(AOA) + \mathrm{rad}(AOA) = \mathrm{rad}(AOA),$$

de modo que $\mathrm{rad}(AOA) = 0$. Logo, pela bijetividade do segundo axioma, não pode existir nenhum outro ângulo, com vértice em O cuja medida seja zero radianos de nenhum dos dois lados de $\overleftrightarrow{OA}$.

Suponhamos, agora, que $B \in \overleftrightarrow{OA} \setminus \overrightarrow{OA}$. Se C é um ponto de um dos lados de $\overleftrightarrow{OA}$, o Primeiro Axioma de Ângulos nos permite afirmar que

$$\mathrm{rad}(AOB) = \mathrm{rad}(AOC) + \mathrm{rad}(COB) \geq \mathrm{rad}(AOC).$$

Mas, por (a), se $C \notin \overrightarrow{OA}$, então $\mathrm{rad}(AOC) > 0$; donde

$$0 < \mathrm{rad}(AOC) \leq \mathrm{rad}(AOB) \leq \pi,$$

para todo C de um dos lados de $\overleftrightarrow{AB}$. Logo,

$$\pi \leq \mathrm{rad}(AOB) \leq \pi, \tag{61}$$

pois, pelo Segundo Axioma de Ângulos, seja qual for o lado de $\overleftrightarrow{AB}$ que estejamos considerando, existe uma semirreta que forma com $\overrightarrow{OA}$ um ângulo igual a π. O item (b) é uma consequência imediata de (61). $\square$

Para podermos distinguir com precisão entre o conjunto de pontos que corresponde a um segmento de reta e seu comprimento, e entre a região do plano que corresponde a um ângulo, e sua medida em radianos, precisamos introduzir as funções d e rad. Tendo passado desta etapa, podemos ser mais flexíveis e simplificar um pouco nossa notação. Faremos isto usando AB para denotar tanto o segmento definido pelos pontos A e B, quanto seu comprimento $d(A,B)$. De maneira semelhante, $\sphericalangle AOB$ será utilizado para denotar um ângulo e sua medida em radianos $\mathrm{rad}(AOB)$.

Nosso próximo teorema, requer uma definição. 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 $A_2, B_2 \in \ell_2$ em semiplanos diferentes relativamente a ℓ_1, como ilustrado na figura 4. Os ângulos $\sphericalangle A_1OA_2$ e $\sphericalangle B_1OB_2$ são *opostos pelo vértice*, e o mesmo se aplica a $\sphericalangle A_1OB_2$ e $\sphericalangle A_2OB_1$.

TEOREMA 3.2. *Ângulos opostos pelo vértice têm a mesma medida.*

DEMONSTRAÇÃO. Se

$$\alpha = \mathrm{rad}(A_2OA_1), \quad \beta = \mathrm{rad}(B_2OB_1) \text{ e } \gamma = \mathrm{rad}(A_1OB_2);$$

como na figura 4, então

$$\alpha + \gamma = \mathrm{rad}(A_2OA_1) + \mathrm{rad}(A_1OB_2) = \mathrm{rad}(A_2OB_2) = \pi,$$

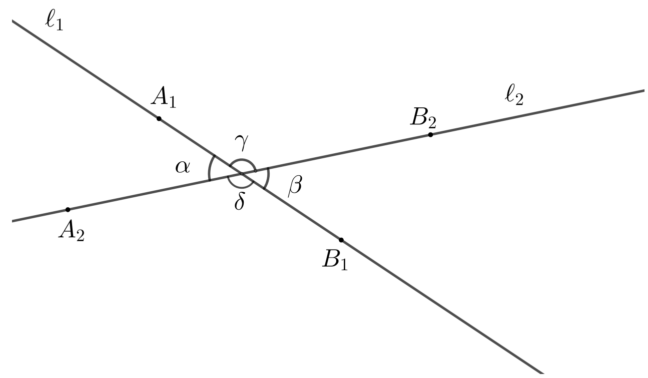


FIGURA 4. Ângulos entre retas que se intersectam no ponto O.

pois A_2, O e B_2 pertencem, todos três, à reta ℓ_2. Analogamente,

$$\gamma + \beta = \mathrm{rad}(A_1OB_2) + \mathrm{rad}(B_2OB_1) = \mathrm{rad}(A_1OB_1) = \pi.$$

Em particular, $\alpha + \gamma = \beta + \gamma$; donde $\alpha = \beta$, como queríamos mostrar. Um argumento semelhante mostra que $\mathrm{rad}(A_1OB_2) = \mathrm{rad}(A_2OB_1)$. $\square$

Com isto podemos definir a noção de perpendicularidade entre retas. Sejam ℓ e ℓ' duas retas que se cortam no ponto O e sejam $P \in \ell$ e $P' \in \ell'$ pontos diferentes de O. As retas ℓ e ℓ' são *perpendiculares* quando $\mathrm{rad}(POP') = \pi/2$. Neste caso, também dizemos que o ângulo entre ela é *reto*.

À primeira vista, esta definição deixa muito a desejar, já que estabelece a perpendicularidade entre retas a partir da escolha de um par de semirretas, uma em cada uma das retas. O problema é que há quatro maneiras possíveis de escolher entre estas semirretas. Assim, para que a definição faça sentido é necessário provar que todas as escolhas produzem ângulos retos. Usando, ainda, a notação da figura 4, suponhamos que α seja um ângulo reto. Como,

$$\alpha + \gamma = \mathrm{rad}(A_2OB_2) = \pi,$$

temos que $\gamma = \pi/2$. Por outro lado, como α e β são opostos pelo vértice, então também medem $\pi/2$. Finalmente, como $\beta + \delta = \pi$, então δ também é um ângulo reto. Portanto, se o ângulo formado por um dos pares de semirretas, contidas uma em cada uma das retas distintas, é reto, o mesmo tem que ser verdade para todos os outros. Com isto estamos prontos para definir o principal objeto de estudo deste capítulo.

Sejam A, B e C três pontos *não* colineares. O *triângulo de vértices* A, B e CC, que denotaremos por $\triangle ABC$, é a união dos segmentos AB, AC e BC. Chamamos estes segmentos de *lados* do $\triangle ABC$. Já $\sphericalangle BAC$, $\sphericalangle ABC$ e $\sphericalangle ACB$ são os *ângulos internos* do $\triangle ABC$ que também denotaremos, simplesmente, por $\sphericalangle A$, $\sphericalangle B$ e $\sphericalangle C$. Frequentemente vamos nos referir a BC como o lado *oposto* ao vértice A e vice-versa. Analogamente, AC e AB são os lados opostos aos ângulos $\sphericalangle B$ e $\sphericalangle C$, respectivamente. Supondo que

$$C \in S_{AB}, \quad B \in S_{AC} \quad \text{e} \quad A \in S_{BC}$$

podemos definir o *interior* do $\triangle ABC$ pela interseção

$$S_{AB} \cap S_{AC} \cap S_{BC}.$$

Encerraremos a seção provando um resultado simples sobre triângulos, que será usado em muitos exercícios. Como aparece como um dos axiomas introduzidos por M. Pasch em suas *Lições de geometria moderna*, este resultado é frequentemente chamado *axioma de Pasch*. Optamos por inverter a ordem usual e tratá-lo como um teorema, porque seu enunciado é bem menos acessível à nossa intuição que as propriedades dos Axiomas de Separação.

TEOREMA 3.3 (Axioma de Pasch). *Seja $\triangle ABC$ um triângulo. Se $B' \in AB$ é diferente de A e B e $\ell \neq \overleftrightarrow{AB}$ é uma reta que contém B', mas não intersecta BC, então ℓ intersecta AC em um ponto diferente de A e C.*

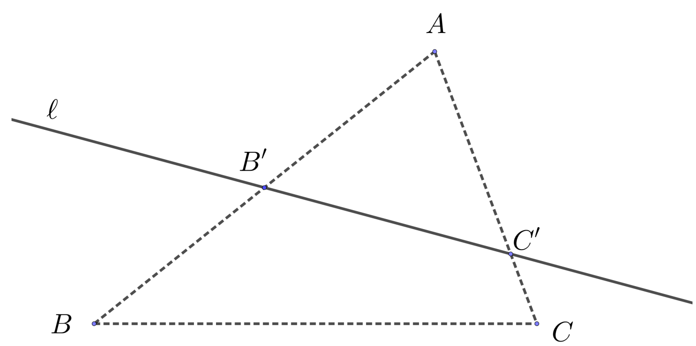


FIGURA 5. Posição dos pontos e retas no Teorema 3.3

DEMONSTRAÇÃO. Como $B' \in \ell \cap AB$ é diferente de A e B, o Segundo Axioma de Separação nos garante que A e B estão em semiplanos distintos relativamente a ℓ. Digamos que $A \in S_\ell$ e $B \in S'_\ell$. Como $\ell \cap BC = \varnothing$ então C está no mesmo semiplano que B, relativamente a ℓ. Logo, $C \in S'_\ell$. Portanto, pelo Segundo Axioma de Separação

$$\{C'\} = AC \cap \ell \neq \varnothing.$$

Finalmente, se $C' = A$, então $\ell = \overleftrightarrow{AB}$, ao passo que, se $C' = B$, então $\ell \cap BC \neq \varnothing$. Como estas afirmações são, ambas, falsas, a demonstração do teorema está completa. $\square$

Por fim, um comentário sobre notação. Até aqui usamos AB para designar o segmento de reta definido pelos pontos A e B e $d(A, B)$ para a distância entre estes pontos. Analogamente, usamos $\sphericalangle POQ$ para denotar o ângulo formado pelas semirretas $\overrightarrow{OP}$ e $\overrightarrow{OQ}$ e $\mathrm{rad}(POQ)$ para a medida deste ângulo em radianos. Ainda que, do ponto de vista formal, estas distinções sejam necessárias, elas levam a uma multiplicação da notação que é um pouco incoveniente. Evitaremos isto, de agora em diante, usando AB para denotar tanto o segmento quanto sua medida e $\sphericalangle POQ$ para denotar tanto o ângulo quanto sua medida. Este é um caso típico de "abuso de notação", que ocorre sempre que abreviamos algo usando uma notação que não é formalmente correta. O risco de confusão, embora exista, é muito pequeno. Por exemplo, ao escrever $P \in AB$, estamos considerando AB como um conjunto e, portanto, neste caso trata-se do segmento AB; já em AB/AC, tanto AB, quanto AC denotam números, caso contrário não poderíamos dividi-los.

4. Paralelas

Neste capítulo começamos nosso estudo das propriedades dos triângulos. Porém, para isto, precisamos de um axioma que nos permite determinar quando duas retas são paralelas. Para tornar o axioma mais fácil de interpretar, é conveniente introduzir a seguinte definição. Sejam

$$\overleftrightarrow{AB}, \quad \overleftrightarrow{A'B'} \text{ e } \overleftrightarrow{AA'}$$

três retas do plano e suponhamos que B e B' estão de um mesmo lado de $\overleftrightarrow{AA'}$. Neste caso diremos que $\sphericalangle BAA'$ e $\sphericalangle B'A'A$ são os *ângulos internos de um mesmo lado* de $\overleftrightarrow{AA'}$, relativamente às retas $\overleftrightarrow{AB}$ e $\overleftrightarrow{A'B'}$. A posição relativa das três retas é ilustrada na figura 6, em que os ângulos internos rotulados como α e α'.

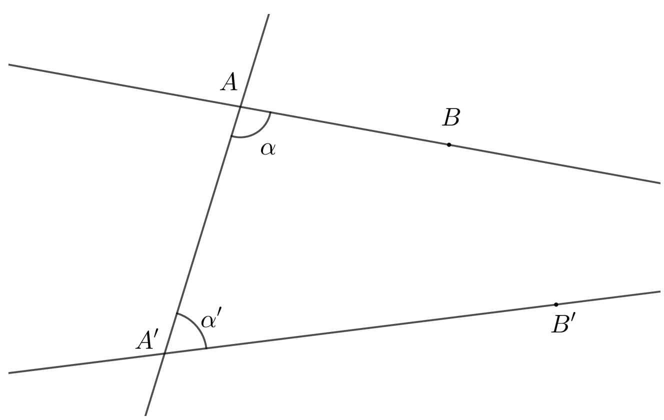


FIGURA 6. Ângulos internos do um mesmo lado de uma reta.

Com isto estamos prontos para enunciar o axioma, cuja inspiração remonta ao famoso quinto postulado dos *Elementos* de Euclides

AXIOMA EUCLIDIANO. *Duas retas ℓ e ℓ' que cortam uma terceira reta r em pontos distintos têm um ponto de intersecção em um dos lados de r se e somente se, a soma dos ângulos internos formados por ℓ e ℓ', deste mesmo lado, é menor do que π.*

Na notação da figura 6, a condição que aparece no axioma implica que se $\alpha + \alpha' < \pi$ então $\overleftrightarrow{AB}$ e $\overleftrightarrow{A'B'}$ têm um ponto de interseção no mesmo lado em que estão B e B'. Como veremos em breve, o axioma não é verdadeiro se os dois ângulos *não* estão em um mesmo lado de $\overleftrightarrow{AA'}$. Este axioma tem várias consequências importantes, a primeira das quais é o seguinte corolário, obtido a partir das contrapositivas das duas implicações do axioma.

COROLÁRIO 4.1. *Duas retas ℓ e ℓ' são paralelas se, e somente se, a soma dos ângulos internos que formam com uma reta transversal r é igual a π de um dos lados de r.*

O segundo corolário estabelece que o paralelismo é transitivo e, quando precisarmos citá-lo, vamos apenas nos referir à transitividade do paralelismo.

COROLÁRIO 4.2. *Duas retas paralelas a uma terceira são paralelas entre si.*

DEMONSTRAÇÃO. Sejam ℓ_1, ℓ_2 e ℓ_3 as três retas e digamos que, tanto ℓ_1 e ℓ_3 são ambas paralelas a ℓ_2. Pelo corolário anterior, qualquer reta r transversal a ℓ_2, também será transversal a ℓ_1 e ℓ_3. Sejam P_1, P_2 e P_3 os pontos de interseção de r com ℓ_1, ℓ_2 e ℓ_3 e

$$Q_1 \in \ell_1, \quad Q_2 \in \ell_2 \;\text{ e }\; Q_3 \in \ell_3$$

pontos do mesmo lado de r, como ilustrado na figura 7.

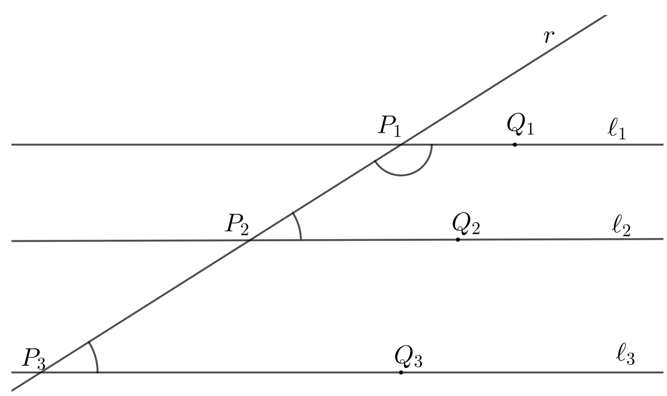


FIGURA 7. Transitividade do paralelismo.

Supondo P_1, P_2 e P_3 aparecem nesta ordem ao longo da reta r, os ângulos internos entre ℓ_1 e ℓ_2 são $\sphericalangle Q_1P_1P_2$ e $\sphericalangle P_1P_2Q_2$, de modo que

$$\sphericalangle Q_1P_1P_2 + \sphericalangle P_1P_2Q_2 = \pi,$$

pelo corolário anterior. Analogamente, os ângulos internos entre ℓ_2 e ℓ_3 são $\sphericalangle Q_2P_2P_3$ e $\sphericalangle P_2P_3Q_3$, donde

$$\sphericalangle Q_2P_2P_3 + \sphericalangle P_2P_3Q_3 = \pi,$$

Somando as duas últimas equações, obtemos

$$(\sphericalangle Q_1P_1P_2 + \sphericalangle P_1P_2Q_2) + (\sphericalangle Q_2P_2P_3 + \sphericalangle P_2P_3Q_3) = 2\pi. \tag{62}$$

Mas

$$\sphericalangle P_1P_2Q_2 + \sphericalangle Q_2P_2P_3 = \pi$$

pois P_1, P_2 e P_3 são colineares. Substituindo isto em (62), obtemos

$$\sphericalangle Q_1P_1P_2 + \sphericalangle P_2P_3Q_3 = \pi;$$

de modo que ℓ_1 e ℓ_3 são paralelas pelo corolário 4.1. O argumento acima só se aplica ao caso em que ℓ_2 está entre ℓ_1 e ℓ_3, nos outros os casos o argumento é ligeiramente diferente, por causa do posicionamento dos ângulos internos; mas os detalhes ficarão por sua conta. □

Note que, se a afirmação de que a soma dos ângulos internos é igual a π vale em um dos lados de r, então também valerá do outro lado. Antes de prosseguir, precisamos dar nomes a alguns dos ângulos formados por duas retas paralelas. Como a descrição formal destes ângulos é bastante prolixa, vamos nos contentar em localizá-los usando a figura 8, na qual r é uma reta transveral às retas paralelas ℓ e ℓ'. Na notação da figura, os ângulos α' e β são chamados de *correspondentes*, ao passo que α' e β' são chamados de *alternos internos*. Note que, ao contrário dos ângulos internos no enunciado do axioma, dois ângulos que são alternos internos estão de lados opostos relativamente à reta r. Esta terminologia se estende, facilmente, aos demais ângulos formados por estas três retas.

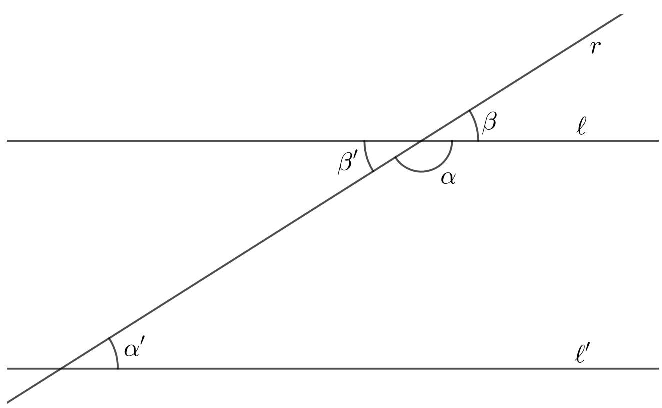


FIGURA 8. Ângulos alternos internos e correspondentes.

TEOREMA 4.3. *Os ângulos corresponentes formados por uma reta transversal a duas paralelas são iguais e o mesmo vale para os ângulos alternos internos.*

DEMONSTRAÇÃO. Usaremos a notação da figura ao longo da demonstração. Como α e α' são ângulos internos de um mesmo lado de uma transversal que corta duas paralelas, então

$$\alpha + \alpha' = \pi$$

pelo Axioma Euclidiano. Como,

$$\alpha + \beta = \pi,$$

pelo teorema 3.1, então $\beta = \alpha'$; o que mostra que ângulos correspondentes são iguais. Para mostrar que alternos internos são iguais basta lembrar que $\beta = \beta'$ pelo teorema 3.2, pois estes ângulos são opostos pelo vértice. $\square$

O próximo teorema garante que sempre há uma, e não mais que uma, reta paralela a uma reta dada reta por pontos fora dela.

TEOREMA 4.4. *Por um ponto fora de uma reta passa uma, e somente uma, paralela à reta dada.*

DEMONSTRAÇÃO. Seja ℓ a reta e P um ponto fora de ℓ. Escolha dois pontos distintos quaisquer $Q, Q' \in \ell$ e seja r a reta definida por P e Q. Pelo Segundo Axioma de Ângulos, podemos construir uma semirreta $\overrightarrow{PP'}$, do mesmo lado de r no qual Q' está, de modo que

$$\sphericalangle P'PQ = \pi - \sphericalangle Q'QP,$$

como ilustrado na figura 9, na qual as duas retas que foram construídas aparecem pontilhadas. Por construção, $\sphericalangle P'PQ$ e $\sphericalangle Q'QP$ são ângulos internos, de um mesmo lado de $\overleftrightarrow{PQ}$, cuja soma é π. Logo, pelo Axioma Euclidiano, as retas $\overleftrightarrow{PP'}$ e $\overleftrightarrow{QQ'}$ são paralelas.

