Teoria dos Conjuntos

18 de Junho de 2017

Conteúdo

Capítulo 1

Introdução

O infinito sempre assombrou os matemáticos e filósofos, estando relacionada aos maiores paradoxos e crises nos fundamentos da matemática, como é o caso dos famosos paradoxos de Zeno, de Eléia, que se baseiam em interpretações tortuosas do conceito de infinitude para "provar" a não existência de movimento.

No desenvolvimento da teoria dos conjuntos, o conceito de infinito desepenha um papel fundamental, sendo responsável por uma das maiores crises filosóficas na história da matemática. Como efeito dessa "crise", tivemos pressupostos matemáticos e filosóficos sendo destruídos, novos questionamentos surgindo, divergências na própria concepção de verdade matemática, novas propostas de formalização da disciplina. Enfim, como aconteceu com as grandezas incomensuráveis e a história do quinto postulado de Euclides (esses dois também estão diretamente relacionados ao conceito de infinitude) os paradoxos da teoria dos conjuntos contribuíram enormemente para o enriquecimento do pensamento matemático.

Vamos refletir um pouco sobre o conceito de conjuntos. Podemos pensar em conjunto como um agrupamento de objetos que compartilham uma propriedade comum [1]. Assim, podemos pensar na palavra *pássaro* como um conjunto de animais que possuem certas características, como o corpo coberto de penas e reprodução ovípora. Cada descrição de objetos, animais ou pessoas nos fornece um conjunto. Por exemplo: *o conjunto dos pássaros que voam*, ou *o conjunto dos alunos da USP*.

Podemos considerar a noção de conjuntos como um dos primeiros conceitos abstratos da mente humana. Um pássaro é um ser vivo que existe independente do pensamento humano (se deixarmos um pouco de lado excesso de divagações filosóficas). Mas a noção de pássaro – o conjunto de todos os pássaros do mundo – é uma noção abstrata, criada pelo nosso raciocínio.

O surgimento dos números naturais – uma das criações mais úteis do pensamento humano e um dos alicerces da matemática – pode ser visto como consequência da noção de conjuntos. Dizem alguns historiadores que, há muitos séculos atrás, antes de existir a contagem, os pastores usavam um saquinho de pedras para não perderem

[1] Isso está longe de ser uma definição. Seria difícil definir conjunto sem usar algum termo como *agrupamento*, *coleção*, ou outro que seja praticamente sinônimo de conjunto. Conjunto deve ser tratado como *conceito primitivo*, que não requer definição. O propósito deste parágrafo é discutirmos um pouco a ideia intuitiva de um conceito que já conhecemos, antes de entrar na abordagem axiomática.

as ovelhas. Quando levavam as ovelhas para pastar, para cada ovelha que passava guardavam uma pedrinha. Na hora de recolher as ovelhas, para cada ovelha que voltava ao curral, retiravam uma pedrinha. Se, no final, não sobrasse nenhuma pedra no saquinho, eles podiam se assegurar de que nenhuma ovelha se perdeu [2].

Dessa forma, estabeleceu-se uma semelhança entre o conjunto das pedras que um pastor levava em um saquinho e o conjunto das ovelhas que eles possuía: ambos têm a mesma *quantidade de elementos*, já que podemos associar biunivocamente cada elemento de um conjunto com um elemento de outro. Com o tempo, as pessoas perceberam que não havia necessidade de usar pedras de verdade, e inventaram um conceito abstrato chamado *número* (os números naturais). Cada número representava uma possível quantidade de elementos de um conjunto. Assim, o número 1 representava todos os conjuntos existentes que possuem 1 elemento. O número 2, todos os conjuntos que possuem 2 elementos, e assim por diante. E como saber a quantidade de elementos que cada conjunto contém? Bastava colocar os números em uma sequência (0, 1, 2, 3 etc.) e a cada elemento novo que colocamos num conjunto avançamos um número na sequência. Por exemplo, se um conjunto não tem nenhum elemento, ele é rotulado pelo número 0. A partir do momento que colocamos alguma coisa nesse conjunto, seu tamanho é "promovido" a 1, e assim por diante.

Nessas "trivialidades primitivas" é impressionante quantos conceitos complexos da matemática moderna estão sendo abordados: funções bijetoras, classe de conjuntos, relação de equivalência, ordinais, cardinais etc. Não é fácil explicar para uma pessoa ainda não familiarizada com a linguagem formal da matemática o que é uma função bijetora. Todavia, uma criança consegue contar o número de brinquedos que tem, e observem como é esse processo de contagem:

- A criança aponta com o dedo para um brinquedo e diz, em voz alta, o número 1.
- A criança aponta para outro brinquedo e diz 2, e assim sucessivamente na sequência dos números.
- A criança toma o cuidado de não apontar duas vezes para o mesmo brinquedo (não "contar repetido").
- A criança toma o cuidade de não deixar de apontar para nenhum brinquedo.
- No final, o último número mencionado pela criança durante a contagem – quando todos os brinquedos já foram contados – é a quantidade de brinquedos que ela possui.

Ou seja, a criança que contou e descobriu que tinha 10 brinquedos soube perfeitamente construir uma bijeção entre o conjunto dos brinquedos e o conjunto dos números que vão de 1 a 10. A cada número ela associou um brinquedo (*função*), números diferentes foram associados a brinquedos diferentes (*função injetora*) e cada brinquedo foi associado a um número (*função bijetora*).

[2] Alguns sustentam essa teoria com base na origem da palavra cálculo, que vem do latim *calculus*, que significa *pedra*.

Se pararmos para pensar sobre esse processo de contagem, surgem algumas questões. Uma possível pergunta é: *será que, no processo de contagem, chegaremos sempre ao mesmo número, indepente da sequência que seguimos?* Ou seja, se a criança começar a contar pelo carrinho e depois ir para o boneco, ou fazer o contrário, chegará no mesmo resultado? Outra questão: *será que os números naturais são suficientes para contar a quantidade de elementos de qualquer conjunto?* Ou seja, sempre haverá um momento em que a contagem para em um determinado número?

É justamente aí que os problemas da infinitude começam. As respostas a essas perguntas são *sim, e sim*, mas apenas para os *conjuntos finitos*. Aliás, essa pode ser justamente a definição de conjunto finito: *quando existe uma bijeção entre ele e o conjunto dos números menores que um determinado número.*

Reparem na sutileza no desenvolvimento do conceito de conjuntos. A princípio, conjunto é um conceito abstrato, mas até agora citamos exemplos de *conjuntos formados por objetos concretos*, como ovelhas, brinquedos e pessoas. Esse tipo de conjunto é sempre finito. Às vezes, não conseguimos contar por impossibilidade física. Quando dizemos, na linguagem natural, que a quantidade de areia na praia e a quantidade de estrelas no céu são *inumeráveis* [3], dizemos que é humanamente impossível contar. Mas existe uma quantidade finita delas. Existe um número natural que representa a quantidade de estrelas no céu, mesmo que nunca venhamos a saber qual é esse número. Até mesmo a quantidade de átomos no universo é finita, por mais que seja espantosamente grande.

Acontece que, a partir do momento que criamos conceitos abstratos – como conjuntos e números – podemos imaginar conjuntos não só de objetos concretos, *mas também de objetos abstratos*. Assim, uma vez que inventamos os números naturais, podemos pensar no *conjunto dos números naturais*. Como, para qualquer número natural, sempre existe um maior, então o conjunto dos números naturais é *infinito*.

O conjunto dos números naturais não é o único conjunto infinito que existe. Temos o conjunto dos pontos de uma reta, o conjunto das retas em um plano, o conjunto das frações, o conjunto dos números reais etc. Mas todos esses conjuntos são fomados por conceitos abstratos, e não por objetos concretos. Não é à toa, portanto, que a ideia de infinitude seja tão difícil de assimilar e, por muitas vezes, traia a nossa intuição e senso comum.

Agora voltemos à primeira questão: será que a ordem que utilizamos para contar as coisas não afeta o resultado? Ora, ninguém havia pensado nessa questão em conjuntos infinitos. Afinal, um conjunto infinito é infinito e pronto. Não tem como contar os elementos de um conjunto infinito. Porém, algumas mentes mais aguçadas ousaram aprofundar-se nas questões filosóficas da infinitude. O cientista italiano Galileu Galilei (1564–1642) decidiu usar a noção de funções bijetoras para comparar conjuntos infinitos, chegando em um resultado bem curioso. Ele considerou a função que associa, a cada número natural, o seu dobro, conforme o diagrama seguinte:

[3] Não confundir com *enumeráveis*, que é um conceito exclusivamente matemático, como será visto daqui a pouco.

$$\begin{array}{ccc} 0 & \longleftrightarrow & 0 \\ 1 & \longleftrightarrow & 2 \\ 2 & \longleftrightarrow & 4 \\ 3 & \longleftrightarrow & 6 \\ & \ldots & \end{array}$$

Com isso, Galilei mostrou que o conjunto dos números naturais "tem o mesmo tamanho" que o conjyunto dos números pares. Na época, isso parecia contradizer o axioma de Euclides que dizia que "o todo é sempre maior que a parte". O conjunto dos números pares é apenas uma parte do conjunto de todos os números naturais, e ainda assim ambos os conjuntos têm o mesmo tamanho, se utilizarmos essa noção de bijeções. Notem que isso só acontece com conjuntos infinitos. Em um conjunto finito, se tirarmos um único elemento já não conseguimos associar biunivocamente os elementos do conjunto reduzido com os do conjunto todo.

O hotel de Hilbert O matemático alemão David Hilbert (1862–1943) deu um exemplo parecido. Se chegamos em um hotel e todos os quartos estão ocupados, então sabemos que não há vaga nesse hotel, a menos que uma família saia. Agora imaginemos um hotel com infinitos quartos – um para cada número natural – sendo que todos estão ocupados. Chega uma nova família querendo se hospedar e o dono não quer despejar nenhum hóspede, mas também não quer recusar quarto para os recém-chegados. Como há infinitos quartos – mesmo que todos ocupados – é fácil resolver o problema. Basta passar cada hóspede para o quarto ao lado. Assim, quem está hospedado no quarto 0 vai para o quarto 1, e do quarto 1 para o 2, e assim por diante, sobrando o quarto 0 para os novos hóspedes.

O problema do dono do hotel parece se complicar quando chega um ônibus com uma infinidade de hóspedes, um hóspede para cada número natural. Mas a solução ainda é simples: ele passa cada hóspede de um quarto para outro cujo número é o dobro do primeiro. Sobra, assim, todos os números ímpares para colocar os novos hóspedes.

E se chegarem infinitos ônibus – cada ônibus marcado por um número natural diferente – com infinitos passageiros cada um – cada passageiro também marcado por um número – poderá ainda o dono do hotel hospedar todo mundo? Sim. E poderá fazê-lo de forma que não fique nenhum quarto vazio. Basta colocar o n-ésimo passageiro do m-ésimo ônibus no quarto $2^n \cdot (m + 1)$ (para simplificar, desta vez assumimos que o hotel está vazio – fica como exercício verificar o que se faria se o hotel estivesse lotado).

O paraíso de Cantor Aparentemente o paradoxo criado por Galilei não causou tanto impacto na matemática e na filosofia, nem foi devidamente explorado durante alguns séculos. Foi só no século XIX que o assunto foi trazido novamente à tona pelo matemático alemão Georg Cantor (1845–1918). Dessa vez, o impacto transformou totalmente o rumo da matemática moderna e deu início à teoria dos conjuntos, que será estudada neste curso.

Cantor não só criou um paradoxo ou uma discussão filosófica através dessa ideia de comparar tamanho de conjuntos infinitos: ele de fato resolveu um problema matemático usando esse conceito. Enquanto outros matemáticos tiveram uma grande dificuldade para provar que números como π e e são transcendentes (isto é, não são raízes de equações polinomiais de coeficientes inteiros), Cantor provou, de maneira relativamente simples, que existem muitos números transcendentes, mesmo sem exibir um sequer. Vamos aqui tratar brevemente dessa demonstração.

O conjunto dos números algébricos (os não transcendentes) aparentemente é muito maior que os números naturais. Para começar, esse engloba todos os racionais, uma vez que a fração $\frac{a}{b}$ é raiz da equação $bx - a$, e quase todos os números reais que conhecemos. Os transcendentes parecem ser estranhas exceções dentro do conjunto dos números reais. Se os irracionais já parecem aberrações, mais ainda os números transcendentes. Pois Cantor provou justamente o contrário: há muito mais números transcendentes do que algébricos. De fato, *o conjunto dos números algébricos tem o mesmo tamanho que o conjunto dos números naturais.*

Estabelecer uma bijeção entre os números naturais e os algébricos não é difícil. Primeiro, precisamos estabelecer uma bijeção entre os números naturais e os polinômios de coeficientes inteiros, ou seja, colocarmo-los numa sequência, como uma fila infinita.

O início da sequência deve ser constituída pelos polinômios de grau 1 e cujos coeficientes têm módulo menor ou igual a 1. Está claro que existe apenas uma quantidade finita desses polinômios. Podemos dispô-los em ordem lexicográfica, como a usada em dicionários, conforme descrevemos abaixo.

$$
\begin{gathered}
-x-1 \\
-x \\
-x+1 \\
x-1 \\
x \\
x+1
\end{gathered}
$$

Continuamos a sequência escrevendo os polinômios de grau menor ou igual a 2, cujos coeficientes têm módulo menor ou igual a 2, e que não estão na lista anterior. Usamos a mesma ordem lexicográfica dos coeficientes, começando com os polinômios de grau menor (ou maior, como queiram). Prosseguimos esse processo para 3, 4 e assim por diante, e isso irá contemplar todos os polinômios de coeficientes inteiros, conforme ilustra o seguinte diagrama:

$$
\begin{array}{ccc}
0 & \longleftrightarrow & -x-1 \\
1 & \longleftrightarrow & -x \\
2 & \longleftrightarrow & -x+1 \\
3 & \longleftrightarrow & x-1 \\
4 & \longleftrightarrow & x \\
5 & \longleftrightarrow & x+1 \\
6 & \longleftrightarrow & -2x-2 \\
7 & \longleftrightarrow & -2x-1 \\
8 & \longleftrightarrow & -2x \\
9 & \longleftrightarrow & -2x+1 \\
10 & \longleftrightarrow & -2x+2 \\
11 & \longleftrightarrow & -x-2 \\
12 & \longleftrightarrow & -x+2 \\
13 & \longleftrightarrow & x-2 \\
14 & \longleftrightarrow & x+2 \\
15 & \longleftrightarrow & 2x-2 \\
16 & \longleftrightarrow & 2x-1 \\
17 & \longleftrightarrow & 2x \\
18 & \longleftrightarrow & 2x+1 \\
19 & \longleftrightarrow & 2x+2 \\
20 & \longleftrightarrow & -2x^2-2x-2 \\
 & \cdots &
\end{array}
$$

Agora, para "colocarmos em fila" os números algébricos basta substituirmos cada polinômio pelas suas raízes (em ordem crescente), suprimindo os que já foram listados. Fazendo assim obtemos:

$$
\begin{array}{ccll}
0 & \longleftrightarrow & -1 & (\text{raiz do polinômio } -x-1) \\
1 & \longleftrightarrow & 0 & (\text{raiz do polinômio } -x) \\
2 & \longleftrightarrow & 1 & (\text{raiz do polinômio } -x+1) \\
3 & \longleftrightarrow & -\frac{1}{2} & (\text{raiz do polinômio } -2x-1) \\
4 & \longleftrightarrow & \frac{1}{2} & (\text{raiz do polinômio } -2x+1) \\
5 & \longleftrightarrow & -2 & (\text{raiz do polinômio } -x-2) \\
6 & \longleftrightarrow & -2 & (\text{raiz do polinômio } -x+2) \\
7 & \longleftrightarrow & \frac{1-\sqrt{3}}{2} & (\text{primeira raiz de } -2x^2-2x+1) \\
8 & \longleftrightarrow & \frac{1+\sqrt{3}}{2} & (\text{segunda raiz de } -2x^2-2x+1) \\
 & \cdots & &
\end{array}
$$

Com isso Cantor mostrou que o conjunto dos números algébricos "tem o mesmo tamanho" que o dos números naturais. Isso significa dizer que o conjunto dos números algébricos é *enumerável*, ou seja, podemos enumerar todos seus elementos numa lista infinita, indexada com os números naturais.

É fácil intuir [4] que um subconjunto infinito de um conjunto enumerável é enumerável. Assim, os conjuntos dos números inteiros, racionais e algébricos são todos enumeráveis.

[4] A demonstração rigorosa desse fato é mais trabalhosa, como veremos posteriormente.

A essa altura começamos a imaginar que todos os conjuntos são enumeráveis. Talvez por isso o aparente paradoxo de Galilei não tenha impactado tanto os matemáticos. Infinito é infinito e parece natural que todos os conjuntos infinitos tenham o mesmo tamanho. Parece que, se nos esforçarmos bem, como fizemos com os números algébricos, conseguimos colocar qualquer conjunto infinito numa sequência bem comportada. Porém, Cantor surpreende a todos ao provar que o conjunto dos números reais *não* é enumerável.

Vejamos a prova de Cantor da não-enumerabilidade dos números reais. Seja f uma função de $\mathbb{N}$ em $\mathbb{R}$. Mostraremos que f não pode ser sobrejetora.

Para cada n natural, consideremos a_n a parte inteira de $f(n)$ e $(a_{nm})_{m\in\mathbb{N}}$ a sequência dos algarismos após a vírgula na representação decimal [5] de $f(n)$.

$$\begin{aligned} f(0) &= a_0, a_{00}, a_{01}, a_{02}, a_{03} \dots \\ f(1) &= a_1, a_{10}, a_{11}, a_{12}, a_{13} \dots \\ f(2) &= a_2, a_{20}, a_{21}, a_{22}, a_{23} \dots \\ f(3) &= a_3, a_{30}, a_{31}, a_{32}, a_{33} \dots \\ &\dots \end{aligned}$$

Agora mostremos que existe um real r que não pertence a essa lista. Definimos r da seguinte forma: a parte inteira pode ser qualquer número (0, por exemplo) e a n-ésima casa decimal de r será 1 se a_{nn} for 0 e será 0 caso contrário. Portanto, para todo n teremos que a n-ésima casa de $f(n)$ difere da n-ésima casa de r, de onde concluímos que r não está na imagem de f.

Ou seja, escolhemos um número real que "evita" a diagonal da matriz infinita formada pelas casas decimais de cada número real da sequência. Essa prova ficou conhecida como *argumento diagonal de Cantor* [6].

Com isso Cantor mostrou que o conjunto dos números reais é *não-enumerável*, isto é, realmente a quantidade de números reais é maior que dos números naturais. Ora, se o conjunto dos números algébricos é enumerável, e o conjunto dos números reais é não-enumerável, concluímos que existem infinitos números reais que não são algébricos.

Concluímos também que há uma bijeção entre os números reais e os transcendentes. De fato, considere em $\mathbb{R}$ uma sequência $(x_n)_{n\in\mathbb{N}}$ de números transcendentes distintos (por exemplo, x_n pode ser $\pi + n$) e $(a_n)_{n\in\mathbb{N}}$ a sequência de todos os números algébricos (lembre-se que os algébricos são enumeráveis). Podemos definir uma função bijetora do conjunto dos números reais nos transcendentes da seguinte forma: cada a_n é mapeado para x_{2n}, cada x_n é mapeado para x_{2n+2}, e os demais números são mapeados para eles mesmos.

A demonstração de Cantor causou uma das maiores controvérsias da história da matemática. Para alguns, essa prova desvirtua o propósito da matemática e perde relação com o mundo real. Para outros, foi uma inovação no pensamento abstrato e um grande passo para a Rainha das Ciências. O matemático francês Henri Poincaré

[5] Aqui assumimos que **a** representação decimal é aquela que nunca utiliza uma dízima de período 9. Ou seja, a representação decimal de 1 que consideraremos é $1,000\dots$, e não $0,999\dots$.

[6] Um argumento semelhante foi usado por Gödel em uma parte crucial da demonstração do Teorema da Incompletude.

(1854–1912) chegou a dizer que "*o cantorismo é uma doença da qual a matemática precisa se curar*", enquanto, por outro lado, David Hilbert reagia às críticas a Cantor dizendo que "*ninguém nos tirará do paraíso criado por Cantor*".

Teoria dos conjuntos nos fundamentos da matemática Apesar dos protestos dos construtivistas e intuicionistas, a matemática moderna cedeu à elegância da teoria dos conjuntos desenvolvida por Cantor, Zermelo, Frankel e Von Neumann, dentre outros, e fez dela um dos pilares – ao lado da lógica de primeira ordem – da sua fundamentação. Pelo menos é o que aceita a maioria da comunidade matemática atual.

A ideia de usar conjuntos para formalização da matemática é definir todos os objetos matemáticos como conjuntos. *Tudo é conjunto.* Cada número natural é um conjunto, uma função é um conjunto, uma relação é um conjunto, os elementos de um conjunto são, eles próprios, conjuntos.

Inicialmente, o conceito de conjuntos estava diretamente ligado a fórmulas da linguagem de primeira ordem com uma variável livre. Por exemplo, a fórmula $\exists y(x = 2 \cdot y)$ tem x como variável livre e, se pensarmos no universo dos números naturais, representa o conjunto dos números pares. Um conjunto, então, é determinado por uma propriedade, conforme explicamos no início deste capítulo.

Gottlob Frege (1848–1925) tentou levar essa ideia adiante, propondo uma formalização da matemática em que lógica e conjuntos eram praticamente indissociáveis. Porém,Bertrand Russell (1872–1970) encontrou uma inconsistência nessa formalização, através do seu famoso paradoxo [7].

Se qualquer propriedade determina um conjunto, então podemos definir um conjunto X como *o conjunto de todos os conjuntos que não pertencem a si mesmos* [8].

Surge a pergunta: X pertence a si mesmo? Se sim, então, pela sua definição, ele não pode pertencer. Se não pertence a si mesmo, novamente usando sua definição, concluímos que ele pertence. Chegamos numa inevitável contradição, que só se resolve não permitindo a existência de tal conjunto.

Isso destrói a proposta de Frege de unificar conjuntos e lógica, relacionando um conjunto com uma sentença que descreve seus elementos. A existência de um conjunto

[7] Esse paradoxo possui uma variança popular conhecido como *paradoxo do barbeiro*, que dizia que havia numa cidade um barbeiro que cortava o cabelo de todas as pessoas que não cortavam seu próprio cabelo, e apemas dessas. Pergunta: quem cortava o cabelo do barbeiro?

[8] Podemos nos perguntar se é possível um conjunto pertencer a si próprio. Nota-se que há uma diferença entre *pertencer* a si próprio e *estar contido* em si próprio. Essa confusão entre as duas relações é muito comum, devido a uma falha clássica do ensino de matemática no nível básico, que será discutida melhor durante o início da disciplina. Um conjunto sempre está contido nele próprio, mas poderá pertencer a si próprio?

Se permitirmos livremente a construção de conjuntos através de uma expressão que descreve todos seus elementos, e ainda utilizarmos a linguagem natural, cheia de auto-referências, podemos definir o *conjunto de todos os objetos que podem ser descritos com menos de vinte palavras.* Certamente esse conjunto, se assim existisse, pertenceria a ele próprio. Ou, um exemplo mais simples, se existir *o conjunto de todos os conjuntos*, ele pertence a si próprio.

Porém, veremos posteriormente que, pelo axioma da regularidade, na teoria axiomática dos conjuntos não é possível um conjunto pertencer a si mesmo. Muito menos existe um conjunto de todos os conjuntos .

precisava ser mostrada, a partir de uma lista de axiomas, não sendo mais sua definição o suficiente para garantir a sua existência.

Ernest Zermelo (1871–1953) e Abraham Fraenkel (1891–1965) foram os responsáveis pela formalização axiomática dos conjuntos, que, em sua homenagem, ficou conhecida como *sistema ZFC*. A letra C vem do inglês *choice*, uma referência ao axioma da escolha, que, pelas polêmicas em torno dele, costuma ser "evitado" por alguns matemáticos. Assim, nos referimos ao sistema ZF quando excluímos o axioma da escolha, e ZFC quando o utilizamos. Vários matemáticos gostam de deixar bem claro quando um resultado usa esse axioma, fazendo bastante esforço para não precisar lançar mão dele.

Embora o sistema ZFC tenha sido criado por Zermelo e Fraenkel, parte da formalização que temos hoje é atribuído a John von Neumann (1903 – 1957), que também teve grande participação na invenção do computador moderno.

A lógica e a teoria dos conjuntos passaram a seguir caminhos separados – porém entrelaçados – na formalização da matemática. O próximo capítulo trata dessa dicotomia.

Exercícios

Os exercícios apresentados neste capítulo são apenas para fins de uma discussão introdutória, sem muita formalização, e usando noções intuitivas de conjuntos e funções.

1. Mostre uma bijeção entre o conjunto dos números inteiros e os naturais.

2. Prove que qualquer subconjunto infinito dos números naturais é enumerável.

3. Na bijeção que construímos entre os números naturais e os polinômios, encontre o polinômio associado ao número 30.

4. Na bijeção que construímos entre os números naturais e os números algébricos, encontre o número natural associado ao número $\sqrt{3}$

5. Suponha que, em um conjunto infinito, existe uma forma de representar cada elemento do conjunto como uma sequência finita de símbolos, dentre um conjunto finito de símbolos. Mostre que esse conjunto é enumerável e use esse resultado diretamente para mostrar que os conjuntos dos números racionais e dos números algébricos são enumeráveis.

Capítulo 2

Lógica de primeira ordem

Antes de falarmos sobre a teoria dos conjuntos, precisamos de algumas noções sobre a lógica de primeira ordem, que será usada em sua formalização.

Há um círculo vicioso entre lógica de primeira ordem e teoria dos conjuntos. A formalização de uma depende da formalização da outra. Seja como for que lidemos com essa dicotomia, em algum momento precisamos apelar para a abordagem intuitiva da outra. Ou seja, podemos desenvolver toda a teoria dos conjuntos de forma axiomática mas utilizando a linguagem natural (tal como Halmos faz em seu livro, e também como é feito nas disciplinas de Análise Real e Álgebra) para, posteriormente, formalizarmo-la com a lógica de primeira ordem (que possui a vantagem de ser muito próxima à argumentação que costumamos fazer na linguagem natural, para provarmos teoremas matemáticos). Ou podemos estudar lógica primeiro, utilizando noções intuitivas de teoria dos conjuntos – tais quais aprendemos no Ensino Médio – para depois desenvolvermos a teoria dos conjuntos axiomaticamente. Seguiremos aqui uma terceira opção: apresentar apenas uma parte da lógica de primeira ordem (a sintaxe) – que requer apenas uma parcela mínima de noções intuitivas de conjuntos e aritmética – para depois formalizar a teoria dos conjuntos com o rigor da lógica.

Podemos separar a lógica de primeira ordem em três aspectos: a linguagem, o sistema de axiomas e a semântica. Os dois primeiros constituem a *sintaxe* da lógica de primeira ordem, que trata da manipulação dos símbolos através de regras bem definidas, livre de contexto e de significado. A *semântica* trata justamente do significado das expressões lógicas. É justamente na semântica que o uso de teoria dos conjuntos é mais evidente e, por essa razão, trataremos aqui apenas da parte sintática, fazendo apenas alguns comentários a respeito da semântica.

A lógica de primeira ordem pode se adaptar a vários contextos, apresentando símbolos específicos de algum assunto que quisermos axiomatizar. Assim, para axiomatizar a aritmética utilizamos alguns símbolos específicos da aritmética, como $+$, $\times$, 0 e 1. Na teoria dos conjuntos, o símbolo específico será o de pertinência ($\in$). Por isso, muitas vezes, em vez de dizermos *a* lógica de primeira ordem, dizemos *uma* lógica de primeira ordem, ou *uma linguagem* de primeira ordem.

Aqui trataremos especificamente da linguagem da teoria dos conjuntos. Não demonstraremos nenhum dos teoremas aqui enunciados [1]. Como referência recomen-

[1] Os teoremas a respeito da lógica de primeira ordem fazem parte do que chamamos de *meta-*

damos o livro *Set Theory and Logic*, de Robert Stoll.

2.1 O alfabeto

Os símbolos utilizados na linguagem da teoria dos conjuntos são os seguintes:

Variáveis: representadas pelas letras minúsculas: $x, y, z, \ldots$. Eventualmente, são indexadas pelos números naturais: $x_1, x_2, x_3, \ldots$.

Conectivos: $\neg$ (negação – "não"), $\rightarrow$ (condicional – "se... então"), $\wedge$ (conjunção – "e"), $\vee$ (disjunção – "ou"), $\leftrightarrow$ (bicondicional – "se, e somente se").

Quantificadores: $\forall$ (quantificador universal – "para todo"), $\exists$ (quantificador existencial – "existe").

Parênteses: são os parênteses esquerdo e direito: (e).

Símbolo de igualdade: $=$

Predicado binário: $\in$ (pertence).

2.2 Fórmulas

Fórmulas são sequências finitas de símbolos do alfabeto que seguem as seguintes regras:

1. Se x e y são variáveis, $x \in y$ e $x = y$ são fórmulas.
2. Se A e B são fórmulas, $\neg(A)$, $(A) \rightarrow (B)$, $(A) \wedge (B)$, $(A) \vee (B)$ e $(A) \leftrightarrow (B)$ são fórmulas;
3. Se A é fórmula e x é uma variável, então $\forall x(A)$ e $\exists x(A)$ são fórmulas.
4. Todas as fórmulas têm uma das formas descritas nos itens 1, 2 e 3.

Por exemplo, pela regra 1, temos que $x \in y$ é uma fórmula. Pela regra 1, $x = z$ também é uma fórmula. A regra 2 nos garante que $(x \in y) \rightarrow (x = z)$ é uma fórmula. Logo, a regra 3 nos garante que $\forall x((x \in y) \rightarrow (x = z))$ é uma fórmula.

---

matemática, isto é, a matemática utilizada para formalizar a matemática. A lógica de primeira ordem é a linguagem utilizada na matemática. Então nos perguntamos qual é a linguagem utilizada quando formalizamos a lógica de primeira ordem. Obviamente, utilizamos a linguagem natural, mas podemos, posteriormente, formalizá-la utilizando a própria ordem de primeira ordem. A essa linguagem que utilizamos para descrever a lógica de primeira ordem chamamos de *metalinguagem*.

Em seu livro *Uma Breve História do Tempo*, Stephen Hawking menciona uma história que serve como uma curiosa alegoria para entendermos o que é metalinguagem e metamatemática: de acordo com algumas pessoas, a Terra era achatada e estava apoiada no casco de uma tartaruga gigante, sendo que essa tartaruga, por sua vez, estava apoiada no casco de uma outra tartaruga gigante, e assim sucessivamente.

De fato, é uma expressão que "faz sentido" (ou seja, entendemos o que ela significa, independente de ser verdadeira ou não). Traduzindo para a linguagem natural, seria o seguinte: "para todo x, se x pertence a y então x é igual a z". Ou, simplesmente, "z é o único elemento de y".

As fórmulas usadas no processo de construção de fórmulas mais complexas são chamadas de *subfórmulas*. Por exemplo, A e B são subfórmulas de $(A) \to (B)$. No caso do nosso exemplo, as subfórmulas de $\forall x((x \in y) \to (x = z))$ são $x \in y$, $x = z$, $(x \in y) \to (x = z)$ e, para alguns efeitos práticos, consideramos a própria fórmula $\forall x((x \in y) \to (x = z))$ como subfórmula dela mesma.

As fórmulas que constam no item 1 são chamadas de *fórmulas atômicas*, porque não podem ser divididas em subfórmulas menores.

2.3 Unicidade de representação

A regra 4 nos diz que as únicas fórmulas são aquelas que se enquadram numa das três anteriores. Ou seja, toda fórmula é da forma $x \in y$, $x = y$, $\neg(A)$, $(A) \to (B)$, $(A) \wedge (B)$, $(A) \vee (B)$, $(A) \leftrightarrow (B)$, $\forall x(A)$ ou $\exists x(A)$, onde x e y são variáveis e A e B são fórmulas. Uma questão importantíssima para evitarmos ambiguidades na limguagem é: toda fórmula pode ser escrita em *apenas uma* dessas maneiras? Isto é, olhando para uma sequência de símbolos que representa uma fórmula, existe apenas uma maneira de lermos essa sequência de símbolos como uma dessas formas?

A resposta é *sim*: se escrevemos uma mesma fórmula (enxergando fórmula como sequência de símbolos) de duas das maneiras escritas acima, tanto o símbolo quanto as variáveis e fórmulas envolvidas são as mesmas, nas duas maneiras. Não demonstraremos isso aqui. Apenas ressaltamos que esse é o papel dos parênteses na fórmula. Por exemplo, se não houvesse parênteses, considere a fórmula $x \in y \to x = z \vee z \in x$. Podemos cosiderá-la como da forma $A \to B$, onde A é a fórmula $x \in y$ e B é a fórmula $x = z \vee z \in x$, ou como da forma $A \vee B$, onde A é a fórmula $x \in y \to x = z$ e B é a fórmula $z \in x$. Assim, sem os parênteses não sabemos se se trata de uma disjunção ou de uma implicação, gerando uma ambiguidade que, inclusive, fará diferença na interpretação da fórmula. Porém, com a regra dos parênteses na formação das fórmulas, ou a escrevemos $(x \in y) \to ((x = z) \vee (z \in x))$ – que não há outra forma de descrevermo-la senão da forma $(A) \to (B)$ – ou escrevemos $((x \in y) \to (x = z)) \vee (z \in x)$ – que é uma fórmula exclusivamente da forma $(A) \vee (B)$.

Há uma notação que dispensa o uso de parênteses e, mesmo assim, é livre de ambiguidades. Chama-se *notação pré-fixada*, ou *notação polonesa*, que consiste em colocar os símbolos na frente das fórmulas e variáveis. Por exemplo, no lugar de $x \in y$ escreveríamos $\in xy$, no lugar de $x = y$ seria $= xy$, em vez de $(A) \wedge (B)$ teríamos $\wedge AB$. As fórmulas que acabamos de escrever ficariam $\to \in xy \vee = xz \in zx$ ou $\vee \to \in xy = xz \in zx$. Essa notação é elegante e evidencia a questão da unicidade, pois basta observarmos o primeiro símbolo para reconhecermos o formato da fórmula. Porém, como o leitor deve ter percebido, a leitura e compreensão das fórmulas escritas nessa notação não são nada intuitivas, e se tornam piores para fórmulas longas [2].