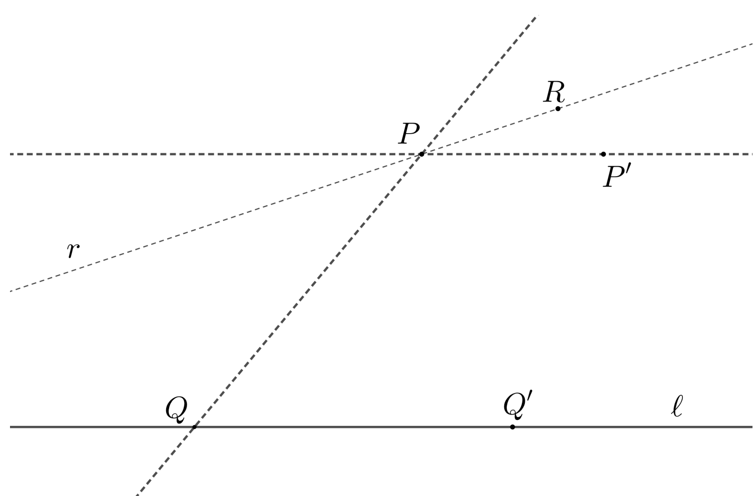


FIGURA 9. Paralela por um ponto fora de uma reta

Para concluir a demonstração do teorema basta mostrar que a paralela é única. Suponhamos, por contradição, que exista uma reta r, diferente de $\overleftrightarrow{PP'}$, que é paralela a ℓ passando por P. Seja R um ponto de r do mesmo lado de Q', relativamente a $\overleftrightarrow{PQ}$, como na figura 9. Pela bijetividade no Segundo Axioma de Ângulos sabemos que $\sphericalangle QPR \neq \sphericalangle QPP'$. Portanto,

$$\sphericalangle QPR + \sphericalangle Q'QP \neq \sphericalangle QPP' + \sphericalangle Q'QP = \pi,$$

de modo que, pelo Axioma Euclidiano, r não pode ser paralela a ℓ. □

Com isto estamos prontos para provar nosso primeiro resultado importante sobre triângulos, o bem conhecido teorema sobre a soma dos ângulos internos de um triângulo. Este é o primeiro resultado em que nos deparamos com a necessidade de fazer uma *construção*; isto é, em que será preciso acrescentar elementos (neste caso, uma reta) aos dados de um teorema para podermos demonstrá-lo.

TEOREMA 4.5. *A soma dos ângulos internos de qualquer triângulo é igual a* π.

DEMONSTRAÇÃO. Considere o triângulo $\triangle ABC$ e digamos que

$$\alpha = \sphericalangle A,\ \beta = \sphericalangle B \ \text{ e } \ \gamma = \sphericalangle C.$$

Usando o teorema 4.4, podemos traçar, pelo vértice A, uma paralela ao lado BC do triângulo, como ilustrado na figura 10.

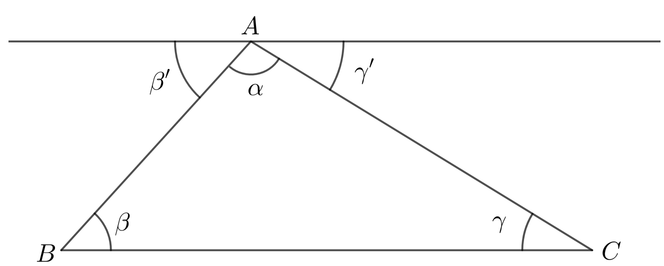


FIGURA 10. Soma dos ângulos internos de um triângulo.

Pelo igualdade dos ângulos alternos internos, temos que $\beta' = \beta$ e $\gamma' = \gamma$. Logo,

$$\alpha + \beta + \gamma = \alpha + \beta' + \gamma' = \pi,$$

pois α, β' e γ' são ângulos consecutivos que ocupam todo um lado de uma reta. □

5. Semelhança de triângulos

Nesta seção introduzimos nosso último axioma, que nos permitirá comparar triângulos entre si, determinando se são ou não semelhantes. Começamos com uma definição. Dois triângulos $\triangle ABC$ e $\triangle A'B'C'$ são *semelhantes* quando

(a) seus ângulos correspondentes são iguais e
(b) seus lados correspondentes são proporcionais.

Trocando em miúdos, (a) e (b) afirmam que dois triângulos são semelhantes se, e somente se, é possível indexar os vértices de um dos triângulos por A, B e C e os vértices do outro por A', B' e C', de modo que

$$\sphericalangle A = \sphericalangle A',\ \sphericalangle B = \sphericalangle B',\ \sphericalangle C = \sphericalangle C' \ \text{ e } \ \frac{AB}{A'B'} = \frac{AC}{A'C'} = \frac{BC}{B'C'}. \tag{63}$$

Ao valor comum destas três frações daremos o nome de *razão de semelhança*. Quando este número é igual a um, os triângulos são *congruentes*. A semelhança entre os triângulos $\triangle ABC$ e $\triangle A'B'C'$ será denotada por

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

e sua congruência por

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

Naturalmente não há nada de especial na escolha de A, B, C e A', B', C' para indexar os vértices dos triângulos; quaisquer outros símbolos poderiam ser usados desde que fique claro a que vértice de um triângulo corresponde um dado vértice do outro triângulo.

Do ponto de vista heurístico, dois triângulos são congruentes quando é possível sobrepô-los de maneira que coincidam exatamente, e são semelhantes quando uma mudança de escala transforma um no outro. Antes de enunciar nosso último axioma, provaremos uma propriedade que é consequência direta da definição de semelhança de triângulos.

TEOREMA 5.1. *Dois triângulos semelhantes a um terceiro são semelhantes entre si.*

DEMONSTRAÇÃO. Por definição, $\triangle ABC \sim \triangle A'B'C'$ equivale a dizer que

$$\sphericalangle A = \sphericalangle A',\ \sphericalangle B = \sphericalangle B',\ \sphericalangle C = \sphericalangle C' \ \text{ e } \ \frac{AB}{A'B'} = \frac{AC}{A'C'} = \frac{BC}{B'C'} = k. \tag{64}$$

Analogamente, se $\triangle A'B'C' \sim \triangle A''B''C''$, então

$$\sphericalangle A' = \sphericalangle A'',\ \sphericalangle B' = \sphericalangle B'',\ \sphericalangle C' = \sphericalangle C'' \ \text{ e } \ \frac{A''B''}{A'B'} = \frac{A''C''}{A'C'} = \frac{B''C''}{B'C'} = k'. \tag{65}$$

Da igualdade dos ângulos em (64) e (65) temos que

$$\sphericalangle A = \sphericalangle A'',\ \sphericalangle B = \sphericalangle B'' \ \text{ e } \ \sphericalangle C = \sphericalangle C''.$$

Por outro lado,

$$\frac{k}{k'} = \frac{AB/A'B'}{A''B''/A'B'} = \frac{AB}{A''B''}$$

e, por um cálculo análogo,

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

o que mostra que os triângulos $\triangle ABC$ e $\triangle A''B''C''$ são semelhantes. □

Podemos interpretar nosso último axioma como especificando condições *mínimas* sob as quais dois triângulos são semelhantes. Contudo, estas não são, de forma alguma, as únicas condições simples sob as quais dois triângulos são semelhantes; ao longo desta seção, várias outras condições serão deduzidos deste axioma.

AXIOMA DE SEMELHANÇA. *Para que os triângulos $\triangle ABC$ e $\triangle A'B'C'$ sejam semelhantes, basta que*

$$\sphericalangle ABC = \sphericalangle A'B'C' \quad e \quad \sphericalangle BAC = \sphericalangle B'A'C'.$$

Este axioma marca mais um ponto em que nossa apresentação da geometria se distancia daquela proposta nos *Elementos*. Em primeiro lugar, Euclides trata a congruência e a semelhança de triângulos de maneira inteiramente independente. A congruência é introduzida no livro I dos *Elementos*, ao passo que a semelhança só aparece no livro VI. Ele é forçado a isto porque precisa tratar a teoria de proporções de Eudoxo, apresentada no livro V dos *Elementos*, antes de poder enunciar resultados equivalentes às nossas equações (63). A bem da verdade, a razão pela qual podemos tratar semelhança e congruência juntas é justamente que os números reais, introduzidos no capítulo 3, desempenham, em nossa exposição, um papel semelhante ao da teoria de proporções nos *Elementos*.

Segundo o Axioma de Semelhança, basta que dois triângulos tenham dois ângulos correspondentes iguais para que seus lados mantenham uma proporção fixa. Tendo isto em vista, é razoável perguntar: se dois triângulos têm dois lados correspondentes com a mesma proporção, isto implica que os dois triângulos sejam semelhantes? A resposta é não. A figura 11 ilustra um contraexemplo: embora $AB = A'B'$ e $AC = A'C'$, os triângulos ABC e $A'B'C'$ claramente não são semelhantes. Analisando a figura, vemos que há duas maneira de expressar o porquê destes triângulos não serem semelhantes. A primeira consiste em afirmar que os ângulos $\sphericalangle BAC$ e $\sphericalangle B'A'C'$ não são iguais; a segunda, que os lados BC e $B'C'$ não mantêm a mesma proporção que os outros dois. O próximo teorema mostra que se o ângulo que delimita os lados proporcionais são iguais, então os triângulos são semelhantes.

TEOREMA 5.2. *Dois triângulos $\triangle ABC$ e $\triangle A'B'C'$ para os quais*

$$\sphericalangle BAC = \sphericalangle B'A'C' \quad e \quad \frac{AB}{A'B'} = \frac{AC}{A'C'},$$

são semelhantes.

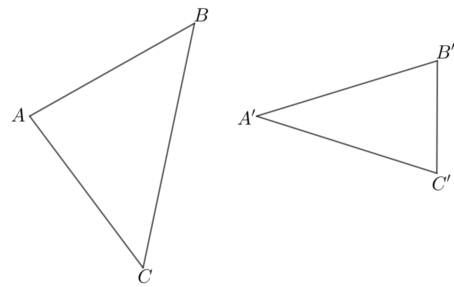


FIGURA 11. Triângulos com dois lados iguais que não são semelhantes.

DEMONSTRAÇÃO. Para provar o teorema, basta mostrar que $\sphericalangle(ABC) = \sphericalangle(A'B'C')$, porque isto nos permite aplicar o Axioma de Semelhança. Suponhamos, por contradição, que isto é falso, e que

$$\beta = \sphericalangle(ABC) \neq \sphericalangle(A'B'C').$$

Seja S o lado da reta $\overleftrightarrow{A'B'}$ que contém C'. O Segundo Axioma de Ângulos nos permite construir em S uma semirreta $\overrightarrow{B'P}$ que forma um ângulo $\beta' = \sphericalangle(ABC)$ com a semirreta $\overrightarrow{A'B'}$, como ilustrado na figura 12.

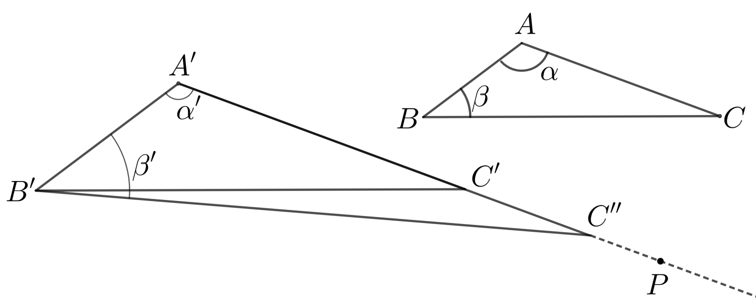


FIGURA 12. O caso LAL de semelhança de triângulos.

Como, por hipótese,

$$\alpha = \sphericalangle BAC = \sphericalangle B'A'C' = \alpha',$$

temos que

$$\alpha' + \beta' = \alpha + \beta < \pi,$$

pelo Axioma Euclidiano, pois as semirretas $\overrightarrow{AB}$ e $\overrightarrow{BC}$ têm o ponto B em comum. Logo, por este mesmo axioma, as semirretas $\overrightarrow{A'C'}$ e $\overrightarrow{B'P}$ se intersectam em um ponto C''. Assim, por construção,

$$\sphericalangle ABC = \sphericalangle A'B'C'' = \beta$$

ao passo que, por hipótese, $\alpha = \alpha'$. Portanto, pelo Axioma de Semelhança, os triângulos $\triangle ABC$ e $\triangle A'B'C''$ são semelhantes. Em particular,

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

Como, por hipótese,

$$\frac{AB}{A'B'} = \frac{AC}{A'C'},$$

podemos deduzir que

$$\frac{AC}{A'C''} = \frac{AC}{A'C'},$$

que só é possível se $A'C'' = A'C'$. Como C e C' estão ambos sobre uma mesma semirreta, isto implica que $C'' = C'$. Portanto,

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

concluindo nossa demonstração. □

É importante ressaltar que, no enunciado do teorema anterior, *os ângulos iguais têm que ser, obrigatoriamente, aqueles que são limitados pelos lados proporcionais*. Por exemplo, na figura 13 os triângulos $\triangle AOB$ e $\triangle AOC$ têm em comum o lado AO, e o ângulo $\sphericalangle OAB$, ao passo que os lados OB e OC têm a mesma medida, porque são raios de uma mesma circunferência. No entanto, estes triângulos claramente não são semelhantes.

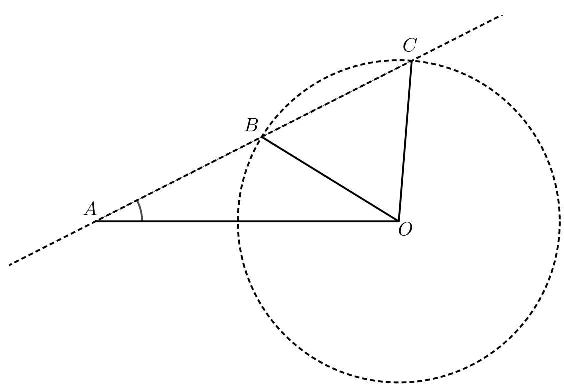


FIGURA 13. Os triângulos $\triangle AOB$ e $\triangle AOC$ não são congruentes.

Resta-nos considerar a segunda maneira pela qual identificamos que os triângulos da figura 11 não podiam ser semelhantes. Faremos isto mostrando que, se os três lados de um triângulo mantêm a mesma proporção com os lados correspondentes de outro triângulo, então estes triângulos são semelhantes. Para isto precisamos de um lema.

LEMA 5.3. *Se dois lados de um triângulo são iguais, então os ângulos opostos a estes lados também são iguais.*

DEMONSTRAÇÃO. Seja $\triangle ABC$ um triângulo cujos lados AB e AC são iguais. Usando o Segundo Axioma de Ângulos, construímos uma semirreta $\overrightarrow{AP}$ que forma um ângulo igual à metade de $\sphericalangle BAC$ do lado de $\overleftrightarrow{AB}$ que contém C; como ilustrado na figura 14.

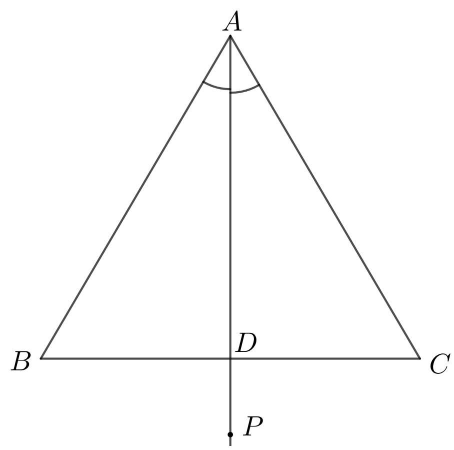


FIGURA 14. Triângulo com lados iguais.

Como a soma dos ângulos internos de um triângulo é igual a π, temos que

$$\frac{1}{2}\sphericalangle BAC + \sphericalangle BCA < \sphericalangle BAC + \sphericalangle BCA + \sphericalangle ABC = \pi,$$

o Axioma Euclidiano garante que $\overrightarrow{AP}$ intersecta BC em um ponto, que chamaremos de D. Mas, os triângulos $\triangle ABD$ e $\triangle ACD$ têm, por construção,

$$\sphericalangle BAD = \sphericalangle CAD$$

ao passo que $AB = AC$, por hipótese, e AD é lado comum a ambos os triângulos. Logo, $\triangle ABD \equiv \triangle ACD$ pelo caso LAL. Portanto,

$$\sphericalangle ABC = \sphericalangle ACD,$$

pois são ângulos opostos ao lado comum AD nos dois triângulos. $\square$

Um triângulo com dois lados iguais é chamado de *isosceles*, porque, em grego, *iso* significa igual e *skelos*, que significa perna. Portanto, o lema 5.3 mostra que um triângulo isósceles também tem dois ângulos iguais. A recíproca do lema também é verdadeira; você encontrará um esboço da demonstração no exercício 11.

TEOREMA 5.4. *Dois triângulos $\triangle ABC$ e $\triangle A'B'C'$ para os quais*

$$\frac{AB}{A'B'} = \frac{AC}{A'C'} = \frac{BC}{B'C'},$$

são semelhantes.

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

$$\frac{AB}{A'B'} = \frac{AC}{A'C'} = \frac{BC}{B'C'}, \tag{66}$$

número que denotaremos por k. Começaremos construindo a semirreta $\overrightarrow{A'P}$ que forma com $\overleftrightarrow{A'B'}$ um ângulo $\beta = \sphericalangle CAB$, do lado da reta $\overleftrightarrow{A'B'}$ oposto àquele onde C' está. Seja $C'' \in \overrightarrow{A'P}$ tal que $A'C'' = k \cdot AC$; como ilustrado na figura 15.

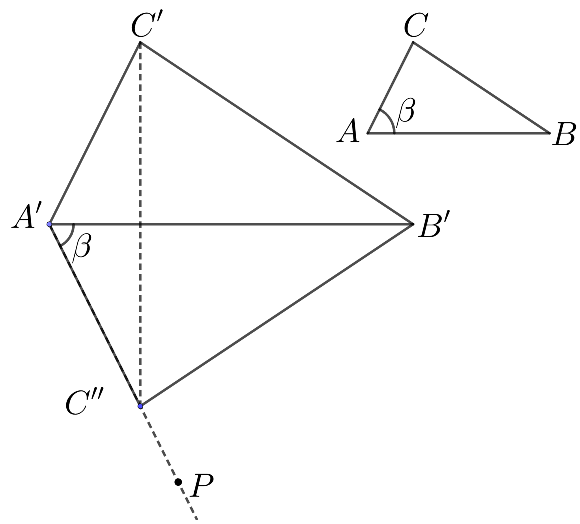


FIGURA 15. Os triângulos $\triangle ABC$, $\triangle A'B'C'$ e $\triangle A'B'C''$.

Pelo teorema 5.2, os triângulos $\triangle ABC$ e $\triangle A'B'C''$ são semelhantes com razão de semelhança igual a k. Podemos concluir disto, e de (66), que

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

Em particular,

$$A'C'' = A'C' \text{ e } B'C'' = B'C', \tag{67}$$

de modo que os triângulos $\triangle C'A'C''$ e $\triangle C'B'C''$ são ambos isósceles. Logo, pelo lema 5.3,

$$\sphericalangle A'C'C'' = \sphericalangle A'C''C' \text{ e } \sphericalangle B'C'C'' = \sphericalangle B'C''C'. \tag{68}$$

Mas, pela aditividade da medida de ângulos,

$$\sphericalangle A'C'B' = \sphericalangle A'C'C'' + \sphericalangle B'C'C'' \text{ e } \sphericalangle A'C''B' = \sphericalangle A'C''C' + \sphericalangle B'C''C'. \tag{69}$$

Combinando (68) com (69), concluímos que

$$\sphericalangle A'C'B' = \sphericalangle A'C''B'.$$

Contudo, aplicando o teorema 5.2 a (67) e levando em conta esta última igualdade, temos que os triângulos $\triangle A'B'C'$ e $\triangle A'B'C''$ são congruentes. Como $\triangle A'B'C'' \sim \triangle ABC$ por construção, podemos concluir que $\triangle A'B'C' \sim \triangle ABC$, como devíamos mostrar. □

Como a congruência de triângulos é um caso especial da semelhança, podemos inferir diretamente do Axioma de Semelhança e dos Teoremas 5.2 e 5.4 os casos três casos de congruência de triângulos.

COROLÁRIO 5.5. *Se em dois triângulos*

Caso ALA: *um lado e dois ângulos correspondentes são iguais, ou*
Caso LAL: *um ângulo e dois lados correspondentes são iguais, ou*
Caso LLL: *os três lados correspondentes são iguais,*

então os triângulos são congruentes.

De agora em diante usaremos estas abreviações sempre que precisarmos aplicar um dos casos de congruência de triângulos. Além disso, embora fosse razoável aplicar estas abreviações também aos casos de semelhança, vamos limitar à tradição de este uso aos casos de congruência, ainda que fosse legítimos aplicar as mesmas abreviações ao Axioma de Semelhança e aos teoremas 5.2 e 5.4.

Encerramos a seção usando semelhança de triângulos para justificar a estratégia usada, supostamente, por Tales para calcular a altura de uma pirâmide. Como vimos na página 4, Tales teria feito isto comparando sua sombra à da pirâmide.

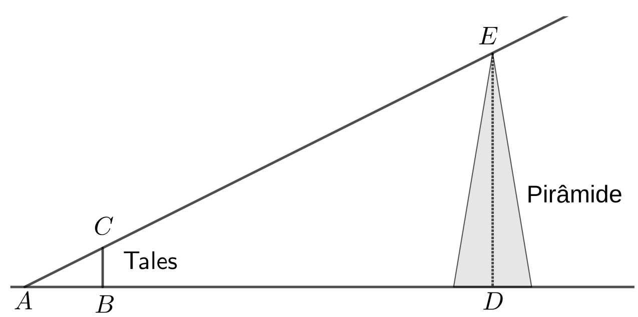


FIGURA 16. Tales e a pirâmide

Na figura 16, o sol está no canto superior direito e a semirreta $\overrightarrow{AE}$ corresponde a um raio solar que toca o ponto mais alto da pirâmide, assim como o alto da cabeça de Tales. Portanto, AD é a sombra da pirâmide e AB a sombra de Tales. Como o ângulo $\sphericalangle CAB$ é comum aos triângulos $\triangle ABC$ e $\triangle ADE$ e como

$$\sphericalangle(CBA) = \sphericalangle(EDA) = \pi/2,$$

podemos concluir pelo Axioma de Semelhança que $\triangle ABC \sim \triangle ADE$. Em particular,

$$\frac{AD}{AB} = \frac{ED}{CB}. \tag{70}$$

Segundo a história relatada por Diógenes Laércio, Tales teria aguardado o momento em que sua sombra era igual à sua altura. Na notação que estamos usando, isto corresponde a dizer que $AB = CB$. Substituindo na equação (70), obtemos $ED = AD$. Logo, bastaria que Tales soubesse o comprimento AD da sombra da pirâmide para determinar a altura ED que estava procurando, uma vez que ambas são iguais.

A semelhança de triângulos também pode ser usada para explicar a maneira como Descartes efetua a multiplicação dos comprimentos de dois segmentos de maneira geométrica, mencionada na seção 2. Digamos que AB e CD sejam os dois segmentos cujos comprimentos queremos multiplicar. Começamos construindo duas semirretas, com origem em O, que formam um ângulo agudo entre si. Sobre uma delas marcamos o segmento $OP = AB$ e, sobre a outra, os segmentos $OQ = CD$ e OU de comprimento uma unidade. Então, ligamos P a U e, pelo ponto Q, construímos uma paralela ao segmento UP. Esta construção é ilustrada na figura 17.

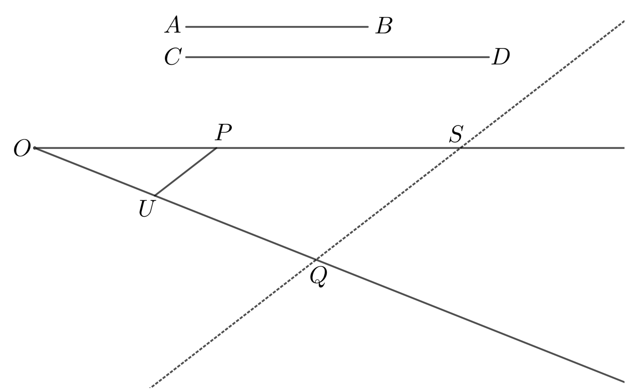


FIGURA 17. Multiplicação de segmentos segundo Descartes.

Como a semirreta $\overrightarrow{OP}$ é tranversal às retas paralelas $\overleftrightarrow{UP}$ e $\overleftrightarrow{SQ}$, a igualdade dos ângulos correspondentes garante que

$$\sphericalangle OPU = \sphericalangle OSQ.$$

Levando em conta que o ângulo $\sphericalangle UOP$ é comum aos triângulos $\triangle OPU$ e $\triangle OSQ$, temos por ALA que estes triângulos são semelhantes. Logo,

$$\frac{OP}{OS} = \frac{OU}{OQ};$$

donde,

$$OS = \frac{OP \cdot OQ}{OU} = OP \cdot OQ,$$

pois estamos supondo que OU é igual a uma unidade no sistema de medida que estamos utilizando. Uma questão óbvia é como faríamos para achar os pontos P e Q sem medir os segmentos e usar uma régua para achar estes pontos. Porque, se sabemos a medida dos segmentos, qual o ponto em fazer este processo complicado para achar seu produto? A resposta é que Descartes usaria um compasso, não só para achar P e Q, mas também para achar a paralela à reta $\overleftrightarrow{UP}$ por Q.

6. Perpendiculares

Vimos, na seção 4, que sempre é possível construir uma paralela a uma reta dada, que passa por um ponto fora dela. Resta-nos examinar como construir perpendiculares. Se o ponto estiver sobre reta, a existência da perpendicular é consequência do Segundo Axioma de Ângulos. Portanto, precisamos apenas considerar o caso em que o ponto está fora da reta.

TEOREMA 6.1. *Dados um ponto e uma reta, existe uma reta que passa pelo ponto dado e é perpendicular à reta dada.*

DEMONSTRAÇÃO. Como dito no parágrafo anterior ao teorema, basta considerar o caso em que o ponto não pertence à reta. Sejam P o ponto e ℓ a reta. Escolha um ponto Q, qualquer, sobre a reta e seja α o menor ângulo entre as retas ℓ e $r = \overleftrightarrow{PQ}$. Se $\alpha = \pi/2$, as retas r e ℓ são perpendiculares e não há mais nada a fazer. Portanto, podemos supor que $\alpha \neq \pi/2$.

Pela definição de ângulo, existe um ponto $Q' \in \ell$ tal que $\alpha = \sphericalangle PQQ'$. Seja S_r o lado da reta r que contém Q'. O Segundo Axioma de Ângulos nos permite construir uma semirreta $\overrightarrow{QT}$ em S_r de modo que $\sphericalangle PQT = 2\alpha$. Temos, então, pela aditividade da medida de ângulos, que

$$\sphericalangle PQQ' + \sphericalangle Q'QT = 2\alpha; \tag{71}$$

donde, $\sphericalangle Q'QT = \alpha$, pois $\sphericalangle PQQ' = \alpha$. Sobre a semirreta $\overrightarrow{QT}$ marcamos o ponto P' de modo que $QP' = QP$.

Como a medida de um ângulo é sempre um número positivo, a equação (71) garante que Q' é um ponto interior ao ângulo $\sphericalangle PQP'$. Logo, P e P' estão de lados opostos da reta ℓ. Assim, pelo Segundo Axioma de Separação, o segmento PP' intersecta ℓ em um ponto, que chamaremos de Q''. Toda esta construção está ilustrada na figura 18. Mas, por construção,

$$PQ = P'Q \quad \text{e} \quad \sphericalangle PQQ' = \alpha = \sphericalangle P'QQ'.$$

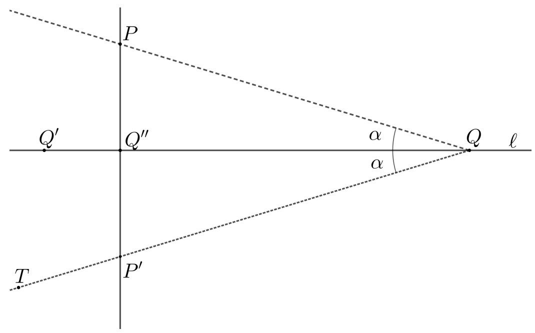


FIGURA 18. Perpendicular por ponto fora da reta.

Além disso, PQ'' é lado comum a $\triangle PQQ''$ e $\triangle P'QQ''$, de modo que estes dois triângulos são congruentes por LAL. Em particular,

$$\sphericalangle PQ''Q = \sphericalangle P'Q''Q.$$

Contudo, P, P' e Q'' pertencem a uma mesma reta, de modo que

$$\pi = \sphericalangle PQ''Q + \sphericalangle P'Q''Q = 2\sphericalangle PQ''Q.$$

Logo, $\overleftrightarrow{PP'}$ é perpendicular à reta ℓ, concluindo a demonstração. □

Este teorema nos permite definir a altura de um triângulo $\triangle ABC$. Começamos usando o teorema para construir a reta perpendicular à reta $\overleftrightarrow{BC}$ pelo ponto A. O ponto H em que estas duas retas se encontram é conhecido como *pé* da perpendicular. A *altura* de um triângulo é igual ao comprimento do segmento AH. É importante lembrar que o segmento correspondente à altura de um triângulo pode estar fora do triângulo. Como mostra a figura 19, isto acontecerá sempre que um dos ângulos internos do triângulo for maior que $\pi/2$.

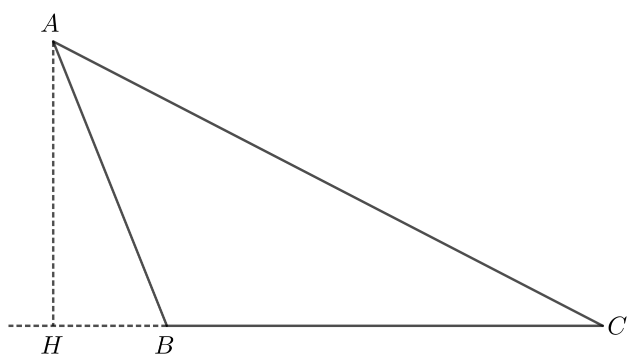


FIGURA 19. O segmento AH é a altura do triângulo $\triangle ABC$.

Podemos provar o teorema de Pitágoras usando o teorema 6.1 e o Axioma de Semelhança. Lembre-se que um *triângulo retângulo* é aquele em que um dos ângulos internos é reto. O lado oposto a um ângulo reto é chamado de *hipotenusa*, os outros dois ângulos são os *catetos* do triângulo retângulo.

TEOREMA DE PITÁGORAS. *Em um triângulo retângulo, o quadrado da hipotenusa é igual à soma dos quadrados dos catetos.*

DEMONSTRAÇÃO. Seja $\triangle ABC$ um triângulo retângulo e digamos que o ângulo reto é $\sphericalangle A$, de modo que a hipotenusa é BC e os catetos são AB e AC. Usando o teorema 6.1, podemos construir a altura de A sobre BC. Denotando por H o pé desta altura, verificamos que

$$\sphericalangle(AHB) = \sphericalangle(AHC) = \pi/2.$$

Temos, assim, três triângulos retângulos, como ilustrado na figura 20.

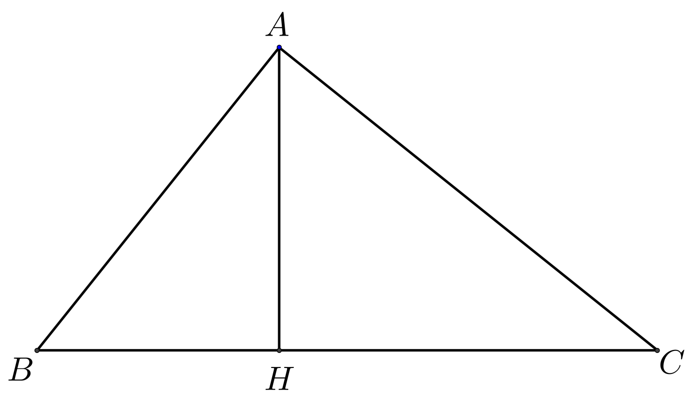


FIGURA 20. Teorema de Pitágoras

Mas, $\triangle ABC$ e $\triangle ABH$ têm dois ângulos iguais, pois

$$\sphericalangle(BAC) = \sphericalangle(AHB) = \pi/2$$

e o ângulo $\sphericalangle ABH$ é comum a estes dois triângulos. Portanto,

$$\frac{AB}{BC} = \frac{BH}{AB}.$$

No lado esquerda desta equação, AB faz o papel de hipotenusa do triângulo $\triangle ABH$, ao passo que no lado direito, o mesmo lado representa um cateto do triângulo $\triangle ABC$. A equação anterior é equivalente a

$$AB^2 = BC \cdot BH. \tag{72}$$

Um argumento análogo mostra que o triângulo $\triangle ACH$ também é semelhante a $\triangle ABC$ e que disto resulta

$$AC^2 = BC \cdot CH. \tag{73}$$

Somando (72) e (73), obtemos

$$AB^2 + AC^2 = BC \cdot BH + BC \cdot CH = BC \cdot (BH + CH).$$

Como $BH + CH = BC$, podemos concluir que

$$AB^2 + AC^2 = BC^2,$$

provando, assim, o teorema. □

É difícil subestimar a importância do teorema de Pitágoras, já que está por trás de aplicações como: o cálculo de distância entre pontos do plano dados em termos de coordenadas, a equação do círculo e a relação básica entre o seno e o cosseno de um ângulo. Basta lembrar dos inúmeros usos das funções trigonométricas em engenharia e física para se dar conta da importância deste teorema.

Exercícios

1. Seja ℓ uma reta e $P \in \ell$. Prove que, se $Q \neq \ell$, então os pontos do segmento PQ, diferentes de P, estão todos do mesmo lado de ℓ em que Q está.
   ESBOÇO DA DEMONSTRAÇÃO: suponhamos, por contradição, que $PQ \setminus \{P\}$ não está todo contido do mesmo lado S de ℓ que contém Q. Neste caso, existe um pontos $R \notin S \cup \ell$ tal que $QR \cap \ell = \{S\}$. Como $R \neq P$, temos que $x_P < x_R < x_S$, de modo que $S \neq P$. Use isto e o Axioma de Incidência para obter uma contradição.

2. Sejam r uma reta, $A \in r$ e S_r um dos semiplanos em que r divide o plano. Mostre, por contradição, que se B é um ponto de S_r, então $AB \subset S_r \cup \{A\}$.

3. Sobre uma semirreta $\overrightarrow{OX}$ marcamos um ponto $A \neq O$ e sobre uma semirreta $\overrightarrow{OX'}$ marcamos $A' \neq O$. Mostre que $AA' \subset \sphericalangle AOA'$.

4. Sobre uma semirreta $\overrightarrow{OX}$ marcamos dois pontos distintos $A \neq O$ e $B \neq O$ e, sobre uma semirreta $\overrightarrow{OX'}$ marcamos A' e B', de modo que $OA = OA'$, $AB = A'B'$, $A \in OB$ e $A' \in OB'$.
   (a) Mostre que O e B estão de lados opostos da reta $\overleftrightarrow{AB'}$.
   (b) Mostre que O e A' estão do mesmo lado da reta $\overleftrightarrow{AB'}$.
   (c) Prove que $\overleftrightarrow{A'B}$ intersecta AB'.
   (d) Use um argumento semelhante para mostrar que $\overleftrightarrow{AB'}$ intersecta $A'B$.
   (e) Prove que os segmentos $A'B$ e AB' têm um ponto de interseção.

5. Sejam r, ℓ e ℓ' três retas. Prove que, se r é transversal a ℓ e ℓ e ℓ' são paralelas, então r também é transversal a ℓ'.
   DICA: Seja P o ponto de interseção entre r e ℓ e suponha, por contradição, que a afirmação é falsa. Use a unicidade no teorema 4.4 para obter uma contradição.

6. Prove que se r é transversal a ℓ e ℓ' e soma dos ângulos internos é π, então ℓ e ℓ' são paralelas.
   DICA: Seja P o ponto de interseção entre r e ℓ e construa uma paralela ℓ'' a ℓ' por P. Mostre que ℓ e ℓ' formam o mesmo ângulo com r e use o segundo axioma de ângulos para concluir que $\ell = \ell''$.

7. Complete os detalhes da demonstração do corolário 4.2.

8. Sejam A, B e C três pontos em uma reta ℓ, de modo que B está entre A e C e sejam P e Q dois pontos de um dos semiplnaos S_ℓ em que ℓ divide o plano. Prove, por contradição, que se $\sphericalangle PAC = \sphericalangle QBC$ então as retas $\overleftrightarrow{PA}$ e $\overleftrightarrow{QB}$ não têm interseção em S_ℓ. O que falta provar para podermos garantir que $\overleftrightarrow{PA}$ e $\overleftrightarrow{QB}$ são paralelas?

9. Prove que, se P e Q são dois pontos no interior de um triângulo, então o segmento PQ também está contido no triângulo.
   DICA: use o Segundo Axioma de Separação três vezes, uma para cada reta suporte dos lados.

10. Seja $ABCD$ um quadrilátero no qual $\sphericalangle ABC = \sphericalangle ADC$. Sabendo-se que as retas AB e DC são paralelas:
    (a) prove que os triângulos $\triangle ABC$ e $\triangle ADC$ são congruentes;
    (b) prove que $ABCD$ é um paralelogramo.

11. Seja $\triangle ABC$ um triângulo. Prove que se $\sphericalangle ABC = \sphericalangle ACB$ então $AB = AC$.
    DICA: seja H o pé da altura de A sobre BC; mostre que $\triangle ABH \equiv \triangle ACH$.

12. Mostre que todo ângulo exterior de um triângulo é a soma de seus ângulos interiores não adjacentes. Por exemplo, na figura 21, θ é o ângulo exterior correspondente ao vértices B e os ângulos interiores não adjacentes a ele são α e γ

13. Seja $\triangle ABC$ um triângulo cujos lados AB e AC são iguais. Mostre que as alturas de B sobre AC e de C sobre AB são iguais.

14. A mediana de ABC pelo vértice A é o segmento de reta que liga A ao ponto médio do lado BC.Mostre que se, em um triângulo $\triangle ABC$, a mediana AM é perpendicular ao lado BC, então o triângulo é isósceles.

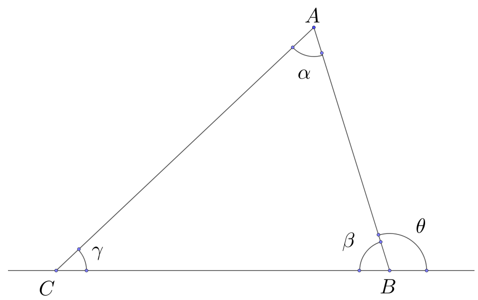


FIGURA 21. Exercício 12.

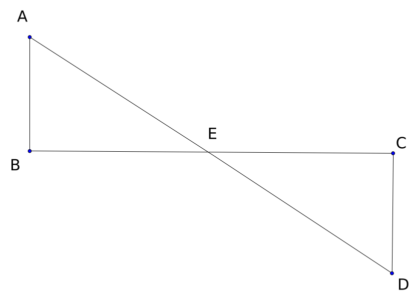


FIGURA 22. Exercício 15.

15. Na figura 22, sabe-se que AD bissecta BC e que os ângulos $\sphericalangle ABE$ e $\sphericalangle DCE$ são retos. Prove que $AE = ED$

16. Seja $\triangle ABC$ um triângulo cujos lados AB e AC são iguais. Sabe-se que E é o ponto médio de BC e que os pontos $F \in AB$ e $G \in BC$ são tais que $\sphericalangle BEF = \sphericalangle CEG$, como ilustrado na figura 23. Prove que $FE = GE$ e que o triângulo $\triangle AFG$ é isósceles.