[2] Quem já usou a calculadora financeira HP12C deve se lembrar que ela usa uma notação seme-

Lembremos que a unicidade de representação se refere às fórmulas como sequências de símbolos, garantindo que existe uma única maneira correta de *ler* essa sequência. Naturalmente, como veremos posteriormente, existem fórmulas distintas que possuem o mesmo *significado*. Fórmulas desse tipo são ditas *equivalentes*, mas não são iguais.

2.4 Omissão de parênteses

Como uma espécie de abuso de notação, às vezes omitimos alguns parênteses desnecessários para a correta compreensão da fórmula. Por exemplo, embora a forma correta seja $(x = y) \wedge (\neg(x \in y))$, podemos escrever simplesmente $(x = y) \wedge \neg(x \in y)$, sem prejuízo da compreensão da fórmula. Outra situação é que evitamos o uso de parênteses é em torno de um quantificador, como no exemplo $\forall x(x \in y) \rightarrow \exists x(x \in y)$.

Em sequências de conjunções – e em sequências de disjunções – também omitimos os parênteses. Por exemplo, podemos escrever simplesmente $(x = y) \vee (x \in y) \vee (y \in x)$. Embora essa notação seja ambígua a respeito do formato – pois, apesar de sabermos que é uma fórmula do tipo $(A) \vee (B)$, não tem como sabermos se A é $x = y$ e B é $(x \in y) \vee (y \in x)$, ou se A é $(x = y) \vee (x \in y)$ e B é $y \in x$ – as duas possíveis formas são logicamente equivalentes, ou seja, expressam o mesmo significado.

2.5 Variáveis livres

Cada lugar que surge uma variável dentro de uma subfórmula atômica de uma fórmula chamamos de *ocorrência* de tal variável. Por exemplo, a fórmula $(x = y) \vee (x \in z)$ apresenta duas ocorrências da variável x, e uma de cada uma das variáveis y e z. Na fórmula $\forall x(x = y)$, não consideramos o primeiro símbolo x como uma ocorrência da variável, pois não está numa subfórmula atômica. Ou seja, não consideramos como ocorrência de uma variável quando tal símbolo está imediatamente após um quantificador.

Dizemos que uma ocorrência de uma variável y numa fórmula A está *no escopo* de uma variável x se a A apresenta uma subfórmula da forma $\forall x(B)$ ou $\exists x(B)$, e essa ocorrência de y está em B. Por exemplo, na fórmula $(x \in y) \wedge \exists x(y = x)$, a segunda ocorrência de y está no escopo da variável x, mas a primeira, não.

Dizemos que uma ocorrência de uma variável x numa fórmula A *é livre* se tal ocorrência não está no escopo dela mesma. Chamamos de *variáveis livres de uma fórmula* A aquelas que apresentam pelo menos uma ocorrência em que é livre. Uma *sentença* é uma fórmula que não apresenta variáveis livres.

Por exemplo, a fórmula $\neg(x \in y)$ (x não pertence a y) apresenta duas variáveis livres: x e y. Não podemos, portanto, julgar tal fórmula como verdadeira ou falsa, pois não conhecemos quem é x ou quem é y. As variáveis correspondem ao pronome, na linguagem cotidiana. Se falarmos *Ele foi à feira*, a pergunta que naturalmente surge é: *Ele quem?* Se falarmos *João foi à feira*, ou *alguém do prédio foi à feira*, ou

lhante, só que pós-fixada, em vez de pré-fixada. Ou seja, nessa calculadora pressionamos primeiro os números (separados pela tecla "enter") e depois pressionamos a operação para obtermos os resultados.

todo mundo do prédio foi à feira, então a frase fica mais completa, e ganha o status de *sentença*, que permite averiguar se a frase é verdadeira ou falsa.

Digamos, então, que acrescentemos um quantificador no nosso exemplo. A fórmula $\forall x \neg(x \in y)$ tem apenas uma variável livre: que é y. A variável x não ocorre livre, pois só ocorre no escopo dela própria. A fórmula significa "para todo x, x não pertence a y", ou, colocada de outra forma, "y não possui elementos", ou, simplesmente "y é um conjunto vazio". Observamos que, para julgarmos a fórmula como verdadeira ou falsa, basta agora conhecermos quem é y. Em outras palavras, a fórmula em questão nos dita uma propriedade a respeito de y, enquanto a fórmula $\neg(x \in y)$ dita uma propriedade a respeito de x e de y.

Se, porém, escrevemos $\exists y \forall x \neg(x \in y)$, não há mais variáveis livres nessa fórmula. Essa é uma *sentença*, cujo significado não depende mais de interpretarmos as variáveis. Essa sentença diz que *existe um conjunto vazio*, que veremos ser verdadeira. Se escrevêssemos $\forall y \forall x \neg(x \in y)$ teríamos um significado totalemnte diferente, que seria *todo conjunto é vazio*. Claramente essa é uma sentença falsa. Mas é uma sentença, pois os símbolos estão dispostos numa ordem que faz sentido e não apresenta variáveis livres.

Se A é uma fórmula e x e y são variáveis, denotamos por A_x^y a fórmula obtida ao substituirmos toda ocorrência livre da variável x pela variável y. Por essa notação, A é sentença se A_x^y é igual a A, para todas variáveis x e y.

Frequentemente denotamos por $P(x)$ uma fórmula que tem x como (única) variável livre, ou por $P(x, y)$ uma fórmula que tem duas variáveis livres, x e y (e analogamente para outras quantidades de variáveis livres). Nesse caso, $P(y)$ denota $P(x)_x^y$.

O motivo de utilizarmos a letra P nessa notação é justamente pelo fato de $P(x)$ designar uma propriedade de x. Veremos mais para frente como criar fórmulas para representar propriedades como "x é um conjunto infinito", ou "x é enumerável".

2.6 Abreviaturas

À medida que desenvolvemos assuntos mais complexos, as fórmulas vão se tornando demasiadamente longas e ilegíveis. Para resolver isso, introduzimos novos símbolos que funcionam como abreviaturas para expressões maiores. O importante é que o processo de conversão da linguagem abreviada para a linguagem da lógica de primeira ordem seja perfeitamente claro.

Comecemos a exemplificar isso com o símbolo de inclusão. Dizemos que x *está contido* em y se todo elemento de x pertence a y. A fórmula para designar inclusão é $\forall z((z \in x) \to (z \in y))$. Observe que essa fórmula tem duas variáveis livres, x e y. Abreviamos essa fórmula como $x \subset y$.

Assim como o símbolo de pertinência, a inclusão é um *predicado binário* (ou *símbolo relacional binário*), pois relaciona uma propriedade entre dois objetos do universo (no caso, o universo dos conjuntos). Poderíamos ter introduzido o símbolo de inclusão entre os *símbolos primitivos*, como o de pertinência. Mas como a inclusão é perfeitamente definível a partir da pertinência e dos demais símbolos lógicos, é tecnicamente mais fácil utilizarmos o símbolo de inclusão apenas como abreviatura.

Outras abreviaturas são um pouco mais sutis na transcrição. Por exemplo, o conjunto vazio é denotado por $\emptyset$. A rigor, para utilizarmos a expressão *o* conjunto vazio e denotá-lo por um símbolo, antes precisaríamos mostrar que ele existe e é único. Aceitemos esse fato, por enquanto, antes de o provarmos num momento oportuno.

Saber utilizar corretamente essa abreviatura requer um pouco mais de atenção. Primeiro notemos que, ao contrário da inclusão, o conjunto vazio não se refere a uma relação entre objetos, mas *a um objeto em particular*, e, ao contrário das variáveis, se refere a *um objeto bem definido*. Corresponde a um nome próprio na linguagem cotidiana. A esse tipo de símbolo, na lógica, chamamos de *constante*.

Assim como as variáveis, as constantes são *termos*, isto é, se referem a objetos do universo. Podemos utilizá-las no lugar de uma variável em fórmulas atômicas. Por exemplo, $\emptyset \in x$ é uma fórmula na linguagem abreviada. Para encontrarmos o correspondente na linguagem original, precisamos *explicar quem é* $\emptyset$. Para isso, tomamos uma variável que não está na fórmula (y, por exemplo) e escrevemos da seguinte forma:

$$\forall y((\forall x \neg(x \in y)) \rightarrow y \in x)$$

Um importante detalhe da fórmula acima é que a ocorrência não-livre da variável x não mantém qualquer relação com a ocorrência livre que ocorre a seguir (se quiserem, podem substituir x por z, tanto na primeira ocorrência, em $x \in y$ quanto após o $\forall$). A fórmula significa, numa interpretação literal, "para todo y, se y não possui elementos, então y é pertence a x", ou, "para todo y, se y é vazio, então y pertence a x", ou, simplesmente, "o conjunto vazio pertence a x". Notem que essa fórmula apresenta x como a única variável livre.

Descrevemos, a seguir, o processo formal dessa abreviatura:

> Seja B a sequência de símbolos obtida ao substituirmos todas as ocorrências livres de uma variável x numa fórmula A pelo símbolo $\emptyset$. Então B designará a fórmula $\forall x((\forall y \neg(y \in x)) \rightarrow (A)$.

Outro exemplo que citaremos aqui é da união de conjuntos. A expressão $x \cup y$ representa o conjunto formado pelos elementos que pertencem x ou a y. Ou seja, $\forall z(z \in x \cup y \leftrightarrow ((z \in x) \vee (z \in y))$.

Desta vez, essa abreviatura trata-se de um *símbolo funcional binário*, pois associa a cada dois objetos do universo um terceiro. Outros exemplos de símbolos funcionais binários são as operações $+$ e $\times$ na aritmética. Eis o detalhamento do processo de abreviatura:

> Sejam A uma fórmula e x, y, z variáveis distintas. Seja B a sequência de símbolos obtida ao substituirmos toda ocorrência livre de z em A por $x \cup y$. Então B designa a fórmula
>
> $$\forall z(\forall w((w \in z) \leftrightarrow ((w \in x) \vee (w \in y))) \rightarrow A)$$

Para algumas finalidades – como no estudo da metamatemática ou na elaboração do sistema de axiomas, como será feito na seção seguinte – convém reduzirmos os

símbolos primitivos ao mínimo possível. A partir de agora, passaremos a considerar como símbolo primitivo da linguagem apenas as variáveis, os parênteses, o símbolo de pertinência $\in$, o símbolo de igualdade $=$, o quantificador universal $\forall$, a negação $\neg$ e a implicação $\rightarrow$.

Definiremos a partir desses símbolos os demais anteriormente descritos: $\vee$, $\wedge$, $\leftrightarrow$ e $\exists$. Eis as regras:

$(A) \vee (B)$ é abreviatura para $(\neg(A)) \rightarrow (B)$;

$(A) \wedge (B)$ é abreviatura para $\neg((\neg(A)) \vee (\neg(B))$;

$(A) \leftrightarrow (B)$ é abreviatura para $((A) \rightarrow (B)) \wedge ((B) \rightarrow (A))$;

$\exists x(A)$ é abreviatura para $\neg(\forall x(\neg(A)))$.

2.7 Símbolos relacionais e funcionais

A linguagem da teoria dos conjuntos possui como símbolos específicos – isto é, símbolo que não é comum a todas as linguagem de primeira ordem – o de pertinência ($\in$) e o de igualdade ($=$). Porém, para facilitar a escrita, introduzimos diversos símbolos novos que podem ser vistos como abreviaturas da linguagem.

Os símbolos específicos de uma linguagem de primeira ordem se dividem em três tipos: as constantes, os símbolos funcionais e os símbolos relacionais.

As constantes se referem a objetos específicos do universo. Por exemplo, $\emptyset$ é uma constante da linguagem estendida da teoria dos conjuntos que significa o conjunto vazio. O símbolo ω é uma constante que irá designar o conjunto dos números naturais. Os símbolos 0 e 1 são constantes da linguagem da aritmética que corresponde aos números naturais zero e um, respectivamente.

Os símbolos funcionais – juntamente com as variáveis e constantes – são usadas para compor os *termos*, que também denotam objetos do universo. Cada símbolo funcional vem acompanhado de um número inteiro positivo chamado de *grau de aridade*, que corresponde ao número de parâmetros. Dizemos que um símbolo funcional é n-ário se tem grau de aridade n. As regras de formação de termos – semelhante a de fórmulas – são:

1. As variáveis são termos;
2. As constantes são termos;
3. Se $t_1, \ldots, t_n$ são termos e F é um símbolo funcional n-ário, então $F(t_1, \ldots, t_n)$ é um termo;
4. Todos os termos têm uma das formas acima.

A composição de símbolos funcionais pode não seguir à risca essas regras. Por exemplo, quando usamos a notação (x, y) para um par ordenado, ou $\{x, y\}$ para um par não-ordenado, estamos construindo um termo a partir de um símbolo funcional

binário. Na aritmética as operações $+$ e $\cdot$ são símbolos funcionais binários, mas escrevemos $x + y$ em vez de $+(x, y)$. A união de dois conjuntos – denotada por $\cup$ – e a intersecção de dois conjuntos – denotada por $\cap$ – são exemplos de símbolos funcionais binários da linguagem estendida da teoria dos conjuntos, e seguem as mesmas regras de formação das operações aritméticas.

Os termos correspondem ao *sujeito* e *predicados* da linguagem natural. Para transformá-los em *oração* – transformando-se numa frase passível de ser julgada como verdadeira ou falsa – precisamos de um *verbo*. Os símbolos que fazem o papel dos verbos são os símbolos relacionais e a igualdade. Assim como os símbolos funcionais, eles também têm um grau de aridade, dependendo do número de parâmetros. Por exemplo, o símbolo $\in$ é um símbolo relacional binário, assim como a relação de inclusão: $\subset$.

Ao contrário dos símbolos funcionais, que podem ser compostos um sobre o outro formando expressões complexas, como $\emptyset \cup \{x, \{\emptyset\}\}$, por exemplo, os símbolos relacionais não podem conter outros símbolos relacionais no seu escopo. Por exemplo, não faz sentido escrevermos $x \subset (y \in z)$, enquanto faz sentido escrevermos $x \cup (y \cap z)$. Porém, eventualmente escrevemos coisas como $x \in y \in z$ como abreviatura de $(x \in y) \wedge (y \in z)$.

Em uma linguagem de primeira ordem com símbolos funcionais e relacionais, na definição de fórmulas adicionamos as seguintes regras, mantendo as anteriores:

1. Se t_1 e t_2 são termos, $t_1 = t_2$ é uma fórmula;
2. Se $t_1, \ldots, t_n$ são termos e R é um símbolo relacional n-ário, então $R(t_1, \ldots, t_n)$ é uma fórmula.

As fórmulas desse tipo são chamadas de *fórmulas atômicas*.

2.8 Notas sobre a semântica

Para explicar a semântica da lógica de primeira ordem – ou seja, o significado das fórmulas – precisamos antes definir (ou trabalhar informalmente com esses conceitos) relações e funções, que será feito no Capítulo 4. Sem aqui pretender entrar muito em detalhes sobre o tema – para isso, consulte um livro de lógica, como [9], ou as minhas notas de aula disponibilizadas no site – vamos apenas esboçar a ideia, assumindo que o leitor tenha familiaridade com a notação intuitiva de produto cartesiano, relações e funções. Se não, poderá voltar a esta seção depois de estudar o Capítulo 4.

Seja $\mathcal{L}$ uma linguagem de primeira ordem. Um *modelo* $\mathcal{M}$ para a linguagem $\mathcal{L}$ é uma estrutura constituída das seguintes componentes:

- Um conjunto não-vazio D, que chamaremos de *domínio*, ou *universo*, de $\mathcal{M}$;
- Para cada símbolo relacional n-ário R uma relação n-ária $R^{\mathcal{M}}$ contida em D^n;
- Para cada constante c um elemento $c^{\mathcal{M}}$ de D;
- Para cada símbolo funcional n-ário F uma função $F^{\mathcal{M}}$ de D^n em D.

Uma *valoração* é uma função σ que associa a cada variável um elemento de D.

Dados um modelo $\mathcal{M}$ e uma atribuição de variáveis σ, a *interpretação de termos* sob a atribuição σ é uma função σ^* que estende a função σ a todos os termos, conforme as seguintes condições:

- Se x é variável $\sigma^*(x) = \sigma(x)$;
- Se F é um símbolo funcional n-ário e $t_1, \ldots, t_n$ são termos, então $\sigma^*(F(t_1, \ldots, t_n)) = F^{\mathcal{M}}(\sigma^*(t_1), \ldots, \sigma^*(t_n))$.

Se $\mathcal{M}$ é um modelo, σ é uma valoração e A é uma fórmula, escrevemos $(\mathcal{M}, \sigma) \models A$ quando *A é verdadeira no modelo $\mathcal{M}$ para uma valoração σ*. A definição precisa deixaremos para o leitor consultar em um dos textos supracitados, embora não é dificil intuir como verificar se uma fórmula é verdadeira em um modelo. Veja que as variáveis não são interpretadas de maneira única no modelo. Por isso a valoração é necessária para julgarmos uma fórmula como verdadeira ou falsa em um modelo, quando essa possui variáveis livres. As sentenças – fórmulas sem variáveis livres – não dependem da valoração: ou são verdadeiras sempre, no modelo, ou são falsas sempre, independente da valoração. Quando uma fórmula é verdadeira em um modelo para toda valoração, escrevemos, simplesmente, $\mathcal{M} \models A$.

Quando $\mathcal{M} \models A$, também dizemos que o modelo $\mathcal{M}$ *satisfaz* a fórmula A.

Consequência sintática e consequência semântica: Sejam $\mathcal{L}$ uma linguagem de primeira ordem e Γ um conjunto de fórmulas de $\mathcal{L}$. Dizemos que uma fórmula A é *consequência sintática* de Γ (que denotaremos por $\Gamma \vdash A$) se existe uma demonstração de A a partir das fórmulas de Γ (consulte [9] ou as minhas notas de aula). Dizemos que uma fórmula A é *consequência semântica* de Γ (que denotaremos por $\Gamma \models A$) se todo modelo que satisfaz todas as fórmulas em Γ também satisfaz A.

Teoremas fundamentais: Os três principais teoremas metamatemáticos a respeito da lógica de primeira ordem são os teoremas da *completude*, da *compacidade* e de *Loweinhein-Skolen*.

O teorema da completude diz que $\Gamma \vdash A$ se, e somente se, $\Gamma \models A$. Ou seja, consequência sintática é equivalente a consequência semântica, provando que o sistema de axiomas que construímos é suficiente para provar tudo que podemos provar pelos argumentos usuais da linguagem cotidiana.

O teorema da compacidade diz que, se para todo Γ' subconjunto finito de Γ existe um modelo que satisfaz todas as fórmulas de Γ', então existe um modelo que satisfaz todas as fórmulas de Γ.

O teorema de Loweinhein-Skolen pode ser enunciado da seguinte maneira: se existe um modelo que satisfaça todas as fórmulas de um conjunto Γ, então, para qualquer conjunto infinito X, existe um modelo cujo domínio é X e que também satisfaz Γ. Em geral, as linguagens de lógica de primeira ordem que utilizamos têm uma quantidade enumerável de símbolos. Senão, precisamos assumir que X tem cardinalidade maior ou igual à cardinalidade do alfabeto. Uma versão do teorema diz que todo modelo possui um modelo equivalente (isto é, ambos possuem as mesmas fórmulas como verdadeiras) cujo domínio é enumerável.

Exercícios

1. Usando a linguagem de primeira ordem da teoria dos conjuntos, escreva fórmulas para representar as seguintes frases:

a) Não existe conjunto de todos os conjuntos.

b) Existe um único conjunto vazio.

c) x é um conjunto unitário.

d) Existe um conjunto que tem como elemento apenas o conjunto vazio.

e) y é o conjunto dos subconjuntos de x.

2. Marque as ocorrências livres de variáveis nas fórmulas abaixo.

a) $(\forall x(x = y)) \to (x \in y)$

b) $\forall x((x = y) \to (x \in y))$

c) $\forall x(x = x) \to (\forall y \exists z(((x = y) \wedge (y = z)) \to \neg(x \in y)))$

d) $\forall x \exists y(\neg(x = y) \wedge \forall z((z \in y) \leftrightarrow \forall w((w \in z) \to (w \in x))))$

e) $(x = y) \to \exists y(x = y)$

3. Na linguagem da aritmética dos números naturais (com os símbolos funcionais $+$ e $\times$ e as constantes 0 e 1) escreva as fórmulas de primeira ordem que correspondem às frases abaixo.

a) x é número primo.

b) x é menor do que y.

c) A soma de dois números ímpares é par.

d) A equação $x^3 + y^3 = z^3$ não tem soluções inteiras positivas.

e) Todo número par maior do que dois pode ser escrito como soma de dois números primos.

4. Julgue se cada uma das fórmulas abaixo é verdadeira em cada um dos seguintes modelos: $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$.

a) $\forall x \forall y existsz(x + y = z)$

b) $\forall x \forall y(\neg(y = 0) \to existsz(x \times y = z))$

c) $\exists x(x \times x = 1 + 1)$

Capítulo 3

Primeiros axiomas

Há três tipos de axiomas no sistema de Zermelo-Frankel. Alguns axiomas – o axioma do vazio e o axioma da infinidade – garantem a existência de conjuntos bem específicos. Outros axiomas – do par, da união, das partes, da escolha, da separação e da substituição – nos permite construir conjuntos a partir de outros. Há outros dois axiomas – da extensão e da regularidade – que nos dizem a respeito da natureza dos conjuntos, ajudando-nos a entender o seu significado.

Neste capítulo apresentamos oito dos dez axiomas usuais do sistema de Zermelo-Frankel. O motivo de deixarmos os dois outros axiomas para depois é que esses precisam de várias definições e resultados ou para defini-los (como é o caso do axioma da escolha) ou para aplicá-los (é o caso do axioma da substituição).

3.1 Axioma da extensão

O *axioma da extensão* nos dá, de certa forma, uma definição de conjuntos.

Axioma 1 *(da extensão) Dois conjuntos são iguais se, e somente se, eles têm os mesmos elementos.*

$$\forall x \forall y((x = y) \leftrightarrow (\forall z(z \in x \leftrightarrow z \in y))$$

Há essencialmente duas maneiras de representar um conjunto: descrevendo os elementos do conjunto através de uma propriedade comum a todos eles ou descrevendo cada elemento, entre chaves e separados por vírgulas. Por exemplo, numa abordagem informal, considere os seguintes "conjuntos":

> {Uruguai, Itália, Alemanha, Brasil, Inglaterra, Argentina, França, Espanha}
>
> Conjunto dos países que já venceram alguma Copa do Mundo de futebol

Ambos os conjuntos possuem os mesmos elementos. Cada elemento do primeiro conjunto também é um elemento do segundo, e vice-versa. Logo, *os dois conjuntos são iguais*, isto é, *são o mesmo conjunto.*

Considere agora o seguinte conjunto:

{Alemanha, Argentina, Brasil, Espanha, França, Inglaterra, Itália, Uruguai, Brasil}

O axioma da extensão nos garante que esse conjunto é o mesmo que o anterior. Ou seja, vale aquela máxima que aprendemos no ensino básico: *em um conjunto não importa a ordem dos elementos nem contamos as repetições.*

3.2 Axioma do vazio

Agora enunciemos um dos axiomas que garante a existência de certo tipo de conjunto.

Axioma 2 (do vazio) *Existe um conjunto vazio.*

$$\exists x \forall y \neg (y \in x)$$

Introduzimos uma outra definição: $\notin$ significa *não pertence*. Ou seja, $x \notin y$ é abreviatura para $\neg(x \in y)$. Assim, o axioma do vazio pode ser reescrito como

$$\exists x \forall y (y \notin x)$$

Na verdade, o axioma do vazio é dispensável, pois veremos que ele pode ser provado a partir do axioma da separação, *desde que assumamos que existe pelo menos um conjunto*. Assim, podemos reescrever o axioma do vazio como *existe um conjunto* [1]

O primeiro teorema que apresentamos, onde aplicamos diretamente o axioma da extensão, é a unicidade do conjunto vazio.

Teorema 3.1 *Existe um único conjunto vazio.*

Demonstração: A existência de um conjunto vazio é ditada pelo axioma do vazio. Mostremos a unicidade. Suponhamos que existem x e y conjuntos vazios diferentes. Pelo axioma da extensão, existe um elemento de x que não pertence a y ou existe um elemento de y que não pertence a x, o que, em ambos os casos, contradiz que x e y são vazios. ■

Como o conjunto vazio é único, podemos adicionar uma constante na linguagem que o represente. O símbolo adotado para o conjunto vazio é $\emptyset$.

Vamos agora definir a relação de inclusão.

Definição 3.2 Dizemos que um conjunto x *está contido* em y (e denotamos por $x \subset y$) se todo elemento de x é um elemento de z. Isto é

$$(x \subset y) \leftrightarrow \forall z((z \in x) \rightarrow (z \in y)).$$

[1] Na verdade, a formulação que aqui apresentamos da lógica de primeira ordem não permite que o domínio (vide a seção sobre semântica) seja vazio. Logo, a rigor, o axioma do vazio – ou da existência de conjuntos – é dispensável. Porém, mantemos esse axioma por motivos históricos e didáticos.

Quando $x \subset y$, também dizemos que x *é um subconjunto de* y.
Através dessa definição, o axioma da extensão pode ser escrito como

$$(x = y) \leftrightarrow ((x \subset y) \wedge (y \subset x))$$

O mesmo argumento do teorema anterior prova o seguinte:

Teorema 3.3 *O conjunto vazio está contido em qualquer conjunto.*

$$\forall x(\emptyset \subset x)$$

Demonstração: Suponha que existe um x tal que $\emptyset$ não está contido em x. Logo, existe y pertencente a $\emptyset$ que não pertence a x, contradizendo a definição do conjunto vazio. ■

Esse tipo de argumento – conhecido como *argumento da vacuidade* – é bastante estranho na linguagem cotidiana mas muito comum na matemática e na lógica. Se eu disser "*toda vez que eu fui para Marte encontrei homenzinhos verdes*" essa frase é logicamente correta, visto que eu nunca fui para Marte.

Uma coisa que precisamos falar sobre as relações de pertinência e inclusão é sobre um mito clássico divulgado nas escolas de ensino fundamental e médio. Dizem alguns que *pertence se usa entre elemento e conjunto* e *contido se usa entre conjuntos*. Esse é um erro grave, que pode causar um vício de aprendizagem que precisa ser derrubado.

O primeiro erro dessa frase é ignorar que existem *conjuntos de conjuntos*. Isto é, como, a princípio (na teoria ingênua dos conjuntos), podemos formar conjuntos de *qualquer tipo de objeto*, nada impede que os próprios elementos dos conjuntos sejam conjuntos. Por exemplo, $\{\emptyset\}$ é um conjunto (assumindo que existe, pois ainda não provamos isso), cujo único elemento é o conjunto vazio. Isto é, podemos afirmar tranquilamente que $\emptyset \in \{\emptyset\}$, não obstante também valha (como acabamos de mostrar) $\emptyset \subset \{\emptyset\}$. Aliás, lembremos que, na teoria dos conjuntos, *tudo é conjunto*, e, portanto, todos elementos de um conjunto são também conjuntos. Por isso definimos o símbolo $\in$ como uma relação entre conjuntos, contrariando totalmente o mito divulgado nas escolas.

O segundo crime desse mito é desvirtuar a real compreensão dos dois símbolos. A frase – destinada a ajudar os alunos a acertarem algumas questões mesmo sem compreendê-las de fato – insinua que a pertinência e a inclusão são ambos símbolos primitivos com significados parecidos, como se essa "regrinha" fosse a única maneira de diferenciarmos os dois símbolos. Ora, na verdade, o símbolo primitivo é o de pertinência, e a inclusão foi definida logicamente a partir desse. Um conjunto x está contido em y se *cada* elemento de x também é um elemento de y.

Para ajudar o leitor a superar eventuais vícios de aprendizagem, preparamos uma lista de exercícios sobre o uso correto desses símbolos. Os exercícios de cada capítulo são poucos e não muito difíceis. É altamente recomendável que o estudante os faça.

A última observação deixada neste capítulo é a seguinte: até agora, tendo em mãos apenas os axiomas da extensão e do vazio, não podemos garantir a existência de qualquer outro conjunto senão o conjunto vazio. Os outros axiomas serão necessários para construirmos toda a teoria dos conjuntos a partir do "nada".

3.3 Axioma do par

O primeiro dos axiomas que usamos para construir outros conjuntos a partir do conjunto vazio é o axioma do par.

Axioma 3 (do par) *Para todos conjuntos x e y existe um conjunto cujos elementos são x e y.*

$$\forall x \forall y \exists z \forall w((w \in z) \leftrightarrow ((w = x) \vee (w = y))$$

Mantendo a notação do capítulo anterior, o axioma do par nos diz que para todos x e y existe o conjunto $\{x, y\}$. Note que trata-se de um par *não-ordenado*, isto é, conforme o axioma da extensão, a ordem dos elementos não importa. O conjunto $\{x, y\}$ é o mesmo que o conjunto $\{y, x\}$. Note também que x e y *não precisam ser distintos*. No caso de x ser igual a y, o conjunto $\{x, y\}$ é igual a $\{x\}$ (aplicando-se o axioma da extensão).

Por exemplo, tomando ambos x e y iguais a $\emptyset$, o axioma do par nos garante a existência do conjunto $\{\emptyset, \emptyset\}$, que, pelo axioma da extensão, é igual ao conjunto $\{\emptyset\}$.

Já vimos no capítulo anterior que $\emptyset$ e $\{\emptyset\}$ são conjuntos diferentes, e, agora, conseguimos provar a existência desse segundo conjunto. Ou seja, até agora já provamos que existem pelo menos dois conjuntos distintos!

Aplicações sucessivas do axioma do par nos fornece uma infinidade de conjuntos (finitos). Podemos provar a existência de $\{\emptyset, \{\emptyset\}\}$, $\{\{\emptyset\}\}$, $\{\{\{\emptyset\}\}\}$ e assim por diante.

3.4 Axioma da união

Enunciemos o próximo axioma, que também pertence ao grupo de *axiomas de construção*.

Axioma 4 (da união) *Para todo conjunto x existe o conjunto de todos os conjuntos que pertencem a algum elemento de x.*

$$\forall x \exists y \forall u((w \in y) \leftrightarrow \exists v((w \in v) \wedge (v \in x)))$$

Repare que o axioma da união não garante, a princípio, a união de dois conjuntos, mas, sim, a *união de uma famíla de conjuntos*. Se pensarmos em um conjunto de conjuntos como uma caixa cheia de pacotes menores, a união desse conjunto de conjuntos corresponde a despejarmos todo o conteúdo dos pacotes menores na caixa maior. Vejamos, como exemplo (assumindo que existe – visto que ainda nem explicamos o que são os números naturais), o seguinte conjunto:

$$\{\{1, 2\}, \{1, 3\}, \{4\}\}$$

A união do conjunto acima é o conjunto formado por todos os números que pertencem a pelo menos um de seus elementos, a saber:

$$\{1, 2, 3, 4\}$$

Em outras palavras, a união de x é *o conjunto dos elementos dos elementos de* x.

Denotamos a união de um conjunto x por $\bigcup x$. Convém ressaltar – e isso vale também para o axioma do par – que o axioma da extensão garante que a união é única. Isto é, dado qualquer conjunto x, não existem dois conjuntos diferentes que, no lugar de y, tornariam a sentença correspondente ao axioma da união verdadeira. O axioma da união determina unicamente um conjunto a partir de x. O artigo definido *o* que consta no enunciado do axioma, em linguagem natural, está bem colocado.

Deixamos ao leitor a tarefa de mostrar as seguintes igualdades:

$\bigcup \emptyset = \emptyset$

$\bigcup \{\emptyset\} = \emptyset$

$\bigcup \{\emptyset, \{\emptyset\}\} = \{\emptyset\}$

$\bigcup \{\{\emptyset\}\} = \{\emptyset\}$

Com o axioma do par e o axioma da união em mãos podemos definir a união de dois conjuntos.

Teorema 3.4 *Dados dois conjuntos x e y existe o conjunto formado por todos os conjuntos que pertencem a x ou a y.*

$$\forall x \forall y \exists z \forall w((w \in z) \leftrightarrow ((w \in x) \vee (w \in y)))$$

Demonstração: Dados dois conjuntos x e y, aplicamos o axioma do para para obtermos o conjunto $\{x, y\}$. Aplicando o axioma da união sobre o conjunto $\{x, y\}$ obtemos o conjunto $z = \bigcup \{x, y\}$. Observe, pela definição da união de uma famíla de conjuntos, que, para todo w, $w \in z$ se, e somente se, existe $u \in \{x, y\}$ tal que $w \in u$. Mas, se $u \in \{x, y\}$, temos que $u = w$ ou $u = y$, provando que z satisfaz o enunciado do teorema. ■

Novamente notamos que a união de dois conjuntos é única, pelo axioma da extensão, o que nos permite introduzir a seguinte definição.

Definição 3.5 *Definimos a* união *de x e y como o conjunto formado por todos os conjuntos que pertencem a x ou a y, e denotaremos esse conjunto por $x \cup y$.*

3.5 Axioma das partes

O próximo axioma traz novamente à tona a ideia de conjunto de conjuntos, e a necessidade de não confundirmos os símbolos de pertinência e inclusão.

Axioma 5 (das partes) *Para todo conjunto x existe o conjunto dos subconjuntos de x.*

$$\forall x \exists y \forall z((z \in y) \leftrightarrow (z \subset x))$$

Se quisermos transformar a fórmula acima sem usar o símbolo de inclusão, basta escolhermos uma variável nova que não consta na fórmula (w, por exemplo) e substituirmos $z \subset x$ pela fórmula $\forall w((w \in z) \to (w \in x))$. É importante que o leitor esteja familiarizado com essas abreviaturas e com o processo de converter essas abreviaturas pela fórmula completa.