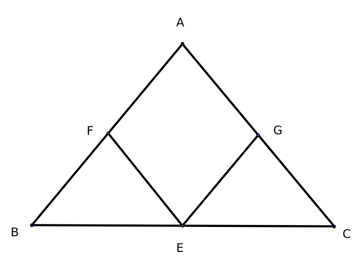


FIGURA 23. Exercício 11

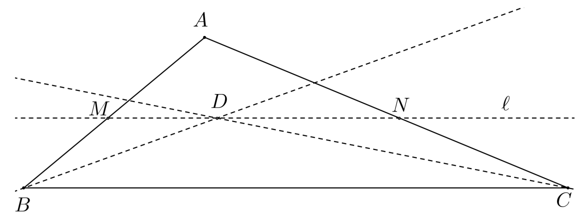


FIGURA 24. Exercício 17

17. Seja D o ponto de interseção das bissetrizes dos ângulos $\sphericalangle B$ e $\sphericalangle C$ de um triângulo $\triangle ABC$. Trace por D uma reta ℓ paralela ao lado BC e digamos que $AB \cap \ell = \{M\}$ e que $AC \cap \ell = \{N\}$, como na figura 24. Prove que $BM + CN = DM + DN$.

18. Sobre uma semirreta $\overrightarrow{OX}$ marcamos dois pontos distintos $A \neq O$ e $B \neq O$ e, sobre uma semirreta $\overrightarrow{OX'}$ marcamos A' e B', de modo que $OA = OA'$ e $AB = A'B'$. Seja I o ponto de interseção entre os segmentos AB' e $A'B$. O objetivo desta questão é provar que a semirreta $\overrightarrow{OI}$ divide o ângulo $\sphericalangle XOX'$ ao meio. Para isto, siga o roteiro abaixo:
    (a) Prove que $\triangle OAB' \equiv \triangle OA'B$ são congruentes.
    (b) Explique porque $\sphericalangle A'IB' = \sphericalangle AIB$.
    (c) Prove que $\triangle AIB \equiv \triangle A'IB'$.
    (d) Prove que $\triangle OIB \equiv \triangle OIB'$.
    (e) Explique porque $\sphericalangle BOI = \sphericalangle B'OI$.

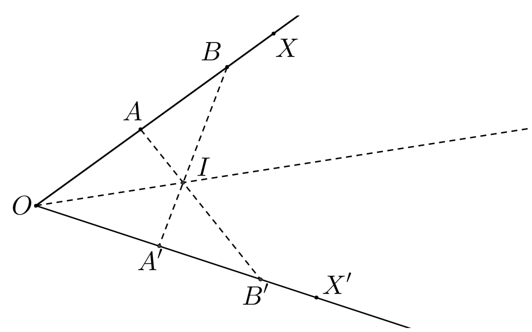


FIGURA 25. Exercício 18

19. Seja $\triangle ABC$ um triângulo no qual $AB = AC$ e seja $D \in BC$ tal que $\sphericalangle BDA = \sphericalangle CDA$. Prove que:
    (a) AD é perpendicular a BC;
    (b) $\sphericalangle BAD = \sphericalangle CAD$;
    (c) $\triangle ABD \equiv \triangle ADC$.

20. Para medir a distância entre as duas ilhas (regiões pintadas de cinza) na figura 26, Hipácia mede as seguintes distâncias $AC = 5m$, $DE = 6m$ e $AE = 15m$. Sabendo-se que os ângulos $\sphericalangle EAC$ e $\sphericalangle DEA$ são retos, calcule a distância entre as ilhas.

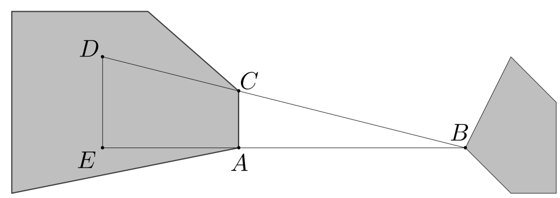


FIGURA 26. Exercício 20

21. O paralelogramo $ABCD$ da figura abaixo tem base igual a $AB = 51$ cm, altura $DE = 15$ cm e lado transversal $AD = 17$ cm. Calcule FC sabendo-se que o ângulo $\sphericalangle DFC$ é reto.

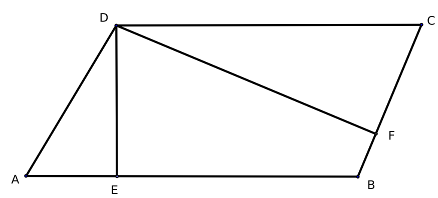


FIGURA 27. Exercício 21.

22. Resolva o nono problema do tablete BM85196, encontrado na Mesopotâmia, segundo o qual

    > uma haste de 30 (unidades) está apoiada em uma parede, de quando a extremidade inferior da haste vai se afastar da parede se sua extremidade superior escorregar 6 unidades?

    Mais detalhes sobre este problema podem ser encontrados em (Melville, 2004, p. 150-151).

CAPÍTULO 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, em vez do nosso seno e cosseno, os gregos utilizavam cordas do círculo. O seno e o cosseno que utilizamos hoje em dia foram introduzidos pelos matemáticos indianos antes de 543 a.C. Neste capítulo apresentamos as principais funções trigonométricas de maneira rigorosa, usando apenas o que aprendemos sobre geometria no capítulo 5.

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

$$\text{sen}(\alpha) = \frac{AB}{OB} \quad \text{e} \quad \cos(\alpha) = \frac{OA}{OB}. \tag{74}$$

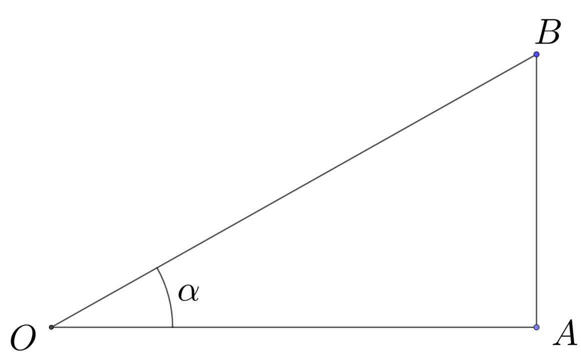


FIGURA 1. Triângulo retângulo com ângulo α

Como,

$$\operatorname{sen}^2(\alpha) + \cos^2(\alpha) = \left(\frac{AB}{OB}\right)^2 + \left(\frac{OA}{OB}\right)^2 = \frac{AB^2 + OA^2}{OB^2},$$

temos, pelo teorema de Pitágoras, que

$$\operatorname{sen}^2(\alpha) + \cos^2(\alpha) = 1. \tag{75}$$

Em particular, se construírmos o triângulo retângulo de modo que sua hipotenusa tenha comprimento igual a 1, o vértice B cai sobre o círculo de raio 1. Isto nos permite redesenhar a figura 1 na forma abaixo. Neste contexto frequentemente nos referimos ao círculo de raio 1 como *o círculo trigonométrico*.

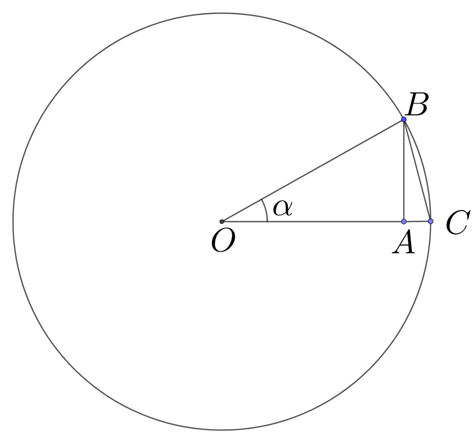


FIGURA 2. Triângulo retângulo inscrito no círculo

Sob a hipótese de que a hipotenusa OB é igual a 1, as fórmulas em (74) resumem-se a

$$\operatorname{sen}(\alpha) = AB \quad \text{e} \quad \cos(\alpha) = OA. \tag{76}$$

Como observado na introdução deste capítulo, o seno e o cosseno foram introduzidos pelos matemáticos indianos; os matemáticos gregos utilizavam a corda BC. Podemos calculá-la, facilmente, usando o seno, o cosseno e o teorema de Pitágoras, pois BC é a hipotenusa do triângulo retângulo $\triangle ABC$ cujos catetos são $\operatorname{sen}(\alpha)$ e $1 - \cos(\alpha)$. Assim,

$$BC^2 = \operatorname{sen}^2(\alpha) + (1 - \cos(\alpha))^2,$$

de modo que, por (75),

$$BC^2 = 2 - 2\cos(\alpha).$$

Uma das vantagens do seno e cosseno sobre as cordas, é que os dois primeiros são mais fáceis de aproximar.

Como, para cada triângulo retângulo de hipotenusa igual a um, existe apenas um cateto oposto ao ângulo $\alpha = \sphericalangle COB$, as fórmulas em (76) defininem o seno e o cosseno como funções de α. Contudo, como os ângulos opostos aos catetos de um triângulo retângulo têm que ter soma igual a $\pi/2$, as fórmulas em (76) só nos permitem calcular o seno e o cosseno de α quando $0 < \alpha < \pi/2$.

Utilizando a notação da figura 2, vemos que, à medida que B desce ao longo do círculo em direção ao ponto C, o segmento BA se aproxima de zero, ao passo que OA se aproxima de $OC = 1$. Isto sugere que

$$\operatorname{sen}(0) = 0 \quad \text{e} \quad \cos(0) = 1.$$

Por outro lado, à medida que B sobe em direção à vertical, OA se aproxima cada vez mais de zero e BA cada vez mais de um; de modo que esperamos que

$$\operatorname{sen}(\pi/2) = 1 \quad \text{e} \quad \cos(\pi/2) = 0.$$

Note que usamos palavras como "sugere" e "esperamos", porque estes argumentos não são, de forma alguma, rigorosos. No próximo capítulo veremos como generalizar as definições do seno e do cosseno de maneira rigorosa, para que possamos tratá-los como funções cujo domínio é todo o conjunto $\mathbb{R}$. Por enquanto vamos nos considerar ambas as funções como tendo domínio no intervalo $[0, \pi/2]$.

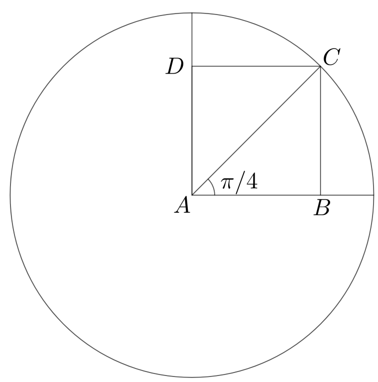


FIGURA 3. Seno e cosseno de $\pi/4$.

Um ângulo cujo seno e cosseno são fáceis de calcular é $\pi/4$ radianos, correspondente a um oitavo de circunferência. Quando $\alpha = \pi/4$, o vértice C cai, exatamente, a meio caminho entre zero e $\pi/2$, como mostra a figura 3. Assim,

$$\sphericalangle BAC = \sphericalangle CAD = \frac{\pi}{4}.$$

Mas,

$$\sphericalangle ADC = \sphericalangle ABC$$

são ambos ângulos retos. Logo, os triângulos $\triangle ABC$ e $\triangle ADC$ são congruentes pelo caso ALA, porque AC é um lado comum a a estes dois triângulos. Em particular,

$$\cos\left(\frac{\pi}{4}\right) = AB = AD = \operatorname{sen}\left(\frac{\pi}{4}\right).$$

Levando isto em conta, obtemos por (75) que

$$2\cos^2\left(\frac{\pi}{4}\right) = 1.$$

Mas o seno e o cosseno, como definimos até aqui, são sempre maiores ou iguais a zero, porque são comprimentos de segmentos. Portanto,

$$\operatorname{sen}\left(\frac{\pi}{4}\right) = \cos\left(\frac{\pi}{4}\right) = \frac{\sqrt{2}}{2}.$$

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 da soma de dois ângulos.

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 4, na qual o ângulo $\sphericalangle OCQ$ é reto. Note que

$$\sphericalangle OEA = \sphericalangle CED,$$

porque são opostos pelo vértice. Como os triângulos $\triangle OBC$ e $\triangle CEQ$ são ambos retângulos, segue-se que

$$\gamma = \sphericalangle CQD = \sphericalangle BOC = \alpha.$$

Isto nos permite calcular o seno e o cosseno de α a partir dos triângulos $\triangle OBC$, $\triangle QDC$ e $\triangle QCE$. Faremos uso disto no cálculo de ambas as fórmulas para soma de ângulos.

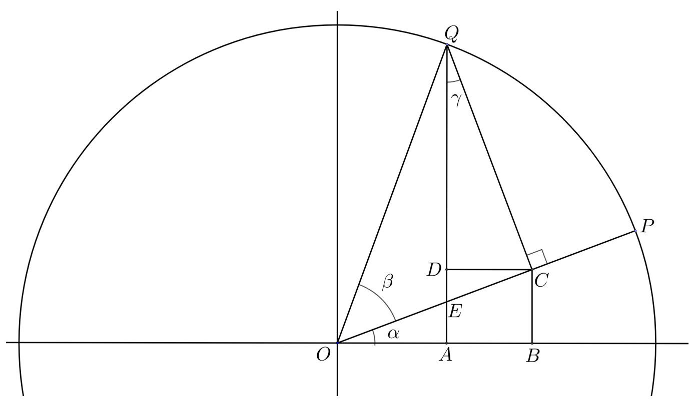


FIGURA 4. Seno e cosseno de uma soma de ângulos.

Começamos calculando

$$\operatorname{sen}(\alpha+\beta) = QA = QD + DA. \tag{77}$$

Do triângulo retângulo $\triangle QDC$,

$$\cos(\alpha) = \cos(\gamma) = \frac{QD}{QC}; \quad \text{isto é,} \quad QD = QC\cos(\alpha).$$

Por outro lado, do triângulo retângulo $\triangle OQC$,

$$\operatorname{sen}(\beta) = \frac{QC}{OQ},$$

donde

$$QC = \operatorname{sen}(\beta), \tag{78}$$

pois $OQ = 1$ é o raio do círculo. Logo,

$$QD = \cos(\alpha)\operatorname{sen}(\beta). \tag{79}$$

Para calcular DA, note, primeiramente, que $DA = CB$, de modo que

$$\operatorname{sen}(\alpha) = \frac{BC}{OC}; \quad \text{isto é,} \quad BC = OC\operatorname{sen}(\alpha).$$

Porém, pelo triângulo $\triangle OQC$,

$$\cos(\beta) = \frac{OC}{OQ} = OC. \tag{80}$$

Assim,

$$BC = \cos(\beta)\operatorname{sen}(\alpha). \tag{81}$$

Combinando (77) com (79) e (81), encontramos

$$\operatorname{sen}(\alpha+\beta) = \cos(\alpha)\operatorname{sen}(\beta) + \cos(\beta)\operatorname{sen}(\alpha), \tag{82}$$

que é a primeira das duas fórmulas que almejamos obter. Um caso particular muito útil é obtido fazendo $\beta = \alpha$ em (82), que nos dá,

$$\operatorname{sen}(2\alpha) = 2\operatorname{sen}(\alpha)\cos(\alpha). \tag{83}$$

O argumento usado para determinar a fórmula do cosseno da soma de dois ângulos é muito semelhante ao anterior. Como a hipotenusa OQ do triângulo $\triangle OQA$ é igual a 1,

$$\cos(\alpha+\beta) = OA = OB - AB. \tag{84}$$

Mas, calculando o cosseno de α relativamente ao triângulo $\triangle OCB$,

$$\cos(\alpha) = \frac{OB}{OC}, \quad \text{isto é,} \quad OB = \cos(\alpha)OC. \tag{85}$$

Calculando, agora, o seno de α relativamente ao triângulo $\triangle DCQ$,

$$\operatorname{sen}(\alpha) = \frac{DC}{QC}, \quad \text{donde,} \quad DC = \operatorname{sen}(\alpha)QC. \tag{86}$$

Lembrando que $AB = DC$, basta combinar as equações (84), (85) e (86) para obter

$$\cos(\alpha + \beta) = \cos(\alpha)OC - \operatorname{sen}(\alpha)QC.$$

Portanto, por (78) e (80),

$$\cos(\alpha + \beta) = \cos(\alpha)\cos(\beta) - \operatorname{sen}(\alpha)\operatorname{sen}(\beta). \tag{87}$$

Substituindo $\beta = \alpha$ em (87),

$$\cos(2\alpha) = \cos^2(\alpha) - \operatorname{sen}^2(\alpha) = 2\cos^2(\alpha) - 1, \tag{88}$$

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 (82), obtendo

$$\operatorname{sen}\left(\frac{2\pi}{3} + \frac{\pi}{3}\right) = \operatorname{sen}\left(\frac{\pi}{3}\right)\cos\left(\frac{2\pi}{3}\right) + \cos\left(\frac{\pi}{3}\right)\operatorname{sen}\left(\frac{2\pi}{3}\right)$$

Mas,

$$\operatorname{sen}\left(\frac{2\pi}{3} + \frac{\pi}{3}\right) = \operatorname{sen}(\pi) = 0;$$

donde

$$\operatorname{sen}\left(\frac{\pi}{3}\right)\cos\left(\frac{2\pi}{3}\right) = -\cos\left(\frac{\pi}{3}\right)\operatorname{sen}\left(\frac{2\pi}{3}\right)$$

Usando, agora, (83) e (88),

$$\operatorname{sen}\left(\frac{\pi}{3}\right)\cdot\left(2\cos^2\left(\frac{\pi}{3}\right) - 1\right) = -\cos\left(\frac{\pi}{3}\right)\cdot 2\operatorname{sen}\left(\frac{\pi}{3}\right)\cos\left(\frac{\pi}{3}\right).$$

Como $0 < \pi/3 < \pi/2$, 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, por causa da definição que estamos utilizando, o seno e o cosseno de $\pi/3$ tem que ser maior ou igual a zero; de modo que

$$\cos\left(\frac{\pi}{3}\right) = \frac{1}{2}.$$

Combinando isto com

$$\operatorname{sen}^2\left(\frac{\pi}{3}\right) + \cos^2\left(\frac{\pi}{3}\right) = 1,$$

Ângulo	0	$\pi/6$	$\pi/4$	$\pi/3$
Seno	0	$1/2$	$\sqrt{2}/2$	$\sqrt{3}/2$
Cosseno	1	$\sqrt{3}/2$	$\sqrt{2}/2$	$1/2$

TABELA 1. Senos e cossenos de alguns ângulos no intervalo $[0, \pi/2]$.

e levando em conta que o seno de $\pi/3$ também é positivo,

$$\operatorname{sen}\left(\frac{\pi}{3}\right) = \frac{\sqrt{3}}{2}.$$

Por outro lado, pela fórmula do cosseno do dobro de um ângulo,

$$\cos\left(\frac{\pi}{3}\right) = \cos\left(\frac{2\pi}{6}\right) = 2\cos^2\left(\frac{\pi}{6}\right) - 1;$$

donde

$$\frac{1}{2} = 2\cos^2\left(\frac{\pi}{6}\right) - 1.$$

Assim

$$\cos^2\left(\frac{\pi}{6}\right) = \frac{3}{4};$$

donde

$$\cos\left(\frac{\pi}{6}\right) = \frac{\sqrt{3}}{2}.$$

Como

$$\operatorname{sen}^2\left(\frac{\pi}{6}\right) + \cos^2\left(\frac{\pi}{6}\right) = 1,$$

obtemos

$$\operatorname{sen}\left(\frac{2\pi}{6}\right) = \sqrt{1 - \cos\left(\frac{\pi}{6}\right)^2} = \sqrt{1 - \frac{3}{4}} = \frac{1}{2}.$$

Na tabela 1 listamos os valores dos senos e cossenos dos ângulos que calculamos até aqui.

3. Tangente

Tendo investigado o seno e o cosseno, passamos a analisar a *tangente* de um ângulo α que é definida por

$$\tan(\alpha) = \frac{\operatorname{sen}(\alpha)}{\cos(\alpha)}.$$

Note que este quociente *não* faz sentido quando $\alpha = \pi/2$, porque o cosseno deste ângulo é nulo. Portanto, sob nossas atuais restrições, a tangente é uma função cujo domínio é o

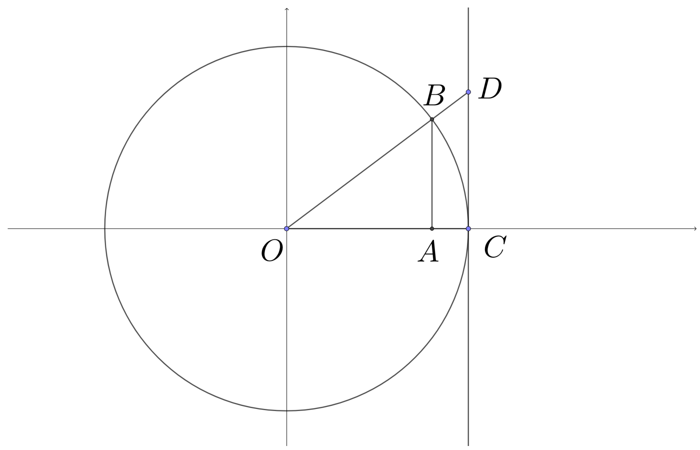


FIGURA 5. Interpretação geométrica da tangente

intervalo $[0, \pi/2)$. Usando os valores dos senos e cossenos calculados nas seções anteriores é facil verificar que

$$\tan(0) = 0, \quad \tan\left(\frac{\pi}{6}\right) = \frac{\sqrt{3}}{3}, \quad \tan\left(\frac{\pi}{4}\right) = 1 \text{ e } \tan\left(\frac{\pi}{3}\right) = \sqrt{3}.$$

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 retângulos $\triangle OBA$ e $\triangle OCD$ da figura 5, 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 $\sphericalangle AOB$ é comum aos triângulos $\triangle OBA$ e $\triangle OCD$, podemos concluir, pelo Axioma de Semelhança, que estes triângulos são semelhantes. Logo,

$$\frac{CD}{1} = \frac{AB}{OA} = \frac{\operatorname{sen}(\alpha)}{\cos(\alpha)};$$

donde,

$$CD = \tan(\alpha).$$

Isto também explica porque a tangente tem este nome, já que a reta DC é tangente à circunferência em C. Para entender 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,

$$OD = \sqrt{1 + DC^2}.$$

Como estamos supondo que $D \neq C$, temos que $DC > 0$, de modo que, pela fórmula anterior, $OD > 1$. 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.

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 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 6.

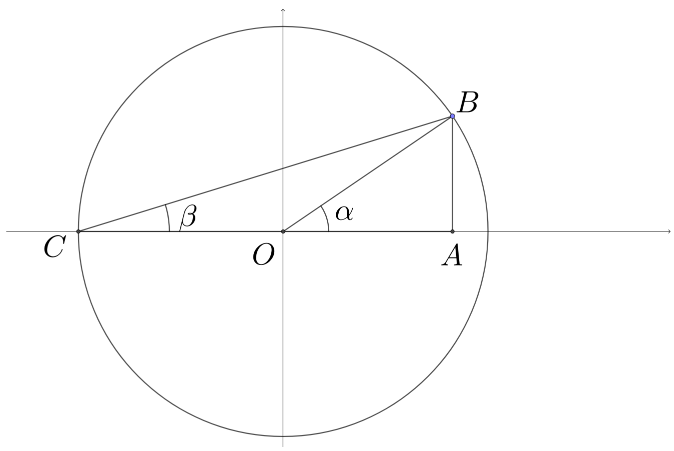


FIGURA 6. Seno e cosseno relativamente à tangente do meio ângulo

A primeira coisa a observar é que o triângulo $\triangle COB$ é isósceles, porque OC e OB são raios do círculo unitário. Logo $\sphericalangle CBO = \beta$ pelo Lema 5.3. O teorema 4.5 nos permite, então, concluir que

$$2\beta + (\pi - \alpha) = \pi;$$

donde

$$\beta = \frac{\alpha}{2}.$$

Calculando a tangente de β relativamente ao triângulo $\triangle ABC$, obtemos

$$\tan(\beta) = \frac{AB}{AC}. \tag{89}$$

Por outro lado, como o círculo tem raio igual a um, o triângulo $\triangle OBA$ nos dá

$$OA = \cos(\alpha) \quad \text{e} \quad AB = \operatorname{sen}(\alpha).$$

Logo,

$$AC = 1 + OA = 1 + \cos(\alpha).$$

Substituindo as expressões para AC e AB em (89),

$$\tan(\beta) = \frac{\operatorname{sen}(\alpha)}{1 + \cos(\alpha)}.$$

Para simplificar a notação, escreveremos

$$s = \operatorname{sen}(\alpha), \quad c = \cos(\alpha) \ \text{ e } \ t = \tan(\beta);$$

de modo que a fórmula anterior toma a forma

$$t = \frac{s}{1+c}. \tag{90}$$

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 (90),

$$s = t(1+c) = t\left(1 + \frac{1-t^2}{1+t^2}\right) = \frac{2t}{1+t^2}.$$

Resumindo o que vimos, temos que

$$\cos(\alpha) = \frac{1-t^2}{1+t^2} \ \text{ e } \ \operatorname{sen}(\alpha) = \frac{2t}{1+t^2}, \ \text{ em que } \ t = \tan\left(\frac{\alpha}{2}\right);$$

desde que $0 \leq \alpha < \pi/2$.

4. Identidades trigonométricas

As funções trigonométricas satisfazem várias relações que, como valem para qualquer valor do ângulo, são conhecidas como *identidades trigonométricas*. A mais óbvia de todas é

$$\operatorname{sen}^2(\alpha) + \cos^2(\alpha) = 1, \tag{91}$$

Uma identidade mais interessante é

$$\operatorname{sen}(\alpha)(1 + \tan(\alpha)) = \tan(\alpha)(\operatorname{sen}(\alpha) + \cos(\alpha)).$$

Para prová-la, substituímos $\tan(\alpha)$ por $\operatorname{sen}(\alpha)/\cos(\alpha)$, do lado esquerdo e usamos a distributividade, obtendo

$$\operatorname{sen}(\alpha)\left(1 + \frac{\operatorname{sen}(\alpha)}{\cos(\alpha)}\right) = \tan(\alpha)(\operatorname{sen}(\alpha) + \cos(\alpha)).$$

que prova a identidade desejada. Nesta demonstração, bastou fazer a substituição da tangente por sua definição em termos do seno e do cosseno, mas raramente estas demonstrações são tão simples.

Considere, por exemplo,

$$(1-\operatorname{sen}(\alpha)+\cos(\alpha))^2 = 2(1-\operatorname{sen}(\alpha))(1+\cos(\alpha)). \tag{92}$$

Aplicando o produto notável referente ao quadrado de uma soma, obtemos

$$(1-\operatorname{sen}(\alpha)+\cos(\alpha))^2 = 1-2(\operatorname{sen}(\alpha)-\cos(\alpha))+(\operatorname{sen}(\alpha)-\cos(\alpha)))^2.$$

Fazendo isto mais uma vez,

$$(1-\operatorname{sen}(\alpha)+\cos(\alpha))^2 = 1-2(\operatorname{sen}(\alpha)-\cos(\alpha))+(\cos^2(\alpha)+\operatorname{sen}^2(\alpha)-2\cos(\alpha)\operatorname{sen}(\alpha)).$$

Usando (91),

$$(1-\operatorname{sen}(\alpha)+\cos(\alpha))^2 = 2-2(\operatorname{sen}(\alpha)-\cos(\alpha)-\cos(\alpha)\operatorname{sen}(\alpha)).$$

Por outro lado, multiplicando os dois termos do lado direito de (92),

$$2(1-\operatorname{sen}(\alpha))(1+\cos(\alpha)) = 2(1-\operatorname{sen}(\alpha)+\cos(\alpha)-\operatorname{sen}(\alpha)\cos(\alpha))$$

que é igual à expressão anterior. Uma maneira alternativa de proceder utiliza as fórmulas do seno e do cosseno relativas à tangente da metade do ângulo, encontradas ao final da seção anterior; isto é

$$\cos(\alpha) = \frac{1-t^2}{1+t^2} \quad \text{e} \quad \operatorname{sen}(\alpha) = \frac{2t}{1+t^2}, \quad \text{em que } t = \tan\left(\frac{\alpha}{2}\right). \tag{93}$$

Substituindo estas expressões em (92), encontramos

$$\left(1-\frac{1-t^2}{1+t^2}+\frac{2t}{1+t^2}\right)^2 = 2\left(1-\frac{1-t^2}{1+t^2}\right)\left(1+\frac{2t}{1+t^2}\right). \tag{94}$$

Como

$$1-\frac{2t}{1+t^2}+\frac{1-t^2}{1+t^2} = \frac{1+t^2+1-t^2-2t}{1+t^2} = \frac{2t(t+1)}{1+t^2}$$

o lado esquerdo de (94) é igual a

$$\left(\frac{2(t-1)}{1+t^2}\right)^2 = \left(\frac{4(t^2-2t+1)}{(1+t^2)^2}\right). \tag{95}$$

Já o lado direito é igual a

$$2\left(1-\frac{2t}{1+t^2}\right)\left(1+\frac{1-t^2}{1+t^2}\right) = 2\left(1-\frac{2t}{1+t^2}+\frac{1-t^2}{1+t^2}-\frac{2t}{1+t^2}\cdot\frac{1-t^2}{1+t^2}\right).$$

Pondo tudo sobre um mesmo denominador, encontramos

$$2\left(\frac{(1+t^2)^2-2t(1+t^2)+(1+t^2)(1-t^2)-2t(1-t^2)}{(1+t^2)^2}\right).$$

Expandindo os produtos notáveis e simplificando, obtemos

$$2\left(\frac{2t^2-4t+2}{(1+t^2)^2}\right),$$

que é igual a (95), como queríamos mostrar. É provável que você esteja achando que este segundo método é tudo, menos mais simples que o outro. Mas a grande vantagem é que pode ser facilmente automatizado, porque só envolve reescrever as funções trigonométricas em termos de seno e cosseno e substituir as expressões (93); o resto dos cálculos é executado muito rápido em um sistema de computação algébrica.

Vejamos mais um exemplo. Desta vez, queremos saber se a igualdade

$$\left(\frac{(\sec^2(-\alpha)-\tan^2(\alpha))}{\tan(\alpha)}\right)\left(\frac{1+\tan(\alpha)}{1+\cot(\alpha)}\right)-2\operatorname{sen}^2(\alpha)=\cos(2\alpha) \tag{96}$$

é ou não verdadeira. A equação envolve algumas funções trigonométricas que ainda não encontramos, porque são extremamente secundárias. Entre elas estão a *secante*

$$\sec(\alpha)=\frac{1}{\cos(\alpha)},$$

a *cossecante*

$$\csc(\alpha)=\frac{1}{\operatorname{sen}(\alpha)}$$

e a *cotangente*

$$\cot(\alpha)=\frac{1}{\tan(\alpha)}=\frac{\cos(\alpha)}{\operatorname{sen}(\alpha)}.$$

Fazendo as devidas substituições do lado esquerdo da igualdade, obtemos

$$\left(\frac{\frac{1}{\cos(\alpha)}-\frac{\operatorname{sen}^2(\alpha)}{\cos^2(\alpha)}}{\frac{\operatorname{sen}^2(\alpha)}{\cos^2(\alpha)}}\right)\left(\frac{1+\frac{\operatorname{sen}(\alpha)}{\cos(\alpha)}}{1+\frac{\cos(\alpha)}{\operatorname{sen}(\alpha)}}\right)-2\operatorname{sen}^2(\alpha).$$

Substituindo nesta expressão as fórmulas para o seno e o cosseno em termos da tangente da metade do ângulo,

$$\left(\frac{\frac{1}{\frac{2t}{1+t^2}}-\frac{(\frac{2t}{1+t^2})^2}{\left(\frac{2t}{1+t^2}\right)^2}}{\frac{\left(\frac{2t}{1+t^2}\right)^2}{\left(\frac{2t}{1+t^2}\right)^2}}\right)\left(\frac{1+\frac{\frac{2t}{1+t^2}}{\frac{2t}{1+t^2}}}{1+\frac{\frac{2t}{1+t^2}}{\frac{2t}{1+t^2}}}\right)-2\left(\frac{2t}{1+t^2}\right)^2.$$

Fazendo os cálculos no computador, obtemos

$$\frac{-t^8-14t^6+8t^4-10t^2+1}{t^8-2t^4+1}.$$

Contudo, usando a fórmula para o cosseno do dobro de um ângulo, o lado direito de (96) é igual a
$$1 - 2\cos^2(\alpha).$$
Substituindo a expressão do cosseno em (93), podemos reescrever o lado direito de (96) na forma
$$\frac{-t^4 + 6t^2 - 1}{t^4 + 2t^2 + 1}$$
o que mostra que a identidade é falsa, pois seus lados esquerdo e direito, expressos como funções racionais de t, são claramente diferentes.

Exercícios

1. Determine as fórmula para a tangente da soma de dois ângulos e do dobro de um ângulo a partir das fórmulas da seção 2.

2. Determine as fórmulas para o seno e o cosseno de 3α a partir do seno e cosseno de α.

3. Use as fórmulas do exercício 2 para calcular o seno e o cosseno de $\pi/3$.

4. Prove, por indução em n, que $|\operatorname{sen}(nx)| \leq n \operatorname{sen}(x)$, para todo $x \in (0, \pi)$.

5. Use o fato de que a soma dos ângulos de um triângulo é igual a π para relacionar o seno e o cosseno de $\pi/6$ com o seno e o cosseno de $\pi/3$.

6. Sabe-se que, para um dado ângulo α, $\operatorname{sen}(\alpha) = -\sqrt{11}/6$ e que $\tan(\alpha) < 0$. Calcule $\cos(\alpha)$ e $\tan(\alpha)$.

7. Calcule a área de um triângulo que tem dois lados de comprimentos 2 e 3 que delimitam um ângulo $\pi/4$.

8. Caminhando ao longo da margem de um rio, você vê um poste de iluminação na outra margem, diretamente oposto ao local onde você está. Depois de caminhar 30 metros em linha reta, você toma o transferidor, que sempre carrega na mochila, e verifica que a reta que liga você ao poste forma um ângulo de $\pi/7$ radianos com o caminho que percorreu desde que viu o poste. Qual a largura do rio no lugar onde se encontra o poste?

9. Se um triângulo retângulo tem um ângulo de $3\pi/8$ e o cateto adjacente a este ângulo mede 5, quanto mede sua hipotenusa? Você pode supor que $\operatorname{sen}(3\pi/8) \approx 0,92$.

10. Prove as seguintes identidades trigonométricas:
    (a) $(1 + \cot^2(\alpha))\operatorname{sen}^2(\alpha) = 1$;
    (b) $(\sec^2(\alpha) - 1)(\csc^2(\alpha) - 1) = 1$;
    (c) $(1 + \tan^2(\alpha))/(\csc^2(\alpha)) = \tan^2(\alpha)$;

11. Quais das seguintes identidades trigonométricas são verdadeiras?
    (a) $\frac{1}{1+\cos(\alpha)} - \frac{1}{1-\cos(-\alpha)} = -2\cot(\alpha)\csc(\alpha)$;
    (b) $\csc^2(\alpha)(1 + \operatorname{sen}^2(\alpha)) = \cot^2(\alpha)$;
    (c) $\left(\frac{\tan(\alpha)}{\sec(\alpha)}\right)\operatorname{sen}(-\alpha) = \cos^2(\alpha)$;
    (d) $\frac{\sec(-\alpha)}{\tan(\alpha)+\cot(\alpha)} = -\operatorname{sen}(-\alpha)$;
    (e) $\frac{1+\operatorname{sen}(\alpha)}{\cos(\alpha)} = \frac{\cos(\alpha)}{1+\operatorname{sen}(-\alpha)}$.

CAPÍTULO 7

Gráficos

Neste capítulo usaremos o que aprendemos sobre funções e geometria para introduzir coordenadas no plano e analisar, de maneira bastante criteriosa, o gráfico de algumas funções que estudamos nos capítulo anteriores.

1. Coordenadas no plano

Na seção 2 do capítulo 5 introduzimos coordenadas em uma reta; nesta seção daremos um passo adiante introduzindo 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 encontrar 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. Não é nada surpreendente que o nome de Descartes apareça neste contexto, pois, como vimos na seção 2 do capítulo 5, ele foi o primeiro a explicitar claramente a relação entre pontos de uma reta e números reais.

Embora o sistema de dois eixos perpendiculares que usamos hoje em dia seja chamado de cartesiano em homenagem a 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 curvas definidas por equações de grau três que identifica em seu artigo.

Usando a terminologia introduzida nos capítulos anteriores, desejamos associar um par de números reais a cada ponto do plano, construindo, para isso, uma função bijetiva do plano no produto cartesiano $\mathbb{R}^2 = \mathbb{R} \times \mathbb{R}$. Para isso, começamos escolhendo uma reta ℓ_1 e um ponto $O \in \ell_1$. Então usamos o Teorema 6.1 para construir uma reta ℓ_2 perpendicular a ℓ_1 pelo ponto O.

Apelando, agora, para o Segundo Axioma de Distância, sabemos que existem funções bijetivas

$$x : \ell_1 \longrightarrow \mathbb{R} \text{ e } y : \ell_2 \longrightarrow \mathbb{R}$$

que satisfazem

$$x(O) = y(O) = 0. \tag{97}$$

Dado um ponto $P \in \mathbb{P}$, usamos o Corolário 4.1 para construir retas r_1 e r_2 que contêm P e são paralelas a ℓ_1 e ℓ_2, respectivamente.

Pela transitividade do paralelismo, r_1 não pode ser paralela a $\ell 1$, nem r_2 a ℓ_2, porque qualquer uma destas duas afirmações implicaria que ℓ_1 e ℓ_2 são paralelas. Se

$$\ell_1 \cap r_1 = \{P_1\} \text{ e } \ell_2 \cap r_2 = \{P_2\}$$

contruímos uma função do plano no produto cartesiano

$$\mathbb{R}^2 = \mathbb{R} \times \mathbb{R},$$

que ao ponto P associa o par ordenado $(x(P), y(P))$.

Utilizando a terminologia usual, diremos que as retas ℓ_1 e ℓ_2 são *eixos coordenados*. Os números reais $x(P)$ e $y(P)$ são as *coordenadas* do ponto P. A coordenada $x(P)$ é a *abscissa* de P, e $y(P)$ é sua *ordenada*. Por (97), os pontos de ℓ_1 têm todos ordenadas iguais a zero, ao passo que os pontos de ℓ_2 têm abscissas iguais a zero. Em particular, o ponto O no qual ℓ_1 intersecta ℓ_2 corresponde ao par ordenado $(0, 0)$.

Reciprocamente, dado um par de números reais (q_1, q_2), sejam

$$Q_1 \in \ell_1 \text{ e } Q_2 \in \ell_2$$

os pontos para os quais

$$x(Q_1) = q_1 \text{ e } y(Q_2) = q_2.$$

O teorema 4.4 nos permite construir retas s_1, paralela a ℓ_1 por Q_2 e s_2, paralela a ℓ_2 por Q_1. Pela transitividade do paralelismo, s_1 não pode ser paralela a s_2, pois isto implicaria que ℓ_1 e ℓ_2 são paralelas. Logo, r_1 intersecta r_2 em um ponto que chamaremos de Q.

As duas construções dadas acima nos permitem ir e voltar em os pontos do plano e o produto cartesiano $\mathbb{R}^2$. Mas ainda nos falta mostrar que uma é inversa da outra. Neste ponto a unicidade da paralela por um ponto, provada no teorema 4.4, vem em nosso socorro. Pois, se $q_1 = x(P)$ e $q_2 = y(P)$, as retas s_1 e s_2, têm que ser iguais às retas r_1 e r_2, cuja interseção é P.

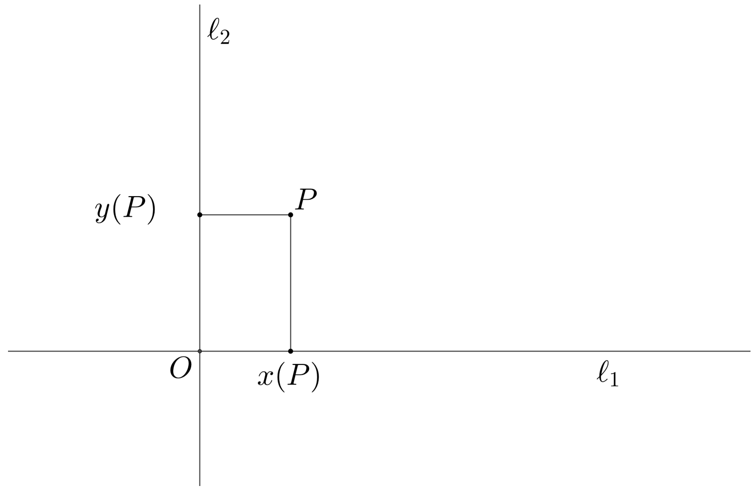


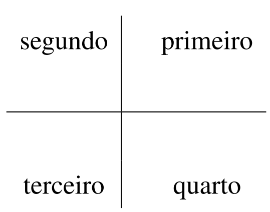
FIGURA 1. O ponto P e suas coordenadas.