O conjunto definido pelo axioma das partes também é único. Introduzimos a seguinte definição:

Definição 3.6 *Definimos o* conjunto das partes *de x como o conjunto dos subconjuntos de x, e denotaremos por $\mathcal{P}(x)$.*

3.6 Axioma da separação

O axioma da separação de certa forma resgata a concepção inicial de Frege de definir um conjunto através de uma fórmula lógica que descreve seus elementos. Mas, para evitar o paradoxo de Russell, na formulação do axioma da separação é necessário estabelecer um conjunto do qual iremos "separar" os elementos que satisfazem uma determinada propriedade.

Assim, para cada fórmula $P(x)$, temos que, para todo conjunto y, existe o conjunto formado por todos $x \in y$ tais que $P(x)$ é verdadeiro.

Formalmente, o axioma da separação é um esquema de axiomas, isto é, uma lista infinita de axiomas, conforme abaixo:

Axioma 6 (Esquema de axiomas da separação) *Para cada fórmula P em que z não ocorre livre a seguinte fórmula é um axioma:*

$$\forall y \exists z \forall x((x \in z) \leftrightarrow ((x \in y) \wedge P))$$

O conjunto z, como no axioma, será denotado por

$$\{x \in y : P(x)\}$$

Notemos que a única restrição sobre a fórmula P é não conter z como variável livre. Essa restrição é necessária porque utilizamos essa variável no axioma para definir o conjunto $\{x \in y : P(x)\}$. Se permitirmos que a mesma variável que define o conjunto dado pelo axioma da separação também ocorra livre em P, poderíamos tomar P como a fórmula $x \notin z$ e teríamos a seguinte instância do axioma da separação:

$$\forall y \exists z \forall x((x \in z) \leftrightarrow ((x \in y) \wedge (x \notin z)))$$

Se tomássemos, por exemplo, $y = \{\emptyset\}$ e $x = \emptyset$, teríamos $x \in y$ verdadeiro e, portanto, teríamos

$$(x \in z) \leftrightarrow (x \notin z)$$

o que é uma contradição.

Não precisamos impor qualquer outra restrição sobre as variáveis livres em P. Em todas as aplicações do axioma da separação, a variável x ocorre livre em P (por

isso utilizamos a notação $P(x)$ para a fórmula P). Mas se x não ocorrer livre em P, isso não causará inconsistência no sistema. Apenas a aplicação do axioma da separação seria trivial, pois o conjunto z seria vazio ou o próprio y (já que a validade de P, nesse caso, não depende da variável x, que não ocorre livre em P).

Podemos ter outras variáveis livres em P além de x. Isso ocorre, por exemplo, na definição de intersecção de conjuntos:

$$a \cap b = \{x \in a : x \in b\}$$

A própria variável y (que reservamos para o – digamos – "conjunto universo") pode ocorrer livre em P, como na seguinte definição:

$$\{x \in y : x \subset y\}$$

Com essa formulação do sistema de Zermelo-Fraenkel o Paradoxo de Russell deixa de ser um paradoxo que leva a uma contradição no sistema e passa a ser um teorema que afirma não haver conjunto de todos os conjuntos.

Teorema 3.7 (Paradoxo de Russell) *Não existe conjunto de todos os conjuntos.*

$$\forall x \exists y (y \notin x)$$

Demonstração: Suponha que exista um conjunto y tal que, para todo x, $x \in y$. Pelo axioma da separação para a fórmula $x \notin x$, existe z tal que, para todo x,

$$(x \in z) \leftrightarrow ((x \in y) \wedge (x \notin x))$$

Como $x \in y$ é verdadeiro para todo x temos que

$$(x \in z) \leftrightarrow (x \notin x))$$

Tomando z no lugar de x temos

$$(z \in z) \leftrightarrow (z \notin z),$$

chegando a uma contradição. ■

O axioma do vazio segue como consequência do axioma da separação, pois, pelos axiomas lógicos podemos provar a sentença $\exists y(y = y)$ (lembre-se que, na definição de semântica da lógica de primeira ordem, no Capítulo 2, exigimos que o domínio de um modelo é não-vazio, o que significa que *existe um conjunto*). Usemos o axioma da separação para esse y e para a fórmula $x \neq x$. Obtemos o conjunto

$$\{x \in y : x \neq x\},$$

que é o conjunto vazio.

A partir do axioma da separação podemos definir as operações conjuntísticas. Comecemos pela intersecção de uma família de conjuntos.

Teorema 3.8 (Intersecção de uma família de conjuntos) *Dado um conjunto não vazio x existe o conjunto formado por todos os conjuntos que pertencem simultaneamente a todos os elementos de x.*

$$\forall x(\exists y(y \in x) \rightarrow \exists y(\forall z((z \in y) \leftrightarrow \forall w((w \in x) \rightarrow (z \in w))))).$$

Denotaremos esse conjunto por $\bigcap x$.

Demonstração: Seja z um elemento de x. Defina o conjunto y como

$$\{v \in z : \forall w((w \in x) \to (v \in w))\}.$$

O axioma da separação garante a existência do conjunto y. Agora verifiquemos que y satisfaz as condições do teorema. Seja $v \in y$. Pela definição de y, para todo $w \in x$ temos $v \in w$. Reciprocamente, se para todo $w \in x$ temos $v \in w$, então, em particular, $v \in z$ e, portanto, $v \in y$. Isso prova que, para todo v, $v \in y$ se, e somente se, $v \in w$, para todo $w \in x$.

■

É bom notar que, diferente da união de uma família de conjuntos, na intersecção precisamos impor a restrição de que a família é não-vazia. A união de uma família vazia é o conjunto vazio. Mas se fizéssemos a intesecção de uma família vazia obteríamos o "conjunto de todos os conjuntos", já que todo conjunto x satisfaz, por vacuidade, a condição "para todo y pertencente ao conjunto vazio $x \in y$".

Agora aplicaremos o axioma da separação para definir diversas operações (símbolos funcionais) binárias entre conjuntos. A saber, são elas:

Intersecção: $x \cap y = \{z \in x : z \in y\}$

Subtração: $x \setminus y = \{z \in x : z \notin y\}$

Diferença simétrica: $x \Delta y = \{z \in x \cup y : z \notin x \cap y\}$

Quando $x \cap y = \emptyset$, dizemos que x e y são *disjuntos*.

A intersecção de conjuntos está relacionado ao operador booleano **e**, pois pertencer a $x \cap y$ significa pertencer a x **e** a y. A união significa **ou**, pois pertencer a $x \cup y$ seginifica pertencer a x **ou** pertencer a x. A diferença simétrica é ou **ou exclusivo** (pertencer a x ou a y, mas não a ambos). A união de uma família de conjuntos está relacionada ao quantificador existencial, pois pertencer a $\bigcup x$ significa pertencer a *algum* elemento de x, enquanto a intersecção de uma família de conjutos representa o quantificador universal, porque pertencer a $\bigcap x$ significa pertencer a *todos* os elementos de x.

Classes de conjuntos. Para contornar a falta do "conjunto de todos os conjuntos" e de outros conjuntos "grandes demais para existirem" (exemplos: o conjunto de todos os conjuntos finitos, o conjunto de todos os conjuntos ao qual o vazio pertence), adotamos a ideia intuitiva de classes. Em outras formalizações da Teoria dos Conjuntos, como as teorias de NGB (Neumann-Gödel-Bernays) e KM (Kelley-Morse), os objetos matemáticos são divididos entre *conjuntos* e *classes*, sendo que todos os conjuntos são classes mas nem todas as classes são conjuntos. As classes que não são conjuntos são chamadas de *classes próprias*. Nessas teorias existe, por exemplo, *a classe de todos os conjuntos*, bem como outras classes derivadas a partir dessa usando o axioma da separação. Claro que não existe *a classe de todas as classes*, pois isso geraria novamente o paradoxo de Russell.

Contudo, mesmo em ZFC não existindo as classes próprias que existem em outras teorias, há uma maneira informal – mas não imprecisa – de falar de classes, de modo

que podemos transcrever qualquer fórmula ou demonstração de uma das teorias NGB ou KM para uma fórmula ou demonstração em ZFC que possui o mesmo significado. Para isso, pensamos nas classes simplesmente como *fórmulas da linguagem de ZFC, escolhendo uma variável livre*. Para explicar o que significa "escolher uma variável livre", considere P a fórmula $x \subset y$. A partir daí derivamos duas classes diferentes, escolhendo uma das variáveis para designar os elementos das classes e a outra como parâmetro: a classe $\{x : x \subset y\}$ (essa classe coincide com o conjunto $\mathcal{P}(y)$) e a classe $\{y : x \subset y\}$ (essa é uma *classe própria*, pois não define um conjunto). As variáveis livres que servem de parâmetro permitem – assim como acontece com o axioma da separação – que definamos uma classe a partir de um ou mais conjuntos fixados. O segundo exemplo ilustra bem o que está sendo dito: para cada conjunto x definimos "a classe de todos os conjuntos que contêm o conjunto x".

O leitor poderá se deparar com livros e artigos que usem a notação de classes, mas poderá facilmente adaptar a escrita para a linguagem de ZFC aqui apresentada. Por exemplo, é comum usar a notação Ord para representar "a classe de todos os ordinais". Nesse caso, sempre que aparecer a fórmula $x \in Ord$ basta pensarmos nela como uma abreviatura da fórmula "x é um ordinal", que será descrita no Capítulo 7.

3.7 Axioma da regularidade

Até agora, todos os axiomas que vimos garantem a construção de alguns conjuntos partindo apenas do conjunto vazio. O próximo axioma garante que *todos* os conjuntos são construídos a partir do vazio. Também irá evitar coisas como $x \in x$ e será útil em teoria dos modelos para fazermos indução sobre a relação de pertinência.

Axioma 7 (da regularidade) *Para todo conjunto x não-vazio existe $y \in x$ tal que $x \cap y = \emptyset$.*

$$\forall x(x \neq \emptyset \rightarrow \exists y(y \in x \wedge x \cap y = \emptyset))$$

O axioma da regularidade parece um pouco artificial e contraintuitivo, mas ele pode ser resumido como: *todo conjunto não-vazio possui um elemento $\in$-minimal*. Isto é, todo conjunto não vazio possui um elemento tal que nenhum outro elemento do conjunto pertence a ele. Perceba a semelhança com a definição de *boa ordem*, que será dada no Capítulo 4.

Teorema 3.9 *Não existem x e y tais que $x \in y$ e $y \in x$.*

Demonstração: Sejam x e y conjuntos quaisquer. Vamos provar que $x \notin y$ ou $y \notin x$.

Usando o axioma do par, tome $z = \{x, y\}$. Como z não é vazio, pelo axioma da regularidade existe $w \in z$ tal que $w \cap z = \emptyset$. Se $w = x$, isso implica que $y \notin x$. Se $w = y$, isso implica que $x \notin y$, provando o teorema. ■

Corolário 3.10 *Não existe x tal que $x \in x$.*

Demonstração: Aplique o teorema anterior para $x = y$. ■

O axioma da regularidade garante que não existe uma sequência infinita decrescente na relação de pertinência. Ou seja, não existe uma sequência da forma $\ldots x_3 \in x_2 \in x_1 \in x_0$. Para se ter uma ideia da demonstração desse fato, supondo, por absurdo, que existem tais conjuntos, considere $x = \{x_0, x_1, x_2, \ldots\}$. Para qualquer $x_n \in x$ temos $x_{n+1} \in x_n \cap x$, contradizendo o axioma da regularidade.

O problema dessa demonstração está na formalização da ideia de sequência, já que ainda não definimos o conjunto dos números naturais nem funções. No momento oportuno apresentaremos esse resultado como exercício.

Concluímos desse resultado que, para qualquer conjunto x, se tomarmos um elemento de x, e um elemento de um elemento de x, e um elemento de um elemento de um elemento de x, assim sucessivamente, chegaremos, após uma quantidade finita de passos, no conjunto vazio.

É bom notar que se, por um lado, não existe uma sequência infinita decrescente, na relação de pertinência, por outro lado, como veremos no próximo capítulo, é possível existir uma sequência infinita crescente. Ou seja, sequências infinitas da forma $x_0 \in x_1 \in x_2 \ldots$ existem (os números naturais, por exemplo).

3.8 Axioma da infinidade

O axioma da infinidade é, ao lado do vazio, um axioma que garante a existência de um conjunto específico. No caso, de um conjunto infinito. Há várias formas de apresentar o axioma da infinidade. Uma delas enuncia a existência do conjunto dos números naturais, conforme a concepção de von Neuman. Outra forma, utilizada aqui, é enunciar a existência de um conjunto do qual deduzimos a existência (e definimos) do conjunto dos números naturais. A terceira simplesmente enuncia a existência de um conjunto infinito (embora ainda não tenhamos definido o que é um conjunto infinito), e a construção do conjunto dos números naturais torna-se um pouco mais complicada e utiliza o axioma da substituição, semelhante ao que será feito na construção dos ordinais.

Na definição dos números naturais atribuída a von Neumann, pensamos em um número natural como o *conjunto dos números naturais menores que ele*. Assim, o 0 é o conjunto dos números naturais menores que 0. Como não existe número natural menor que 0, então 0 será representado pelo conjunto vazio. O número 1 é o conjunto formado pelos números menores que 1. Ou seja, 1 é o conjunto $\{0\}$, que é igual a $\{\emptyset\}$. O número 2 é o conjunto $\{0, 1\}$, ou seja, o conjunto $\{\emptyset, \{\emptyset\}\}$, e assim por diante.

Note que o número 3, que é o conjunto $\{0, 1, 2\}$, pode ser escrito como $\{0, 1\} \cup \{2\}$, assim como $1 = \emptyset \cup \{0\}$ e $2 = \{0\} \cup \{1\}$. Ou seja, o sucessor de um número natural n é o resultado de acrescentarmos o próprio n ao conjunto n. Isto é, $n + 1 = n \cup \{n\}$. Isso justifica a seguinte definição de sucessor:

Definição 3.11 Dado um conjunto x, definimos x^+ como $x \cup \{x\}$. Isto é,

$$\forall y(y \in x^+ \leftrightarrow (y \in x \vee y = x))$$

Quando um conjunto possui o vazio como elemento, e é fechado pela operação de sucessor, então dizemos que tal conjunto é *indutivo*, conforme segue a definição.

Definição 3.12 Dizemos que um conjunto x é *indutivo* se, e somente se, $\emptyset \in x$ e, para todo y, se $y \in x$ então $y^+ \in x$.

O axioma da infinidade garante a existência de algum conjunto indutivo.

Axioma 8 (da infinidade) *Existe um conjunto indutivo.*

$$\exists x(\emptyset \in x \forall y(y \in x \rightarrow y^+ \in x))$$

Note que um conjunto indutivo precisa possuir o vazio e todos os sucessores obtidos a partir do vazio. Ou seja, um conjunto indutivo precisa conter o conjunto dos números naturais (conforme será provado no teorema 3.14, parte (b)), mas pode ter elementos a mais. Usando o teorema 3.8 e os axiomas da separação, das partes e da infinidade, definimos o conjunto dos números naturais da seguinte forma:

Definição 3.13 Definimos o *conjunto dos números naturais* – que será denotado por ω – como o seguinte conjunto:

$$\omega = \bigcap \{x \in \mathcal{P}(I) : x \text{ é indutivo}\},$$

onde I é o conjunto indutivo determinado pelo axioma da infinidade.

Notemos que a intersecção é permitida porque a família de subconjunts de I que são indutivos não é vazia, dado que pelo menos o próprio conjunto I é indutivo. Agora, resta-nos mostrar que o próprio conjunto ω é indutivo, e que segue da definição que ele é o menor conjunto indutivo que existe. Fica como exercício provar – a partir do teorema seguinte – que a definição de ω não depende da escolha de I.

Teorema 3.14 (a) *ω é um conjunto indutivo.*

(b) *Se A é um conjunto indutivo então $\omega \subset A$.*

Demonstração: Seja I o conjunto indutivo dado pelo axioma da infinidade. Vamos provar que ω é indutivo. Primeiro, provemos que $\emptyset \in \omega$. De fato, se A é um subconjunto de I que é indutivo, então $\emptyset \in A$. Logo $\emptyset$ pertence à intersecção de todos os subconjuntos indutivos de I. Agora, suponha que $x \in \omega$. Isso significa que $x \in A$, para todo A suconjunto indutivo de I. Mas isso implica que $x^+ \in A$, para todo $A \subset I$ indutivo. Logo, $x^+ \in A$, provando a parte (a) do teorema.

Agora provemos a parte (b). Seja A um conjunto indutivo. Repetindo o argumento do parágrafo anterior, concluímos que $A \cap I$ é indutivo. Como $A \cap I \subset I$, temos, pela definição de ω, que todo elemento de ω também pertence a $A \cap I$. Ou seja, $\omega \subset A \cap I$ e, portanto, $\omega \subset A$. ■

Observe que segue da demonstração do Teorema 3.14, que a definição de ω independe da escolha do conjunto indutivo I. Para verificarmos isso, tome J qualquer

outro conjunto indutivo e defina $A = \bigcap\{x \in \mathcal{P}(J) : J \text{ é indutivo}\}$. A demonstração do Teorema 3.14 pode ser aplicada para A, no lugar de ω, e concluímos que A é indutivo e, pelo item (b) (aplicado duas vezes) temos que $A \subset \omega$ e $\omega \subset A$. Do axioma da extensão segue, portanto, que $A = \omega$.

Agora verificaremos por que convém chamarmos ω de "conjunto dos números naturais". Primeiro, vamos enunciar os axiomas de Peano, sobre números naturais. Adotamos como conceitos primitivos *zero* e *sucessor de*. São esses os axiomas:

1. Zero é um número natural.
2. O sucessor de um número natural é um número natural.
3. Números naturais distintos nunca têm o mesmo sucessor.
4. Zero não é sucessor de qualquer número natural.
5. Se uma propriedade vale para zero e, valendo para um dado número natural, também vale para o seu sucessor imediato, então essa propriedade para todos os números naturais.

O quinto axioma de Peano é o que conhecemos como *princípio da indução finita*.

Uma formalização precisa dos axiomas de Peano, usando lógica de primeira ordem, é a seguinte: introduzimos 0 (zero) como uma constante e s (sucessor de) como um símbolo funcional unário da linguagem. O primeiro e o segundo axioma tornam-se desnecessários. O terceiro e o quarto axioma são respectivamente $\forall x \forall y(\neg(x = y) \to \neg(s(x) = s(y)))$ e $\forall x(\neg(s(x) = 0))$. O quinto axioma torna-se um esquema de axiomas, em que, para cada fórmula P, a fórmula

$$(P_x^0 \wedge \forall x(P \to P_x^{s(x)}) \to \forall x P$$

é um axioma.

O próximo teorema diz que o conjunto ω serve como domínio de um modelo para os axiomas de Peano, interpretando 0 como $\emptyset$ e $s(n)$ como n^+.

Teorema 3.15 *O conjunto ω satisfaz os axiomas de Peano, identificando "zero" com o conjunto vazio e o sucessor de x com x^+.*

Demonstração: Os dois primeiros axiomas seguem do fato de ω ser indutivo. Para provarmos o terceiro axioma, suponhamos, por absurdo, que $x \neq y$ e $x^+ = y^+$. Temos que $x \in x^+$, logo, pela hipótese, $x \in y^+$. Como $y^+ = y \cup y$, e $x \neq y$, então $x \in y$. Analogamente provamos que $y \in x$, contradizendo o axioma da regularidade [2](Teorema 3.9).

O quarto axioma segue do fato que $x \in x^+$, logo, não podemos ter, para nenhum x, $x^+ = \emptyset$.

[2]Na verdade, o uso do axioma da regularidade facilita a demonstração, mas não é necessário, uma vez que o conjunto ω e seus elementos satisfazem o axioma da regularidade, sem precisarmos assumi-lo. Voltaremos nesse assunto quando falarmos sobre números ordinais.

Para provarmos o princípio da indução finita, seja P uma fórmula tal que $P_x^{\emptyset}$ e $\forall x(P \to P_x^{x^+})$ são verdadeiros. Usando o axioma da separação, considere A o conjunto $\{x \in \omega : P\}$. Pela hipótese sobre P é fácil verificar que A é indutivo. Logo, pelo Teorema 3.14, parte (b), temos que $\omega \subset A$, provando que todo elemento de ω satisfaz P. ■

Exercícios

1. Usando o axioma da extensão, prove que $\emptyset$ e $\{\emptyset\}$ são conjuntos diferentes.

2. Para cada par de conjuntos abaixo, decida qual(is) dos símbolos $\in$ e $\subset$ tornam a fórmula verdadeira (assumindo que esses conjuntos existen). Lembre-se de que a resposta também pode ser ambos os símbolos ou nenhum deles. Justifique cada resposta.

(a) $\{\emptyset\} \ldots \{\emptyset, \{\emptyset\}\}$

(b) $\{\emptyset\} \ldots \{\{\emptyset\}\}$

(c) $\{1, 2, 3\} \ldots \{\{1\}, \{2\}, \{3\}\}$

(d) $\{1, 2, 3\} \ldots \{\{1\}, \{1, 2\}, \{1, 2, 3\}\}$

(e) $\{1, 2\} \ldots \{1, \{1\}, 2, \{2\}, \{3\}\}$

(f) $\{\{1\}, \{2\}\} \ldots \{\{1, 2\}\}$

3. Seja x o conjunto $\{\emptyset, \{\emptyset\}, \emptyset, \{\emptyset, \{\emptyset\}\}\}$

(a) Quantos elementos tem o conjunto x?

(b) Descreva todos os subconjuntos de x.

(c) Descreva, utilizando chaves e vírgula, o conjunto de todos os subconjuntos de x.

(d) Quantos elementos o conjunto dos subconjuntos de x possui?

(e) Prove que o conjunto x existe.

4. Prove que, para todos conjuntos x, y

(a) $x \subset x$;

(b) $x \in y$ se, e somente se, $\{x\} \subset y$;

(c) $\bigcup \mathcal{P}(x) = x$;

(d) se $x \subset y$, então $\bigcup x \subset \bigcup y$.

5. Escreva uma fórmula de primeira ordem, na linguagem da teoria dos conjuntos, com quatro variáveis livres, que represente o conjunto $\{x, y, z\}$.

6. Escreva os seguintes conjuntos, listando seus elementos entre chaves.

(a) $\bigcup\{\{0,1\},\{\{1\}\},\{1,2\},\{\{1,2\}\}\}$;

(b) $\mathcal{P}(\{\emptyset,\{\emptyset\}\})$.

7. Prove que não existe o conjunto de todos os conjuntos unitários.

Dica: Assuma, por absurdo, a existência do conjunto de todos os conjuntos unitários e prove a existência do conjunto de todos os conjuntos.

8. Prove que, para todo conjunto X existe o conjunto

$$\{\{x\} : x \in X\}$$

9. Sendo x um conjunto não vazio, prove que

(a) $\forall y(y \in x \to (\bigcap x \subset y))$;

(b) $x \subset y \to \bigcap y \subset \bigcap x$.

10. Escreva na linguagem da lógica de primeira ordem, sem abreviaturas, a seguinte fórmula:

$$x \in \bigcup\bigcap(y \cup (w \setminus z))$$

11. Usando o axioma da regularidade, prove que:

(a) não existem x, y, z tais que $x \in y$, $y \in z$ e $z \in x$;

(b) não existem w, x, y, z tais que $w \in x$, $x \in y$, $y \in z$ e $z \in w$.

12. Dizemos que um conjunto x é *transitivo* se todo elemento de x está contido em x. Prove que:

(a) para todo conjunto transitivo x, se $x \neq \emptyset$ então $\emptyset \in x$.;

(b) ω é transitivo.

13. Prove que não existe x tal que $\mathcal{P}(x) = x$.

14. Prove que existe um modelo para teoria dos conjuntos em que valem os axiomas do par, da união e das partes, mas não valem os axiomas do vazio e da regularidade.

Dica: Considere um modelo formado por um único elemento x tal que $x \in x$.

15. Prove, a partir dos axiomas de Peano, os seguintes teoremas:

(a) Todo número natural é diferente do seu sucessor.

(b) Zero é o único número natural que não é sucessor de algum número natural.

16. Prove que:

(a) para todo $n \in \omega$, $\emptyset \in n$ ou $\emptyset = n$;

(b) para todos $n, m \in \omega$, se $m \in n$ então $m \subset n$.

17. A união de dois conjuntos indutivos é necessariamente um conjunto indutivo? Justifique sua resposta.

Capítulo 4

Produto cartesiano, relações e funções

As definições neste e no próximo capítulo são fundamentais para enunciarmos o axioma da escolha. Começamos definindo pares ordenados, produto cartesiano e relações.

4.1 Pares ordenados

O axioma do par nos garante construirmos, a partir de dois conjuntos a e b, o par $\{a, b\}$. Porém, nessa definição de par a ordem dos elementos não importa, de modo que $\{a, b\} = \{b, a\}$. Na definição de par ordenado, a igualdade só vale quando a ordem é a mesma.

Definição 4.1 Dados dois conjuntos a e b, definimos o *par ordenado* (a, b) como o conjunto $\{\{a\}, \{a, b\}\}$. Ou seja,

$$\forall x(x \in (a, b) \leftrightarrow \forall y((y \in x \leftrightarrow y = a) \vee (y \in x \leftrightarrow (y = a \vee y = b))))$$

É fácil verificar que o par ordenado entre quaisquer conjuntos existe (aplicando três vezes o axioma do par: uma para formar o conjuntoi $\{a\}$, outra para o conjunto $\{a, b\}$ e outra para o conjunto $\{\{a\}, \{a, b\}\}$) e é único (aplicação padrão do axioma da extensão).

Assim, podemos introduzir a notação (a, b) como mais um símbolo funcional binário na nossa linguagem estendida da teoria dos conjuntos (ou mais uma abreviatura).

Notemos que, quando $a = b$, o par ordenado (a, b) é igual ao conjunto $\{\{a\}\}$.

Teorema 4.2 *Dois pares ordenados (a, b) e (c, d) são iguais se, e somente se, $a = c$ e $b = d$.*

Demonstração: Um dos lados da equivalência é trivial: se $a = c$ e $b = d$ então os pares ordenados (a, b) e (c, d) são iguais. Mostraremos o outro lado.

Suponha que $(a, b) = (c, d)$. Como $\{a\} \in (a, b)$ temos que $\{a\} \in (c, d)$. Logo $\{a\} = \{c\}$ ou $\{a\} = \{c, d\}$. Em ambos os casos temos que $a = c$.

Para provarmos que $b = d$, separemos em dois casos. No primeiro caso, supomos que $a = b$, o que implica que $(a, b) = \{\{b\}\}$. Teremos que $\{c, d\} \in (a, b)$ e, portanto, $\{c, d\} = \{b\}$, provando que $b = d$. No segundo caso, supomos que $a \neq b$. Como $\{a, b\} \in (c, d)$ temos $\{a, b\} = \{c\}$ ou $\{a, b\} = \{c, d\}$. Como $\{c\} \subset \{c, d\}$, em ambos os casos o axioma da extensão garante que $b \in \{c, d\}$. Não podemos ter $b = c$, pois provamos que $a = c$ e assumimos que $a \neq b$. Portanto, $b = d$.

■

4.2 Produto cartesiano

O próximo teorema nos garante a existência do produto cartesiano entre dois conjuntos.

Teorema 4.3 *Dados dois conjuntos A e B, existe o conjunto de todos os pares ordenados (a, b) que satisfazem $a \in A$ e $b \in B$.*

Demonstração: Usando os axiomas do par, da união, das partes e da separação, definimos o conjunto

$$X = \{x \in \mathcal{P}(\mathcal{P}(A \cup B)) : \exists a \exists b(a \in A \wedge b \in B \wedge x = (a, b))\}$$

Para verificarmos que X atende as condições do teorema, só resta verificarmos que todo par ordenado (a, b), onde $a \in A$ e $b \in B$, pertence a $\mathcal{P}(\mathcal{P}(A \cup B))$.

De fato, $\{\{a\}, \{a, b\}\} \in \mathcal{P}(\mathcal{P}(A \cup B))$ é equivalente a $\{\{a\}, \{a, b\}\} \subset \mathcal{P}(A \cup B)$, que ocorre se, e somente se, $\{a\} \in \mathcal{P}(A \cup B)$ e $\{a, b\} \in \mathcal{P}(A \cup B)$, o que é verdade, pois $\{a\} \subset A \cup B$ e $\{a, b\} \subset A \cup B$. ■

O conjunto estabelecido pelo Teorema 4.3 é chamado de *produto cartesiano* de A e B, e será denotado por $A \times B$. Introduzimos essa notação como outra abreviatura, desempenhando o papel de um símbolo funcional binário.

Quando A e B são iguais, utilizamos uma notação especial: denotamos o produto cartesiano $A \times A$ por A^2.

4.3 Relações

Uma relação é um subconjunto de um produto cartesiano.

Definição 4.4 Dizemos que R é uma *relação* (ou *relação binária*) entre A e B se é um subconjunto de $A \times B$. Quando R é uma relação, utilizamos a notação xRy como abreviatura de $(x, y) \in R$.

Nas próximas seções estudaremos algumas propriedades das relações, bem como algumas relações especiais, como as *funções*, as *relações de equivalência* e as *relações de ordem*.

4.4 Funções

Uma função de A em B é uma relação que associa a cada elemento de A um único elemento de B. Posto isso formalmente temos a seguinte definição:

Definição 4.5 Dizemos que uma relação F entre A e B é uma *função de A em B* se para todo $x \in A$ existe um único $y \in B$ tal que $(x, y) \in F$. Isto é, F é uma função de A em B se a seguinte fórmula é verdadeira:

$$(F \subset A \times B) \wedge \forall x(x \in A \rightarrow \exists y((x, y) \in F)) \wedge \forall x \forall y \forall z(((x, y) \in F \wedge (x, z) \in F) \rightarrow (y = z))$$

. Notemos que a fórmula dada é uma conjunção de três subfórmulas. A primeira diz que uma função de A em B é uma relação entre A e B. Ou seja, para todo par ordenado $(x, y) \in f$ temos $x \in A$ e $y \in B$. A segunda subfórmula diz que todo elemento de A é contemplada pela função F (quando não exigimos essa condição, dizemos que f é uma *função parcial* de A em B). Finalmente, a terceira subfórmula nos diz que a função só relaciona um elmento de B, para cada elemento de A.

Denotamos por ${}^{A}B$ o conjunto das funções de A em B. Deixamos como exercício ao leitor provar a existência de ${}^{A}B$, pois é uma simples aplicação do axioma da separação. Introduzimos, assim, mais um símbolo funcional binário na linguagem [1].

Mantendo a tradição, usaremos preferencialmente letras minúsculas para denotar funções.

Se f é uma função de A em B, dizemos que A é o *domínio de f* – que será denotado por $dom(f)$ – e o conjunto $\{b \in B : \exists a((a, b) \in f)\}$ é chamado de *imagem de f* – que será denotada por $im(f)$.

Normalmente se utiliza o termo *contradomínio* de uma função para designar o conjunto B, quando a função é de A em B. Todavia, esse termo não é muito adequado na definição aqui adotada de função, já que, dada uma função f, não é possível "recuperar" o contradomínio. Por exemplo, se tomarmos o conjunto (supondo que já temos construídos os números reais) $\{(x, y) \in \mathbb{R}^2 : y = x^2\}$, esse pode tanto ser visto como uma função de $\mathbb{R}$ em $\mathbb{R}$ quanto uma função de $\mathbb{R}$ em $\mathbb{R}_+$ (os reais não-negativos).

Por outro lado, essa ambiguidade não existe ao definirmos o domínio e a imagem a partir da função. É possível "recuperar" o domínio e a imagem de uma função. Abaixo seguem as definições do domínio e imagem a partir da função, e a tarefa de mostrar que essas definições cumprem o prometido é deixada ao leitor:

$$dom(f) = \{a \in \bigcup\bigcup f : \exists b((a, b) \in f)\}$$

$$im(f) = \{b \in \bigcup\bigcup f : \exists a((a, b) \in f)\}$$

Nessas definições é bom notar em como os axiomas do par e das partes "empacotam" os conjuntos, enquanto o axioma da união "desempacota".

Também notamos que as mesmas definições podem ser aplicadas para relações binárias quaisquer.

[1] A rigor, não há símbolo algum na notação ${}^{A}B$. Mas mantemos a nomenclatura por uma questão de conveniência, para se adequar à definição de linguagem de primeira ordem.

Como uma função associa a cada elemento do domínio um único elemento da imagem, podemos introduzir a seguinte notação: se (x, y) pertence a uma função f, denotamos y por $f(x)$. Essa notação só é possível, pois, para $x \in dom(f)$, existe um único y satisfazendo $(x, y) \in f$. Porém, precisamos ser mais cautelosos com essa notação do que somos com outras como a do par $(\{a, b\})$, da união de dois conjuntos $(a \cup b)$ e do par ordenado. Isso porque, enquanto as outras notações valem para quaisquer termos, $f(x)$ *só está bem definido quando* f *é uma função e* x *pertence ao domínio de* f. Logo, não podemos desavisadamente introduzir essa notação como um símbolo funcional binário da linguagem, pois $f(x)$ não está definido para quaisquer conjuntos f e x.

Outra notação que podemos introduzir, que aparece na linguagem cotidiana da matemática, é $f : A \longrightarrow B$ para designar que f *é uma função de* A *em* B, ou, em outras palavras (ou melhor, símbolos), $f \in^A B$.

Suponha que f é uma função de A em B e que C é um subconjunto de A. Definimos

$$f|C = (C \times B) \cap f$$

a *restrição de* f *ao conjunto* C. Fica como exercício ao leitor mostrar que $f|C$ é uma função de C em B.

Dizemos que uma função $f : A \longrightarrow B$ é *injetora* se, para todo $x, y \in A$ temos que, se $x \neq y$, então $f(x) \neq f(y)$. Ou seja, quando dois elementos distintos do domínio nunca são mapeados para o mesmo elemento da imagem. Dizemos que f é *sobrejetora em relação a* B se para todo $y \in B$ existe $x \in A$ tal que $f(x) = y$. Ou seja, quando B é a imagem de f. A necessidade de relativizarmos a B a definição de sobrejetora vem daquele problema anteriormente mencionado, sobre a impossibilidade de "recuperarmos" o contra-domínio de uma função. Quando está claro no contexto qual contradomínio está sendo considerado (quando, por exemplo, escrevemos que "f é uma função de A em B") dizemos apenas que a função é sobrejetora, mas é necessária uma cautela extra para esse tipo de nomenclatura.

Uma função $f : A \longrightarrow B$ é bijetora (ou *bijetora* em relação a B) quando é injetora e sobrejetora (em relação a B). Nesse caso também dizemos que A é uma *bijeção* entre A e B. No capítulo sobre conjuntos equipotentes discutiremos melhor a propriedade de existir uma bijeção entre dois conjuntos (lembram-se da introdução, sobre como comparar tamanhos de conjuntos infinitos?)

Ainda há algumas definições a serem introduzidas, com as quais o estudante de matemática deve estar bem acostumado. Se f e g são funções, e $im(g) \subset dom(f)$, então definimos a *função composta* de f e g da seguinte forma:

$$f \circ g = \{(x, z) \in dom(g) \times im(f) : \exists y((x, y) \in g \wedge (y, z) \in f\}$$

Novamente, é preciso tomar cuidado com essa notação, pois ela só faz sentido dentro das hipóteses estritas apresentadas acima.

4.5 Relações de ordem

Definição 4.6 Uma relação $\leq \subset X \times X$ é chamada de *ordem* em X se satisfaz as seguintes propriedades, para todos $x, y, z \in X$:

- Reflexividade: $x \leq x$;
- Transitividade: se $x \leq y$ e $y \leq z$ então $x \leq z$.
- Antissimetria: se $x \leq y$ e $y \leq x$ então $x = y$;

Chamamos de *conjunto ordenado* um par $(X, \leq)$, onde $\leq$ é uma ordem em X, e dizemos que X é o *domínio* da ordem $\leq$.

Uma relação de ordem também é chamada de *ordem parcial*, para diferenciar da ordem total, que veremos daqui a pouco.

Um exemplo de ordem em um conjunto X é a relação de inclusão. Isto é, o conjunto $\{(x, y) \in X \times X : x \subset y\}$. De fato, todo conjunto está contido nele mesmo, se x está contido em y e y está contido em z então x está contido em z, e o axioma da extensão nos garante que $x = y$ toda vez que x está contido em y e y está contido em x. Por abuso de notação, usaremos, eventualmente, o símbolo $\subset$ para designar a relação de inclusão, como conjunto de pares ordenados.

Veremos que toda relação de ordem pode ser vista como uma relação de inclusão. Para explicar o que isso significa, introduzimos a seguinte definição:

Definição 4.7 Sejam $\leq_1$ e $\leq_2$ duas ordens em X_1 e X_2, respectivamente. Dizemos que $\leq_1$ e $\leq_2$ são *ordens isomorfas* (ou que *os conjuntos ordenados* $(X_1, \leq_1)$ *e* $(X_2, \leq_2)$ *são isomorfos*) se existe uma função $f : X_1 \longrightarrow X_2$ bijetora em X_2 tal que $x \leq_1 y$ se, e somente se, $f(x) \leq_2 f(y)$.

Nesse caso, dizemos que a função f é um *isomorfismo de ordens*.

O próximo resultado diz que toda ordem é isomorfa à relação de inclusão sobre algum conjunto.

Teorema 4.8 *Seja* $(X, \leq)$ *um conjunto ordenado. Existe um conjunto ordenado* $(Y, \preceq)$ *isomorfo a* $(X, \leq)$ *tal que*

$$\preceq = \{(x, y) \in Y \times Y : x \subset y\}$$

Demonstração: Defina $f : X \longrightarrow \mathcal{P}(X)$ como

$$f(x) = \{y \in X : y \leq x\}$$

Tome Y a imagem de f. Mostraremos que f é injetora, o que basta para provarmos que é bijetora em Y.

Suponha que $f(x) = f(y)$. Pela reflexividade, como $x \leq x$ e $y \leq y$, temos $x \in f(x)$ e $y \in f(y)$. Como $f(x)$ e $f(y)$ são iguais, temos $x \in f(y)$ e $y \in f(x)$. Pela definição de f isso nos dá $x \leq y$ e $y \leq x$, que, pela antissimetria, implica que $x = y$, provando que f é bijetora em Y.

Agora resta-nos mostrar que $x \leq y$ se, e somente se, $f(x) \subset f(y)$. Suponha que $x \leq y$. Seja $z \in f(x)$. Temos que $z \leq x$ e, por transitividade, $z \leq y$. Logo, $z \in f(y)$. Reciprocamente, suponha que $f(x) \subset f(y)$. Como $x \in f(x)$, temos $x \in f(y)$, o que significa que $x \leq y$. ■

Listamos agora uma série de definições que serão usadas durante o livro.

Definição 4.9 Seja $\leq$ uma relação de ordem em um conjunto X. Para todo $x \in X$ e todo $S \subset X$ não-vazio dizemos que

- x é *limitante superior* de S se $y \leq x$, para todo $y \in S$;
- x é *limitante inferior* de S se $x \leq y$, para todo $y \in S$;
- S é *limitado superiormente* se possui um limitante superior;
- S é *limitado inferiormente* se possui um limitante inferior;
- x é *máximo* de S se $x \in S$ e $y \leq x$, para todo $y \in S$;
- x é *mínimo* de S se $x \in S$ e $x \leq y$, para todo $y \in S$;
- x é *maximal* se não existe $y \in X$ tal que $x \neq y$ e $x < y$;
- x é *minimal* se não existe $y \in X$ tal que $x \neq z$ e $y < x$;
- x é *supremo* de S se x é o mínimo dos limitantes superior de S;
- x é *ínfimo* de S se x é o máximo dos limitantes inferior de S;
- S é uma *cadeia* se, para todos $y, z \in S$ temos $y \leq z$ ou $z \leq y$.

Essas definições dependem da ordem. Portanto, quando não estiver claro no contexto qual é a ordem que estamos considerando sobre o conjunto X, escrevemos $\leq$-*máximo*, $\leq$-*maximal* etc.

Notemos – pela definição e pela antissimetria da relação de ordem – que nem sempre um conjunto possui um elemento máximo, mas, se possuir, esse é único. O mesmo vale para mínimo, supremo e ínfimo. Porém, podemos ter vários limitantes superiores e inferiores de um conjunto e elementos maximais e minimais da ordem.

Agora podemos enunciar os principais tipos de ordem usados na matemática:

Definição 4.10 Dizemos que uma ordem $\leq$ sobre um conjunto X é uma(um):

- *ordem total* (ou *ordem linear*) se, para todos $x, y \in X$ temos $x \leq y$ ou $y \leq x$;
- *boa ordem* se todo subconjunto não-vazio de X possui elemento mínimo;
- *árvore* se, para todo $x \in X$, o conjunto $\{y \in X : y \leq x\}$ é uma cadeia em X;
- *reticulado* se, para todos $x, y \in X$, o conjunto $\{x, y\}$ possui supremo e ínfimo.

Aplicamos os termos acima também para o conjunto ordenado $(X, \leq)$ e, por abuso de notação, para o domínio X.

Uma ordem total tem esse nome porque todos os elementos do domínio podem ser comparados. Também a chamamos de ordem linear porque podemos visualizar todos os elementos da ordem como se estivessem numa mesma reta. As ordens usuais nos números naturais, inteiros, racionais e reais são exemplos de ordens totais.

Nota-se que toda boa ordem também é uma ordem total, uma vez que o conjunto $\{x, y\}$ tem mínimo, o que nos dá $x \leq y$ ou $y \leq x$.

Uma árvore é uma ordem que pode "bifurcar", mas nunca "juntar", como na copa de uma árvore, em que o tronco se ramifica em galhos, que se ramificam em galhos menores, mas os galhos nunca se reajuntam. Além das numerosas aplicações em teoria dos conjuntos, as árvores são usadas em computação e em teoria dos jogos. Por exemplo, as possíveis sequências de jogadas a partir de uma posição numa partida de xadrez formam uma árvore, que um programa de computador (ou o cérebro humano, de uma maneira mais intuitiva) analisará para poder decidir o melhor lance.

Uma ordem total é uma árvore, já que todo o conjunto é uma cadeia e, portanto, todos seus subconjuntos são cadeias.

Se considerarmos a ordem da inclusão em uma família de conjuntos fechada pelas operações de união e intersecção, essa ordem será um reticulado, onde o ínfimo de $\{x, y\}$ é $x \cap y$, e o supremo é $x \cup y$. Esse tipo de ordem é particularmente interessante nos estudos de álgebras de Boole. O reticulado é um pouco mais geral, pois temos as operações de supremo e ínfimo (que correspondem às operações booleanas "e" e "ou") mas não precisamos do complemento (correspondente à operação booleana "não").

Também é evidente que toda ordem total é um reticulado, já que o próprio x e o próprio y serão um deles o ínfimo e o outro o supremo do conjunto $\{x, y\}$.

Neste livro, nosso foco será no conceito de boa ordem, pela sua importância no estudo de ordinais e cardinais. Um boa ordem $(X, \leq)$ também é chamada de *conjunto bem-ordenado*.

O próximo teorema é bastante importante para falarmos sobre ordem em ω e discutir o conceito de infinitude.

Teorema 4.11 *Em ω, vale as seguintes afirmações:*

(a) *Para todos $n, m \in \omega$, temos $n \in m$ ou $m \in n$ ou $n = m$;*

(b) *Para todos $n, m \in \omega$ temos $n \subset m$ se, e somente se, $n \in m$ ou $n = m$;*

(c) $\bigcup \omega = \omega$;

(d) $(\omega, \subset)$ *é bem-ordenado.*

Demonstração: A parte (a) utiliza uma indução dupla que precisa ser analisada com cuidado. Considere a fórmula de duas variáveis livres

$$P(n, m) \equiv ((n \in \omega) \wedge (m \in \omega)) \rightarrow ((n \in m) \vee (n = m) \vee (m \in n))$$

Para provar a parte (a) do teorema mostraremos, por indução em n, a fórmula

$$\forall n \forall m P(n, m) \tag{1}$$

O passo inicial da indução é mostrar que

$$\forall m P(0, m) \tag{2}$$

e o passo indutivo é

$$\forall m P(n, m) \to \forall m P(n^+, m) \tag{3}$$

No entanto, cada uma das fórmulas acima será provada por indução em m. O passo inicial de (2) é a fórmula $P(0, 0)$, que é obviamente verdadeira, pois $0 = 0$. O passo indutivo é mostrar que

$$\forall m(P(0, m) \to P(0, m^+)) \tag{4}$$

De fato, assumindo $P(0, m)$ temos $0 \in m$, $0 = m$ ou $m \in 0$. O terceiro caso não é possível, visto que 0 é o conjunto vazio. Se $m = 0$ temos que $m^+ = \{0\}$ e, portanto, $0 \in m^+$. Se $0 \in m$ temos $0 \in m^+$, pois $m \subset m \cup \{m\}$, concluindo a prova de (2).

Para provar (3), antes provaremos, por indução em n, que, para todo $n \in \omega$ vale

$$\forall m(m \in n \to (m^+ \in n \vee m^+ = n)) \tag{5}$$

De fato, se $n = 0$ a afirmação é trivial, já que a premissa $m \in 0$ é falsa. Suponha que (5) vale para n e provaremos para n^+. Suponha que $m \in n^+$. Isso significa que $m \in n$ ou $m = n$. O primeiro caso, pela hipótese de indução, implica que $m^+ \in n$ ou $m^+ = n$, o que garante que $m^+ \in n^+$. No segundo caso temos $m^+ = n^+$, provando (5).

Voltemos à demonstração de (3). Fixe $n \in \omega$ e suponha valer $P(n, m)$, para todo m. Se $m \in n$, temos $m \in n^+$. Se $m = n$ temos $m \in n^+$. Se $n \in m$, por (5) concluímos que $n^+ \in m$ ou $n^+ = m$, provando (3) e concluindo o item (a) do teorema.

Para o item (b), começaremos provando, por indução em n, que para todo número natural n

$$\forall m(m \in n \to m \subset n) \tag{6}$$

De fato, (6) vale trivialmente para 0. Supondo que (6) vale para n, seja $m \in n^+$. Nesse caso temos $m \in n$ ou $m = n$, o que, em ambos os casos, implica que $m \subset n$.

Agora precisamos mostrar a recíproca. Sejam $m, n \in \omega$ tais que $m \subset n$ e $m \neq n$. Mostraremos que $m \in n$. Suponha que seja falso. Pelo item (a) isso implica que $n \in m$ e, por (6), concluímos que $n \subset m$, contradizendo que $m \subset n$ e $m \neq n$. Provamos assim o item (b).

A parte (c) é basicamente a transitividade de ω. Suponha que $x \in \bigcup \omega$. Ou seja, $x \in n$, para algum $n \in \omega$. Mostraremos que $x \in \omega$. Para isso, provaremos por indução em n que

$$n \in \omega \to n \subset \omega \tag{7}$$

Para $n = 0$ a afirmação é trivial. Supondo que vale para $n \in \omega$ e $n \subset n$, como $n^+ = n \cup \{n\}$ temos que $n^+ \subset \omega$, provando (7).

Reciprocamente, seja $n \in \omega$. Temos $n \in n^+$ e $n^+ \in \omega$, provando que $n \in \bigcup \omega$ e concluindo a prova do item (c).

Provemos agora a parte (d) do teorema. Primeiro provaremos, por indução em n, que todo natural n é bem-ordenado com a ordem da inclusão. O passo inicial $n = 0$ é trivial, já que 0 não contém subconjunto não-vazio. Supondo que n é bem-ordenado, considere S um subconjunto não-vazio de n^+. Seja $S' = S \setminus \{n\}$. Observe que $S' \subset n$. Se $S' = \emptyset$, então $S = \{n\}$, que possui n como elemento mínimo. Se $S' \neq \emptyset$, pela hipótese indutiva existe m que é o mínimo de S'. Como $m \in S'$, temos que $m \in n$. Logo, pelo item (b), $m \subset n$, provando que m é o mínimo também de S.

Seja agora $S \subset \omega$ não-vazio e fixe $n_0 \in S$. Em particular $S \cap n_0^+ \neq \emptyset$ e, portanto, sendo $S \cap n_0^+$ um subconjunto de n_0^+, possui um elemento mínimo. Seja m o mínimo de $S \cap n^0$ e provemos que m é o mínimo de S. Seja $n \in S$. Pelo item (a), temos $n \in n_0$, $n = n_0$ ou $n_0 \in n$. Nos dois primeiros casos, de $n \in S \cap n_0$ segue que $m \subset n$, pois m é o mínimo de $S \cap n_0$. No terceiro caso, como $m \in n_0$ e $n_0 \in n$, do item (b) segue que $m \subset n_0$ e $n_0 \subset n$, de onde concluímos que $m \subset n$, provando que $(\omega, \subset)$ é bem-ordenado.

■

4.6 Relação de equivalência

Para construirmos o conjunto dos números inteiros a partir do conjunto dos números naturais, e o conjunto dos números racionais a partir do conjunto dos número inteiros, precisamos, antes, desenvolver o conceito de relação de equivalência.

Definição 4.12 Dizemos que uma relação $R \subset X \times X$ é uma *relação de equivalência em* X se satisfaz as seguintes propriedades, para todos $x, y, z \in X$:

- Reflexividade: xRx;
- Simetria: se xRy então yRx;
- Transitividade: se xRy e yRz então xRz.

Definimos o *conjunto das classes de equivalência de* R como

$$X/R = \{Y \in \mathcal{P}X : \exists x \forall y (y \in Y \leftrightarrow xRy)\}$$

Os elementos de X/R são, obviamente, chamados de *classes de equivalência*, também denotado do seguinte modo:

$$X/R = \{[x] : x \in X\}$$

onde

$$[x] = \{y \in X : xRy\}$$

Teorema 4.13 *Seja R uma relação de equivalência em um conjunto X. As seguintes afirmações são verdadeiras:*

(a) $\bigcup X/R = X$;

(b) $\emptyset \notin X/R$;

(c) *Para todos* $Y, Z \in X/R$, *se* $Y \neq Z$ *então* $Y \cap Z = \emptyset$;

(d) *Se* $x \in Y$ *e todo* $Y \in X/R$, *para todo* $y \in X$ *temos que* xRy *se, e somente se,* $y \in Y$.

Demonstração: Usaremos a notação $[x]$ para o conjunto $\{y \in X : xRy\}$.

Dado $x \in X$, temos que $x \in [x]$, uma vez que, pela propriedade reflexiva, xRx. Isso prova (a). Como todo elemento de X/R é da forma $[x]$, para algum $x \in X$, isso prova também a parte (b)

Para provar (c), assumindo que Y e Z são dois elementos de X/R que não são disjuntos, mostraremos que $Y = Z$. Sejam $x \in Y \cap Z$ e $y_0, z_0 \in X$ tais que $Y = [y_0]$ e $Z = [z_0]$. Dado $y \in Y$, temos, por definição, que y_0Ry. Logo, pela simetria, yRy_0. Mas como $x \in Y$, temos y_0Rx. Pela transitividade temos yRx. Mas, como $x \in Z$, temos z_0Rx e, pela simetria, xRz_0. Logo, a transitividade nos dá yRz_0 e, novamente pela simetria, z_0Ry, o que prova que $y \in Z$. Isso conclui que $Y \subset Z$ e um argumento análogo mostra que $Z \subset Y$, provando que $Y = Z$.

Mostremos a parte (d). Se $Y \in X/R$, existe $y_0 \in X$ tal que $Y = [y_0]$. Como $x \in Y$, temos que y_0Rx e, portanto, xRy_0. Se yRx, por transitividade e simetria temos yRy_0 e y_0Ry, de onde temos que $y \in Y$. Por outro lado, se $y \in Y$, temos y_0Ry e, portanto, xRy, concluindo a prova do teorema. ■

Em outras palavras, o Teorema 4.13 parte (d) nos diz que duas classes de equivalência $[x]$ e $[y]$ são iguais se, e somente se, xRy.

4.7 Teorema da recursão

Diversas construções e teoremas conjuntísticos usam funções especiais de domínio ω que, para serem definidas, precisam do teorema da recursão, enunciado a seguir:

Teorema 4.14 (da recursão) *Sejam* X *um conjunto,* x *um elemento de* X *e* g *uma função de* X *em* X. *Então existe uma única função* f *de* ω *em* ω *tal que*

- $f(0) = x$;
- $f(n^+) = g(f(n))$, *para todo* $n \in \omega$.

Demonstração: Usando o axioma da separação, defina o conjunto

$$\mathcal{C} = \{R \in \mathcal{P}(\omega \times X) : (0, x) \in R \wedge \forall n \forall y((n, y) \in R \rightarrow (n^+, g(y))) \in R\}.$$

Claramente $\omega \times X \in \mathcal{C}$. Logo, $\mathcal{C}$ é não-vazio. Podemos, portanto, definir o conjunto

$$f = \bigcap \mathcal{C}$$

Precisamos provar que f é uma função e que satisfaz a condição para pertencer a $\mathcal{C}$.

Afirmação 1: $f \in \mathcal{C}$

O procedimento da demonstração da afirmação 1 é análogo à demonstração que ω é um conjunto indutivo. Como $(0, x) \in R$, para todo $R \in \mathcal{C}$, então $(0, x) \in f$. Se $(n, y) \in f$, então $(n, y) \in R$, para todo $R \in \mathcal{C}$. Logo, pela hipótese sobre os elementos de $\mathcal{C}$, $(n^+, g(y)) \in R$, para todo $R \in \mathcal{C}$. Logo, $(n^+, g(y)) \in f$, concluindo a prova da afirmação.

Afirmação 2: f é uma função de domínio ω

Vamos provar, por indução, que para todo $n \in \omega$ vale a fórmula $P(n)$, definida abaixo:

$$P(n) \equiv \exists y((n, y) \in f) \wedge \forall y \forall z(((n, y) \in R \wedge (n, z) \in R) \rightarrow (y = z))$$

Vamos provar $P(0)$. Pela afirmação 1, $(0, x) \in f$. Vamos provar que, se $(0, y) \in f$, então $y = x$. Suponha, por absurdo, que existe $y \neq x$ tal que $(0, y) \in f$. Considere $R = f \setminus \{(0, y)\}$. Vamos verificar que $R \in \mathcal{C}$. De fato, $(0, x) \in R$, pois $(0, x) \in f$ e $x \neq y$. Se $(n, y) \in R$, então $(n, y) \in f$, pois $R \subset f$. Logo, $(n^+, g(y)) \in f$ (pela afirmação 1). Como $n^+ \neq 0$ (axioma 4 de Peano), temos que $(n^+, g(y)) \in f$ é diferente de $(0, y)$ e, portanto, pertence a R.

Portanto, concluímos que $R \in \mathcal{C}$, o que implica que $f \subset R$. Como $R \subset f$, temos $f = R$, absurdo, pois $(0, y) \in f$ e $(0, y) \notin R$.

Vamos agora provar que $P(n)$ implica $P(n^+)$.

Assumindo $P(n)$ como verdadeiro, temos que existe y tal que $(n, y) \in f$. Logo, como $f \in \mathcal{C}$, temos que $(n^+, g(y) \in f$, provando a "primeira parte" de $P(n^+)$.

Agora supomos, por absurdo, que existe $z \neq g(y)$ tal que $(n^+, z) \in f$. Defina $R = f \setminus \{(n^+, z)\}$. Vamos verificar que $R \in \mathcal{C}$,

Como $n^+ \neq 0$, continuamos tendo $(0, x) \in R$. Suponha que $(m, v) \in R$. Como $f \in \mathcal{C}$ e $R \subset f$ temos que $(m^+, g(v)) \in R$. Se $m \neq n$, o axioma 3 de Peano nos garante que $m^+ \neq n^+$, logo, $(m^+, g(v)) \neq (n^+, z)$, provando que $(m^+, g(v)) \in R$. Se $m = n$, pela hipótese indutiva $P(n)$ temos que $v = y$ (pois $(n, y) \in f$), e já vimos que $(n^+, g(y) \in f$. Como $z \neq g(y)$, também temos que $(n^+, g(y) \in R$. Provamos, com isso, que $R \in \mathcal{C}$ o que novamente contradiz com o fato de R estar contido propriamente em f. Isso conclui a demonstração da afirmação 2.

Das afirmações 1 e 2 segue imediatamente o teorema. Sendo f uma função de domínio ω e satisfazendo as condições da família de conjuntos $\mathcal{C}$, temos que $(0, x) \in f$, o que significa que $f(0) = x$. Como, para todo n, temos, pela própria definição de função, $(n, f(n)) \in f$, da afirmação 1 segue que $(n^+, g(f(n)) \in f$, o que significa que $f(n^+) = g(f(n))$.

A unicidade da função f pode ser provada por indução. Suponha que existe h satisfazendo as mesmas condições do teorema estabelecidas para f. Temos que $f(0) = h(0)$, pois ambos são iguais a x. Se $f(n) = h(n)$, então $g(f(n)) = g(h(n))$, e ambos são iguais a $f(n^+)$ e $h(n^+)$. Logo, por indução, $f = h$.

■

4.8 Aritmética dos números naturais

Já definimos ω como o conjunto dos números naturais, e mostramos que ele satisfaz os axiomas de Peano. Falta definir a aritmética, ou seja, as duas funções de $\omega \times \omega$ em ω que correspondem às operações de soma e produto.

A ideia geral da definição da soma é utilizar o teorema da recursão para definir, para cada número natural m, uma função $s_m : \omega \longrightarrow \omega$ tal que

$$s_m(0) = m$$

$$s_m(n^+) = (s_m(n))^+$$

e definimos $m + n$ como $s_m(n)$. Utilizando novamente o teorema da recursão e a definição das funções acima podemos definir, para cada número natural m, uma função $p_m : \omega \longrightarrow \omega$ tal que

$$p_m(0) = 0$$

$$p_m(n^+) = (p_m(n))^+$$

e definimos $m \cdot n$ como $p_m(n)$.

Essa definição de soma e produto ainda precisa ser melhor justificada, para podemos construí-la axiomaticamente. Façamos isso.

Teorema 4.15 *Existe uma função s de ω em ${}^{\omega}\omega$ tal que, para todo $n, m \in \omega$, $s(m)(0) = m$ e $s(m)(n^+) = (s(m)(n))^+$.*

Demonstração: Usando o axioma da separação defina

$$s = \{(m, f) \in \omega \times^{\omega} \omega : \forall n((f(0) = m) \wedge (f(n^+) = (f(n))^+))\}$$

Pelo teorema da recursão, utilizando-o para a função $g = \{(n, n^+) : n \in \omega\}$, para cada m existe uma única f satisfazendo as condições descritas na definição de s. Logo, s é uma função.

■

Definição 4.16 *Definimos a* operação de soma em ω *como a função* $+ : \omega \times \omega \longrightarrow \omega$ *dada por* $+((m, n)) = s(m)(n)$. *Denotamos* $+((m, n))$ *por* $m + n$.

Teorema 4.17 *Existe uma função p de ω em ${}^{\omega}\omega$ tal que, para todo $n, m \in \omega$, $p(m)(0) = 0$ e $p(m)(n^+) = p(m)(n) + m$.*

Demonstração: Usando o axioma da separação defina

$$p = \{(m, f) \in \omega \times^{\omega} \omega : \forall n((f(0) = 0) \wedge (f(n^+) = (f(n) + m)))\}$$

Tomando a função $g = \{(i, j) \in \omega\omega : i + m = j\}$, o teorema da recursão garante que p é uma função. ■

Definição 4.18 *Definimos a* operação de soma em ω *como a função* $\cdot : \omega \times \omega \longrightarrow \omega$ *dada por* $\cdot((m, n)) = p(m)(n)$. *Denotamos* $\cdot((m, n))$ *por* $m \cdot n$.

Da definição de soma e produto seguem os seguintes axiomas da aritmética de Peano, quando adicionamos os símbolos funcionais binários $+$ e $\cdot$ à linguagem da aritmética:

$$m + 0 = m$$
$$m + n^+ = (m + n)^+$$
$$m \cdot 0 = 0$$
$$m \cdot n^+ = (m \cdot n) + n$$

Eventualmente usaremos a notação xy para representar $x \cdot y$.

Exercícios

1. Encontre uma definição alternativa para par ordenado de modo que o Teorema 4.2 continue valendo. Justifique.

2. Prove que $A \times B = \emptyset$ se, e somente se $A = \emptyset$ ou $B = \emptyset$.

3. Prove que, se $A \subset C$ e $B \subset D$, então $A \times B \subset C \times D$.

4. Vale a recíproca do exercício 3? Justifique.

5. Descreva todos os elementos de $\mathcal{P}(2 \times 2)$.

6. Escreva uma fórmula de primeira ordem, de três variáveis livres, sem abreviaturas da linguagem de teoria dos conjuntos, que significa "x é uma função de y em z".

7. Sendo f e g funções tais que a imagem de g é igual ao domínio de f, prove que $f \circ g$ é injetora se, e somente se, f e g são injetoras. Dê um contraexemplo que prova que isso não acontece se assumirmos apenas que a imagem de g está contida no domínio de f.

8. Em quais condições temos ${}^AB \subset {}^CD$? Justifique.

9. Dada uma relação R, definimos a *inversa* de R – que será denotada por R^{-1} – como o conjunto $\{(y, x) : (x, y) \in R\}$. Com base nisso, prove as seguintes asserções:

(a) Para toda relação R existe R^{-1}.

(b) Se f é uma função, f^{-1} é uma função se, e somente se, f é injetora.

(c) Se f e g são funções injetoras tais que $im(g) \subset dom(f)$, então $(f \circ g)^{-1} = g^{-1} \circ f^{-1}$.

10. Prove que existe uma função injetora de ω em ω que não é sobrejetora (em relação a ω).

11. Usando o axioma da regularidade, prove que, para todo conjunto x, não existe uma função f de ω em x tal que $f(n^+) \in f(n)$, para todo $n \in \omega$.

12. Seja X um conjunto e sejam x_0 e y_0 dois elementos distintos de X. Considere a seguinte relação em X:

$$R = \{(x, y) \in X \times X : x = y\} \cup \{(x_0, y_0), (y_0, x_0)\}$$

(a) Prove que R é uma relação de equivalência em X.

(b) Descreva os elementos de X/R.

13. Considere C um conjunto não-vazio de conjuntos não-vazios tal que, para todos x e y pertencentes a C, se $x \neq y$ então $x \cap y = \emptyset$. Seja $X = \bigcup C$. Defina em X a relação:

$$R = \{(x, y) \in X^2 : \exists z(z \in C \wedge x \in z \wedge y \in z)\}$$

(a) Prove que R é uma relação de equivalência em X.

(b) Mostre que $C = X/R$.

(c) Prove que duas relações de equivalência diferentes possuem classes de equivalências diferentes.

14. Como fica uma relação de equivalência sobre $\emptyset$? Ela satisfaz o Teorema 4.13?

15. Considere X o conjunto das funções f tais que $dom(f) \in \omega$ e $im(f) \subset \omega$. Prove que $(X, \subset)$ é uma árvore.

16. Dê exemplos ou prove que não existe:

(a) Uma ordem total que não é uma boa ordem;

(b) Uma árvore que não é uma ordem total;

(c) Um reticulado que não é árvore;

(d) Uma árvore que é um reticulado mas não é totalmente ordenado.

17. Prove o *princípio da indução transfinita*. Ou seja, suponha que $(X, \leq)$ é um conjunto bem-ordenado e $P(x)$ é uma propriedade de primeira ordem tal que:

- para todo $x \in X$, se vale $P(y)$ para todo $y \in X$ tal que $y \neq x$ e $y \leq x$, então vale $P(x)$.

Prove que $P(x)$ é verdadeiro para todo $x \in X$.

Capítulo 5

Axioma da escolha e suas aplicações

Neste capítulo apresentamos o Axioma da Escolha, suas equivalências e consequências, servindo de base para discutirmos ordinais e cardinalidade de conjuntos.

5.1 Axioma da escolha

O axioma da escolha enuncia que, dada uma família de conjuntos não-vazios, existe uma função que a cada conjunto pertencente a essa família seleciona um elemento desse conjunto.

Axioma 9 (da escolha) *Para todo conjunto x de conjuntos não-vazios existe uma função $f : x \longrightarrow \bigcup x$ tal que, para todo $y \in x$, $f(y) \in y$.*

$$\forall x(\emptyset \notin x \rightarrow \exists f((f \text{ é função}) \wedge (dom(f) = x) \wedge \forall y(y \in x \rightarrow f(y) \in y)))$$

A função f garantida pelo axioma da escolha é chamada de *função de escolha*.

Esse é certamente o axioma mais controverso da teoria dos conjuntos, rejeitado por algumas correntes filosóficas da matemática, como os construtivistas. Por isso alguns matemáticos preferem tomar um cuidado especial quando utilizam o axioma da escolha, evitando-o a todo custo, referindo-se por ZF ao sistema de axiomas de Zermelo e Fraenkel *sem* o axioma da escolha e por ZFC ao sistema ZF *com* o axioma da escolha (a letra C vem de *choice*, da sigla em inglês).

Para entendermos melhor por que esse axioma é tão controverso, precisamos entender para quais construções ele é necessário. Ou seja, precisamos entender para quais conjuntos x a existência da função de escolha depende do axioma da escolha e para quais podemos prová-la em ZF.

Primeiro notemos que, se x for finito (isto é, se existe uma função bijetora entre x e um número natural) então a existência de uma função de escolha é garantida pelos outros axiomas. Por exemplo: se x é o conjunto $\{a, b, c\}$, sendo seus três elementos não-vazios, sabemos que existem $a_0 \in a$, $b_0 \in b$ e $c_0 \in c$. Usando sucessivas vezes o axioma do par, da união, das partes e da separação (como fizemos quando mostramos

a existência de pares ordenados) construímos o conjunto $\{(a, a_0), (b, b_0), (c, c_o)\}$, que é precisamente uma função de escolha no conjunto x.

Formalizando o argumento geral, temos o seguinte: provaremos por indução em n que, dados x um conjunto de conjuntos não-vazios, n um número natural e s uma função bijetora de n em x, existe uma função de escolha em x. Se $n = 0$, x necessariamente será o conjunto vazio e, portanto, o conjunto vazio é uma função de escolha em x (verifiquem que, de acordo com a definição dada neste livro, $\emptyset$ é uma função de $\emptyset$ em $\emptyset$). Suponha que a hipótese de indução é verdadeira para algum natural n, e provaremos para n^+. Sejam x um conjunto de conjuntos não-vazios e s uma função bijetora de n^+ em x. Como $n^+ = n \cup \{n\}$, defina t a restrição de s a n, isto é, $t = \{(m, s(m) : m \in n\}$. Defina $y = im(t)$. Claramente t é uma bijeção de n em y. Logo, pela hipótese indutiva, existe $g : y \longrightarrow \bigcup y$ tal que $g(z) \in z$, para todo $z \in y$. Como $s(n) \neq \emptyset$, pois $s(n) \in x$, existe $a \in s(n)$. Defina $f = g \cup \{(s(n), a)\}$. Como $x = y \cup \{s(n)\}$ é fácil verificar que f é uma função de escolha em x.

Ou seja, se substituirmos "para todo x" pela expressão "para todo x finito" no enunciado do axioma da escolha, teremos um teorema que é válido em ZF.

Outro caso bem significativo em que não precisamos lançar mão do axioma da escolha para provarmos a existência de uma função de escolha é quando existe uma fórmula que desempenha esse papel de selecionar exatamente um elemento de cada conjunto que pertence a x.

De fato, suponha que existe uma fórmula $P(y, z)$ tal que, para todo $y \in x$, existe um único z em y para o qual $P(y, z)$ é verdadeira. Isto é, suponha que existe uma fórmula P para a qual conseguimos provar que

$$\forall y(y \in x \rightarrow \exists! z(z \in y \wedge P)),$$

onde o símbolo $\exists!$ é definido da seguinte forma:

$$\exists! z A \equiv \exists z(A \wedge \forall w(A_z^w \rightarrow (z = w)))$$

Nesse caso, provamos a existência da função de escolha usando o axioma da separação:

$$f = \{(y, z) \in x \times \bigcup x : (z \in y) \wedge P\}$$

Por exemplo, suponha que x é um conjunto formado por subconjuntos não-vazios de ω. Vimos em um exercício anterior que ω é bem-ordenado pela relação de inclusão (que coincide com a relação de ordem usual dos números naturais), o que significa que cada subconjunto não-vazio de ω possui um único elemento que está contido em todos os demais (isto é, o *mínimo* desse subconjunto). Logo, podemos definir a seguinte função de escolha

$$f = \{(y, n) \in x \times \omega : (n \in y) \wedge \forall m(m \in y \rightarrow n \subset m)\}$$

O fato de $(\omega, \subset)$ ser bem ordenado garante que f é uma função cujo domínio é x, e é claramente uma função de escolha.