Portanto, a função que a cada ponto do plano associa seu par de coordenadas em $\mathbb{R}^2$ é bijetiva. Para simplificar a notação, escreveremos, de agora em diante,

$$P = (x(P), y(P))$$

em vez de dizer que $x(P)$ é a abscissa e $y(P)$ a ordenada do ponto P; que é mais um caso típico de abuso de notação.

Sabemos pelo Primeiro Axioma de Separação, que cada reta subdivide o plano em dois semiplanos. Como os dois eixos coordenados se cortam no ponto $O = (0, 0)$, eles dividem o plano em quatro partes, conhecidas como *quadrantes*. Seguindo a convenção usual, o eixo das abscissas será sempre desenhado ao longo da horizontal, com as coordenadas dos pontos crescendo da esquerda para a direita. Com isso o eixo das ordenadas será sempre vertical; além disso, suporemos que suas coordenadas crescem de baixo cima. Sob estas convenções, a posição de cada quadrante e os sinais das coordenadas dos seus pontos são como indicados na tabela 1.



Quadrante	Sinal da abscissa	Sinal da ordenada
Primeiro	positivo	positivo
Segundo	negativo	positivo
Terceiro	negativo	negativo
Quarto	positivo	negativo

TABELA 1. Os quadrantes e os sinais das coordenadas dos seus pontos

Tendo introduzido coordenadas no plano, podemos definir o *gráfico* de uma função f de $X \subset \mathbb{R}$ em $\mathbb{R}$, que é o conjunto

$$\Gamma_f = \{(x, f(x)) \in \mathbb{R}^2 \mid x \in X\}. \tag{98}$$

Note que, pela definição de função, há apenas um ponto em Γ_f para cada $x \in X$. É por isto que podemos representar o gráfico da função f escrevendo apenas $y = f(x)$; isto é, expressando de que maneira a ordenada de um ponto de Γ_f depende de 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 $\mathbf{k}$ que leva qualquer número real em 1 é

$$\Gamma_{\mathbf{k}} = \{(x, 1) \mid x \in \mathbb{R}\},$$

cujos pontos formam uma reta paralela ao eixo, como ilustrado na figura 2.

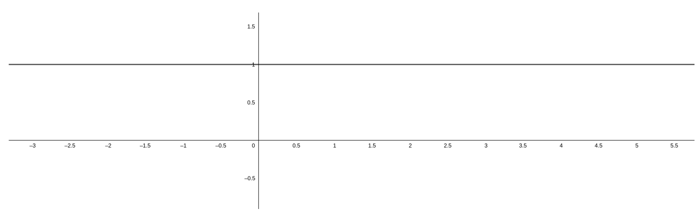


FIGURA 2. Representação geométrica do gráfico da função constante.

Nas próximas duas seções usaremos nosso conhecimento de geometria para descrever completamente os conjuntos de pontos correspondentes aos gráficos das funções lineares e quadráticas. Antes, porém, é necessário chamar sua atenção de que nem sempre é possível fazer o esboço do gráfico de uma função. Por exemplo, o gráfico da função de Dirichlet, definida em $\mathbb{R}$ pela regra

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

é o conjunto

$$\Gamma_{\mathcal{D}} = \{(x, 1) \mid x \in \mathbb{Q}\} \cup \{(x, 0) \mid x \notin \mathbb{Q}\}.$$

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 diminuir a escala do gráfico porque, se fizéssemos isto, continuaríamos tendo que enfrentar o mesmo problema. Em outras palavras, se usássemos o computador para desenhar seu gráfico, veríamos duas retas paralelas horizontais: uma com ordenada 0, 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.

2. Gráficos de funções lineares

Sejam $a \neq 0$ e b números reais e seja $\lambda : \mathbb{R} \longrightarrow \mathbb{R}$ a função linear definida pela regra

$$\lambda(x) = ax + b. \tag{99}$$

Os números a e b são chamados, respectivamente, de *coeficiente angular* e *coeficiente linear* de λ.

Como aprendemos no ensino fundamental, o gráfico de uma função linear é uma reta. Para provar isto, basta mostrar que se

$$P_0 = (x_0, y_0), \quad P_2 = (x_2, y_2) \tag{100}$$

pertencem ao gráfico de λ e $x_0 < x_1 < x_2$ então

$$(x_1, y_1) \in P_0P_1 \iff y_1 = \lambda(x_1).$$

Começamos traçando uma paralela r ao eixo horizontal por P_0 e paralelas ao eixo vertical pelos pontos $(x_1, 0)$ e $(x_2, 0)$. Denotaremos por Q_1 e Q_2 as interseções de r com as as duas retas verticais, como ilustrado na figura 3.

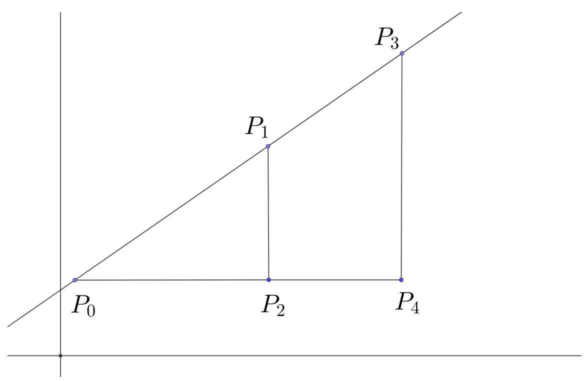


FIGURA 3. Achando a equação da reta.

Os triângulos retângulos $\triangle P_0P_1Q_1$ e $\triangle P_0P_2Q_2$ são semelhantes porque, além do ângulo reto, têm o ângulo $\sphericalangle P_1P_0Q_1$ em comum. Portanto,

$$\frac{P_1Q_1}{Q_1P_0} = \frac{P_2Q_2}{Q_2P_0}. \tag{101}$$

Usando as coordenadas de cada um destes pontos dadas acima, temos que

$$Q_1 = (x_1, y_0) \text{ e } Q_2 = (x_2, y_0)$$

de modo que

$$P_1Q_1 = y_1 - y_0, \quad P_2Q_2 = y_2 - y_0, \quad Q_1P_0 = x_1 - x_0 \text{ e } Q_2P_0 = x_2 - x_0;$$

que nos permite reescrever (101) como

$$\frac{y_1 - y_0}{x_1 - x_0} = \frac{y_2 - y_0}{x_2 - x_0}. \tag{102}$$

Mas, para que P_0 e P_2 sejam pontos do gráfico de λ é necessário que

$$y_0 = \lambda(x_0) = ax_0 + b \text{ e } y_2 = \lambda(x_2) = ax_2 + b;$$

donde

$$a = \frac{y_2 - y_0}{x_2 - x_0}. \tag{103}$$

Substituindo em (102), obtemos

$$y_1 = a(x_1 - x_0) + y_0 = ax_1 + b,$$

pois

$$y_0 - ax_0 = \lambda(x_0) - ax_0 = b. \tag{104}$$

Compo isto mostra que $y_1 = \lambda(x_1)$, mostramos que um ponto de abscissa x_1 que pertence à reta $\overleftrightarrow{P_0P_2}$ se, e somente se, está no gráfico de λ.

Reciprocamente, se as coordenadas dos pontos P_0 e P_2 são conhecidas, podemos usar (103) e (104) para achar o coeficiente angular e o coeficiente linear de λ. Em particular, dois pontos do plano que tenham abscissas diferentes pertencem ao gráfico de uma função linear. A restrição sobre as abscissas é necessária porque, caso contrário, os dois pontos pertencem ao gráfico de nenhuma função.

Por exemplo, quando $P_0 = (1,2)$ e $P_2 = (2,-1)$, as fórmulas em (103) e (**??**) nos dão

$$a = \frac{2-(-1)}{1-2} = -3 \text{ e } b = 2 - (-3 \cdot 1) = 5,$$

de modo que a função linear cujo gráfico corresponde à reta $\overleftrightarrow{P_0P_2}$ é $\lambda(x) = -3x + 5$. O gráfico correspondente é o da figura 4.

Tendo identificado o gráfico da função linear $\lambda(x) = ax + b$ como sendo uma reta, parece natural nos perguntarmos qual a função linear $\mu(x) = a'x + b'$ cujo gráfico é a reta perpendicular a Γ_λ pelo ponto $P_0 \in \Gamma_\lambda$. Mais uma vez, teremos que apelar para o que aprendemos sobre a geometria dos triângulos. Na figura 5, os pontos

$$P_0 = (x_0, y_0) \text{ e } P_2 = (x_2, y_2)$$

pertencem à reta Γ_λ; ao passo que P_1 está na interseção de Γ_λ com a reta horizontal que contém P_2. Como P_3 pertence a esta última reta e também à paralela ao eixo vertical pelo ponto P_0. Assim, $P_3 = (x_0, y_2)$.

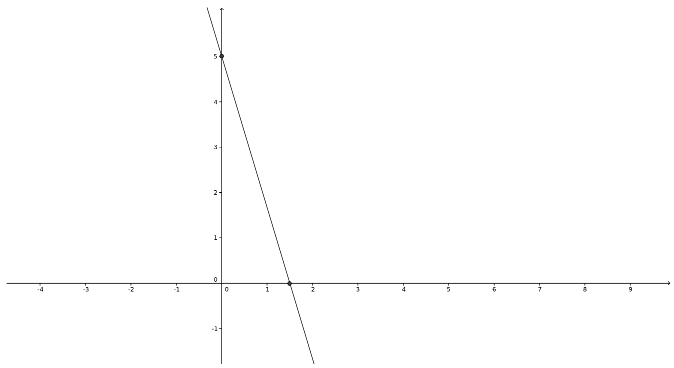


FIGURA 4. Esboço do gráfico de $\lambda(x) = -3x + 5$.

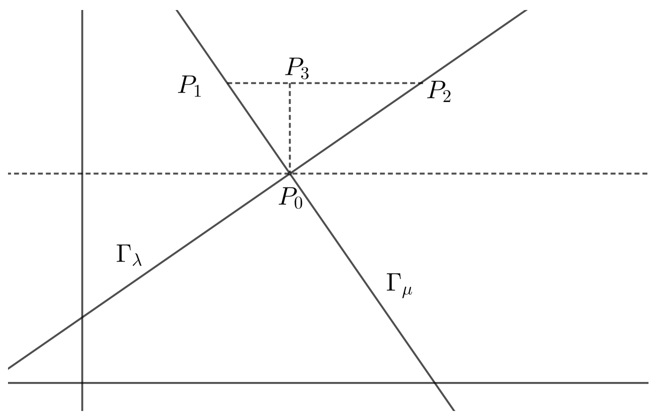


FIGURA 5. Calculando a reta perpendicular a Γ_λ.

Como o triângulo retângulo $P_0P_1P_2$ tem um ângulo agudo em comum com os triângulos $\triangle P_0P_1P_3$ e $\triangle P_0P_2P_3$, ambos também retângulos, o Axioma de Semelhança nos permite concluir que que estes três triângulos são semelhantes. Mas, de

$$\triangle P_0P_1P_3 \sim \triangle P_0P_2P_3$$

temos que

$$\frac{P_3P_0}{P_3P_1} = \frac{P_2P_3}{P_3P_0}. \tag{105}$$

Usando as coordenadas dos pontos, temos que

$$P_3P_0 = y_2 - y_0, \quad P_3P_1 = x_0 - x_1, \quad \text{e} \quad P_2P_3 = x_2 - x_0.$$

Substituindo estes valores em (105),

$$-a' = \frac{y_2 - y_0}{x_0 - x_1} = \frac{x_2 - x_0}{y_2 - y_0} = \frac{1}{a}. \tag{106}$$

em que a primeira e a última igualdades são consequências de (103).

$$\mu(x) = -\frac{1}{a}x + b'.$$

Como esta reta também passa por P_0, sabemos que

$$y_0 = \mu(x_0) = -\frac{1}{a}x_0 + b',$$

de modo que

$$b' = y_0 - \frac{1}{a}x_0,$$

que nos dá uma descrição completa da reta perpendicular a Γ_λ.

Voltando ao exemplo que apresentamos acima, a reta ℓ perpendicular ao gráfico de $\lambda(x) = -3x + 5$ tem coeficiente angular $1/3$. Portanto, se (x, y) pertence a ℓ, então

$$y = \frac{1}{3}x + b'$$

para algum número real b'. Em particular, qualquer reta perpendicular a Γ_λ tem coeficiente angular igual a $1/3$. Para fixar o valor de b' precisamos escolher um ponto pelo qual a reta passa. Apesar do argumento acima pressupor que o ponto pertence à reta, podemos usar qualquer ponto do plano para achar b'. Ao fazer isto estamos escolhendo uma reta, perpendicular a Γ_λ, que passa por aquele ponto. Por exemplo, supondo que a perpendicular passe por $(6, -7)$, devemos ter que

$$-7 = \frac{1}{3} \cdot 6 + b',$$

que nos dá $b' = -9$. Portanto, a reta perpendicular a $y = -3x + 5$ pelo ponto $(6, -7)$ tem equação $y = (1/3)x - 9$.

3. Gráficos de funções quadráticas

Nesta seção analisaremos em detalhe o gráfico Γ_f da função quadrática $q(x) = ax^2 + bx + c$ no caso em que $a > 0$. O caso em que $a < 0$ pode ser tratado de maneira análoga e os detalhes ficarão por sua conta. Em seções anteriores já identificamos algumas propriedades desta função. Por exemplo, na seção 3 do capítulo 3, mostramos que $q(x) = 0$ quando

$$x = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a} \tag{107}$$

e na seção 3 do capítulo 4, mostramos que

$$\operatorname{im}(q) = [-(b^2 - 4ac)/4a, +\infty), \tag{108}$$

quando $a > 0$. Do ponto de vista do gráfico

$$\Gamma_q = \{(x, q(x)) \mid x \in \mathbb{R}\}$$

os pontos cujas abscissas são dadas por (107) são aqueles em que o gráfico intersecta o eixo OX; ao passo que (108) nos permite afirmar que o gráfico está todo acima da reta horizontal $y = -(b^2 - 4ac)/4a$.

Outra propriedade da função quadrática que provamos na seção 3 do capítulo 4, é que esta função nunca é injetiva quando seu domínio é tomado como sendo todo o conjunto dos números reais. A primeira coisa que faremos é refinar este resultado, provando que Γ_q tem um eixo vertical. Em outras palavras, dado $y_0 \in \operatorname{im}(q)$, existe uma reta vertical $x = x_0$ e um número real d, tal que

$$y_0 = q(x_0 - d) = q(x_0 + d). \tag{109}$$

Como $y_0 = q(x)$ é equivalente a

$$ax^2 + bx + (c - y_0) = 0;$$

temos, pela fórmula usual, que

$$x = \frac{-b \pm \sqrt{b^2 - 4a(c - y_0)}}{2a}.$$

Portanto,

$$\frac{-b}{2a} - \frac{\sqrt{b^2 - 4a(c - y_0)}}{2a} \quad \text{e} \quad \frac{-b}{2a} + \frac{\sqrt{b^2 - 4a(c - y_0)}}{2a}$$

são as abscissas dos pontos cuja ordenada é y_0. Logo, os pontos de Γ_q são simétricos relativamente à reta $x = -b/2a$. Além disso, como

$$q\left(\frac{-b}{2a}\right) = -\frac{b^2 - 4ac}{4a};$$

o que já aprendemos sobre a imagem de q nos permite afirmar que

$$\left(\frac{-b}{2a}, -\frac{b^2 - 4ac}{4a}\right)$$

é o ponto mais baixo de Γ_q, que chamamos de *vértice* da parábola.

Aplicando o que fizemos até aqui à função $q_0(x) = x^2 - x - 2$, vemos que seu gráfico é simétrico ao eixo $x = 1/2$, com vértice no ponto

$$\left(\frac{1}{2}, -\frac{9}{4}\right).$$

Além disso, como as raízes de $x^2 - x - 2 = 0$ são -1 e 2, o gráfico cortará o eixo das abscissas nos pontos $(-1, 0)$ e $(2, 0)$.

O que ainda não é, de forma alguma, claro é como Γ_q se comporta de cada lado do eixo. Em princípio ele poderia zigue-zaguear para cima e para baixo, desde que os mesmos zigue-zagues aparecessem refletidos de um lado e do do outro do eixo. Na verdade isto não acontece, porque, quando $a > 0$, a função $q(x) = ax^2 + bx + c$ é *decrescente* à esquerda do eixo e *crescente* à sua direita. Podemos enunciar isto de maneira mais precisa afirmando que

$$\frac{-b}{2a} < x_1 < x_2 \Longrightarrow q(x_1) < q(x_2).$$

ao passo que,

$$x_1 < x_2 < \frac{-b}{2a} \Longrightarrow q(x_1) > q(x_2);$$

A demonstração destes fatos depende da diferença

$$q(x_2) - q(x_1) = (ax_2^2 + bx_2 + c) - (ax_1^2 + bx_1 + c) = a(x_2^2 - x_1^2) + b(x_2 - x_1). \tag{110}$$

Lembrando que

$$x_2^2 - x_1^2 = (x_2 - x_1)(x_2 + x_1)$$

podemos reescrever (110) na forma

$$q(x_2) - q(x_1) = (a(x_2 + x_1) + b)(x_2 - x_1). \tag{111}$$

Quando

$$\frac{-b}{2a} < x_1 < x_2$$

temos que

$$x_1 + x_2 > 2 \cdot \left(\frac{-b}{2a}\right) = \frac{-b}{a},$$

de forma que

$$a(x_2 + x_1) + b > a \cdot \frac{-b}{a} + b = 0 \quad \text{e} \quad x_2 - x_1 > 0.$$

Logo, neste caso, os dois termos do produto no lado direito de (111) são positivos. Assim, por fechamento,

$$q(x_2) - q(x_1) > 0 \Longleftrightarrow q(x_2) > q(x_1).$$

Por outro lado, quando

$$x_1 < x_2 < \frac{-b}{2a}$$

obtemos

$$a(x_2 + x_1) + b < 0 \quad \text{e} \quad x_2 - x_1 > 0;$$

de modo que, desta vez, um dos termos do lado lado direito de (111) é positivo e o outro é negativo. Assim, neste caso,

$$q(x_2) - q(x_1) < 0 \Longleftrightarrow q(x_2) < q(x_1),$$

o que conclui nossa verificação de que a função decresce à esquerda de $x = -b/2a$ e cresce à sua direita.

Embora o que já vimos nos dê uma boa ideia de como deve ser o gráfico de uma parábola, ainda não é suficiente para garantir que Γ_q não seja formado, por exemplo, por uma sucessão de segmentos interligados entre si, como na figura 6. Para mostrar que esta figura não pode ser a representação correta de uma parábola precisamos investigar para suas tangentes.

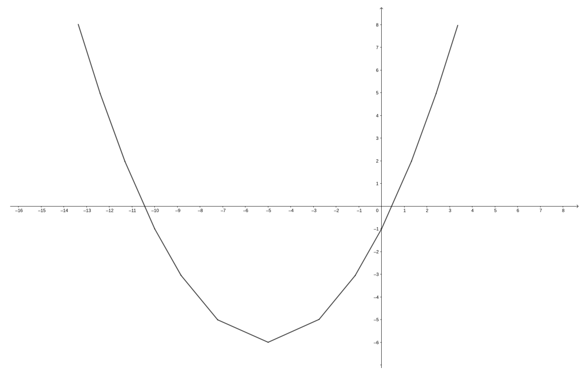


FIGURA 6. Será este o gráfico de uma função quadrática?

Quando estudamos círculos no ensino médio, a tangente é definida como uma reta que contém apenas um ponto do círculo. Esta definição não é adequada para a parábola, porque qualquer reta vertical satisfaz esta propriedade, mas não corresponde à nossa ideia intuitiva de tangente. Afinal, em latim, o verbo *tangere* significa tocar, mas as retas verticais cortam o gráfico, em vez de apenas tocá-lo. Por isso, diremos que um ponto da parábola tem uma tangente bem definida se há apenas uma reta que não é vertical e contém apenas este ponto. Como ilustrado na figura 7 isto não aplica quando a curva tem quinas. Portanto, para mostrar que o gráfico de uma função quadrática não pode ter quinas como o da figura 6 basta mostrar que tem tangente em todos os seus pontos.

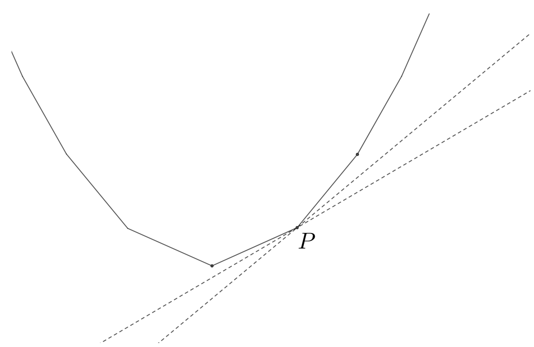


FIGURA 7. Retas que tocam uma curva com quinas

Para calcular a tangente ao gráfico Γ_q de $q(x) = ax^2 + bx + c$ no ponto $P_0 = (x_0, q(x_0))$, construímos uma reta ℓ_α de equação

$$y = \alpha(x - x_0) + q(x_0),$$

que passa pelo ponto P_0 e escolhemos sua inclinação α de modo que o único ponto em comum com Γ_q seja P_0. Note que não precisamos nos preocupar com as retas verticais, porque não podem ser representadas por uma equação como a de ℓ_α. Como a equação que descreve o gráfico da parábola é $y = q(x)$, a interseção $\Gamma_f \cap \ell_\alpha$ é obtida igualando os valores de y nas duas equações e determinando as raízes do polinômio

$$\alpha(x - x_0) + q(x_0) = q(x)$$

assim obtido. Substituindo a fórmula para $q(x)$ na igualdade acima e cancelando o c dos dois lados da expressão, encontramos

$$\alpha(x - x_0) + ax_0^2 + bx_0 = ax^2 + bx;$$

que podemos reescrever na forma

$$ax^2 + (b - \alpha)x + \alpha x_0 - (ax_0^2 + bx_0) = 0.$$

As raízes desta equação determinam as abscissas dos pontos de interseção desejados. Nossa análise anterior mostra que haverá apenas um ponto de interseção quando o discriminante

$$(b - \alpha)^2 - 4a(\alpha x_0 - (ax_0^2 + bx_0)) = (b - \alpha)^2 + 4a(b - \alpha)x_0 + 4a^2x_0^2$$

for igual a zero. Contudo,

$$(b - \alpha)^2 + 4a(b - \alpha)x_0 + 4a^2x_0^2 = (b - \alpha + 2ax_0)^2$$

de modo que o discriminante só será nulo quando

$$\alpha = 2ax_0 + b.$$

Logo, cada ponto de Γ_f tem apenas uma tangente, garantindo que o gráfico não pode ter quinas. O seguinte teorema resume tudo o que aprendemos sobre a função quadrática.

TEOREMA 3.1. *Quando $a > 0$, o gráfico da função quadrática $q(x) = ax^2 + bx + c$:*

1. *é simétrico relativamente à reta $x = -b/2a$;*
2. *corta o eixo das abscissas nos pontos dados pela fórmula* (40) *quando $\Delta \geq 0$;*
3. *está todo acima da reta $y = -\Delta/4a$;*
4. *decresce à esquerda de $x = -b/2a$ e cresce à sua direita;*
5. *tem uma única tangente no ponto $(x_0, q(x_0))$ cuja inclinação é igual a $2ax_0 + b$.*

Aplicando tudo o que fizemos à função quadrática $q_0(x) = x^2 - x - 2$, obtemos o gráfico ilustrado na figura 8.

Por outro lado, a função quadrática

$$q_1(x) = -4x^2 + 4x - 2.$$

tem vértice no 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

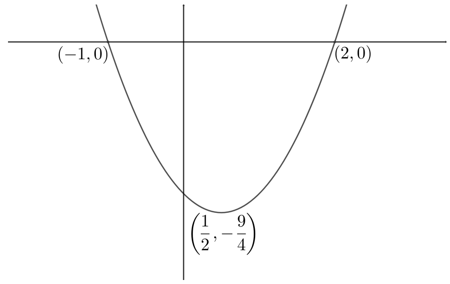


FIGURA 8. Gráfico de $q_2(x) = x^2 - x - 2$.

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 9.

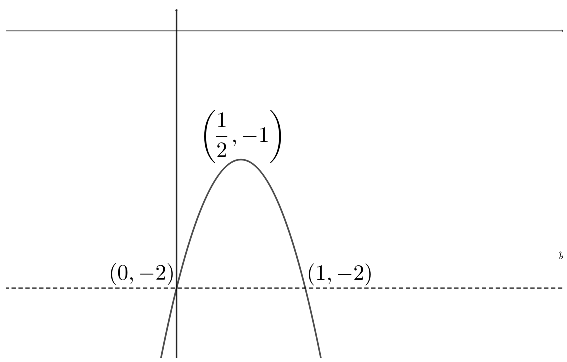


FIGURA 9. Gráfico da função $q_3(x) = -4x^2 + 4x - 2$.

Finalmente, não custa alertá-lo de que nem todo gráfico que parece uma parábola é, de fato o gráfico de uma função quadrática. Um exemplo é o gráfico da função $g(x) = x^4$, ilustrado na figura 10. Além de mais achatado que uma parábola, no intervalo $[0,1]$, o gráfico desta função sobe de maneiras mais íngreme que uma parábola fora deste intervalo.

4. Funções trigonométricas

No capítulo 6 definimos as funções trigonométrica e estudamos suas propriedades usando o que aprendemos sobre geometria plana. Por causa disto, o domínio das funções que definimos estavam limitados ao intervalo $[0,\pi/2]$. Mesmo os valores do seno e do cosseno em 0 e em $\pi/2$ foram obtidos através de um argumento informal. Nesta seção, veremos como estender estas definições de modo a definir o seno e o cosseno para qualquer número real.

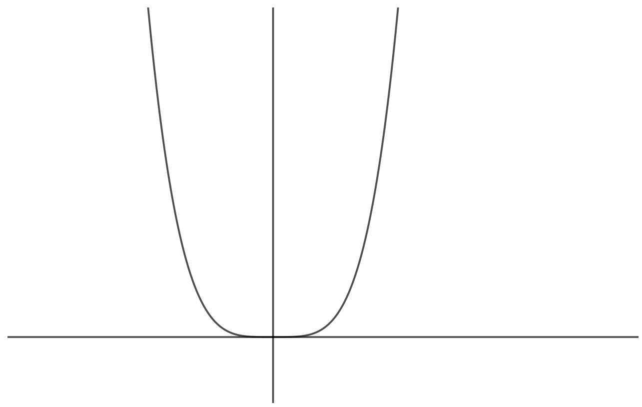


FIGURA 10. Gráfico de $g(x) = x^4$

Seguindo o que já havíamos feito no capítulo 6 posicionaremos o centro do círculo de raio 1 na origem, como ilustrado na figura 11. Com isto, o cosseno OA e o seno AB do ângulo α correspondem à abscissa e ordenada do vértice B do triângulo $\triangle OAB$. Podemos, então, redefinir, o seno e o cosseno de $\alpha = \sphericalangle COB$ como sendo a ordenada e a abscissa do ponto B. Em outras palavras, se $B = (b_1, b_2)$, então

$$\operatorname{sen}(\alpha) = b_2 \quad \text{e} \quad \cos(\alpha) = b_1.$$

Como $\alpha = 0$ se, e somente se, $B = (1, 0)$, segue-se desta nova definição que

$$\operatorname{sen}(0) = 0 \quad \text{e} \quad \cos(0) = 1.$$

Analogamente, $\alpha = \pi/2$ quando $B = (0, 1)$, de modo que

$$\operatorname{sen}(\pi/2) = 1 \quad \text{e} \quad \cos(\pi/2) = 0.$$

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 conforme à tabela 1 da página 157. Por exemplo, se $\pi/2 < \alpha < \pi$, como na figura 12, o seno BC continuará sendo positivo, ao passo que o cosseno 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

$$\operatorname{sen}(\pi) = 0 \quad \text{e} \quad \cos(\pi) = -1.$$

Por outro lado, quando $\alpha = 3\pi/2$, temos que $B = (0, -1)$, o que nos dá

$$\operatorname{sen}\left(\frac{3\pi}{2}\right) = -1 \quad \text{e} \quad \cos\left(\frac{3\pi}{2}\right) = 0.$$

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

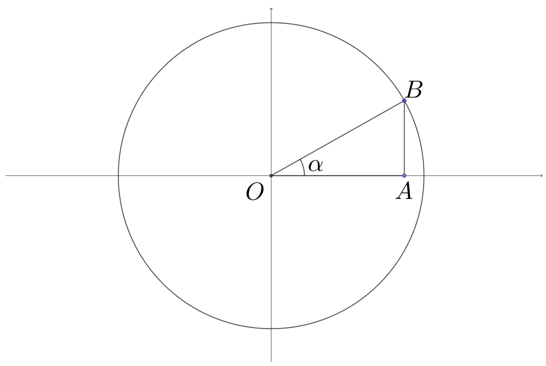


FIGURA 11. $0 < \alpha < \pi/2$

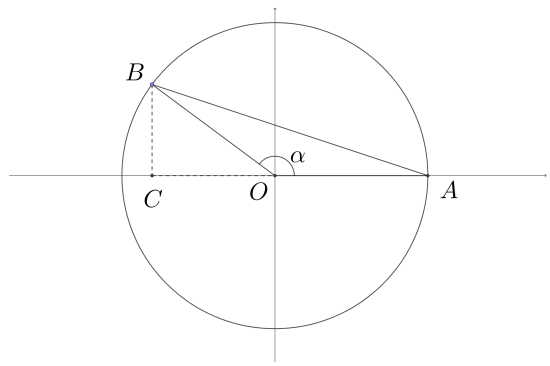


FIGURA 12. $\pi/2 < \alpha < \pi$

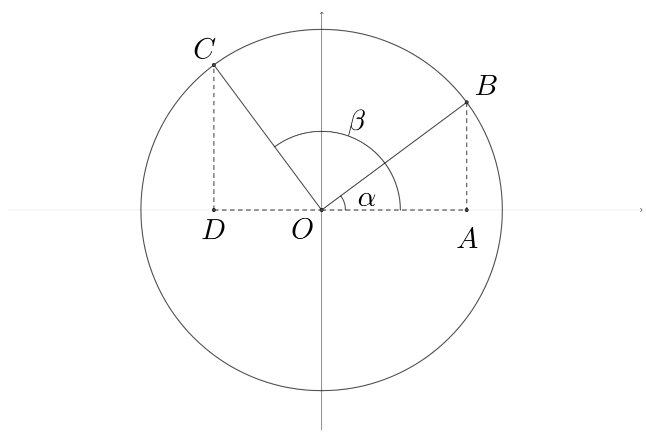


FIGURA 13. $\beta - \alpha = \pi/2$

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,

$$\operatorname{sen}(\alpha + 2\pi) = \operatorname{sen}(\alpha) \quad \text{e} \quad \cos(\alpha + 2\pi) = \cos(\alpha).$$

Outra relação muito útil é aquela entre os senos e cossenos de ângulos que diferem por $\pi/2$ radianos. Suponhamos que, como na figura 13, $\beta = \alpha + \pi/2$. Como

$$\beta + \sphericalangle DOC = \pi,$$

temos que

$$\alpha + \sphericalangle DOC = \frac{\pi}{2}. \tag{112}$$

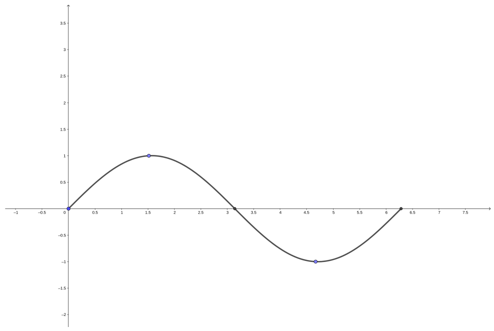


FIGURA 14. O seno de 0 a 2π.

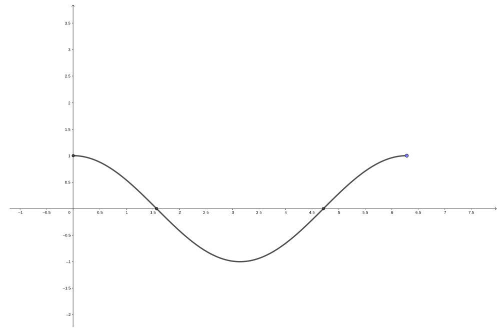


FIGURA 15. O cosseno de 0 a 2π.

Mas os ângulos em $\sphericalangle OAB$ e $\sphericalangle ODC$ são iguais a $\pi/2$ e a soma dos ângulos em um triângulo é igual a π, de modo que, por (112),

$$\sphericalangle OBA = \sphericalangle DOC \ \text{ e } \ \sphericalangle AOB = \sphericalangle OCD.$$

Combinando isto com

$$OC = OB,$$

pois ambos são raios do mesmo círculo, temos, pelo caso ALA, que os triângulos $\triangle OBA$ e $\triangle ODC$ são congruentes. Logo,

$$OD = AB \ \text{ e } \ DC = OA,$$

já que os vértices O, A e B de um dos triângulos correspondem, respectivamente, a C, D e O do outro. Portanto,

$$\operatorname{sen}(\alpha) = AB = -OD = -\cos(\beta) \ \text{ e } \ \cos(\alpha) = OA = CD = \operatorname{sen}(\beta).$$

Em outras palavras,

$$\operatorname{sen}\left(\alpha + \frac{\pi}{2}\right) = \cos(\alpha) \ \text{ e } \ \cos\left(\alpha + \frac{\pi}{2}\right) = -\operatorname{sen}(\alpha). \tag{113}$$

Usando os valores dos senos e cossenos dos ângulos que já calculamos neste e no capítulo 6, poríamos fazer um esboço grosseiro dos gráficos dos senos e cossenos. Porém, garantir que seus gráficos são como apresentados nas figuras 14 e 15, requer o uso de derivadas e vai ter que esperar que você aprenda um pouco de cálculo.

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 neste círculo, tanto no sentido anti-horário, quanto no sentido horário. Ao fazer isto, suporemos que, a cada volta, o ângulo aumenta de 2π. Por outro lado, como observamos anteriormente, 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. Portanto, admitindo que o ângulo tome qualquer valor real, os gráficos destas funções são obtidos copiando e colando os gráficos das figuras 14 e 15 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 16. 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.

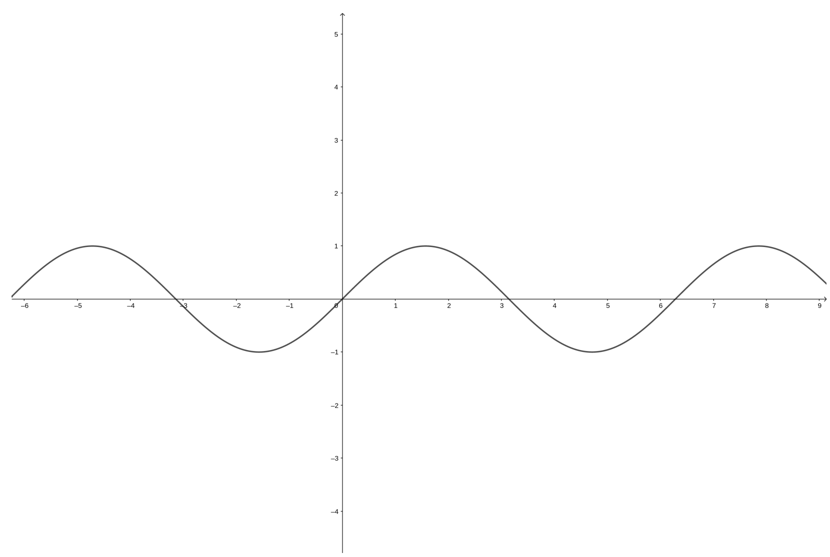


FIGURA 16. O seno como função periódica

A definição geométrica da tangente de α 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 $-\pi/2$. O gráfico da tangente, no intervalo $(-\pi/2, \pi/2)$, é ilustrado na figura 17.

5. Inversas de funções reais

Encerraremos o capítulo mostrando como usar métodos puramente geométricos para obter o gráfico da inversa de uma função inversível. Começaremos investigando como determinar se uma função real é injetiva ou sobrejetiva considerando apenas seu gráfico.

Seja $f : X \longrightarrow \mathbb{R}$ uma função cujo domínio é um subconjunto dos números reais. Como consequência das condições que definem uma função, qualquer reta vertical que

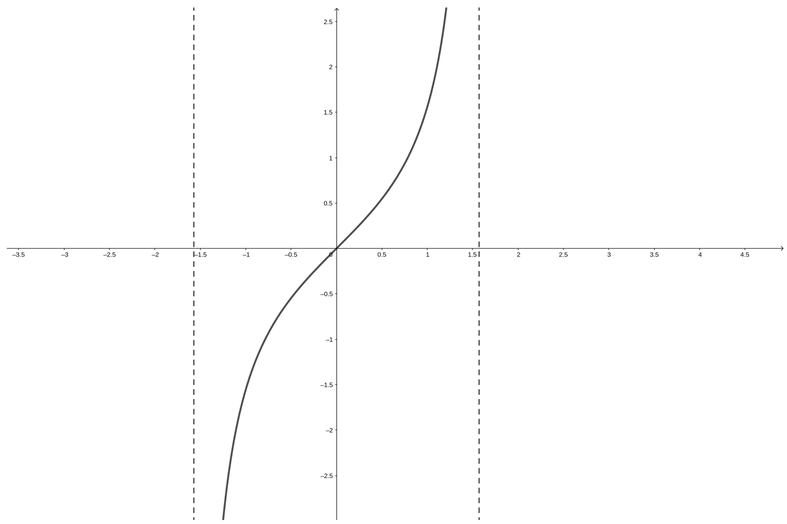


FIGURA 17. Gráfico da tangente no intervalo $(-\pi/2, \pi/2)$.

intersecte o eixo das abscissas em um ponto de X corta o gráfico de f em, exatamente, um ponto. Escolhendo, agora, um número real s e traçando uma reta horizontal ℓ que corte o eixo das ordenadas no ponto $(0, s)$ temos duas possibilidades. Se ℓ cortar o gráfico de f em (r, s), então

$$(r, s) \in \Gamma_f \iff f(r) = s$$

e vice-versa. Logo,

> a função f será sobrejetiva se, e somente se, *toda* reta horizontal corta seu gráfico em algum ponto.

Por outro lado, f não pode ser injetiva se $r \neq r'$ são elementos de X tais que

$$(r, s), (r', s) \in \ell \cap \Gamma_f \iff f(r) = s = f(r'),$$

e vice-versa. Assim,

> f é injetiva se, e somente se, *nenhuma* reta horizontal intersecta o gráfico de f em mais de um ponto.

Combinado estas duas equivalências, podemos concluir que

> f é bijetiva se, e somente se, *toda* reta horizontal intersecta o gráfico de f em *exatamente* um ponto.

Quando o gráfico de uma função inversível 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 se $X \subset \mathbb{R}$ e $f : X \to \mathbb{R}$ é uma função, então

$$(x_0, f(x_0)) \in \Gamma_f \iff (f(x_0), x_0) \in \Gamma_{f^{-1}}.$$

Supondo que $f(x_0) \neq x_0$, a reta que passa por $(x_0, f(x_0))$ e $(f(x_0), x_0)$ tem coeficiente angular igual a

$$\frac{f(x_0) - x_0}{x_0 - f(x_0)} = -1,$$

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 $y = x$.

O primeiro caso interessante ao qual este método pode ser aplicado é o da função quadrática. Contudo, como vimos na seção 3 do capítulo 4, para que uma função quadrática se torne inversível é necessário restringir seu domínio e contradomínio. Por exemplo, podemos fazer com que a função que leva x em x^2 se torne bijetiva, restringindo seu domínio a $[0, +\infty)$ e seu contradomínio a $[0, +\infty)$. Outra opção seria, sob o mesmo contradomínio, restringir o domínio a $(-\infty, 0]$. Denotando a função definida no primeiro caso por p_3; isto é considerando

$$p_3 : [0, +\infty) \to [0, +\infty)$$

definida pela regra $p_3(x) = x^2$, obtemos o gráfico à esquerda da figura 18. Espelhando este gráfico relativamente à reta $y = x$ obtemos o gráfico da função

$$p_3^{-1} : [0, +\infty) \to [0, +\infty)$$

definida por $p_3^{-1}(x) = \sqrt{x}$, ilustrado à direita da figura 18, sem a necessidade de calcular explicitamente nenhum de seus pontos.