Vimos, portanto, dois casos particulares do axioma da escolha que são teoremas de ZF. Então surge a pergunta: *quando precisamos do axioma da escolha para provar*

a existência de uma função de escolha em x? A resposta é: *quando x é infinito e não existe uma maneira explícita e bem determinada de escolher um único elemento de cada elemento de x.*

Bertrand Russell forneceu uma comparação bastante interessante e curiosa para explicar o axioma da escolha: *para escolhermos uma meia de cada par de meias, dentre uma coleção infinita de pares de meias, precisamos usar o axioma da escolha; se forem sapatos, não precisamos.*Isso porque, no caso dos sapatos, podemos escolher o pé direito de cada par, e, no caso das meias, os pés de cada par são indistinguíveis.

Portanto, os objetos matemáticos cuja existências dependem do axioma da escolha não podem ser construídos explicitamente, de forma que possamos determinar precisamente quais são os seus elementos. Quando dizemos que há uma função de escolha em um conjunto x e, para isso, foi essencial o uso do axioma da escolha, isso significa que há, na verdade, uma infinidade de possíveis funções de escolha, e que não podemos precisar qual função nós estamos considerando. Tais objetos matemáticos são ditos *não-construtíveis* e, para alguns matemáticos, nada vale provarmos a existência de um objeto que não conseguimos explicar exatamente quem ele é.

Porém é certo que, desde o início, como mostramos na introdução, a teoria dos conjuntos não agradou os construtivistas. A prova de Cantor de que existem muitos números transcendentes independe do axioma da escolha e, mesmo assim, é altamente não-construtiva. Ainda assim, há muitos que aceitam ZF como algo suficientemente construtivo, mas recusam trabalhar em ZFC.

Um dos resultados dependentes do axioma da escolha e que mais agravaram a polêmica em torno dele é o paradoxo de Banach-Tarski: existe uma forma de particionar uma bola no espaço em uma quantidade finita de partes e remontar essas partes para formar duas bolas disjuntas, idênticas à primeira. Para muitos, esse resultado apenas prova que não existe uma medida universal finitamente aditiva em $\mathbb{R}^3$. Para outros, no entanto, essa é uma evidência de que as aplicações do axioma da escolha são inúteis, sem nenhuma conexão com a realidade.

Por outro lado, muitos resultados importantes da matemática dependem do axioma da escolha, como a existência de uma base em qualquer espaço vetorial e o Teorema de Hahn-Banach. Mas a maior aplicação do axioma da escolha é na teoria dos cardinais. O fato de podermos atribuir a qualquer conjunto um "tamanho" – ao qual chamamos de cardinalidade – depende do axioma da escolha.

5.2 Lema de Zorn

Feita essa pequena discussão filosófica sobre o axioma da escolha, enunciamos, agora, suas principais aplicações. Começamos mostrando a forma equivalente ao axioma da escolha mais utilizada na matemática.

Teorema 5.1 (Lema de Zorn) *Se $(X, \leq)$ é uma ordem parcial em que toda cadeia admite limitante superior, então $(X, \leq)$ admite um elemento maximal.*

Demonstração: Primeiro vamos discutir um pouco a ideia intuitiva desse teorema (que, por motivos históricos, recebeu essa alcunha de *lema*). Suponha que $(X\, leq)$ não admita um elemento maximal. Notemos que a hipótese do teorema implica que X é não-vazio (por quê?). Tomamos, então, algum $x_0 \in X$. Como x_0 não é maximal, encontramos algum x_1 estritamente maior que x_0. Da mesma forma podemos encontrar algum x_2 maior que x_1 e assim por diante (aqui podemos imaginar que o axioma da escolha é necessário para tomarmos sempre um elemento maior do que outro). Após chegarmos em infinitos elementos de X através desse processo, notamos que esses formam uma cadeia, e, então, pela hipótese, tomamos y um limitante superior dessa cadeia, e iniciamos novamente o processo. A ideia intuitiva é que, em algum momento, esse processo *tem que parar*, chegando num elemento maximal. Como, infelizmente, não tem como formalizarmos essa ideia, não nos resta outra solução a não ser procurar uma demonstração rigorosa, que é árdua, trabalhosa e pouco intuitiva. A discussão precedente só serve para dar ao leitor uma vaga noção sobre o que significa o lema de Zorn e por quê ele vale.

Vamos à demonstração formal, que é adaptada do livro de Halmos, que, por sua vez, atribui a Zermelo a criação dessa prova.

Começamos definindo $\overline{X}$ o conjunto das cadeias em X, ordenado pela inclusão. Mostraremos que $\overline{X}$ tem um elemento maximal, e isso será suficiente para mostrar que X tem um elemento maximal, conforme a seguinte afirmação:

> **Afirmação 1:** Se $\overline{X}$ possui um elemento maximal então X possui um elemento maximal.

De fato, suponha que A é um elemento maximal de $\overline{X}$. Pela hipótese sobre X, seja $x \in X$ um limitante superior de A, ou seja, $a \leq x$ para todo $a \in A$. Temos que $x \in A$ pois, caso contrário, teríamos que $A \cup \{x\}$ seria uma cadeia que contém propriamente A, contradizendo a maximalidade de A. Temos que x é maximal em X, pois, se existisse $y \in X$ tal que $x \leq y$ e $x \neq y$ teríamos novamente que $A \cup \{y\}$ seria uma cadeia maior que A. Isso conclui a prova da afirmação.

> **Afirmação 2:** Se C é uma cadeia em $\overline{X}$ então $\bigcup C \in \overline{X}$.

Como $\bigcup C$ é claramente um subconjunto de X, para mostrarmos a afirmação basta provarmos que $\bigcup C$ é uma cadeia em X. Sejam a e b pertencentes a $\bigcup C$. Sejam $A, B \in C$ tais que $a \in A$ e $b \in B$. Como C é uma cadeia, temos que $A \subset B$ ou $B \subset A$, o que significa que $a, b \in A$ ou $a, b \in B$. Como $C \subset \overline{X}$, tanto A quanto B são cadeias, o que significa que $a \leq b$ ou $b \leq a$.

Seja f uma função de escolha em $\mathcal{P}(X) \setminus \{\emptyset\}$. Definimos uma função $s : \overline{X} \longrightarrow \overline{X}$ como

$$s(A) = \begin{cases} A \cup \{f(\{x \in X \setminus A : A \cup \{x\} \in \overline{X}\})\} \text{ , se } A \text{ não é maximal;} \\ A \text{ , se } A \text{ é maximal;} \end{cases}$$

A função s faz o seguinte: se A é uma cadeia não-maximal, s estende A acrescentando-lhe um único elemento. Se A é uma cadeia maximal, $s(A) = A$. Se A é uma cadeia

não-maximal, existirá $x \notin A$ tal que $A \cup \{x\}$ é uma cadeia, pois o subconjunto de uma cadeia é uma cadeia. Reparem a necessidade de usar o axioma da escolha para podermos escolher um elemento para estender a cadeia A.

Com essa definição e pela afirmação 1, nossa tarefa de demonstrar o lema de Zorn se reduz, agora, à tarefa de mostrar que existe $A \in \overline{X}$ tal que $s(A) = A$.

Antes de prosseguirmos a demonstração, precisamos de mais algumas definições. Dizemos que um subconjunto T de $\overline{X}$ é uma *torre* se satisfaz as seguintes condições:

- $\emptyset \in T$;
- se $A \in T$ então $s(A) \in T$;
- se C é uma cadeia em $(T, \subset)$ então $\bigcup C \in T$.

Existe pelo menos uma torre, pois claramente $\overline{X}$ é uma. Logo, podemos introduzir a seguinte definição:

$$\overline{X}_0 = \bigcap \{T \subset \overline{X} : T \text{ é uma torre}\}.$$

Afirmação 3: $\overline{X}_0$ é uma torre e está contida em qualquer outra torre.

Deixamos a cargo do leitor provar essa afirmação, que é bem semelhante à demonstração de que ω é um conjunto indutivo. Pela minimalidade de $\overline{X}_0$ iremos fazer algumas provas utilizando uma espécie de indução, onde s desempenha o papel de sucessor. Na verdade, pela terceira condição sobre torres, essa indução mais se aproxima da indução transfinita, que veremos posteriormente.

Nosso próximo objetivo será mostrar que $\overline{X}_0$ é uma cadeia em $\overline{X}$. Feito isso, não teremos dificuldades em mostrar que $\bigcup \overline{X}_0$ é maximal em $\overline{X}$, isto é, é uma cadeia em X que não está contida propriamente em nenhuma outra cadeia. Pela afirmação 1 isso será suficiente para provarmos o lema de Zorn.

Dizemos que um elemento C de $\overline{X}_0$ é *comparável* se, para todo $A \in \overline{X}_0$, temos $A \subset C$ ou $C \subset A$. Mostrar que $\overline{X}_0$ é uma cadeia é o mesmo que mostrar que todo elemento de $\overline{X}_0$ é comparável.

Introduzimos agora mais uma definição provisória (a última!): uma função $g : \overline{X}_0 \longrightarrow \mathcal{P}(\overline{X}_0)$ dada por

$$g(C) = \{A \in \overline{X}_0 : (A \subset C) \vee (s(C) \subset A)\}$$

Se o leitor teve paciência de acompanhar até aqui, anime-se, pois a demonstração está chegando ao fim. Faltam ainda mais algumas afirmações.

Afirmação 4: Se C é comparável então $g(C) = \overline{X}_0$.

A prova dessa afirmação usa uma espécie de indução, como dissemos anteriormente. Precisamos apenas mostrar que $g(C)$ é uma torre e seguirá da afirmação 3 que $g(C) = \overline{X}_0$.

Está claro que $\emptyset \in g(C)$, pois $\emptyset \subset C$. Seja S uma cadeia em $g(C)$. Temos duas possibilidades: ou todo $A \in S$ está contido em C ou existe pelo menos um $A \in S$ tal que $s(C) \subset A$. No primeiro caso, temos $\bigcup S \subset C$ e, portanto, $\bigcup S \in g(C)$. No segundo caso, como $A \subset \bigcup S$, temos $s(C) \subset \bigcup S$ e, novamente, $\bigcup S \in g(C)$. Para mostrar que $g(C)$ é torre só falta mostrar que, se $A \in s(C)$ então $s(A) \in g(C)$.

Seja $A \in g(C)$. Temos três casos. Ou $A = C$, ou A está contido propriamente em C ou $s(C) \subset A$.

No primeiro caso, temos $s(A) = s(C)$. Em particular, $s(C) \subset s(A)$, o que prova que $s(A) \in g(C)$.

No segundo caso, supomos que A está contido propriamente em C. Como C é comparável, temos $C \subset s(A)$ ou $s(A) \subset C$. Se $s(A) \subset C$ temos $s(A) \in g(C)$. Assumimos, então, que $C \subset s(A)$. Se $C = s(A)$ caímos no caso $s(A) \subset C$. Se $C \neq s(A)$ existe $x \in s(A) \setminus C$. Mas, pela hipótese de A estar contido propriamente em C, existe $y \in C \setminus A$. Portanto, x e y são elementos distintos (pois um pertence a C e outro não) de $s(A) \setminus A$, contradizendo que $s(A)$ tem, no máximo, um elemento que não pertence a A.

No terceiro caso, se $s(C) \subset A$, como $A \subset s(A)$ temos $s(C) \subset s(A)$, o que nos dá $s(A) \in g(C)$. Concluímos, assim, a prova da afirmação.

Afirmação 5: $\overline{X}_0$ é uma cadeia em $\overline{X}$.

Vamos provar "por indução" que todo elemento de $\overline{X}_0$ é comparável. Ou seja, mostraremos que o conjunto dos elementos comparáveis de $\overline{X}_0$ é uma torre e, portanto, coincide com todo o conjunto $\overline{X}_0$.

Como $\emptyset \subset A$, para todo A, temos $\emptyset$ é comparável. Seja S uma cadeia em $\overline{X}_0$ formada de elementos comparáveis. Mostraremos que $\bigcup S$ é comparável. De fato, seja $A \in \overline{X}_0$. Se existe $C \in S$ tal que $A \subset C$, temos, em particular, $A \subset \bigcup S$. Caso contrário, como todo elemento de S é comparável, temos $C \subset A$, para todo $C \in S$, o que nos dá $\bigcup S \subset A$.

Falta mostrar que, se C é comparável, $s(C)$ é comparável. Seja $A \in \overline{X}_0$. Pela afirmação 4 temos que $A \in g(C)$. Ou seja, $A \subset C$ ou $s(C) \subset A$. Como $C \subset s(C)$, temos $A \subset s(C)$ ou $s(C) \subset A$, provando que $s(C)$ é comparável.

Isso conclui que o conjunto dos elementos de $\overline{X}_0$ é uma torre, provando a afirmação.

Afirmação 6: $\bigcup \overline{X}_0$ é maximal em $\overline{X}$.

Seja $C = \bigcup \overline{X}_0$. Provemos que $s(C) = C$. Como, pela afirmação 5, $\overline{X}_0$ é uma cadeia, a afirmação 3 – que diz que $\overline{X}_0$ é uma torre – nos garante que $C \in \overline{X}_0$. Portanto, novamente pela afirmação 3, $s(C) \in \overline{X}_0$. Isso implica que $s(C) \subset \bigcup \overline{X}_0$. Ou seja, $s(C) \subset C$. Como $C \subset s(C)$ concluímos que $s(C) = C$, provando a afirmação.

Portanto $\overline{X}$ tem um elemento maximal e, pela afirmação 1, X também possui, provando o lema de Zorn.

■

5.3 Princípio da Boa Ordem

Como uma consequência simples do lema de Zorn, mostramos que todo conjunto pode ser bem-ordenado. Embora a prova detalhada desse resultado seja um pouco longa, são argumentos bem comuns e corriqueiros, sem tantos "truques" como na demonstração do lema de Zorn.

Teorema 5.2 (Princípio da Boa Ordem) *Para todo conjunto X existe uma relação $\leq$ tal que $(X, \leq)$ é uma boa ordem.*

Demonstração: A demonstração do princípio da boa ordem é uma aplicação *standard* do lema de Zorn. Diversos resultados clássicos da matemática – como a existência de base em espaços vetoriais e o teorema de Hahn-Banach – utilizam argumentos bem parecidos. A ideia é simples: se quisermos mostrar que uma propriedade vale para um conjunto X, consideramos todos a ordem parcial constituída dos subconjuntos de X que satisfazem tal propriedade (no caso, ser bem-ordenado). Verificamos que a hipótese do lema de Zorn é atendida e tomamos Y um elemento maximal dessa ordem parcial. Se Y não for todo o conjunto X, mostramos que esse pode ser estendido um pouco mais, contradizendo sua maximalidade.

Quando a propriedade que queremos mostrar para X envolve alguma estrutura – neste caso, uma ordem – é natural que, nessa ordem parcial que criamos, consideremos algo a mais que os subconjuntos de Y. No caso deste teorema, o domínio da ordem parcial é formada pelos conjuntos bem-ordenados $(Y, \leq)$ tais que $Y \subset X$, e na definição da ordem, precisamos respeitar a compatibilidade entre esses conjuntos ordenados.

Vamos à demonstração.

Definimos uma ordem parcial $(\overline{X}, \preceq)$ da seguinte forma: $\overline{X}$ é o conjunto de todos os conjuntos bem-ordenados $(Y, \leq)$ tais que $Y \subset X$, e $(Y_1, \leq_1) \preceq (Y_2, \leq_2)$ se, e somente se, as seguintes condições são satisfeitas:

1. $Y_1 \subset Y_2$;
2. $x \leq_1 y$ se, e somente se, $x \leq_2 y$, para todos $x, y \in Y_1$;
3. se $x \in Y_1$ e $y \in Y_2 \setminus Y_1$ então $x \leq_2 y$.

Fica como exercíco ao leitor mostrar que $(\overline{X}, \preceq)$ é um conjunto ordenado. Provaremos que ele satisfaz a hipótese do lema de Zorn.

Seja S uma cadeia em $\overline{X}$. Definimos

$$Y = \bigcup \{Y' : \exists \leq' : (Y', \leq') \in S\}$$

e

$$\leq = \bigcup \{\leq' : \exists Y' : (Y', \leq') \in S\}$$

Afirmação: $(Y, \leq) \in \overline{X}$ e é um limitante superior de S.

Para provar a afirmação, primeiro verifiquemos que $\leq$ é uma boa ordem sobre X. Como S é uma cadeia, dados $x, y, z \in Y$ existe $(Y', \leq') \in S$ tal que $x, y, z \in Y'$ e, para todos $u, v \in Y'$, temo $u \leq v$ se, e somente se, $u \leq' v$. Portanto, as propriedades de ordem são satisfeitas para $\leq$, pois são satisfeitas para $\leq'$. Portanto, $\leq$ é uma ordem.

Para verificar que $\leq$ é uma boa ordem, considere $Z \subset Y$ um conjunto não-vazio. Portanto, existe $(Y_1, \leq_1) \in S$ tal que $Z \cap Y_1 \neq \emptyset$. Por hipótese, existe $z \in Z \cap Y_1$ que é mínimo, em relação à ordem $\leq_1$. Vamos mostrar que também é o mínimo de Z, em relação a $\leq$.

Suponhamos, por absurdo, que existe $w \in Z$ tal que $w \neq z$ e $w \leq z$. Como z é mínimo de $Z \cap Y_1$, temos que $w \notin Y_1$ Tome $(Y_2, \leq_2)$ tal que $w \in Y_2$. Como S é uma cadeia, vale $(Y_2, \leq_2) \preceq (Y_1, \leq_1)$ ou $(Y_1, \leq_1) \preceq (Y_2, \leq_2)$. Mas o primeiro caso não é possível, pois $w \in Y_2 \setminus Y_1$.

Temos, então, $(Y_1, \leq_1) \preceq (Y_2, \leq_2)$. Da condição 3 da ordem $\preceq$ segue que $z \leq_2 w$. Porém, como $w \leq z$, da definição de $\leq$, do fato de S ser uma cadeia e da condição 2 da ordem $\preceq$ seguem que $w \leq_2 z$ (deixamos os detalhes dessa passagem como exercício ao leitor). Portanto, a antissimetria de $\leq_2$, nos dá que $w = z$, contradizendo nossa hipótese e provando a afirmação.

Agora, aplicamos o lema de Zorn para obter $(Y, \leq)$ maximal em $\overline{X}$. Tudo que precisamos para concluir o teorema é provar que $Y = X$. De fato, suponha que $Y \neq X$. Tome $x \in X \setminus Y$. Considere $Y' = Y \cup \{x\}$ e defina uma ordem $\leq'$ em Y' acrescentando a condição $y \leq x$, para todo $y \in Y$. Isto é, $\leq' = \leq \cup \{(y, x) : y \in Y\}$. Claramente $(Y', \leq')$ é um conjunto bem-ordenado, diferente de $(Y, \leq)$ e tal que $(Y \leq) \preceq (Y', \leq')$, contradizendo a maximalidade de $(Y, \leq)$.

■

Os dois teoremas anteriores são, na verdade, formas equivalentes ao axioma da escolha, como mostra o seguinte resultado:

Teorema 5.3 *Em ZF, são equivalentes:*

(a) *Axioma da escolha;*

(b) *Lema de Zorn;*

(c) *Princípio da boa ordem.*

Demonstração: Já provamos que (a) implica (b) e que (b) implica (c), lembrando que a demonstração do princípio da boa ordem não utiliza diretamente o axioma da escolha, mas apenas o lema de Zorn. Resta mostrar que (c) implica (a), cuja ideia da demonstração já foi discutida no início deste capítulo.

Seja X um conjunto de conjuntos não-vazios. Aplicando o princípio da boa ordem, considere $\leq$ uma boa ordem no conjunto $\bigcup X$. Definiremos uma função de escolha que a cada elemento x de X associa o mínimo de x, isto é:

$$f = \{(x, y) \in X \times \bigcup X \ : (y \in x) \wedge \forall z (z \in \bigcup X \rightarrow y \leq z\}$$

Pela propriedade de boa ordem e pelo fato de $\emptyset \notin X$, para todo $x \in X$ existe $y \in \bigcup x$ tal que $(x, y) \in f$. A unicidade do elemento mínimo, como já foi discutido anteriormente, segue da antissimetria da ordem (se y e z fossem "dois mínimos", teríamos $y \leq z$ e $z \leq y$, o que implica que $y = z$).

Portanto f é uma função, e é justamente uma função de escolha em X. ■

5.4 Comparabilidade de conjuntos por funções injetoras

Definição 5.4 Dizemos que um conjunto Y *domina* um conjunto X se existe uma função injetora de X em Y. Dizemos que Y *domina estritamente* X se Y domina X mas X não domina Y. Denotamos por $X \preceq Y$ quando Y domina X e $X \prec Y$ quando Y domina estritamente X.

É fácil verificar que $X \preceq Y$ e $Y \preceq Z$ implicam que $X \preceq Z$. Também é imediato que $X \preceq X$.

Lema 5.5 *Sejam X e Y conjuntos não-vazios. Temos que $X \preceq Y$ se, e somente se, existe uma função sobrejetora de Y em X.*

Demonstração: Suponha que existe $f : X \longrightarrow Y$ injetora. Tome $x_0 \in X$ um elemento qualquer. Defina $g : Y \longrightarrow X$ como $g(y) = f^{-1}(y)$, se $y \in im(f)$ (lembrando que f é bijetora sobre sua imagem) e $g(y) = x_0$ se $y \in Y \setminus im(f)$.

Suponha agora que existe $g : Y \longrightarrow X$ sobrejetora. Considere a função $h : X \longrightarrow \mathcal{P}(Y)$ dada por

$$h(x) = \{y \in Y : g(y) = x\}$$

Como g é sobrejetora, $h(x) \neq \emptyset$, para todo $x \in X$. Usando o axioma da escolha defina uma função $s : im(h) \longrightarrow Y$ tal que $s(A) \in A$, para todo $A \in im(h)$. Defina a função $f : X \longrightarrow Y$ por

$$f(x) = s(h(x))$$

Notemos que $h(x) \cap h(x') = \emptyset$, sempre que $x \neq x'$. Logo, f é injetora, provando o que queríamos. ■

Teorema 5.6 *Para todos conjuntos X e Y, ou $X \preceq Y$ ou $Y \preceq X$.*

Demonstração: Suponha que não ocorra $Y \preceq X$. Pelo Lema 5.5 isso significa que não existe uma função sobrejetora de X em Y. Mostraremos que $X \preceq Y$. Podemos assumir que nenhum dos conjuntos X ou Y é vazio, pois, nesse caso, o vazio seria uma função injetora de um em outro.

Seja $\mathcal{F} = \{f \subset X \times Y : f \text{ é função injetora}\}$ e considere a ordema dada pela inclusão ($f \leq g$ se, e somente se, $f \subset g$). Note que $\mathcal{F}$ é não-vazio, pois, como ambos os conjuntos são não-vazios, temos que $\{(x, y)\} \in \mathcal{F}$, onde $x \in X$ e $y \in Y$. Vamos verificar que $(\mathcal{F}, \subset)$ satisfaz as hipóteses do Lema de Zorn.

Seja $\mathcal{C} \subset \mathcal{F}$ uma cadeia. Vejamos que $\bigcup \mathcal{C}$ é uma função injetora contida em $X \times Y$. Se ambos (x, y) e (x, z) pertencem a $\bigcup \mathcal{C}$, existem $f, g \in \mathcal{C}$ tais que $(x, y) \in f$ e $(x, z) \in g$. Como $\mathcal{C}$ é uma cadeia, temos que ou $f \subset g$ ou $g \subset f$. Assumimos, sem perda de generalidade, que $f \subset g$. Logo, ambos (x, y) e (x, z) pertencem a g e, como g é uma função, temos $y = z$. Concluímos que $\bigcup \mathcal{C}$ é uma função e, com um argumento análogo, também podemos concluir que é uma função injetora. Logo, $\bigcup \mathcal{C} \in \mathcal{F}$ e claramente é um limitante superior de $\mathcal{C}$. Provamos, assim, que toda cadeia em $(\mathcal{F}, \subset)$ possui limitante superior, e podemos aplicar o Lema de Zorn para achar um elemento maximal.

Seja $f \in \mathcal{F}$ um elemento maximal. Seja $Z \subset X$ o domínio de f. Mostraremos que $Z = X$.

Se f é sobrejetora em relação a Y, f pode ser estendida a uma função sobrejetora de X em Y (basta definir $\tilde{f}(x) = f(x)$, se $x \in Z$, e $\tilde{f}(x)$ como qualquer valor fixado em Y caso $x \in X \setminus Z$). Pelo Lema 5.5, isso implica que existe uma função injetora de Y em X, isto é, que $Y \preceq X$, o que assumimos não ocorrer. Logo, f não é sobrejetora, isto é, existe $y \in Y$ que não pertence à imagem de f. Suponha que $Z \neq X$. Tome $x \in X \setminus Z$ e defina $g = f \cup \{(x, y)\}$. Temos que g é claramente uma função injetora contida em $X \times Y$ e que estende f, contradizendo que fé maximal em $\mathcal{F}$.

Portanto, f é uma função injetora de X em Y, provando que $X \preceq Y$.

■

Ressaltamos que o uso do axioma da escolha é necessário. De fato, o teorema da comparabilidade dos conjuntos é equivalente ao axioma da escolha, em ZF.

Exercícios

1. Discuta a seguinte afirmação: *sempre que a existência de uma função de escolha sobre um conjunto vale em ZFC mas não é assegurada em ZF, temos, em ZFC, mais de uma função de escolha sobre esse conjunto.*

2. Seja X um conjunto não-vazio tal que, para todo $n \in \omega$, não existe uma função de domínio n e imagem X. Prove que existe uma função injetora de ω em X.

3. Na demonstração do princípio da boa ordem, por que assumimos a condição 3 da ordem $\preceq$? A afirmação contida na demonstração seria verdadeira ou falsa, se tirássemos essa condição? Justifique.

4. Prove que todo espaço vetorial sobre $\mathbb{R}$ possui uma base (algébrica).

5. Prove em ZF (sem assumir o axioma da escolha) que $\omega \times 2$ e $\omega \times \omega$ podem ser bem-ordenados.

6. Prove o axioma da escolha diretamente do Lema de Zorn (sem usar o princípio da boa ordem).

7. Seja $F : A \longrightarrow B$ uma função e suponha que $\emptyset \notin B$. Prove que existe uma função $f : A \longrightarrow \bigcup B$ tal que $f(a) \in F(a)$, para todo $a \in A$.

8. Prove que a função identidade é o único isomorfismo de um conjunto bem-ordenado nele mesmo.

9. Considerando o sistema ZFC sem o axioma da regularidade, prove que o axioma da regularidade é equivalente à seguinte sentença: não existe uma função f de domínio ω tal que $f(n^+) \in f(n)$, para todo $n \in \omega$.

Capítulo 6

Conjuntos equipotentes

Definição 6.1 Dizemos que dois conjuntos X e Y são *equipotentes* se existe uma função bijetora de X em Y. Usamos a notação $X \equiv Y$ para denotar que X e Y são equipotentes.

Está claro que $X \equiv X$ e que $X \equiv Y$ se, e somente se, $Y \equiv X$. Também é fácil verificar (pois a composta de funções bijetoras é bijetora) que $X \equiv Y$ e $Y \equiv Z$ implica $X \equiv Z$. Ou seja, $\equiv$ é, de certa forma, uma relação de equivalência sobre a classe de todos os conjuntos. É claro que, como não existe conjunto de todos os conjuntos, não podemos considerar $\equiv$ como uma relação (a menos quando o restringimos a uma família particular de conjuntos), mas, sim, como um símbolo relacional binário que adicionamos à linguagem, que satisfaz as propriedades de uma relação de equivalência (reflexividade, simetria e transitividade).

6.1 O Teorema de Cantor-Schröder-Bernstein

Já vimos no Capítulo 5 a noção de um conjunto ser "menor" que outro em quantidade de elementos, e introduzimos a notação $X \preceq Y$ quando Y *domina* X, isto é, quando existe uma função injetora de X em Y (ou, equivalentemente, quando existe uma função sobrejetora de Y em X). Naturalmente, como uma função bijetora, tal como sua inversa, é tanto injetora quanto sobrejetora, temos que $X \equiv Y$ implica $X \preceq Y$ e $Y \preceq X$. A pergunta que surge é: se $X \preceq Y$ e $Y \preceq X$ então $X \equiv Y$? É de se esperar que isso ocorra, se entendemos que a relação entre conjuntos $\equiv$ traduz a ideia de "conjuntos do mesmo tamanho" e $X \preceq Y$ a ideia de "X tem tamanho menor ou igual a Y". E de fato o resultado vale, e prová-lo é o objetivo desta seção. Antes, porém, precisamos de um lema.

Lema 6.2 (teorema do ponto fixo de Tarski) *Seja F uma função de $\mathcal{P}(X)$ em $\mathcal{P}(X)$ tal que $A \subset B \subset X$ implica $F(A) \subset F(B)$. Então existe $Z \subset X$ tal que $F(Z) = Z$.*

Demonstração: Sendo F e X como na hipótese do lema, considere o conjunto

$$\mathcal{S} = \{Y \in \mathcal{P}(X) : Y \subset F(Y)\}$$

e tome

$$Z = \bigcup \mathcal{S}.$$

Mostraremos que $F(Z) = Z$. Primeiro vejamos que $Z \subset F(Z)$.

Seja $x \in Z$. Temos $x \in Y$, para algum $Y \in \mathcal{S}$. Logo, $Y \subset F(Y)$ e, portanto $x \in F(Y)$. Como $Y \subset Z$, por hipótese sobre F temos $F(Y) \subset F(Z)$ e, portanto, $x \in F(Z)$.

Reciprocamente, mostraremos que $F(Z) \subset Z$. Como $Z \subset F(Z)$, pela hipótese sobre F temos $F(Z) \subset F(F(Z))$, o que significa que $F(Z) \in \mathcal{S}$. Logo, $F(Z) \subset \bigcup S = Z$, concluindo a demonstração do teorema. ■

Teorema 6.3 (Cantor-Schröder-Bernstein) *Se $X \preceq Y$ e $Y \preceq X$ então $X \equiv Y$.*

Demonstração: Sejam $g : X \longrightarrow Y$ e $h : Y \longrightarrow X$ funções injetoras. Mostraremos que existe $f : X \longrightarrow Y$ bijetora.

A ideia da demonstração é dividir X em duas partes, X_1 e X_2, e Y em duas partes, Y_1 e Y_2, de modo que g restrita a X_1 seja sobrejetora em relação a Y_1 e h restrita a Y_2 seja sobrejetora em relação a Y_1. Em seguida, basta "colar" as funções g restrita a X_1 e a inversa de h restrita a Y_2. Usaremos o teorema do ponto fixo de Tarski para achar as partições de X e Y.

Usaremos a notação $g[A]$ para denotar o conjunto $im(g|A)$, e o mesmo também para a função h.

Defina a função $F : \mathcal{P}(X) \longrightarrow \mathcal{P}(X)$ dada por

$$F(A) = X \setminus h[Y \setminus g[A]]$$

Notemos que, se $A \subset B$, $g[A] \subset g[B]$. Logo, $Y \setminus g[B] \subset Y \setminus g[A]$ e $h[Y \setminus g[B]] \subset h[Y \setminus g[A]]$, concluindo que $F(A) \subset F(B)$.

Portanto, F satisfaz a hipótese do teorema do ponto fixo de Tarski. Tome $X_1 \subset X$ tal que $F(X_1) = X_1$. Defina $Y_1 = g[X_1]$, $X_2 = X \setminus X_1$ e $Y_2 = Y \setminus Y_1$.

Como $F(X_1) = X_1$, temos

$$X \setminus h[Y \setminus g[X_1]] = X_1,$$

o que implica que

$$h[Y \setminus g[X_1]] = X \setminus X_1.$$

Isto é, $h[Y_2] = X_2$. Com isso, concluímos que $(h|Y_2)^{-1}$ é uma função bijetora de X_2 em Y_2. Como $g|X_1$ é uma função bijetora de X_1 em Y_1, temos que

$$f = (g|X_1) \cup (h|Y_2)^{-1}.$$

é uma função bijetora de X em Y.

■

6.2 Conjuntos finitos

Definição 6.4 Dizemos que um conjunto é *finito* se é equipotente a algum $n \in \omega$. Dizemos que um conjunto é *infinito* se não é finito.

Teorema 6.5 *São equivalentes:*

(a) *X é infinito;*

(b) *$\omega \preceq X$;*

(c) *Existe $Y \subset X$ tal que $Y \neq X$ e $Y \equiv X$.*

Demonstração: Para provar que (a) implica (b) a ideia é simples, mas a formalização é um pouco complicada: construímos uma função injetora $h : \omega \longrightarrow X$ recursivamente de modo que $h(n^+)$ seja um elemento de $X \setminus \{h(0), \ldots, h(n)\}$. O item (a) garante que esse conjunto não é vazio. Porém, para formalizar essa ideia precisamos adaptar o teorema da recursão usual para fazermos uma *recursão completa*. Isso significa que precisamos de certa forma "memorizar" todos os valores anteriores da função antes de definirmos para o próximo número natural.

Seja $s : \mathcal{P}(X) \setminus \{\emptyset\} \longrightarrow X$ uma função de escolha (ou seja, $s(Y) \in Y$, para todo $Y \subset X$ não-vazio).

Seja Y o conjunto de todas as funções injetoras que têm como domínio um número natural e imagem contida em X. A saber,

$$Y = \{f \subset \omega \times X : (\text{f é função injetora}) \wedge (dom(f) \in \omega)\}$$

Tome $y_0 = \emptyset$ e $g : Y \longrightarrow Y$ a função definida por

$$g(f) = f \cup \{(dom(f), s(X \setminus im(f))\}$$

Isto é, se f é uma função de domínio n, $g(f)$ é uma função f' de domínio n^+ definida da seguinte forma: $f'(k) = f(k)$, para $k \in n$, e $f'(n) = s(X \setminus im(f))$. Lembre-se de que $X \setminus im(f)$ é não-vazio pela hipótese (a), que garante que f não é sobrejetora em relação a X, e $s(X \setminus im(f))$ é um elemento de $X \setminus im(f)$, garantindo que $f'(n)$ não pertence à imagem de f.

Pelo teorema da recursão, existe uma função $F : \omega \longrightarrow Y$ tal que $F(0) = y_0$ e $F(n^+) = g(F(n))$.

Ou seja, cada $F(n)$ é um "pedaço" da função h, que queremos definir, restrita a n. Definimos

$$h = \bigcup im(F)$$

Para ficar mais clara a definição de h, uma outra definição equivalente a essa seria: h é uma função de ω em X tal que $h(n) = f(n)$, tomando $f = F(n^+)$.

A função h é injetora. De fato, se $n \neq m$, podemos assumir, sem perda de generalidade, que $m \in n$. Sejam $f_1 = F(m^+)$ e $f_2 = F(n^+)$. É fácil verificar, por indução, que $F(m^+) \subset F(n)$. Como, pela construção, $f_2(n) \notin imF(n)$, temos que $f_1(m) \neq f_2(n)$. Logo, $h(m) \neq h(n)$.

Para mostrar que (b) implica (c), seja f uma função injetora de ω em X. Tome $Y = X \setminus \{f(0)\}$ e considere

$$g = \{(f(n), f(n^+)) : n \in \omega\} \cup \{(x, x) : x \in X \setminus im(f)\}$$

Como f é injetora e $n \neq n^+$ é fácil verificar que g é injetora. Vimos que todo número natural diferente de 0 é sucessor de alguém. Portanto, o único elemento da imagem de f que não é da forma $f(n^+)$, para algum $n \in \omega$, é $f(0)$. Logo, g é sobrejetora em Y.

Agora mostraremos que (c) implica (a). Para isso mostraremos a contrapositiva: se X é finito então X não é equipotente a algum subconjunto próprio. Compondo com bijeções, é suficiente provarmos que, para todo $n \in \omega$ e $S \subset n$, se $S \neq n$ então S não é equipotente a n. Provaremos esse resultado por indução.

Suponha que o resultado seja verdadeiro para subconjuntos de n. Seja $S \subset n^+$ tal que $S \neq n^+$. Suponha, por absurdo, que existe $f : n^+ \longrightarrow S$ bijetora. Podemos assumir, sem perda de generalidade, que $n \notin S$. De fato, se tivermos $n \in S$, troque S por $\tilde{S} = (S \setminus n) \cup \{k\}$, para algum $k \in n^+ \setminus S$, e f por $\tilde{f}$ tal que $\tilde{f}(m) = k$, se $f(m) = n$, e $\tilde{f}(m) = f(m)$, se $f(m) \neq n$. É fácil ver que $\tilde{f}$ é uma função bijetora de n^+ em $\tilde{S}$.

Uma vez assumido que $m \notin S$, consideramos $f' = f|n$ e $S' = Imf'$. Claramente $S' \subset n$ e $S' \neq n$, pois, como f é injetora e $Imf = S \subset n$, temos $f(n) \in n \setminus S'$. Portanto, S' é um subconjunto próprio de n equipotente a n, contrariando a hipótese indutiva.

■

Corolário 6.6 *Se X é finito e $Y \subset X$ então Y é finito.*

Demonstração: Mostraremos a contrapositiva. Se Y é infinito e $Y \subset X$, pelo Teorema 6.5, parte (b), temos $\omega \preceq Y$ e, portanto, $\omega \preceq X$, visto que uma função injetora de ω em Y também é uma função injetora de ω em X. Concluímos que X é infinito. ■

Corolário 6.7 *X é finito se, e somente se, existe $n \in \omega$ tal que $X \preceq n$.*

Demonstração: Uma das implicações é trivial, já que $X \equiv n$ implica $X \preceq n$. Suponha que $X \preceq n$ e tome $f : X \longrightarrow n$ uma função injetora. Temos que a imagem de f é um subconjunto de n e, portanto, pelo Corolário 6.6, é finita. Como f é injetora, temos $X \equiv Im(f)$. Logo, X é finito. ■

6.3 Conjuntos enumeráveis

A equivalência entre os itens (a) e (b) do Teorema 6.5 nos dizem, de certa forma, que o conjunto dos números naturais é o *menor* conjunto infinito que existe (se é que existem *infinitos maiores que outros* – responderemos essa pergunta apenas no final do capítulo).

Definição 6.8 Um conjunto X é *enumerável* se $X \preceq \omega$.

Note que X é enumerável se, e somente se, é finito ou equipotente a ω (exercício).

Mostraremos que o produto cartesiano de conjuntos enumeráveis é enumerável. Mas primeiro mostraremos que isso vale para ω.

Lema 6.9 *$\omega \times \omega$ é equipotente a ω.*

Demonstração: Seja $X = \{(m, n) \in \omega \times \omega : m \leq n\}$.

Afirmação 1 *$X \preceq \omega$.*

Para mostrar isso basta construirmos uma função sobrejetora de ω em X. Usaremos recursão. Seja $g : X \longrightarrow X$ a seguinte função: $g(m, n) = (m^+, n)$, se $m < n$, e $g(n, n) = (0, n^+)$. Note que g está bem definida de X em X, pois $m < n$ implica $m^+ \leq n$ (veja a demonstração do Teorema 4.11). Pelo Teorema da Recursão existe uma função $f : \omega \to X$ tal que $f(0) = (0, 0)$ e $f(n^+) = g(f(n))$.

Mostraremos que f é sobrejetora em relação a X. Ou seja, provaremos, por indução dupla, a seguinte sentença:

Para todo $n \in \omega$, para todo $m \in \omega$, se $m \leq n$ então $(m, n) \in Im(f)$.

Seja $P(n, m)$ a fórmula $(m > n) \vee (m, n) \in Im(f)$. Note que a sentença acima é equivalente a

$$\forall n(n \in \omega \to \forall m(m \in \omega \to P(n, m))).$$

Claramente vale $P(m, 0)$, para todo $m \in \omega$. De fato, como $m \leq 0$ se, e somente se, $m = 0$, precisamos mostrar apenas que $(0, 0) \in Im(f)$, o que é verdade, pois definimos f de modo que $f(0) = (0, 0)$.

Suponha que, para um n fixado, vale $P(m, n)$, para todo $m \in \omega$. Mostaremos, por indução em m, que vale $P(m, n^+)$, para todo $m \in \omega$. De fato, $P(0, n^+)$ segue da hipótese $P(n, n)$. Com efeito, como $n \leq n$ temos que $P(n, n)$ implica $(n, n) \in Im(f)$. Seja $k \in \omega$ tal que $f(k) = (n, n)$. Temos que $f(k^+) = g(n, n) = (0, n^+)$, provando que $(0, n^+) \in Im(f)$ e, portanto, vale $P(0, n^+)$. Da mesma forma, se vale $P(m, n^+)$, para $m < n^+$, temos que $f(k) = (m, n^+)$, para algum $k \in \omega$, e $f(k^+) = (m^+, n^+)$, provando $P(m^+, n^+)$. Se $m \geq n^+$, temos $P(m^+, n^+)$ automaticamente verdadeiro, pois $m^+ > n^+$, nesse caso.

Concluímos que f é sobrejetora e provamos, assim, a afirmação 1.

Afirmação 2 *$\omega \times \omega \preceq \omega \times 2$.*

Seja f a função construída na afirmação 1. Se x é um par ordenado (a, b), denotaremos por x^{-1} o par (b, a). Com essa notação, definamos $h : \omega \times 2 \longrightarrow \omega \times \omega$ como $h(n, 0) = f(n)$ e $h(n, 1) = (f(n))^{-1}$. Seja $(a, b) \in \omega \times \omega$. Se $a \leq b$ existe $n \in \omega$ tal que $f(n) = (a, b)$ e, nesse caso, $h(n, 0) = (a, b)$. Se não vale $a \leq b$, vale $b < a$ e, portanto, existe $n \in \omega$ tal que $f(n) = (b, a)$ e, portanto, $h(n, 1) = (a, b)$. Concluímos que h é sobrejetora, provando a afirmação.

Afirmação 3 $\omega \times 2 \preceq \omega$.

Seja $g : \omega \times 2 \longrightarrow \omega \times 2$ definida como $g(n,0) = (n,1)$ e $g(n,1) = (n^+,0)$, para todo $n \in \omega$. Pelo Teorema da Recursão existe $f : \omega \to \omega \times 2$ tal que $f(0) = (0,0)$ e $f(k^+) = g(f(k))$, para todo $k \in \omega$. Mostremos, por indução em n, que para todo $n \in \omega$ temos $(n,0),(n,1) \in Im(f)$, provando que f é sobrejetora em relação a $\omega \times 2$. De fato, $f(0) = (0,0)$ e $f(1) = (0,1)$, provando que a propriedade é verdadeira para $n = 0$. Supondo que $(n,0),(n,1) \in Im(f)$, tome k tal que $f(k) = (n,1)$. Teremos $f(k^+) = (n^+,0)$ e $f((k^+)^+) = (n^+,1)$, provando que ambos os pares pertencem à imagem de f e concluindo a prova da afirmação.

Das afirmações 2 e 3 segue que $\omega \times \omega \preceq \omega$. É fácil verificar que $\omega \preceq \omega \times \omega$, bastando tomar a função que associa n ao par ordenado $(n,0)$. Portanto, segue do Teorema de Cantor-Schröder-Bernstein que $\omega \times \omega \equiv \omega$.

■

A demonstração do lema anterior poderia ser bem mais simples se assumíssemos alguns fatos elementares – mas não demonstrados neste livro – de aritmética. Por exemplo, assumindo conhecidos a definição de potência e o Teorema Fundamental da Aritmética, a função $f(n,m) = 2^n \cdot 3^m$ é claramente uma função injetora de $\omega \times \omega$ em ω. Ou poderíamos construir uma bijeção diretamente, sem usar o Teorema de Cantor-Schröder-Bernstein, bastando definir a função $f(n,m) = 2^m \cdot (2m+1) - 1$. Mas optamos por uma demonstração que usa apenas o aspecto conjuntístico do conjunto dos números naturais.

Teorema 6.10 *Se A e B são conjuntos infinitos enumeráveis, $A \times B$ é infinito enumerável.*

Demonstração: Segue do Lema 6.9, pois, tomando bijeções $f : A \longrightarrow \omega$, $g : B \longrightarrow \omega$ e $h : \omega \times \omega \longrightarrow \omega$, definimos uma função $F : A \times B \longrightarrow \omega$ como $F(a,b) = h(f(a),g(b))$. Vejamos que F é bijetora. Sejam $(a,b) \neq (a',b')$ dois elementos distintos de $A \times B$. Se $a \neq a'$, pela injetividade de f temos $f(a) \neq f(a')$. Em particular, $(f(a),g(b)) \neq (f(a'),g(b'))$ e, pela injetividade de h, concluímos que $F(a,b) \neq F(a',b')$. Analogamente ocorre se $b \neq b'$, provando que F é injetora. Para a sobrejetividade, tome $n \in \omega$. Como h é sobrejetora existe $(i,j) \in \omega \times \omega$ tal que $h(i,j) = n$. Como f e g são ambas sobrejetoras existem $a \in A$ e $b \in B$ tais que $f(a) = i$ e $g(b) = j$. Temos que $F(a,b) = n$, concluindo que F é bijetora.

■

Corolário 6.11 *Se A e B são enumeráveis, então $A \times B$ é enumerável.*

Demonstração: Se A e B são enumeráveis, temos $A \preceq \omega$ e $B \preceq \omega$. Repetindo o argumento acima concluímos que $A \times B \preceq \omega \times \omega$ e, portanto, $A \times B \preceq \omega$. ■

O próximo teorema diz que a união enumerável de conjuntos enumeráveis é enumerável.

Teorema 6.12 *Se $\mathcal{F}$ é enumerável e todo elemento de $\mathcal{F}$ é enumerável, então $\bigcup \mathcal{F}$ é enumerável.*

Demonstração: Seja $\alpha \in \omega \cup \{\omega\}$ tal que α é equipotente a $\mathcal{F}$ (existe, por hipótese de que essa família é enumerável). Tome $f : \alpha \longrightarrow \mathcal{F}$ uma função bijetora. Defina $X = \bigcup \mathcal{F}$.

O que faremos agora é usar o axioma da escolha para escolher, para cada elemento de $\mathcal{F}$, uma bijeção com algum número natural ou com ω. Assim, poderemos organizar os elementos de X como em uma espécie de matriz eventualmente infinita. Formalizar esse argumento requer algum cuidado.

Seja $F : \alpha \longrightarrow \mathcal{P}(\mathcal{P}((\omega \cup \{\omega\}) \times X))$ a função definida como: $F(n)$ é o conjunto das funções de domínio pertencente a $\omega \times \{\omega\}$ e bijetoras sobre $f(n)$. Pela hipótese de que todos os elementos de $\mathcal{F}$ são finitos ou enumeráveis sabemos que $F(n)$ é não-vazio, para todo $n \in \alpha$. Tome s uma função de escolha na imagem de S. Isto é, $s(A) \in A$, para todo $A \in im(F)$. Defina $g : \alpha \longrightarrow \mathcal{P}((\omega \cup \{\omega\}) \times X)$ como $g(n) = s(F(n))$.

Ou seja, $g(n)$ é uma função bijetora com domínio ω ou $k \in \omega$, e imagem $f(n)$.

Definimos $h : \omega \times \omega \longrightarrow X$ da seguinte forma: se $n \in \alpha$ e $m \in dom(g(n))$ definimos $h(n, m) = g(n)(m)$ (informalmente, $h(n, m)$ é o "m-ésimo elemento do n-ésimo conjunto da família $\mathcal{F}$"). Se $n \notin \alpha$ ou $n \in \alpha$ mas $m \notin dom(g(n))$, definimos $h(n, m) = h(0, 0)$ (assumimos que $\mathcal{F}$ é não-vazio e que pelo menos um dos elementos de $\mathcal{F}$ é não-vazio – caso contrário, a conclusão do teorema é trivial).

Vejamos que h é sobrejetora. Seja $x \in X$. Por definição da união, existe $A \in \mathcal{F}$ tal que $x \in A$. Como f é sobrejetora em $\mathcal{F}$, existe $n \in \alpha$ tal que $f(n) = A$. Observe que $\alpha \subset \omega$ e, portanto, $n \in \omega$. Como $g(n)$ é sobrejetora em relação a $f(n)$, existe $m \in dom(g)$ tal que $g(m) = x$. Note mais uma vez que $m \in \omega$. Logo, $h(n, m) = g(n)(m) = x$, como queríamos.

Provamos, assim, que $X \preceq \omega \times \omega$, concluindo, pelo Lema 6.10, que X é enumerável. ■

Corolário 6.13 *Se A e B são enumeráveis, então $A \cup B$ é enumerável.*

Demonstração: Aplique o Teorema 6.12 para $\mathcal{F} = \{A, B\}$. ■

6.4 Comparação entre conjuntos infinitos

Faremos agora uma série de resultados que irão culminar em um dos teoremas mais importantes para calcular cardinalidade de conjuntos: quando X é infinito, $X \times X$ é equipotente a X.

Lema 6.14 *Se A é um conjunto infinito e B é enumerável, então $A \cup B$ é equipotente a A.*

Demonstração: Primero observamos que suconjunto de conjunto enumerável é enumerável, pois a restrição de uma função injetora com imagem contida em ω também é uma função injetora com imagem contida em ω. Por esse motivo, substituindo eventualmente B por $B \setminus A$, podemos assumir que $A \cap B = \emptyset$.

Sejam $g : \omega \to A$ e $h : B \to \omega$ injetoras. Seja $A' = Im(g)$.

Notemos que $A' \cup B \equiv A'$. De fato, como $A' \equiv \omega$, precisamos apenas provar que $A' \cup B \equiv \omega$. Como A' (e, portanto, $A' \cup B$) é infinito e vimos que $\omega \equiv \omega \times 2$, usando o Teorema de Cantor-Schröder-Bersntein basta provarmos que $A' \cup B \preceq \omega \times 2$. Para isso, tomamos a função que associa $a \in A'$ a $(g^{-1}(a), 0)$ e $b \in B$ a $(h(b), 1)$. Essa função está bem definida, visto que $A' \cup B = \emptyset$, e é claramente injetora.

Fixe $f' : A' \to A' \cup B$ uma função bijetora. Defina $f : A \to A \cup B$ como $f(x) = x$, se $x \in A \setminus A'$, e $f(x) = f'(x)$, se $x \in A$. É fácil verificar que f é bijetora, concluindo a demonstração. ■

Lema 6.15 *Se X é infinito, então $X \times \{0, 1\}$ é equipotente a X.*

Demonstração: Seja $\mathcal{F}$ o conjunto de todas as relações $f \subset (X \times \{0, 1\}) \times X$ tais que existe um conjunto $Y \subset X$ para o qual f é uma função injetora de $Y \times \{0, 1\}$ em Y. Esse conjunto é não-vazio porque $\emptyset \in \mathcal{F}$, como uma função de $\emptyset \times \{0, 1\}$ em $\emptyset$. De modo análogo ao que foi feito no Teorema 5.6 usando o Lema de Zorn concluímos que $\mathcal{F}$ tem um elemento maximal em relação à ordem dada pela inclusão.

Seja f esse elemento maximal e $Y \times \{0, 1\}$ o seu domínio. Pelo Teorema de Cantor-Schröder-Bernstein isso implica que $Y \times \{0, 1\}$ é equipotente a Y. Portanto, para mostrarmos que $X \times \{0, 1\}$ é equipotente a X basta provarmos que X é equipotente a Y.

Suponha que X não seja equipotente a Y. Defina $Z = X \setminus Y$. Pelo Lema 6.14 temos que Z é infinito. Logo, $\omega \preceq Z$. Seja W a imagem dessa função injetora em Z. Ou seja, W é um subconjunto enumerável de Z. Pelo Teorema 6.10, $W \times \{0, 1\}$ é equipotente a W. Seja $g : W \times \{0, 1\} \longrightarrow W$ uma função bijetora. Defina $\tilde{f} = f \cup g$. Temos que $\tilde{f}$ é uma função injetora de $(Y \cup W) \times \{0, 1\}$ em $Y \cup W$, contradizendo que f é maximal em $\mathcal{F}$. ■

Lema 6.16 *Se X é infinito e $X = Y \cup Z$, então $X \equiv Y$ ou $X \equiv Z$.*

Demonstração: Primeiro podemos assumir que tanto Y quanto Z são infinitos, pois, se um dele for finito, o lema segue imediatamente do Lema 6.14. Usando o Teorema 5.6 podemos assumir, sem perda de generalidade, que $Y \preceq Z$. Mostraremos que $Z \equiv X$.

Seja $g : Y \longrightarrow Z$ injetora. Defina $f : X \longrightarrow Z \times \{0, 1\}$ como $f(x) = (x, 0)$, se $x \in Z$, e $f(x) = (g(x), 1)$, se $x \in Y \setminus Z$. Claramente f é uma função injetora. Logo, $X \preceq Z \times \{0, 1\}$. Como, pelo Lema 6.15, temos $Z \times \{0, 1\} \equiv Z$, segue que $X \preceq Z$. A outra direção, $Z \preceq X$, é trivial, bastando tomar a função identidade. Portanto, pelo Teorema de Cantor-Schröder-Bernstein temos $X \equiv Z$. ■

Teorema 6.17 *Se X é infinito então $X \times X \equiv X$.*

Demonstração: A demonstração é semelhante à do Lema 6.15. Seja $\mathcal{F}$ o conjunto de todas as funções injetoras tais que existe $Y \subset X$ infinito tal que o domínio da função é $Y \times Y$ e a imagem está contida em Y. Como todo conjunto infinito contém um subconjunto equipotente a ω, e vimos que $\omega \equiv \omega \times \omega$, é fácil observar que $\mathcal{F} \neq \emptyset$ e

que $(\mathcal{F}, \subset)$ satisfaz as hipóteses do Lema de Zorn. Seja $f : Y \times Y \longrightarrow Y$ um elemento maximal de $\mathcal{F}$ em relação à inclusão. Mostremos que $Y \equiv X$. Isso será suficiente para provar que $X \times X \preceq X$ e, como a outra direção é trivial, segue do Teorema de Cantor-Schröder-Bernstein que ambos os conjuntos são equipotentes.

Suponha que X não seja equipotente a Y. Defina $Z = X \setminus Y$. Como $X = Y \cup Z$, pelo Lema 6.16 temos que $X \equiv Z$. Como $Y \preceq X$ temos $Y \preceq Z$. Seja W a imagem de uma função injetora de Y em Z. Temos $Y \equiv W$ e, como $Y \times Y \equiv Y$, concluímos que $Y \times W$, $W \times Y$ e $W \times W$ são todos equipotentes a $Y \times Y$. Logo, pelo Lema 6.16, e usando que Y é infinito, $(Y \times W) \cup (W \times Y) \cup (W \times W)$ é equipotente a $Y \times Y$ e, portanto, a $W \times W$.

Por questão de simplicidade, chamemos o conjunto $(Y \times W) \cup (W \times Y) \cup (W \times W)$ de S e $\tilde{Y}$ o conjunto $Y \cup W$. Seja $g : S \longrightarrow W$ injetora. Observe que $\tilde{Y} \times \tilde{Y} = (Y \times Y) \cup S$. e que ambos os conjuntos são disjuntos. Portanto, definindo $\tilde{f} = f \cup g$ temos que essa é uma função injetora de $\tilde{Y} \times \tilde{Y}$ em $\tilde{Y}$, contradizendo a maximalidade de f. ■

6.5 Conjuntos não-enumeráveis: Teorema de Cantor

O Teorema 6.5 e o Lema 6.7 dizem muito sobre a classificação dos conjunto pela quantidade de elementos. Pelo Lema 6.7 e pelo fato de $(\omega, \subset)$ ser bem ordenado, sabemos dois números naturais diferentes nunca são equipotentes. A equivalência entre os itens (a) e (b) do Teorema 6.5 prova que ω é infinito, e é bom ressaltar que isso vale em ZF, visto que as demonstrações de que (b) implica (c) e (c) implica (a) não usam o axioma da escolha.

A equivalência entre os itens (a) e (b) no Teorema 6.5 nos diz que, em certo sentido, os conjuntos enumeráveis são os "menores infinitos" que existem. Também notemos que os únicos conjuntos que ω domina estritamente são os finitos.

Fica então a questão: existe algum conjunto que não seja nem finito nem enumerável? Essa pergunta foi respondida por Cantor quando mostrou a não-enumerabilidade do conjunto dos números reais. Na verdade, com um pouquinho de trabalho em análise real e usando o Teorema de Cantor-Schröder-Bernstein, não é difícil mostrar que o conjunto dos números reais é equipotente a $\mathcal{P}(\omega)$. O mesmo argumento que prova que existe uma quantidade não-enumerável de números reais pode ser generalizada para mostrar que o conjunto das partes de X é sempre estritamente maior que X, como enunciamos a seguir. Reparem a semelhança do argumento utilizado por Cantor com o Paradoxo de Russell.

Teorema 6.18 (Cantor) *O conjunto $\mathcal{P}(X)$ domina estritamente X.*

Demonstração: A função que associa $x \in X$ a $\{x\} \in \mathcal{P}(X)$ é injetora, pelo axioma da extensão, e prova que $\mathcal{P}(X)$ domina X. Provaremos que X não domina $\mathcal{P}(X)$, mostrando que não existe função sobrejetora de X em $\mathcal{P}(X)$, o que é suficiente, pelo Lema 5.5.

Seja f uma função de X em $\mathcal{P}(X)$. Considere

$$Z = \{x \in X : x \notin f(x)\}.$$

Vamos mostrar que $Z \notin im(f)$.

Suponha que existe $z \in X$ tal que $f(z) = Z$. Se $z \in Z$ então, por definição, $z \notin f(z)$, o que significa que $z \notin Z$. Se $z \notin Z$ isso significa que $z \notin f(z)$, o que implica que $z \in Z$. Chegamos, assim, numa contradição. ■

A pergunta natural a fazer depois de vermos o enunciado do Teorema de Cantor é sobre a existência de alguma coisa intermediária entre ω e $\mathcal{P}(\omega)$, ou, mais geralmente, entre X e $\mathcal{P}(X)$. Essa conjectura de que não existe nada entre ω e $\mathcal{P}(\omega)$ é conhecida como *hipótese do contínuo* e foi colocada por Hilbert no topo dos problemas mais importantes na virada do século XIX para o século XX.

> **Hipótese do contínuo:** Se $\omega \preceq X$ e $X \preceq \mathcal{P}(\omega)$ então X é equipotente a ω ou a $\mathcal{P}(\omega)$.

Esse problema foi provado ser *independente* de ZFC, isto é, não pode ser provado nem refutado utilizando os axiomas usuais de teoria dos conjuntos, a menos que ZFC seja inconsistente. Pelo teorema da completude da lógica de primeira ordem, isso significa que – caso ZFC seja consistente – existem um modelo para teoria dos conjuntos que satisfaz os axiomas de ZFC e a hipótese do contínuo, e outro modelo para teoria dos conjuntos que satisfaz os axiomas de ZFC e a negação da hipótese do contínuo.

Como dissemos, o problema foi postado por Hilbert em 1900 em sua lista dos 22 problemas mais importantes do século XIX. A consistência da hipótese do contínuo só foi mostrada em 1940 por Kurt Gödel, e a consistência da negação da hipótese do contínuo foi provada em 1964 por Paul Cohen.

A seguinte generalização da hipótese do contínuo também foi provada ser independente de ZFC.

> **Hipótese generalizado do contínuo:** Para todos conjuntos infinitos X e Y, se $X \preceq Y$ e $Y \preceq \mathcal{P}(X)$ então $Y \equiv X$ ou $Y \equiv \mathcal{P}(X)$.

Cardinalidade de conjuntos: Embora ainda não tenhamos definido números cardinais, já podemos falar informalmente sobre *cardinalidade* de conjuntos. Dizemos que um conjunto X *tem cardinalidade menor ou igual à cardinalidade de* Y se Y domina X ($X \preceq Y$). Se Y domina estritamente X – isto é, existe uma função injetora de X em Y mas não existe uma função sobrejetora de X em Y – dizemos que X *tem cardinalidade menor que a cardinalidade de* Y. Finalmente, dizemos que X *tem a mesma cardinalidade de* Y se são equipotentes. Essa nomenclatura faz sentido graças aos Teoremas 5.6 e 6.3. Para os conjuntos finitos a cardinalidade pode ser (e assim será) indicada por números naturais. Ou seja, um conjunto equipotente a $n \in \omega$ será dito ter cardinalidade n. Um conjunto enumerável será dito ter cardinalidade ω, ou cardinalidade enumerável. A Hipótese do Contínuo pode ser reescrito como: a primeira cardinalidade não enumerável é a cardinalidade de $\mathcal{P}(\omega)$.

No Capítulo 7 introduziremos os números cardinais, que – a exemplo do conjunto ω para os conjuntos enumeráveis e dos elementos de ω para os conjuntos finitos (esses são casos particulares de cardinais) – serão usados para indicar a cardinalidade dos conjuntos, mesmo os não-enumeráveis.

Exercícios

1. Mostre que um conjunto X é infinito se, e somente se, existe uma boa ordem em X em relação a qual X não possui máximo.

2. Prove que X é infinito se, e somente se, $n \preceq X$, para todo $n \in \omega$.

3. Seja X um conjunto finito tal que todos seus elementos também são finitos. Prove que $\bigcup X$ é finito.

4. Seja X finito, não vazio e não unitário. Prove que X não é equipotente a $X \times X$.

5. Prove que a união e a intersecção de dois conjuntos finitos são finitas.

6. Seja X infinito tal que $x \preceq X$, para todo $x \in X$. Prove que $\bigcup X \preceq X$.

7. Prove que $\mathcal{P}(X)$ é finito se, e somente se, X é finito.

8. Prove que, se A e B são infinitos, então $A \times B$ é equipotente a A ou a B.

9. Complete a prova do Lema 6.15, provando que $\mathcal{F}$ satisfaz as hipóteses do Lema de Zorn.

10. Sejam A e B conjuntos tais que $B \preceq A$, A é infinito e B tem pelo menos dois elementos distintos. Considere X o conjunto de todas as funções de domínio A e imagem contida em B. Prove que X é equipotente a $\mathcal{P}(A)$.

11. Prove que, em ZF, o axioma da escolha é equivalente à seguinte proposição: se X é um conjunto infinito tal que todo $x \in X$ possui pelo menos dois elementos distintos e Y é o conjunto das funções $f : X \longrightarrow \bigcup X$ tais que $f(x) \in x$, para todo $x \in X$, então Y é não enumerável.

Capítulo 7

Ordinais

Os conjuntos bem-ordenados desempenham um papel importante na Teoria dos Conjuntos, especialmente após termos provado – usando o axioma da escolha – que todo conjunto pode ser bem-ordenado. Vimos até agora resultados suficientes para termos uma ideia dessa importância. Reparamos também que conjuntos bem-ordenados diferentes podem ser "parecidos", no sentido de serem ordem-isomorfos. Os ordinais serão construídos como "representantes especiais" dos conjuntos bem-ordenados.

7.1 Axioma da Substituição

Encerramos finalmente a lista de axiomas de ZFC apresentando o Axioma da Substituição, cuja principal utilidade é a construção dos ordinais.

Axioma 10 (da substituição) *Seja $P(x, y)$ uma fórmua e suponha que, para todo x, y, z, tem-se que $P(x, y)$ e $P(x, z)$ implicam $y = z$. Então, para todo conjunto X, existe o conjunto*

$$\{y : \exists x(x \in X \wedge P(x, y)\}.$$

A condição sobre a fórmula P diz que, para todo x, existe *no máximo um* y para o qual $P(x, y)$ vale. Ou seja, P exerce o papel de uma função parcial em X, e o axioma da substituição garante que existe a imagem dessa "função".

Para simplificar a notação, introduzimos alguns símbolos lógicos que serão utilizados neste capítulo. O símbolo $\exists'$ significa "existe no máximo um" e é definido da seguinte forma:

$$\exists' x P \equiv \forall y(P_x^y \to (x = y))$$

O símbolo $\exists!$ significa "existe um único" e é definido como

$$\exists! x P \equiv (\exists x P) \wedge (\exists' x P)$$

Formalmente, utilizando essa notação, o esquema de axiomas da substituição diz que para toda fórmula P em que v não ocorre livre a seguinte fórmula é um axioma:

$$\forall x \exists' y P \to \forall X \exists v \forall y((y \in v) \leftrightarrow \exists x(x \in X \wedge P))$$

O motivo da restrição de v não ocorrer livre em P é o mesmo que foi discutido no axioma da separação: reservamos a variável v para definir o conjunto que o axioma constrói, e a ocorrência livre de v em P poderia resultar em um paradoxo.

Informalmente, podemos dizer que o axioma da substituição garante a existência do conjunto

$$\{F(x) : x \in X\}$$

para um dado conjunto X fixado e uma "fórmula funcional" F. Usando abreviaturas, podemos escrever F omitindo a variável y, assim como fazemos com funções. Por exemplo, podemos representar o conjunto dos números positivos, usando essa notação e o axioma da substituição, como o conjunto

$$\{n^+ : n \in \omega\}.$$

Note que usamos no lugar de $F(x)$ a expressão x^+, que é uma abreviatura (em forma de "fórmula funcional") da fórmula

$$\forall z((z \in y) \leftrightarrow ((z \in x) \vee (z = x))$$

No exemplo acima, o uso do axioma da substituição é desnecessário, porque o mesmo conjunto poderia ser construído usando o axioma da separação. No entanto, veremos exemplos de conjuntos cuja existência depende do axioma da substituição. Isso ocorre quando não conseguimos estabelecer um "conjunto universo" para a "imagem" da fórmula funcional F.

Já a recíproca é verdadeira: podemos suprimir o axioma da separação da lista de axiomas de ZFC, e prová-lo como teorema, a partir do axioma da substituição. Para isso basta tomarmos a fórmula $P(x) \wedge (x = y)$, escolhendo y uma variável. O axioma da substituição nos garante que existe o conjunto

$$\{y : \exists x(x \in X \wedge P(x) \wedge (x = y)\},$$

o que coincide com o conjunto

$$\{x \in X : P(x)\}.$$

7.2 Teorema da Recursão Transfinita

Como vimos no Teorema 4.14, podemos definir uma função cujo domínio é ω recursivamente, isto é, de modo que o valor da função em cada natural depende do valor da mesma nos anteriores. Uma versão semelhante pode ser feita tomando qualquer conjunto bem-ordenado no lugar de ω (o que chamamos de *teorema da recursão transfinita*). Porém, o problema da demonstração do Teorema 4.14 é que, para essa funcionar, precisamos ter pré-fixado um contradomínio. Quando queremos definir uma função recursivamente mas não sabemos se há um conjunto que conterá a imagem dessa função (ou seja, o contra-domínio é, a priori, a *classe de todos os conjuntos*) precisamos usar o axioma da substituição para garantir que essa imagem existe. Claro que, como não temos definido seu contra-domínio, nas hipóteses temos, no lugar de uma função, uma *fórmula funcional*, tal qual enunciada no axioma da substituição.

Teorema 7.1 (recursão transfinita) *Seja $F(x,y)$ uma fórmula tal que $\forall x \exists! y F(x,y)$ seja verdadeira. Seja $(X, \leq)$ um conjunto bem-ordenado. Existe uma única função f cujo domínio é X e que satisfaz, para todo $x \in X$,*

$$F(f|\overleftarrow{x}, f(x)),$$

onde $\overleftarrow{x}$ é definido como $\{y \in X \setminus \{x\} : y \leq x\}$

Demonstração: Considere $G(x,f)$ a seguinte fórmula:

$$(x \in X) \wedge (f \text{ é função}) \wedge (dom(f) = \overleftarrow{x} \cup \{x\}) \wedge \forall y (y \leq x \rightarrow F(f|\overleftarrow{y}, f(y)))$$

Afirmação 1: Para todo $x \in X$, se valem $G(x,f)$ e $G(x,g)$ então $f = g$.