Um exemplo cuja análise requer um pouco mais de cuidado é o da função quadrática $q(x) = x^2 - 4x + 3$. A parábola correspondente a esta função é simétrica relativamente à reta $y = 2$ e tem seu ponto mais baixo em $y = -1$. Portanto, as restrições com os maiores domínios e contradomínios possíveis são

$$q_- : (-\infty, 2] \to [1, +\infty) \ \text{ e } \ q_+ : [2, +\infty) \to [1, +\infty) \tag{114}$$

ambas definidas pela mesma regra que $q(x)$; isto é,

$$q_-(x) = x^2 - 4x + 3 \ \text{ e } \ q_+(x) = x^2 - 4x + 3.$$

O gráfico de q_- e de sua inversa aparecem desenhados na figura 19.

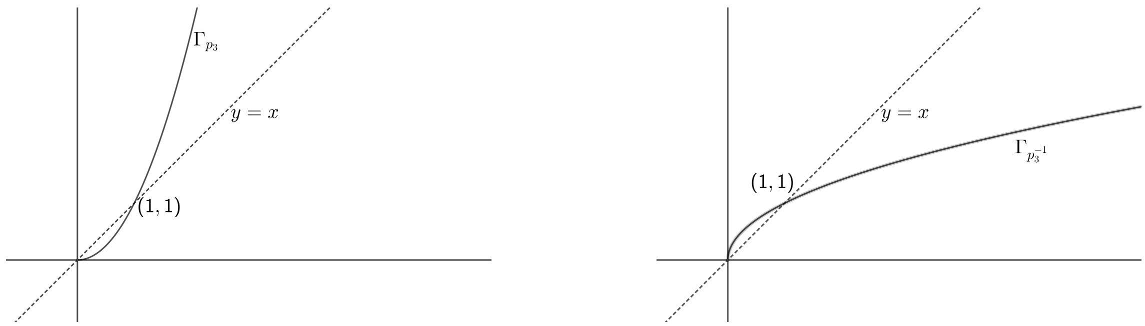


FIGURA 18. Gráficos de p_3 e p_3^{-1}.

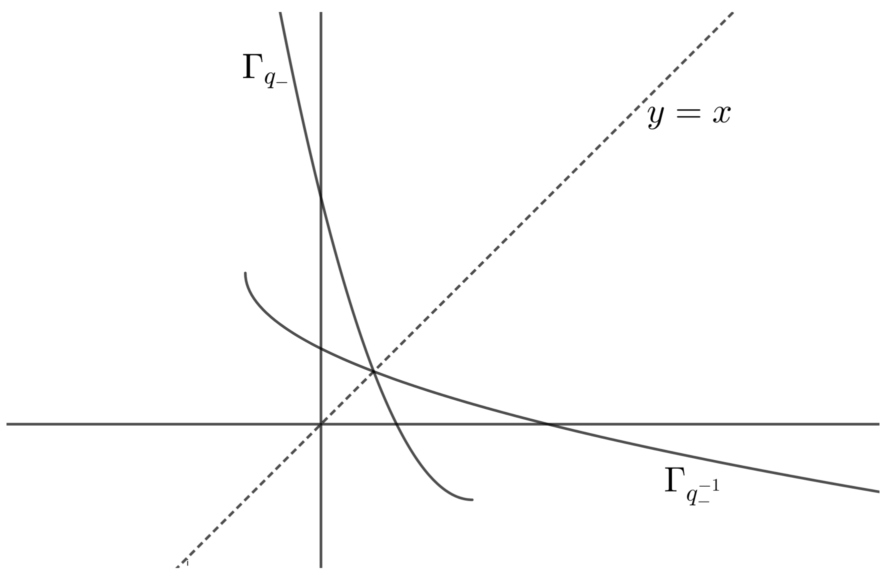


FIGURA 19. Gráficos das funções q_- e sua inversa.

Um exemplo mais divertido é dado pelo gráfico da função

$$f : \mathbb{R} \longrightarrow \mathbb{R}$$

definida pela regra $f(x) = x^3$ e de sua inversa, ambos ilustrados na figura 20.

Encerraremos esboçando o gráfico das funções inversas do seno, cosseno e tangente. Como o seno e o cosseno têm como imagem o 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, como fizemos para a função quadrática.

Para o domínio do seno, escolheremos o intervalo $[-\pi/2, \pi/2]$ porque, neste intervalo, além do seno ser uma função injetiva, seu gráfico é simétrico em relação aos eixos, o que

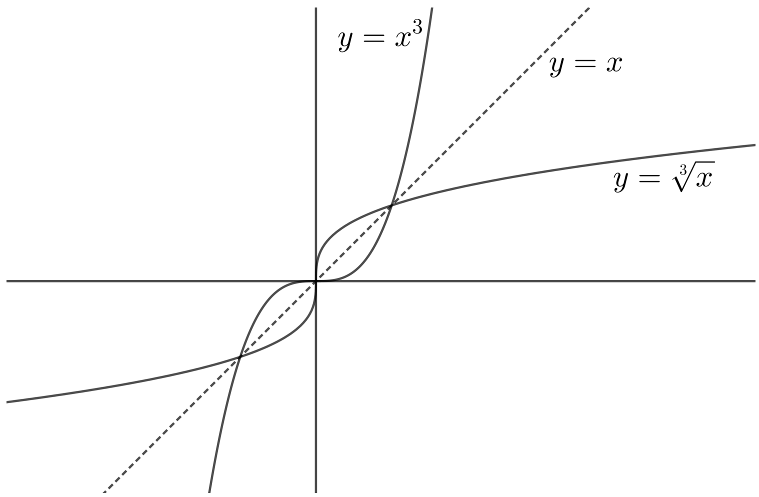


FIGURA 20. Gráficos das funções p_4 e p_4^{-1}.

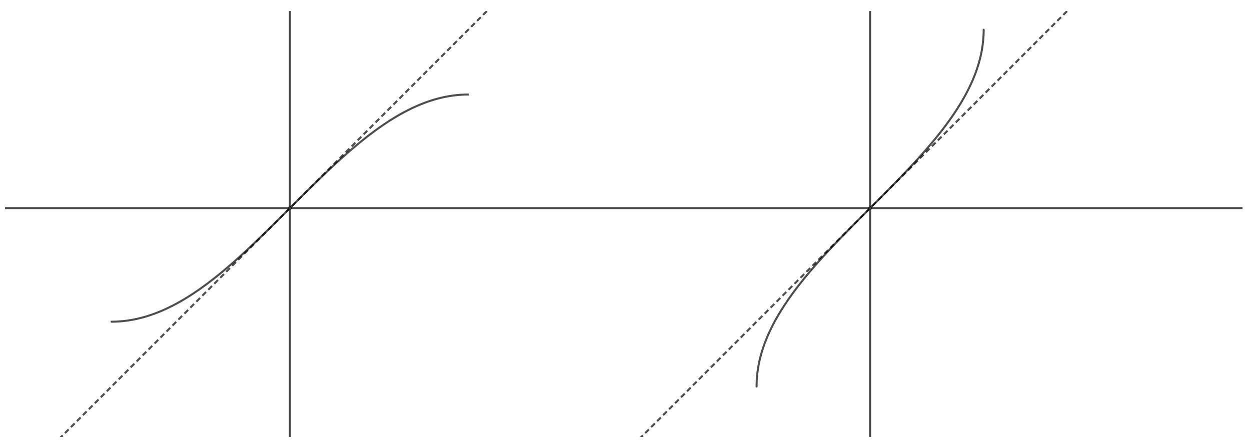


FIGURA 21. Gráficos do seno em $[-\pi/2, \pi/2]$ e do arco-seno.

torna mais fácil refleti-lo relativamente à reta $y = x$. Com isso, podemos construir

$$\operatorname{sen}^{-1} : [-1, 1] \longrightarrow [-\pi/2, \pi/2],$$

que é conhecida como *arco-seno*, e usualmente denotada por arcsen. Na figura 21 temos, à esquerda o gráfico do seno restrito ao intervalo $[-\pi/2, \pi/2]$ e à direita o gráfico do arco-seno, sobreposto aos gráficos 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 injetiva 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

$$\arccos = \cos^{-1} : [-1, 1] \longrightarrow [0, \pi].$$

Na figura 22 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.

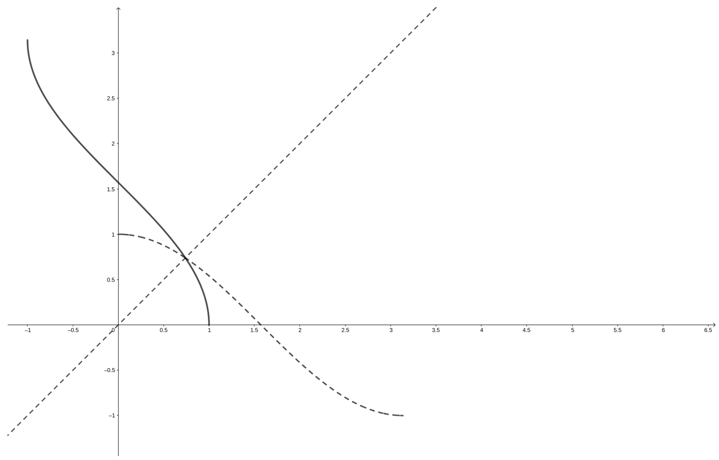


FIGURA 22. Gráficos do cosseno em $[0, \pi]$ e do arco-cosseno.

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 17 relativamente à reta $y = x$, como ilustrado na figura 23, na qual as retas horizontais têm equações $y = -\pi/2$ e $y = \pi/2$.

Exercícios

1. Determine o maior subconjunto de $\mathbb{R}$ para o qual cada uma das regras abaixo define uma função:

$$f_1(x) = x/(x^2 + 1), \quad \text{e} \quad f_2(x) = \sqrt{x}/(x^2 - 2).$$

   Você deve expressar cada um destes domínios como um intervalo ou união de intervalos.

2. Calcule as imagens das funções f_1 e f_2 definidas abaixo e determine quais destas funções são sobrejetivas:

$$\begin{aligned} f_1 : (-\infty, 0] &\longrightarrow \mathbb{R} \\ x &\longmapsto 1/(x^2 + 1) \end{aligned} \qquad\qquad \begin{aligned} f_2 : \mathbb{R} \setminus \{0\} &\longrightarrow \mathbb{R} \\ x &\longmapsto (x^2 - 4)/x \end{aligned}$$

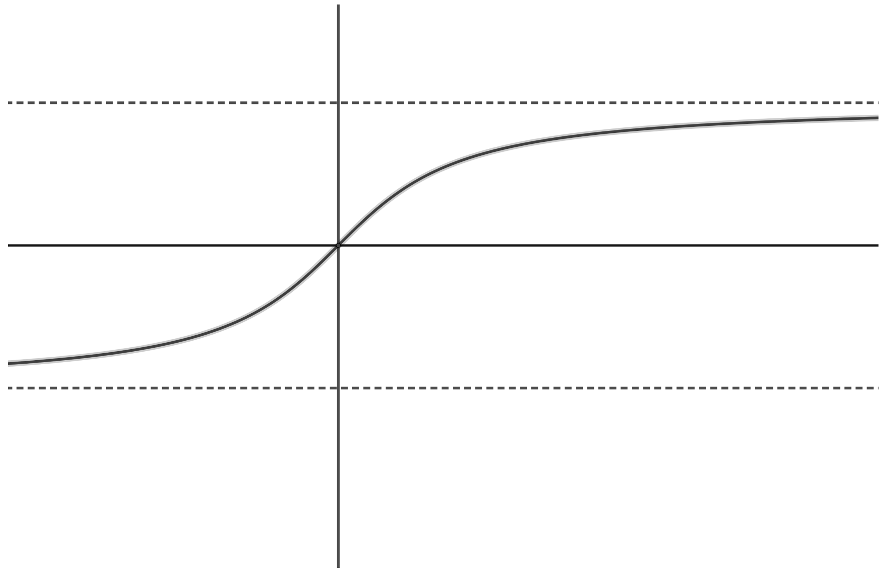


FIGURA 23. Gráfico do arco-tangente no intervalo $(-\pi/2, \pi/2)$.

3. Determine quais das funções da questão 2 são injetivas e determine a inversa daquelas que forem bijetivas.

4. Esboce os gráficos das funções abaixo.
   (a) $f_1(x) = 3x - 7$;
   (b) $f_2(x) = x^2 - 4$;
   (c) $f_3(x) = 2x^2 + 2x - 12$;
   (d) $f_4(x) = -3x^2 + 12x + 15$;
   (e) $f_5(x) = |2x - 1|$;
   (f) $f_6(x) = \begin{cases} 2x - 1 & \text{se} \quad x < 0 \\ 3x - 1 & \text{se} \quad x \geq 0 \end{cases}$
   (g) $f_7(x) = \begin{cases} -2x & \text{se} \quad x < 0 \\ x^2 & \text{se} \quad x \geq 0 \end{cases}$
   (h) $f_8(x) = \begin{cases} -1 & \text{se} \quad x < -1 \\ 7 & \text{se} \quad x \geq 1 \end{cases}$

5. Ache as equações das retas que passam pelos pontos:
   (a) $(1, 2)$ e $(3, -2)$;
   (b) $(-3, 0)$ e $(-1, 1)$;
   (c) $(-1, 2)$ e $(-3, -1)$;
   (d) $(-3, 0)$ e $(-1, 0)$.

6. Ache as retas perpendiculares às retas do exercício anterior.

7. Prove que as retas $y = ax + b$ e $y = a'x + b'$ são paralelas se, e somente se, $a = a'$.

8. Determine o eixo e o vértice do gráfico da função quadrática $q(x) = ax^2 + bx + c$, quando $a, b, c \in \mathbb{R}$ e $a < 0$.

9. Determine a tangente ao gráfico da parábola $q(x) = ax^2 + bx + c$ no ponto de abscissa $x = x_0$, quando $a, b, c \in \mathbb{R}$ e $a < 0$

10. Ache o eixo, o vértice, as raízes e esboce os gráficos das seguintes funções quadráticas:
    (a) $q_1(x) = x^2 - 4x + 3$;
    (b) $q_2(x) = 2x^2 - x - 3$;
    (c) $q_3(x) = -x^2 + x + 2$;
    (d) $q_4(x) = x^2 + x + 2$.

11. Determine os pontos de interseção entre os gráficos das funções dos itens abaixo:
    (a) $q_1(x) = x^2 - 4x + 3$ e $y = x - 1$;
    (b) $q_2(x) = 2x^2 - x - 3$ e $y = -x - 2$;
    (c) $q_3(x) = -x^2 + x + 2$ e $y = 3x + 1$;
    (d) $q_4(x) = x^2 + x + 2$ e $y = -1$.

12. Calcule as retas tangentes e as retas normais às parábolas dos itens abaixo, nos pontos dados:
    (a) $q_1(x) = x^2 - 4x + 3$ no ponto $(2, -1)$;
    (b) $q_2(x) = 2x^2 - x - 3$ no ponto $(1, -2)$;
    (c) $q_3(x) = -x^2 + x + 2$ no ponto $(2, 0)$;
    (d) $q_4(x) = x^2 + x + 2$ no ponto $(3, 14)$.
    Uma reta é *normal* a uma curva em um dado ponto quando ela é perpendicular à tangente à curva naquele ponto.

13. Seja α um ângulo entre 0 e $\pi/2$.
    (a) Calcule o seno, o cosseno e a tangente de $\pi/2 - \alpha$ em função do cosseno, do seno e da tangente de α.
    (b) Se $\beta + \gamma = \pi/2$, qual a relação entre $\tan(\beta)$ e $\tan(\gamma)$?
    Você deve justificar suas respostas para seno e cosseno usando o círculo trigonométrico.

14. Calcule o seno, o cosseno e a tangente de $-\alpha$ a partir do seno, do cosseno e da tangente de α.

15. Calcule o seno, o cosseno e a tangente de $\pi - \alpha$ a partir do seno, do cosseno e da tangente de α.

16. Use o exercício 14 e as fórmulas deduzidas da seção 2 para encontrar as expressões para $\operatorname{sen}(\alpha - \beta)$ e $\cos(\alpha - \beta)$ em função dos senos e cossenos de α e β.

17. Determine se as funções seno, cosseno e tangente são pares, ímpares, ou nenhuma das duas.

Bibliography

Astronomical cuneiform texts. I, II, III. 5. 1983. (Sources in the History of Mathematics and Physical Sciences). Babylonian ephemerides of the Seleucid period for the motion of the sun, the moon, and the planets, Reprint of the 1955 original.

Landmark writings in western mathematics 1640–1940. 2005.

Aristotle . Posterior analytics. 1960. (Loeb Classical Library). Edited and translated from the Greek by Hugh Tredennick.

Aristotle . Metaphysics I–IX. 1989. (Loeb Classical Library). Translated from the Greek by Hugh Tredennick.

Birkhoff George D. A set of postulates for plane geometry, based on scale and protractor // Ann. of Math. (2). 1932. 33, 2. 329–345.

Birkhoff George David, Beatley Ralph. Basic geometry. 1959. 294. 3rd ed.

Bourbaki Nicolas. Functions of a real variable. 2004. (Elements of Mathematics (Berlin)). Elementary theory, Translated from the 1976 French original [MR0580296] by Philip Spain.

Davis Martin. The universal computer. 2000. The road from Leibniz to Turing.

Dedekind Richard. Essays on the theory of numbers. I:Continuity and irrational numbers. II:The nature and meaning of numbers. 1963. Authorized translation by Wooster Woodruff Beman.

Descartes Réne. Key Philosophical Writings. 1997. (Wordsworth Classics of World Literature).

Devlin Keith. The joy of sets. 1993. Second. (Undergraduate Texts in Mathematics). Fundamentals of contemporary set theory.

Diadochi Procli. Primum Euclidis Elementorum Librum Comentarii. 1873.

Dieudonné Jean. Abrégé d'histoire des mathématiques 1700-1900. Nouv. version. (A short history of mathematics 1700-1900). 1986.

Dugac Pierre. Richard Dedekind et les fondements des mathématiques. 1976. (Collection des Travaux de l'Académie Internationale d'Histoire des Sciences. [Collection of the Transactions of the International Academy of the History of Science]). Avec de nombreux textes inédits, Préface de Jean Dieudonné, L'Histoire des Sciences: Textes et Études. [History of Science: Texts and Studies].

Fowler David, Robson Eleanor. Square root approximations in old Babylonian mathematics: YBC 7289 in context // Historia Math. 1998. 25, 4. 366–378.

Grzegorczyk Andrzej. An Outline of Mathematical Logic. 1974.

Hilbert David. The Foundations of geometry. 1910. second.

Kline Morris. Mathematical thought from ancient to modern times. Vol. 2. 1990. Second. i–xx, 391–812 and i–xxii.

Knorr Wilbur Richard. On the early history of axiomatics: the interaction of mathematics and philosophy in Greek antiquity // Classics in the History of Greek Mathematics. 2004. 81–109.

Laertius Diogenes. Lives of eminent philosophers. 2006. (Loeb Classical Library). Translated from the Greek and with an introduction and notes by R. D. Hicks.

Landau Edmund. Foundations of Analysis. The Arithmetic of Whole, Rational, Irrational and Complex Numbers. 1951. Translated by F. Steinhardt.

Melville Duncan J. Poles and walls in Mesopotamia and Egypt // Historia Mathematica. 2004. 31, 2. 148–162.

Netz Reviel, Noel William. The Archimedes codex. 2008. Revealing the secrets of the world's greatest palimpsest.

Pasch Moritz. Vorlesungen über neuere Geometrie. 1903.

Plofker Kim. Mathematics in India. 2009.

Proclus . A commentary on the first book of Euclid's *Elements*. 1992. (Princeton Paperbacks). Translated from the Greek and with an introduction and notes by Glenn R. Morrow, Reprint of the 1970 edition, With a foreword by Ian Mueller.

Russell Bertrand, Whitehead Alfred North. Principia Mathematica. 1910.

Thomas Ivor. Greek Mathematical Works I. 1991. (Loeb Classical Library).

Wardhaugh Benjamin. Encounters with Euclid—how an ancient Greek geometry text shaped the world. 2021. Originally published in 2020 by Harper Collins under the title ıt The book of wonders: the many lives of Euclid's Elements.