Provemos a afirmação 1 por indução em x. Suponha que a afirmação vale para todo $y \in \overleftarrow{x}$. Se $G(x,f)$ e $G(x,g)$ são verdadeiras, está claro, pela definição de $G(x,f)$, que $G(y, f|(\overleftarrow{y} \cup \{y\}))$ e $G(y, g|(\overleftarrow{y} \cup \{y\}))$ também valem, para todo $y < x$. Logo, pela hipótese de indução, temos, para todo $y < x$.

$$f|(\overleftarrow{y} \cup \{y\}) = g|(\overleftarrow{y} \cup \{y\})$$

Em particular, $f(y) = g(y)$, para todo $y \in \overleftarrow{x}$.

Portanto, das hipóteses $G(x,f)$ e $G(x,g)$ seguem que

$$F(f|\overleftarrow{x}, f(x)) \wedge F(f|\overleftarrow{x}, g(x)),$$

o que implica, pela hipótese sobre $F(x,y)$, que $f(x) = g(x)$, concluindo que $f = g$.

Afirmação 2: Para todos $x, y \in X$, se $y \leq x$ e valem $G(x,f)$ e $G(y,g)$, então $f|(\overleftarrow{y} \cup \{y\}) = g$.

Nessas hipóteses, está claro, pela definição de G, que $G(y, f|(\overleftarrow{x} \cup \{x\})$ é verdadeira. Portanto, da afirmação 1 segue que $f|(\overleftarrow{x} \cup \{x\}) = g$.

Afirmação 3: Para todo $x \in X$ existe f tal que $G(x,f)$.

Suponha, por indução transfinita, que a afirmação seja verdadeira para todo $y < x$. Considere o conjunto

$$Z = \{g : \exists y (y \in \overleftarrow{x} \wedge G(y,g))\}$$

A existência do conjunto Z é assegurada pelo axioma da substituição, lembrando que a afirmação 1 nos grante que G satisfaz as hipóteses do axioma da substituição.

As afirmações 2 e 3 provam que $\bigcup Z$ é uma função cujo domínio é $\overleftarrow{x}$. Pela condição sobre F, sabemos que existe um único t tal que $F(\bigcup Z, t)$ é verdadeira.

Defina

$$f = (\bigcup Z) \cup \{(x, t)\}$$

Como $f|\overleftarrow{x} = \bigcup Z$ e $t = f(x)$, está claro que

$$F(f|\overleftarrow{x}, f(x))$$

Pela definição de Z e pela afirmação 2 temos que $G(y, f|\overleftarrow{y} \cup \{y\})$ vale, para todo $y \in \overleftarrow{x}$. Em particular, para todo $y \in \overleftarrow{x}$ temos

$$F(f|\overleftarrow{y}, f(y))$$

o que conclui a afirmação.

Se $(X, \leq)$ possui máximo, então as afirmações 1 e 3 já provam o teorema, pois basta tomar fa única função tal que $G(x, f)$ é verdadeira, onde x é o máximo de X. Se não possui máximo, temos duas maneiras de encerrar a prova. A primeira, repetimos o argumento usado na afirmação 3, usando o axioma da substituição para definir como f a união de todas as funções g que satisfazem $G(x, g)$, para algum $x \in X$.

Outra maneira é acrescentarmos um máximo ao conjunto $(X, \leq)$, obtendo um conjunto bem-ordenado $(X', \leq')$ onde $X' = X \cup \{x'\}$ e $x \leq' x'$ para todo $x \in X$. Como mostramos que existe uma única f tal que $G(x', f)$ vale, é fácil ver que $f|X$ satisfaz as condições do teorema

■

Uma das aplicações do Teorema 7.1 é a definição do fecho transitivo de um conjunto. Dizemos que y é o *fecho transitivo* de x se y é transitivo, x está contido em y e, para qualquer conjunto transitivo z, se $x \subset z$ então $y \subset z$. Ou seja, o fecho transitivo de x é o menor conjunto transitivo que contém y. Está claro que o fecho transtivio, quando existe, é único. A existência segue do teorema anterior.

Corolário 7.2 *Para todo x existe o fecho transitivo de y.*

Demonstração: Usando o Teorema 7.1, defina f de domínio ω tal que $f(0) = x$ e $f(n^+) = \bigcup f(n)$. Deixamos como exercício ao leitor completar os detalhes da demonstração da existência da função f, utilizando uma fórmula F adequada. Basta lembrar que, para provar a recursão finita (como é utilizada aqui) a partir da recursão transfinita, usamos o fato de que todo número natural diferente de 0 é sucessor de alguém.

Tome $y = \bigcup im(f)$. Mostraremos que y é o fecho transitivo de x.

Está claro que $x \subset y$, pois $x \in im(f)$. Se $z \in y$, existe $n \in \omega$ tal que $z \in f(n)$. Logo, $z \subset \bigcup f(n) = f(n^+)$. Portanto, $z \subset y$.

Agora suponha que existe um conjunto transitivo z tal que $x \subset z$. Vamos mostrar que $y \subset z$. Para isso, basta mostrar que $f(n) \subset z$, para todo $n \in \omega$. Mas notemos que, pela transitividade, se $w \subset z$ temos $\bigcup w \in z$. Assim, como $x \subset z$, por indução provamos que $f(n) \subset z$, para todo $n \in \omega$.

■

7.3 Ordinais

Dentre os conjuntos bem-ordenados que vimos até agora, podemos notar que o conjunto ω, ordenado com a inclusão, possui propriedades bastante especiais, operacionalmente úteis. Primeiro porque ω é transitivo, isto é, todos seus elementos são também subconjuntos de ω. Segundo porque a relação de pertinência coincide com a inclusão própria. Isso implica que $\overleftarrow{n}$ (o conjunto dos elementos menores que n) coincide com o próprio conjunto n, o que frequentemente facilita a notação, como no caso do teorema da recursão.

Observamos que os elementos de ω satisfazem essas mesmas propriedades, e o mesmo vale para ω^+ (lembrando que ω^+ é o conjunto $\omega \cup \{\omega\}$, que também é bem-ordenado pela inclusão. Podemos construir outros desses "conjuntos bem-ordenados especiais" tomando $(\omega^+)^+$, $((\omega^+)^+)^+$ e assim por diante. O próximo passo seria tomar a união de todos esses "sucessores" de ω, mas isso exige uma atenção maior, e é justamente nesse ponto que entra o uso do axioma da substituição, como veremos em breve.

Esses conjuntos que são transitivos e bem-ordenados com a relação de pertinência são chamados de *ordinais*, e costumam ser representados por letras gregas.

Definição 7.3 Dizemos que um conjunto α é um *ordinal* se satisfaz as seguintes condições:

- se $\beta \in \alpha$ então $\beta \subset \alpha$ (α é *transitivo*);
- a relação $\{(x, y) \in \alpha^2 : (x \in y) \vee (x = y)\}$ é uma boa ordem em α (α é *bem-ordenado pela pertinência*).

Do Teorema 4.11 segue imediatamente o seguinte resultado:

Lema 7.4 *ω é um ordinal.*

Sempre quando falarmos de um ordinal estamos nos referindo, implicitamente, da ordem dada na definição, que, como veremos pelos itens (a) e (c) do próximo teorema, coincide com a ordem da inclusão.

Do lema anterior e do item (a) do teorema seguinte segue que os números naturais, são, eles próprios, ordinais.

Teorema 7.5 *Seja α um ordinal.*

(a) *Se $\beta \in \alpha$ então β é um ordinal;*

(b) *α^+ é um ordinal;*

(c) *Se $\beta \subset \alpha$, $\beta \neq \alpha$ e β é transitivo, então $\beta \in \alpha$;*

(d) *Se $\beta \in \alpha$ então $\beta^+ \in \alpha$ ou $\beta^+ = \alpha$.*

Demonstração: Seja $\beta \in \alpha$. Pela transitividade de α, temos $\beta \subset \alpha$ e, portanto, β é bem-ordenado pela pertinência.

Vamos mostrar que β é transitivo. Sejam $\gamma \in \beta$ e $\delta \in \gamma$. Pela transitividade de α temos $\gamma \in \alpha$ e $\delta \in \alpha$. Pela boa ordem de α temos $\delta \in \beta$ ou $\delta = \beta$ ou $\beta \in \delta$. Mas as duas últimas possibilidades contradizem o axioma da regularidade, sobrando, então, que $\delta \in \beta$.

Concluímos, assim, a prova de (a). Provemos (b). Primeiro vejamos que α^+ é transitivo. De fato, se $\beta \in \alpha^+$, então $\beta \in \alpha$ ou $\beta = \alpha$. Em ambos os casos (usando a transitividade de α no primeiro) temos $\beta \subset \alpha$. Mas $\alpha \subset \alpha^+$, concluindo que $\beta \subset \alpha^+$.

Agora mostraremos que α^+ é bem ordenado pela pertinência. Seja $A \subset \alpha^+$ não-vazio. Se $A = \{\alpha\}$, então claramente α é o mínimo de A. Senão, $A \cap \alpha$ é não-vazio e está contido em α. Seja β o mínimo de $A \cap \alpha$. Temos que $\beta \in \alpha$, logo, β é o mínimo também de A.

Provemos o item (c). Seja γ o mínimo do conjunto $\{x \in \alpha : x \notin \beta\}$, que é não-vazio pela hipótese $\beta \neq \alpha$. Mostraremos que $\gamma = \beta$. De fato, se $\delta \in \gamma$, então, pela transitividade de α, $\delta \in \alpha$. Mas, pela definição de γ, isso significa que $\gamma \in \beta$, pois γ é menor (na ordem da pertinência) que o mínimo dos elementos de α que não pertencem a β.

Reciprocamente, se $\delta \in \beta$, como $\beta \subset \alpha$, temos que $\delta \in \alpha$. Como α é bem-ordenado, temos que $\delta \in \gamma$ ou $\delta = \gamma$ ou $\gamma \in \delta$. Mas, se $\delta = \gamma$ ou $\gamma \in \delta$, pela transitividade de β isso implica que $\gamma \in \beta$, contradizendo a definição de γ.

Isso prova que $\beta = \gamma$ e, portanto, $\beta \in \alpha$, provando o item (c).

O item (d) segue facilmente dos itens anteriores. De fato, se $\beta \in \alpha$, pela transitividade de α temos que $\beta \subset \alpha$ e, portanto, $\beta^+ \subset \alpha$. Pelo item (a), β é um ordinal, e, por (b), β^+ também é um ordinal. Em particular, β^+ é transitivo. Logo, se $\beta^+ \neq \alpha$, do item (c) segue que $\beta^+ \in \alpha$.

■

O Teorema 7.5, especialmente os itens (b) e (d), motiva a seguinte definição:

Definição 7.6 Dizemos que um ordinal α é um *ordinal sucessor* se existe β tal que $\alpha = \beta^+$. Caso contrário, dizemos que α é um *ordinal limite*.

Do Teorema 7.5, item (d), segue que α é um ordinal limite se, e somente se, $\beta^+ \in \alpha$, sempre que $\beta \in \alpha$.

Teorema 7.7 *Para todos ordinais α e β temos $\alpha \in \beta$, $\alpha = \beta$ ou $\beta \in \alpha$.*

Demonstração: Primeiro provaremos que $\alpha \cap \beta$ é um ordinal. Para isso basta ver que $\alpha \cap \beta$ é transitivo, pois ser bem-ordenado pela pertinência segue de $\alpha \cap \beta \subset \alpha$. Se $\gamma \in \alpha \cap \beta$ então, pela transitividade de α e β temos $\gamma \subset \alpha$ e $\gamma \subset \beta$, o que implica que $\gamma \subset \alpha \cap \beta$.

Como $\alpha \cap \beta \subset \alpha$ e $\alpha \cap \beta \subset \beta$, do Teorema 7.5, item (c), segue que, se $\alpha \cap \beta$ não for igual a α ou a β temos $\alpha \cap \beta \in \alpha$ e $\alpha \cap \beta \in \beta$. Nesse caso, $\alpha \cap \beta \in \alpha \cap \beta$, contradizendo o axioma da regularidade.

Se $\alpha \cap \beta = \alpha$, isso significa que $\alpha \subset \beta$, o que implica, pelo Teorema 7.5 item (c), que $\alpha \in \beta$ ou $\alpha = \beta$. Se $\alpha \cap \beta = \beta$, temos que $\beta \subset \alpha$ e, portanto, $\beta \in \alpha$ ou $\beta = \alpha$, concluindo a prova do teorema.

■

Corolário 7.8 *Um ordinal α é infinito se, e somente se, $\alpha = \omega$ ou $\omega \in \alpha$.*

Demonstração: Se $\omega \in \alpha$ então $\omega \subset \alpha$. Logo, como subconjunto de conjunto finito é finito (consequência do Lema 6.7) e ω é infinito, temos que α é infinito.

Reciprocamente, se α é infinito então $\alpha \notin \omega$. Pelo Teorema 7.7 isso implica que $\alpha = \omega$ ou $\omega \in \alpha$, provando o corolário.

■

Teorema 7.9 *Se A é um conjunto de ordinais então $\bigcup A$ é um ordinal.*

Demonstração: Mostremos que $\bigcup A$ é transitivo. Seja $x \in \bigcup A$. Isso significa que existe $\alpha \in A$ tal que $x \in \alpha$. Mas, como, por hipótese, α é transitivo, temos que $x \subset \alpha$ e, portanto, como $\alpha \subset \bigcup A$, temos $x \subset \bigcup A$.

Falta provar que $\bigcup A$ é bem ordenado pela pertinência. Seja $S \subset \bigcup A$ não-vazio. Tome $\alpha \in A$ tal que $S \cap \alpha \neq \emptyset$. Como α é bem-ordenado pela pertinência, existe $\gamma \in S \cap \alpha$ tal que $\gamma \in \delta$, para todo $\delta \in S \cap \alpha$ diferente de γ.

Seja $\beta \in S$. Pelo Teorema 7.7 temos que $\beta \in \alpha$, $\beta = \alpha$ ou $\alpha \in \beta$. No primeiro e segundo caso temos $\gamma \in \beta$ ou $\gamma = \beta$ pela minimalidade de γ em $S \cap \alpha$. No segundo caso, temos $\gamma \in beta$, pois $\gamma \in \alpha$. No terceiro caso, de $\gamma \in \alpha$ e $\alpha \in \beta$ segue que $\gamma \in \beta$, pela transitividade de β.

Logo, β é o mínimo de S, provando que $\bigcup A$ é bem-ordenado. ■

Corolário 7.10 *Não existe o conjunto de todos os ordinais.*

Demonstração: Suponha que exista A o conjunto de todos os ordinais. Tome $\alpha = \bigcup A$, que é ordinal, pelo Teorema 7.9. Pelo Teorema 7.5 temos que α^+ é um ordinal, logo $\alpha^+ \in A$. Como $\alpha \in \alpha^+$, temos $\alpha \in \bigcup A$. Logo, $\alpha \in \alpha$, contradizendo o axioma da regularidade. ■

Corolário 7.11 *Um conjunto transitivo de ordinais é um ordinal.*

Demonstração: Seja X um conjunto transitivo formado por ordinais. Mostraremos que X é bem-ordenado pela relação de pertinência.

Pela transitividade de X temos que $\bigcup X \subset X$ e, pelo Teorema 7.9, $\bigcup X$ é bem-ordenado pela pertinência.

Mostremos que $X \setminus \bigcup X$ é vazio ou unitário. De fato, suponha que existem x e y distintos que pertencem a X mas não pertencem a $\bigcup X$, pelo Teorema 7.7 temos $x \in y$ ou $y \in x$. Mas $x \in y$ implica que $x \in \bigcup X$, e $y \in x$ implica que $y \in \bigcup X$, contradizendo que ambos não pertencem a $\bigcup X$.

Isso também mostra que, se $x \in X \setminus \bigcup X$, então $y \in x$, para todo $y \in \bigcup X$.

Portanto, se existe $x \in X \setminus \bigcup X$, temos que $X = \bigcup X \cup \{x\}$, que é bem-ordenado pela pertinência, sendo x seu elemento máximo. Se não existe $x \in X \setminus \bigcup X$, temos que $X = \bigcup X$, que é um ordinal e, portanto, é bem-ordenado. ■

Falaremos agora sobre isomorfismos de ordem e de ordinais. Começamos provando que ordinais diferentes não são isomorfos.

Lema 7.12 *Seja f um isomorfismo entre dois ordinais α e β. Então $\alpha = \beta$ e f é a identidade.*

Demonstração: Seja $f : \alpha \longrightarrow \beta$ um isomorfismo de ordem. Podemos assumir que $\alpha \leq \beta$, já que a existência de um isomorfismo de α em β é equivalente à existência de um isomorfismo de β em α. Provamos, por indução, que $f(\gamma) = \gamma$, para todo $\gamma \in \alpha$. De fato, se isso não for verdade, seja γ o menor elemento de α tal que $f(\gamma) \neq \gamma$. Temos que $f(\gamma) < \gamma$ ou $f(\gamma) > \gamma$.

No primeiro caso, por hipótese, teríamos $f(f(\gamma)) = f(\gamma)$, contradizendo que f é injetora.

No segundo caso, como f é sobrejetora, existe $\gamma' \in \alpha$ tal que $f(\gamma') = \gamma$. Mas, pela hipótese, não podemos ter $\gamma' < \gamma$, pois teríamos $f(\gamma') = \gamma'$. Logo, $\gamma < \gamma'$ e $f(\gamma') < f(\gamma)$, contradizendo que f é um isomorfismo de ordem.

Concluímos, portanto, que f é a identidade e que, portanto, $\alpha = \beta$. ■

O próximo teorema é uma das principais aplicações do axioma da substituição, e mostra que toda boa-ordem pode ser representada por um ordinal.

Teorema 7.13 (Princípio da contagem) *Para todo conjunto bem-ordenado $(X, \leq)$ existe um único ordinal α tal que $(\alpha, \subset)$ é isomorfo a $(X, \leq)$.*

Demonstração: A unicidade segue do Lema 7.12, uma vez que composição e inversa de isomorfismos são isomorfismos e, portanto, se um mesmo conjunto bem-ordenado for igual a dois ordinais, eles serão isomorfos entre si e, consequentemente, iguais. Mostraremos a existência, usando o Teorema da Recursão Transfinita.

Seja $F(f, \alpha)$ a fórmula

$$((f \text{ é função}) \rightarrow (\alpha = im(f))) \wedge ((f \text{ não é função}) \rightarrow (\alpha = \emptyset))$$

Está claro que para todo f existe um único α tal que $F(f, \alpha)$. Portanto, podemos usar o Teorema 7.1 para determinar uma função f com domínio X tal que, para todo $x \in X$,

$$F(f|\overset{\leftarrow}{x}, f(x))$$

Mostraremos que f é um isomorfismo de X em sua imagem, e que a imagem de f é um ordinal.

A ideia da construção dessa função f através da fórmula F lembra o conceito intuitivo de ordinal, como "conjunto dos ordinais menores do que ele".

Afirmação 1: Para todo $x \in X$, $f(x)$ é um ordinal.

Provaremos essa afirmação por indução em x. Pelo Corolário 7.11 basta mostrarmos que $f(x)$ é transitivo.

Sejam $z \in f(x)$ e $w \in z$. Como $f(x) = im(f|\overleftarrow{x})$, existe $y < x$ tal que $z \in f(y)$. Pela hipótese de indução, $f(y)$ é ordinal e, portanto, transitivo. Logo $w \in f(y)$. Mas, como $f(y) = im(f|\overleftarrow{y})$, temos que $f(y) \subset f(x)$, provando que $w \in f(x)$.

Afirmação 2: A imagem de f é um ordinal.

Pela afirmação 1, $im(f)$ é um conjunto de ordinais. Analogamente à prova da afirmação 1, podemos mostrar que $im(f)$ é transitiva e, portanto, um ordinal.

Afirmação 3: A função f é um isomorfismo de X em sua imagem.

Se $y < x$, então $f(y) \in im(f|\overleftarrow{x})$ e, portanto, $f(y) \in f(x)$. Reciprocamente, se $f(y) \in f(x)$, isso implica que $f(y) \neq f(x)$ e que $f(x) \notin f(y)$, o que não permite que tenhamos $x \leq y$. Como boas ordens são também ordens totais, isso implica que $y < x$. Disso segue também a injetividade de f, e a sobrejetividade é imediata, pois estamos considerando f em relação à sua imagem. Isso conclui o teorema.

■

7.4 Aritmética dos ordinais

A aritmética de números ordinais estende a definição de aritmética de números inteiros. A diferença fundamental é que, enquanto nos números inteiros, precisamos apenas definir as operações para n^+ uma vez que essas estão definidas para n, na aritmética ordinal precisamos cuidar dos ordinais limites.

Soma de ordinais

Iniciamos com a definição de soma de números ordinais. Mas antes, vamos estabelecer uma notação. Se A é um conjunto de ordinais, chamamos de *supremo* de A – que será denotado por $supA$ – o menor ordinal que é maior ou igual a todo elemento de A. É fácil verificar que o supremo de A sempre existe e é igual a $\bigcup A$. Preferimos, no entanto, em algumas ocasiões, a notação $supA$ quando essa condizer melhor com o contexto utilizado no momento. Particularmente, quando estamos interessados mais na estrutura de ordem dos ordinais do que da sua construção conjuntística. Ou seja, embora $\alpha \in \beta$ e $\alpha < \beta$representam a mesma coisa, quando se trata de ordinais preferimos a notação $<$ quando não nos preocupamos em relembrar a forma como foram construídos os ordinais. Nessa mesma ocasião, preferimos a notação do supremo à da união.

Definição 7.14 Dados dois ordinais α e β definimos a soma $\alpha + \beta$ através das seguintes regras:

- $\alpha + 0 = \alpha$.

- $\alpha + (\beta^+) = (\alpha + \beta)^+$.

- Se β é ordinal limite, então $\alpha + \beta = sup\{\alpha + \gamma : \gamma < \beta\}$.

A definição acima precisa ser melhor explicada. Primeiro notamos que, como não existe o conjunto de todos os ordinais, a soma de ordinais não pode ser considerada uma função, como no caso dos números naturais. Em vez disso, deve ser pensada como mais um símbolo funcional binário ou, mais precisamente, como uma fórmula de três variáveis livres que diz "x, y e z são ordinais tais que z é a soma de x e y".

Precisamos, portanto, definir uma fórmula $F(x, y, z)$ que satisfaz as condições da Definição 7.14, tratando expressões da forma "$x + y = z$" como abreviaturas para $F(x, y, z)$. Além disso, precisamos provar que a cada x e y ordinais existe um único ordinal z tal que $F(x, y, z)$ é verdadeira, e que a fórmula F é única, no sentido de que, se $F'(x, y, z)$ também satisfaz a Definição 7.14,, então vale $F(x, y, z) \longleftrightarrow F'(x, y, z)$.

Para cada α e β ordinais, pelo Teorema 7.1 existe uma única única função

$$f_\alpha : \beta^+ \longrightarrow im(f_\alpha)$$

satisfazendo

$$G(f_\alpha|\gamma, f_\alpha(\gamma)),$$

para cada $\gamma < \beta^+$, onde $G(f, \gamma)$ é a fórmula dada pelas seguintes condições:

- Se $dom(f) = 0$ então $\gamma = \alpha$.

- Se $dom(f) = \delta^+$ então $\gamma = (f(\delta))^+$.

- Se $dom(f)$ é um ordinal limite δ, então $\gamma = sup\{f(\xi) : \xi < \delta\}$.

Definimos $\alpha + \beta$ como $f_\alpha(\beta)$. Ou seja, $F(x, y, z)$ é a fórmula "x, y e z são ordinais e existe uma função f satisfazendo $f(y) = z$ e $G(f|w, f(w))$ para todo $w \in y^+$". Deixamos ao leitor provar que essa fórmula é de fato funcional na classe dos ordinais (isto é, dados ordinais x e y existe um único ordinal z satisfazendo $F(x, y, z)$.

Lema 7.15 *Para todos ordinais α, β e γ valem*

(a) $\alpha + 1 = \alpha^+$;

(b) $\gamma + \beta < \gamma + \alpha$ *se, e somente se,* $\beta < \alpha$;

(c) $(\alpha + \beta) + \gamma = \alpha + (\beta + \gamma)$;

(d) *Se* $\beta < \alpha$ *então existe um único ordinal* ξ *tal que* $\beta + \xi = \alpha$;

(e) *Existe um único ordinal limite* δ *e um único natural* $n \in \omega$ *tais que* $\alpha = \delta + n$.

Demonstração: O item (a) é imediato, pois $\alpha + 1 = \alpha + 0^+ = (\alpha + 0)^+ = \alpha^+$.

Para o item (b) mostraremos por indução em α que $\beta < \alpha$ implica $\gamma + \beta < \gamma + \alpha$. Para $\alpha = 0$ é trivial. Suponha que a implicação vale para α e todo $\beta < \alpha$. Se $\beta < \alpha^+$, temos duas opções: $\beta < \alpha$ ou $\beta = \alpha$. Se $\beta < \alpha$, pela hipótese indutiva temos

$$\gamma + \beta < \gamma + \alpha < (\gamma + \alpha)^+ = \gamma + \alpha^+$$

Agora assumimos que β é um ordinal limite e que a afirmação vale para todo $\delta < \alpha$. Seja $\beta < \alpha$. Como α é limite, $\beta^+ < \alpha$. Portanto, pela hipótese indutiva,

$$\gamma + \beta < \gamma + \beta^+ \leq sup\{\gamma + \delta : \delta < \alpha\} = \gamma + \alpha$$

A recíproca do que acabamos de mostrar, conforme consta no enunciado do item (b), segue da comparabilidade dos ordinais. Se $\gamma + \beta < \gamma + \alpha$, obviamente não podemos ter $\alpha = \beta$. Logo, se não vale $\beta < \alpha$, vale $\alpha < \beta$, o que, segundo mostramos, implica que $\gamma + \alpha < \gamma + \beta$, contradizendo que a desigualdade inversa vale.

Mostremos o item (c) por indução em γ. Se a associatividade vale para um γ fixo, e todos α e β, temos

$$\alpha + (\beta + \gamma^+) = \alpha + (\beta + \gamma)^+ = (\alpha + (\beta + \gamma))^+ = ((\alpha + \beta) + \gamma)^+$$

$$= (\alpha + (\beta + \gamma))^+ = \alpha + (\beta + \gamma)^+ = \alpha + (\beta + \gamma^+)$$

Para o caso limite, suponha, por hipótese indutiva, que $(\alpha + \beta) + \delta = \alpha + (\beta + \delta)$, para todos α e β e todo $\delta < \gamma$, onde γ é um ordinal limite. Temos que

$$(\alpha + \beta) + \gamma = sup\{(\alpha + \beta) + \delta : \delta < \gamma\} = sup\{\alpha + (\beta + \delta) : \delta < \gamma\}$$

Para concluir o item (c) basta mostrarmos que

$$(*) \qquad sup\{\alpha + (\beta + \delta) : \delta < \gamma\} = sup\{\alpha + \xi : \xi < \beta + \gamma\}$$

uma vez que a expressão do lado direito é, por definição, $\alpha + (\beta + \gamma)$.

Para mostrar $(*)$ basta mostrar que cada elemento de um dos dois conjuntos é limitado por um elemento do outro. Se $\xi = \beta + \delta$, para $\delta < \gamma$, então, pelo item (b), $\xi < \beta + \gamma$, o que mostra que o supremo do lado esquerdo, em $(*)$, é menor que o do lado direito. Por outro lado, se $\xi < \beta + \gamma$, como γ é limite, temos que existe $\delta < \gamma$ tal que $\xi \leq \beta + \delta$. Logo, pelo item (b), $\alpha + \xi \leq \alpha + (\beta + \delta)$, provando $(*)$ e concluindo o item (c).

O item (d) será provado por indução em α. Suponha que o item (d) vale para α e todo $\beta < \alpha$. Seja $\beta < \alpha^+$. Se $\beta = \alpha$, pelo item (a) temos $\alpha = \beta + 1$. Se $\beta < \alpha$, temos, pela hipótese de indução, que existe ξ tal que $\beta + \xi = \alpha$. Logo, $\beta + \xi^+ = \alpha^+$.

Suponha agora que α é um ordinal limite e que, para todos $\beta < \delta < \alpha$, existe um único η tal que $\beta + \eta = \delta$. Deixamos como exercício ao leitor provar que $\eta \leq \delta$. Defina

$$\xi = sup\{\eta : \beta + \eta < \alpha\}$$

Primeiro notemos que ξ é um ordinal limite. De fato, se ξ fosse um ordinal sucessor, ξ seria o maior ordinal tal que $\beta + \xi < \alpha$, absurdo, pois, pelo fato de α ser ordinal limite, $\beta + \xi^+ = (\beta + \xi)^+ < \alpha$. Também é imediato que $\xi \neq 0$, uma vez que $\beta + 1 < \alpha$.

Mostraremos que

$$(**) \qquad \beta + \xi = \alpha$$

Para isso, seja $\delta < \beta + \xi$. Tomr $\eta' < \xi$ tal que $\delta \leq \beta + \eta'$. Pela definição de ξ, existe $\eta < \xi$ tal que $\eta' \leq \eta$ e $\beta + \eta < \alpha$. Pelo item (b), $\delta \leq \beta + \eta$ e, portanto, $\delta < \alpha$.

Reciprocamente, suponha que $\delta < \alpha$. Se $\delta \leq \beta$, temos $\delta < \beta + \xi$, já que $\xi \neq 0$. Se $\beta < \delta$, pela hipótese indutiva existe η tal que $\beta + \eta = \delta$, o que significa que $\eta < \xi$ e, portanto, $\beta + \eta < \beta + \xi$.

Isso prova que $\beta + \xi = \alpha$. A unicidade de ξ é imediata do item (b).

Resta mostrarmos o item (e). Defina

$$\delta = sup\{\eta \leq \alpha : \eta \text{ é ordinal limite } \}$$

Claramente $\delta \leq \alpha$. Se $\delta = \alpha$, temos que $\delta + 0 = \alpha$, o que satisfaz o item (e). Se $\delta < \alpha$, existe um único ordinal n tal que $\delta + n = \alpha$. Vamos provar que $n \in \omega$.

Se $n \notin \omega$, temos $n = \omega$ ou $\omega < n$. Em ambos casos temos $\delta + \omega \leq \delta + n = \alpha$, pelo item (b). Como $\delta < \delta + \omega$ e $\delta + \omega$ é um ordinal limite (verifique!), isso contradiz a definição de δ como supremo dos ordinais limites menores que α.

Para a unicidade, suponha que $\alpha = \delta' + m$, onde δ' é um ordinal limite e $m \in \omega$. Vamos provar que $\delta = \delta'$, isto é, δ' também é o supremo dos ordinais limites menores ou iguais a α. De fato, está claro que $\delta' \leq \alpha$. Seja η um ordinal limite tal que $\delta' < \eta$. Mostraremos que $\alpha < \eta$. Para isso, basta mostrarmos por indução que $\delta' + k < \eta$, para todo $k \in \omega$. Para $k = 0$ temos $\delta' + 0 = \delta' < \eta$. Se $\delta' + k < \eta$, temos que $\delta' + k^+ = (\delta' + k)^+$, que é estritamente menor que η, uma vez que η é um ordinal limite. Portanto, $\alpha < \eta$ e $\delta' = \delta$.

Pelo item (b) é imediato que $m = n$, concluindo a prova do item (e). ■

Notemos que a soma de ordinais não é comutativa. Temos, por exemplo, $2 + \omega = \omega$, que é diferente de $\omega + 2$.

Pelo item (b) do lema e pela propriedade de comparabilidade dos ordinais, vale a lei do cancelamento à esquerda, ou seja, se $\gamma + \beta = \gamma + \alpha$ então $\beta = \alpha$. Observe que, pelo exemplo dado no parágrafo acima, sabemos que não vale o cancelamento à direita. Em particular, o item (b) não vale se γ estiver à direita de β e α.

O item (d) nos fornece uma espécie de subtração de ordinais, enquanto o item (e) nos permite decompor qualquer ordinal de uma forma interessante, que nos ajuda não apenas a compreender a noção intuitiva de ordinais, mas também é útil em algumas demonstrações, como veremos posteriormente.

O próximo teorema nos dá uma caracterização bastante interessante de soma de ordinais.

Teorema 7.16 *Sejam $(X_1, \leq_1)$ e $(X_2, \leq_2)$ dois conjuntos bem-ordenados isomorfos aos ordinais α e β, respectivamente. Suponha que $X_1 \cap X_2 = \emptyset$ e considere $X = X_1 \cup X_2$ ordenado com os elementos de X_1 antes dos elementos de X_2. Isto é, definimos em X a ordem*

$$\leq = \leq_1 \cup \leq_2 \cup X_1 \times X_2$$

Então $(X, \leq)$ é isomorfo a $\alpha + \beta$.

Demonstração: Vamos provar o teorema por indução em β. Para $\beta = 0$ é trivial, pois teremos $X_2 = \emptyset$. Suponha que, para um determinado β, o teorema vale para qualquer $(X_1, \leq_1)$ isomorfo a algum ordinal α e qualquer $(X_2, \leq_2)$ isomorfo a β. Sejam $(X_1, \leq_1)$ e $(X_2, \leq_2)$ isomorfos a α e β^+, respectivamente. Isso significa que X_2 possui um máximo y e, se tomarmos,$X_2' = X_2 \setminus \{y\}$, temos que X_2' é isomorfo a β. Pela hipótese de indução, $X_1 \cup X_2'$ é isomorfo a $\alpha + \beta$. Mas y é claramente o máximo de $X_1 \cup X_2$, que é, portanto, isomorfo a $(\alpha + \beta)^+$, que é igual a $\alpha + \beta^+$.

Para o caso β limite, supomos que o teorema vale sempre que X_1 é isomorfo a algum ordinal α e X_2 é isomorfo a algum ordinal $\gamma < \beta$.

Seja δ o único ordinal isomorfo a $X_1 \cup X_2$ com a ordem acima descrita (que é claramente uma boa ordem). Seja f o isomorfismo entre $X_1 \cup X_2$ e δ.

Como β é limite, X_2 não tem máximo. Logo, δ é um ordinal limite. Para cada $x \in X_2$, a restrição de f a $X_1 \cup \overleftarrow{x}$ é um isomorfismo sobre $f(x)$. Pela hipótese indutiva e pelo Lema 7.12 temos $f(x) = \alpha + \gamma$, onde $\gamma < \beta$ é isomorfo a $\overleftarrow{x}$. Logo,

$$\delta = sup\{\alpha + \gamma : \gamma < \beta\} = \alpha + \beta,$$

provando o teorema. ■

Multiplicação de ordinais

Agora definiremos a multiplicação de ordinais e, desta vez, deixamos a formalização por conta do leitor.

Definição 7.17 Dados dois ordinais α e β definimos o produto $\alpha \cdot \beta$ através das seguintes regras:

- $\alpha \cdot 0 = 0$.
- $\alpha \cdot (\beta^+) = (\alpha \cdot \beta) + \alpha$.
- Se β é ordinal limite, então $\alpha \cdot \beta = sup\{\alpha \cdot \gamma : \gamma < \beta\}$.

Como de costume, denotamos o produto $\alpha \cdot \beta$ por $\alpha\beta$. Para vários resultados do Lema 7.15 vale algo similar na multiplicação de ordinais. Verificaremos o análogo para multiplicação do item (b) do Lema 7.15.

Lema 7.18 *Se $\gamma \neq 0$ e $\beta < \alpha$ então $\gamma\beta < \gamma\alpha$.*

Demonstração: Provaremos por indução em α. Para $\alpha = 0$ o lema vale por vacuidade. Suponha que o lema é verdadeiro para α. Seja $\beta < \alpha^+$. Temos $\beta \leq \alpha$. Por hipótese de indução,

$$\gamma\beta \leq \gamma\alpha < \gamma\alpha + \gamma = \gamma\alpha^+.$$

Seja α um ordinal limite e suponha que o lema vale para todo $\delta < \alpha$, no lugar de α. Seja $\beta < \alpha$. Temos que $\beta^+ < \alpha$ e, portanto,

$$\gamma\beta < \gamma\beta + \gamma = \gamma\beta^+ \leq sup\{\gamma\xi : \xi < \alpha\} = \gamma\alpha$$

■

Se, por um lado, a soma de ordinais representa a união disjunta de conjuntos bem-ordenados, a multiplicação representa o produto cartesiano, com a ordem antilexicográfica, conforme o seguinte teorema.

Teorema 7.19 *Sejam $(X, \leq_1)$ e $(Y, \leq_2)$ conjuntos bem-ordenados isomorfos a α e β, respectivamente. Considere $X \times Y$ ordenado pela ordem antilexicográfica, isto é,*

$$(x, y) \leq (x', y') \longleftrightarrow (y <_2 y' \vee (y = y' \wedge x \leq_1 x')).$$

Então $(X \times Y, \leq)$ é isomorfo a $\alpha\beta$.

Demonstração: Podemos assumir que X e Y são os próprios ordinais α e β. Também assumimos que nenhum deles é o ordinal 0, visto que, nesse caso, $\alpha \times \beta$ e $\alpha\beta$ são ambos o conjunto vazio. Provaremos por indução em β que $\alpha \times \beta$, com a ordem antilexicográfica, é isomorfo a $\alpha\beta$.

Defina uma função f de domínio $\alpha \times \beta$ como

$$f(\xi, \eta) = (\alpha\eta) + \xi$$

Como fazemos com os números naturais, escreveremos simplesmente $\alpha\eta + \xi$ no lugar de $(\alpha\eta) + \xi$.

Afirmação 1: f é um isomorfismo sobre a imagem.

Suponha que $(\xi, \eta) < (\xi', \eta')$ na ordem antilexicográfica. Mostraremos que $\alpha\eta + \xi < \alpha\eta' + \xi'$.

Temos dois casos. No primeiro, $\eta = \eta'$ e $\xi < \xi'$, o que implica imediatamente, pelo Lema 7.15 item (b), que $\xi < \xi'$.

Vamos analisar o caso em que $\eta < \eta'$. Pelos Lema 7.15 item (b) e pelo Lema 7.18 temos

$$\alpha\eta + \xi < \alpha\eta + \alpha = \alpha\eta^+ \leq \alpha\eta' \leq \alpha\eta' + \xi'.$$

Afirmação 2: A imagem de f é $\alpha\beta$.

Provaremos a afirmação por indução em β. Suponha que $\beta = \gamma^+$ e que f restrita a $\alpha \times \gamma$ é sobrejetora em $\alpha\gamma$.

Seja $\delta < \alpha\beta$. Se $\delta < \alpha\gamma$, pela hipótese indutiva existe $(\xi, \eta) \in \alpha \times \gamma$ tal que $\delta = \alpha\eta + \xi$, provando que δ pertence à imagem de f. Se $\alpha\gamma \leq \delta$, pelo item (d) do Lema 7.15 existe ξ (possivelmente 0) tal que $\delta = \alpha\gamma + \xi$. Como $\delta < \alpha\beta = \alpha\gamma + \alpha$, pelo Lema 7.15, item (b), temos que $\xi < \alpha$.

Suponhamos que β é um ordinal limite e que f restrita a $\alpha \times \gamma$ é sobrejetora em $\alpha\gamma$, para todo $\gamma < \beta$.

Seja $\delta < \alpha\beta$. Pela definição de $\alpha\beta$ existe $\gamma < \beta$ tal que $\delta \leq \alpha\gamma$. Em particular, $\delta < \alpha\gamma^+$ e, como β é limite, $\gamma^+ < \beta$. Pela hipótese indutiva existem $\xi \in \alpha$ e $\eta \in \beta$ tais que $\delta = \alpha\eta + \xi$, provando a afirmação e o teorema. ■

Corolário 7.20 *Para todos ordinais α, β e γ vale $(\alpha\beta)\gamma = \alpha(\beta\gamma)$.*

Demonstração: Existe uma bijeção natural entre $(\alpha \times \beta) \times \gamma$ e $\alpha \times (\beta \times \gamma)$. É fácil verificar que essa bijeção preserva a ordem antilexicográfica. Logo, pelo Teorema 7.19 e pelo Lema 7.12 temos $(\alpha\beta)\gamma = \alpha(\beta\gamma)$. Deixamos os detalhes para o leitor completar. ■

Exercícios

1. Prove a existência de um conjunto indutivo ao qual ω pertence. Discuta o uso do axioma da substituição.

2. Enuncie o princípio da recursão finita para classes, e demonstre-o a partir do princípio da recursão transfinita.

3. Dê um exemplo de um conjunto transitivo infinito que não contém ω como subconjunto. Justifique sua resposta.

4. Considere a seguinte sentença: *existe um conjunto bem-ordenado que não possui máximo*. Prove que, mediante os demais axiomas de ZFC, essa sentença é equivalente ao axioma da infinidade. Discuta se essa equivalência vale também em ZF.

5. Prove que são equivalentes, para um ordinal α:

(a) α é um ordinal limite;

(b) α não possui máximo;

(c) $\alpha = \bigcup \alpha$;

(d) $\bigcup \alpha \notin \alpha$.

6. Se α é um ordinal sucessor, prove que $\alpha = (\bigcup \alpha)^+$.

7. Para todo ordinal α, prove que existe um ordinal limite β tal que $\alpha \in \beta$.

8. Em ZF, prove que o axioma da escolha é equivalente à seguinte sentença: *todo conjunto é equipotente a algum ordinal.*

9. Prove que $\alpha + \beta$ é um ordinal sucessor se, e somente se, β é sucessor.

10. Prove que a multiplicação de ordinais não é comutativa.

11. Exiba subconjuntos de $\mathbb{Q}$ isomorfos (com a ordem herdada da ordem dos racionais) aos seguintes ordinais:

(a) $\omega \cdot 2$

(b) $\omega \cdot \omega$

Capítulo 8

Cardinais

Aprendemos na gramática da língua portuguesa que existem dois tipos de numerais: os ordinais, usados para contagem (primeiro, segundo, terceiro, etc.), e os cardinais, usados para expressar uma quantidade (um, dois, três, etc.). Algo semelhante acontece na matemática. Vimos nos capítulos anteriores que os ordinais são utilizados para representar os conjuntos bem-ordenados, que extrapolam a ideia da contagem. Para representar a quantidade de elementos de um conjunto, utilizamos os cardinais. No entanto, na matemática tratamos de conjuntos infinitos, em que – diferentemente do que acontece com os números naturais – os cardinais de fato assumem um papel diferente dos ordinais.

8.1 Cardinais

Vimos que todo conjunto é equipotente a um ordinal. Mas, quando se trata de conjuntos infinitos, há vários ordinais equipotentes a um mesmo conjunto. Por exemplo, os conjuntos infinitos enumeráveis são equipotentes a ω, mas também são equipotentes a ω^+ e a $\omega + \omega$. Portanto, para representarmos o "tamanho" de um conjunto X escolheremos o *menor* ordinal equipotente a X, evitando, assim, ambiguidade na definição. Dessa forma, nem todos os ordinais serão usados para representar cardinalidade. Apenas aqueles que são os "primeiros" representantes de uma cardinalidade. É fácil ver, pela propriedade de transitividade da equipotência, que a condição necessária e suficiente para que um ordinal seja um desses "ordinais especiais" é que ele não seja equipotente a outro ordinal menor. Isso motiva a seguinte definição:

Definição 8.1 Um ordinal κ é um *cardinal* se não existe $\alpha < \kappa$ equipotente a κ.

Lema 8.2 *Todos os elementos de ω^+ são cardinais.*

Demonstração: Segue do Lema 6.7 e do Teorema 6.5, já que esses provam que um natural n não é equipotente a $m < n$, nem ω é equipotente a qualquer número natural. ■

Teorema 8.3 *Para qualquer conjunto X existe um único cardinal κ tal que κ é equipotente a X.*

Demonstração: Pelo Princípio da Boa Ordem existe uma boa ordem $\leq$ sobre X. Pelo teorema da contagem existe um ordinal α isomorfo a $(X, \leq)$. Em particular, α é equipotente a X. Seja

$$\kappa = min\{\beta < \alpha^+ : \beta \approx X\}$$

Como o próprio α satisfaz $\alpha \approx X$ temos que κ está bem definido. Está claro que κ é equipotente a X. Se existisse $\beta < \kappa$ equipotente a κ, teríamos β equipotente a X, contradizendo a minimalidade de κ.

Se tivermos κ e κ' cardinais diferentes equipotentes a X, teríamos κ e κ' equipotentes, entre si. Pelo Teorema 7.7, temos $\kappa < \kappa'$ ou $\kappa' < \kappa$, o que contradiz que ambos são cardinais. ■

Definição 8.4 Seja X um conjunto. Denotamos por $|X|$ o único cardinal equipotente a X, que será chamado de *cardinalidade de* X.

Lema 8.5 *Sejam X e Y conjuntos:*

(a) $|X| = |Y|$ *se, e somente se,* $X \equiv Y$*;*

(b) $|X| \leq |Y|$ *se, e somente se,* $X \preceq Y$*.*

Demonstração: Item (a) segue imediatamente da definição de cardinalidade e do Teorema 8.3. Para o item (b), sejam $\alpha = |X|$ e $\beta = |Y|$. Se $\alpha \leq \beta$ temos $\alpha \subset \beta$. Logo, sendo $f : \alpha \longrightarrow X$ e $g : \beta \longrightarrow Y$ funções bijetoras, temos que $g \circ f^{-1}$ está bem definida e é uma função injetora de X em Y. Reciprocamente, suponhamos por absurdo que $X \preceq Y$ e não vale $|X| \leq |Y|$. Pelo Teorema 7.7 temos $|Y| < |X|$ e, portanto, $Y \preceq X$, de onde concluímos (Teorema de Cantor-Schröder-Bernstein) que $X \equiv Y$. Isso contradiz com a hipótese de que $|X| \neq |Y|$ (Lema 8.5). ■

Lema 8.6 *Se X e Y são conjuntos infinitos, então $|X \cup Y| = |X \times Y|$ e é o máximo entre $|X|$ e $|Y|$.*

Demonstração: Segue dos Lemas 6.16, 6.17 (vide Exercício 6 do Capítulo 6) e do Lema 8.5. ■

Lema 8.7 *Sejam X um conjunto de cardinais e $\kappa = \bigcup X$. Então:*

(a) κ *é um cardinal;*

(b) *Se* $\alpha \in X$ *então* $\alpha \leq \kappa$*;*

(c) *Se β é um cardinal e $\alpha \leq \beta$, para todo $\alpha \in X$, então $\beta \leq \kappa$.*

Demonstração: Mostraremos apenas o item (a), pois os itens (b) e (c) seguem do fato da ordem dos cardinais ser a ordem da inclusão.

Se X possui máximo, então κ é o máximo de X e, nesse caso, temos $\kappa \in X$ e, portanto, κ é um cardinal. Assumimos, então, que X não possui máximo.

Pelo Teorema 7.9 sabemos que κ é um ordinal. Suponha que existe $\lambda < \kappa$ tal que λ é equipotente a κ. Como X não tem máximo, existem $\xi, \eta \in X$ tais que $\lambda < \xi < \eta \leq \lambda$. Pelo Teorema de Cantor-Schröeder-Berstein temos que ξ e η são ambos equipotentes a λ (e a κ), contradizendo que eles são cardinais. ■

Definição 8.8 Se X é um conjunto de cardinais, definimos o *supremo* de X como $\bigcup X$.

Observe que, dado um cardinal κ, sempre temos $\kappa = \bigcup X$, onde X é o conjunto dos cardinais menores ou iguais a κ. Portanto, considerando que conjuntos de cardinais são sempre bem ordenados, a seguinte definição faz sentido:

Definição 8.9 Definimos a *cofinalidade* de um cardinal infinito κ como o menor cardinal λ para o qual existe um conjunto $A \subseteq \kappa$ tal que $|A| = \lambda$ e $\kappa = \bigcup A$. Denotamos por $cf(\kappa)$ a cardinalidade de κ.

Como $\bigcup \kappa = \kappa$, visto que todo cardinal infinito é um ordinal limite, temos que a cofinalidade de κ está bem definida e é no máximo κ.

Dado um cardinal κ, pelo Teorema de Cantor temos $\kappa < |\mathcal{P}(\kappa)|$ e, portanto, o conjunto $\{\alpha \in |\mathcal{P}(\kappa)|^+ : \alpha \text{ é um cardinal e } \kappa < \alpha\}$ é não-vazio e possui um elemento mínimo. Isso justifica a seguinte definição:

Definição 8.10 Definimos o *cardinal sucessor* de um cardinal κ – denotado por κ_+ – como o menor cardinal maior do que κ.

Observe que κ_+ – o cardinal sucessor de κ – é diferente (exceto no caso finito) de κ^+, que é o *ordinal* sucessor de κ. O primeiro é o menor *cardinal* maior do que κ, o seguindo é o maior *ordinal* maior do que κ.

Definição 8.11 Dizemos que um cardinal κ é:

(a) *sucessor* se existe um cardinal α tal que κ é o cardinal sucessor de α;

(b) *limite* se não é sucessor;

(c) *regular* se a cofinalidade de κ é igual a κ;

(d) *singular* se não é regular.

Note que um cardinal sucessor infinito é sempre um ordinal limite. Muito cuidado, quando usamos ou lemos essa nomenclatura, para perceber no contexto se estamos tratando o conjunto como ordinal ou como cardinal.

Definição 8.12 Sejam λ e κ ordinais. Dizemos que um subconjunto X de κ é *ilimitado em* κ (ou *cofinal em* κ) se para todo $\beta < \kappa$ existe $\alpha \in X$ tal que $\beta < \alpha$. Dizemos que $f : \lambda \longrightarrow \kappa$ é *ilimitada* (em κ) se a imagem de f é ilimitada em κ. Dizemos que $f : \lambda \longrightarrow \kappa$ é *crescente* se $\alpha < \beta < \lambda$ implica $f(\alpha) < f(\beta)$.

Lema 8.13 *Sejam λ e κ cardinais infinitos. São equivalentes:*

(a) $cf(\kappa) = \lambda$*;*

(b) λ *é o menor cardinal tal que existe um subconjunto A de κ ilimitado em κ tal que $|A| = \lambda$;*

(c) λ *é o menor cardinal tal que existe uma função $f : \lambda \longrightarrow \kappa$ ilimitada;*

(d) λ *é o menor cardinal tal que existe uma função $f : \lambda \longrightarrow \kappa$ crescente e ilimitada.*

Demonstração: A equivalência entre (a) e (b) segue do fato que, para $A \subset \kappa$, temos que $\bigcup A = \kappa$ se, e somente se, A é ilimitado em κ. Isso segue diretamente da definição, lembrando que a ordem estrita coincide com a pertinência.

Sejam $\lambda = cf(\kappa)$ e λ' como no item (c). Tome $A \subset \kappa$ ilimitado de cardinalidade λ e $f : \lambda \to A$ sobrejetora. Temos que f é uma função ilimitada em κ, provando que $\lambda' \leq \lambda$. Reciprocamente, se $f : \lambda' \longrightarrow \kappa$ é ilimitada, a imagem de f é um subconjunto de κ ilimitado e de cardinalidade menor ou igual a λ', mostrando que $\lambda \leq \lambda'$. Isso prova a equivalência entre (a) e (c).

Provemos a equivalência entre (c) e (d). Sejam λ e λ' os cardinais satisfazendo as condições (c) e (d), respectivamente, e provemos que $\lambda = \lambda'$. Claramente, $\lambda \leq \lambda'$. Para a outra direção provaremos que existe uma função $f : \lambda \longrightarrow \kappa$ crescente e limitada.

Dada $f : \lambda \longrightarrow \kappa$ ilimitada, iremos definir $f' : \lambda' \longrightarrow \kappa$ crescente e ilimitada. Iremos definir f' por recursão, deixando os detalhes da formalização desta vez por conta do leitor. Suponhamos que temos definida $f'|_\alpha$, para $\alpha < \lambda$. Seja

$$A = Im(f'|_\alpha) \cup Im(f|_{\alpha^+}).$$

Temos que $|A| \leq |\alpha| < \lambda$. Portanto, pela equivalência entre (b) e (c), A é limitado (não ilimitado) em κ. Defina $f'(\alpha)$ como o menor $\gamma \in \kappa$ tal que $\beta < \gamma$, para todo $\gamma \in A$. Pela construção, temos que f' é crescente e $f'(\alpha) > f(\alpha)$, para todo $\alpha \in \lambda$. Portanto, f' é ilimitada em κ. ■

Teorema 8.14 *Seja κ um cardinal.*

(a) *Se κ é sucessor, então κ é regular.*

(b) $cf(\kappa)$ *é regular.*

Demonstração: Suponha que $\kappa = \gamma_+$ e seja $\lambda < \kappa$ cardinal. Em particular, $\lambda \leq \gamma$. Seja $A \subset \kappa$ tal que $|A| = \lambda$. Mostraremos que $\bigcup A$ tem cardinalidade menor ou igual a γ, concluindo que $\bigcup A \neq \kappa$.

A demonstração que se segue se assemelha à do Teorema 6.12 e emitiremos alguns detalhes referente ao uso do axioma da escolha. Seja $f : \lambda \longrightarrow A$ sobrejetora. Como $A \subset \kappa$, cada elemento de A tem cardinalidade menor do que κ e, em particular, menor ou igual a γ. Usando o axioma da escolha tomamos uma função F de domínio λ tal que $F(\alpha)$ é uma função sobrejetora de γ em $f(\alpha)$.

Defina $g : \lambda \times \gamma \longrightarrow \bigcup A$ como $g(\alpha, \beta) = F(\alpha)(\beta)$. Primeiro observamos que g está bem definida, pois $F(\alpha)(\beta) \in f(\alpha)$ e $f(\alpha) \in A$. Logo, a imagem de g está contida em $\bigcup A$. Mostremos que g é sobrejetora.

Seja $x \in \bigcup A$. Seja $y \in A$ tal que $x \in y$. Como f é sobrejetora em A, existe $\alpha \in \lambda$ tal que $f(\alpha) = y$. Como $F(\alpha)$ é sobrejetora em $f(\alpha)$ e $x \in f(\alpha)$, existe $\beta \in \gamma$ tal que $F(\alpha)(\beta) = x$. Logo, $g(\alpha, \beta) = x$.

Portanto, usando o Teorema 6.17 e que $\lambda \leq \gamma$, concluímos que

$$\bigcup A \preceq \lambda \times \gamma \preceq \gamma \times \gamma \preceq \gamma,$$

provando que $|\bigcup A| \leq \gamma$ e concluindo o item (a) do teorema.

Provemos o item (b). Sejam $\lambda = cf(\kappa)$ e $\gamma = cf(\lambda)$. Provaremos que $\gamma = \lambda$. Sejam $f : \lambda \longrightarrow \kappa$ e $g : \gamma \longrightarrow \lambda$ funções crescentes e ilimitadas respectivamente em κ e λ (existem, pelo item (d) do Lema 8.13). Tome $h = f \circ g : \gamma \longrightarrow \kappa$. Vamos provar que h é ilimitada em κ. Seja $\alpha \in \kappa$. Como f é ilimitada, existe $\xi \in \lambda$ tal que $f(\xi) > \alpha$. Como g é ilimitada, existe $\beta \in \gamma$ tal que $g(\beta) > \xi$. Como f é crescente, $f(g(\beta)) > f(\xi) > \alpha$. Portanto, $h(\beta) > \alpha$, provando que h é ilimitada em κ.

Logo, por 8.13, (c), temos que $cf(\kappa) \leq \gamma$. Como $\gamma \leq \lambda$, temos $\gamma = \lambda$, como queríamos. ■

8.2 Usando os ordinais para enumerar os cardinais

Veremos a seguir que podemos "bem ordenar" a classe de todos os cardinais através dos ordinais. Para cada ordinal α, podemos pensar no "α-ésimo" cardinal infinito. Reciprocamente, todo cardinal é o "α-ésimo" cardinal, para algum ordinal α. Para representar os cardinais enumerando-os através dos ordinais, utiliza-se a primeira letra do alfabeto hebraico, que é a letra $\aleph$ (leia-se *álef*).

Definição 8.15 Para cada ordinal α definimos o ordinal $\aleph_\alpha$ recursivamente da seguinte forma:

- $\aleph_0 = \omega$;
- $\aleph_{\alpha^+}$ é o cardinal sucessor de $\aleph_\alpha$;
- Se α é limite, $\aleph_\alpha = \bigcup\{\aleph_\beta : \beta < \alpha\}$.

É também muito comum usar a notação ω_α, com a mesma definição de $\aleph_\alpha$, para $\alpha > 0$. Não se usa a notação ω_0. Para o primeiro cardinal infinito escrevemos ω ou $\aleph_0$.

Teorema 8.16 **(a)** $\aleph_\alpha$ *está bem definido e é um cardinal, para todo ordinal* α*;*

(b) $\aleph_\beta < \aleph_\alpha$ *se, e somente se,* $\beta < \alpha$*;*

(c) *Para todo cardinal* κ *infinito existe um único ordinal* α *tal que* $\kappa = \aleph_\alpha$*.*

Demonstração: Para a parte (a), usamos o Teorema da Recursão. Formalmente, como enunciamos o Teorema da Recursão de modo a obter uma função definida em um conjunto, não em uma classe própria (como a classe dos ordinais), fazemos o seguinte: para cada ordinal γ definimos uma função f_γ de domínio γ de modo que $f_\gamma(\alpha) = f_{\gamma'}(\alpha)$, sempre que $\alpha \in \gamma \cap \gamma'$. E então definimos $\aleph_\alpha$ como $f_{\alpha^+}(\alpha)$. Fixemos γ e chamemos f_γ simplesmente de f.

Supondo que temos definido $f|\alpha$ – cuja imagem é um conjunto de cardinais – iremos definir $f(\alpha)$ (ou seja, definiremos agora a fórmula funcional $F(x, y)$ do enunciado do Teorema da Recursão). Se $\alpha = 0$ definimos $f(\alpha) = \omega$. Se α é um ordinal limite definimos $f(\alpha) = \bigcup im(f|\alpha)$. Pelo Lema 8.7 temos que $f(\alpha)$ é, de fato, um cardinal. Se α é um ordinal sucessor, existe um único β tal que $\alpha = \beta^+$. Defina $f(\alpha) = f(\beta)^+$.

Pela construção é fácil verificar que, se $\gamma' > \gamma$ então $f_{\gamma'}|\gamma = f_\gamma$. Basta provarmos por indução em α que $f_{\gamma'}(\alpha) = f_\gamma(\alpha)$, para todo $\alpha < \gamma$. Provamos, assim, a parte (a). A parte (b) do teorema é imediata da construção e deixamos a cargo do leitor. Mostremos o item (c).

Por indução é fácil verificar que $\kappa \leq \aleph_\kappa$, para todo ordinal κ[1]. Portanto o conjunto $Y = \{\beta \in \kappa^+ : \kappa \leq \aleph_\beta\}$ é não vazio, visto que $\kappa \in Y$. Considerando κ um cardinal infinito, fixe α o mínimo de Y. Suponha $\kappa < \aleph_\alpha$. Como κ é infinito, temos $\kappa \geq \omega$ e, portanto, $\alpha \neq 0$. Consideremos dois casos: α ordinal sucessor e α ordinal limite. Se $\alpha = \beta^+$, temos, por definição de $\aleph_\alpha$, $\kappa \leq \aleph_\beta$ e, portanto, $\beta \in Y$, contradizendo a minimalidade de α. Se α é limite, de $\kappa < \aleph_\alpha$ (ou seja, $\kappa \in \aleph_\alpha$), segue que existe $\beta < \alpha$ tal que $\kappa < \aleph_\beta$ (pois $\aleph_\alpha = \bigcup\{\aleph_\beta : \beta < \alpha\}$), chegando novamente a uma contradição. Provamos que $\aleph_\alpha = \kappa$. ■

8.3 Aritmética cardinal

Começamos esta seção com um alerta: a notação utilizada aqui para as operações entre cardinais será a mesma que usamos para ordinais, mesmo tendo significado diferente. O contexto dirá se estamos considerando a operação entre cardinais ou entre ordinais.

Antes de definirmos soma e produto de cardinais, precisamos de um lema que garantirá a boa definição das operações.

Lema 8.17 *Sejam* A*,* A'*,* B *e* B' *conjuntos tais que* A *é equipotente a* A' *e* B *é equipotente a* B'*. Temos que*

(a) $A \times B$ *é equipotente a* $A' \times B'$*.*

[1] Reflita: pode acontecer a igualdade?

(b) *Se $A \cap B = \emptyset$ e $A' \cap B' = \emptyset$, então $A \cup B$ é equipotente a $A' \cup B'$.*

(c) *Para todos cardinais κ e λ existem A e B disjuntos tais que $|A| = \kappa$ e $|B| = \lambda$.*

Demonstração: Sejam $f : A \longrightarrow A'$ e $g : B \longrightarrow B'$ funções bijetoras. Definimos $h_1 : A \times B \longrightarrow A' \times B'$ como $h(a, b) = (f(a), f(b))$. É fácil verificar que h_1 é bijetora. Se A e B são como as hipóteses do item (b), também é fácil ver que $f \cup g$ é uma função bijetora de $A \cup B$ em $A' \cup B'$.

Para o item (c) basta tomarmos $A = \kappa \times \{0\}$ e $B = \lambda \times \{1\}$. ■

Definição 8.18 (Soma de cardinais) Sejam κ e λ cardinais. Definimos a soma $\kappa + \lambda$ como a cardinalidade de $A \cup B$, onde $|A| = \kappa$, $|B| = \lambda$ e $A \cap B = \emptyset$.

Observe que a soma de cardinais está bem definida graças ao Lema 8.17.

Definição 8.19 (Multiplicação de cardinais) Sejam κ e λ cardinais. Definimos o produto $\kappa \cdot \lambda$ como a cardinalidade de $A \times B$, onde $|A| = \kappa$ e $|B| = \lambda$.

Poderíamos também definir $\kappa \times \lambda$ como $|\kappa \times \lambda|$, e a definição acima passaria a ser uma consequência do Lema 8.17.

Definição 8.20 (Exponenciação de cardinais) Sejam κ e λ cardinais. Definimos a potência κ^λ como $|{}^\lambda\kappa|$, isto é, a cardinalidade do conjunto de funções de domínio λ e imagem contida em κ.

Lema 8.21 *Sejam κ, λ e θ cardinais. Considerando a exponenciação entre cardinais, temos as seguintes igualdades:*

(a) $(\kappa^\lambda)^\theta = \kappa^{(\lambda \cdot \theta)}$;

(b) $2^\kappa = |\mathcal{P}(\kappa)|$.

Demonstração: Para provarmos a igualdade entre dois cardinais basta provarmos que esses são equipotentes. Pela definição das operações entre cardinais, para provarmos a parte (a) basta provarmos que existe uma bijeção F entre ${}^\theta({}^\lambda 2)$ e ${}^{\lambda \times \theta}\kappa$. Definimos $F(f)(\alpha, \beta) = f(\alpha)(\beta)$. Note que essa função está bem definida e é a bijeção procurada.

Para a parte (b) basta considerarmos a bijeção clássica entre ${}^\kappa 2$ e $\mathcal{P}(\kappa)$ dada por $F(f) = \{x \in \kappa : f(x) = 1\}]$. ■

Teorema 8.22 *Sejam κ e λ cardinais.*

(a) *Se κ e λ são cardinais finitos, então $\kappa + \lambda$, $\kappa \cdot \lambda$ e κ^λ coincidem com a operação correspondente em ordinais;*

(b) *Se κ ou λ é infinito, e ambos são maiores que 0, então $\kappa + \lambda$ e $\kappa \cdot \lambda$ são ambos iguais ao máximo entre κ e λ.*

(c) *Se $2 \leq \kappa \leq \lambda$ e λ é infinito então $\kappa^\lambda = 2^\lambda$.*

Demonstração: O item (b) segue facilmente do Lema 6.16 e do Teorema 6.17. Para a parte (c), a desigualdade $2^\lambda \leq \kappa^\lambda$ é trivial. A outra desigualdade segue do item (b) e do Lema 8.21, conforme mostramos a seguir:

$$\kappa^\lambda \leq (2^\kappa)^\lambda \leq 2^{\kappa\times\lambda} \leq 2^\lambda.$$

O mais complicado é provar a parte (a). Para isso precisamos provar que, dados $n, m \in \omega$, temos que $(m\times\{0\})\cup(n\times\{1\})$ é equipotente a $m+n$ e $m\times n$ é equipotente a $m \cdot n$, tomando essas operações nos ordinais. Um detalhe importante é que todos os números naturais são cardinais, como visto no Lema 8.2. Isso é essencial na prova que faremos a seguir, pois usaremos o tempo todo que as operações entre números naturais (vistos como ordinais) sempre resultam em cardinais.

Provaremos a primeira parte por indução em n. Claramente $(m\times\{0\})\cup(\emptyset\times\{1\})$ é equipotente a $m+0$. Suponha que $(m \times \{0\}) \cup (n \times \{1\})$ é equipotente a $m+n$ e seja f uma função bijetora entre esses dois conjuntos. Defina $f' : f \cup \{((n,1),n)\}$. É fácil ver que f' é uma função bijetora de $(m \times \{0\}) \cup (n^+ \times \{1\})$ em $m+n^+$.

Para a segunda parte usaremos novamente indução em n. Seja $f : m\times n \longrightarrow m\cdot n$ bijetora. Defina $f' : m \times n^+ \longrightarrow m \cdot n^+$ como $f'(k,l) = f(k,l)$, se $(k,l) \in m \times n$, e $f'(k,n) = (m \cdot n) + k$, se $k \in m$. Deixamos como exercício provar que f' é bijetora. ■

É imediato do Teorema 8.22 que a soma e produto de cardinais são comutativos e associativos.

A seguir, enunciamos uma definição que, na verdade, nada mais é que uma notação, bastante comum na linguagem cotidiana da matemática, mas que exige um certo cuidado em termos de formalização.

Definição 8.23 Uma *sequência indexada em I* é uma função cujo domínio é I. Denotamos por $(x_i)_{i\in I}$ a sequência $\{(i, x_i) : i \in I\}$.

Essa definição facilita introduzirmos a notação para produto cartesiano infinito.

Definição 8.24 Seja $(X_i)_{i\in I}$ uma sequência indexada em I. Definimos o *produto cartesiano de* $(X_i)_{i\in I}$ como o conjunto das funções f de domínio I tais que $f(i) \in X_i$, para todo i. Denotaremos tal produto por $\prod_{i\in I} X_i$.

Definimos a *união* da sequência $(X_i)_{i\in I}$ como a união de sua imagem, e a denotamos por $\bigcup_{i\in I} X_i$.

Com essas definições, falaremos agora de soma e produto infinitos de cardinais. Antes, como fizemos no caso da soma e produto de dois cardinais, enunciaremos um lema.

Lema 8.25 *Sejam $(A_i)_{i\in I}$ e $(A'_i)_{i\in I}$ duas sequências tais que A_i é equipotente a A'_i, para todo i.*

$|\prod_{i\in I} A_i| = |\prod_{i\in I} A'_i|$;

Se $A_i \cap A_j = \emptyset$ e $A'_i \cap A'_j = \emptyset$, para todos $i, j \in I$ tais que $i \neq j$, então $|\bigcup_{i\in I} A_i| = |\bigcup_{i\in I} A'_i|$.

Demonstração: A demonstração desse lema é um argumento padrão como usado no Lema 8.17 e deixaremos como exercício ao leitor. ■

Definição 8.26 *Seja $(\kappa_i)_{i\in I}$ uma sequência de cardinais. Seja $(A_i)_{i\in I}$ uma sequência de conjuntos tal que $|A_i| = \kappa_i$, para cada $i \in I$. Definimos*

(a) $\sum_{i\in I} \kappa_i = |\bigcup_{i\in I} A_i|$, *se* $A_i \cap A_j = \emptyset$, *quando* $i \neq j$.

(b) $\prod_{i\in I} \kappa_i = |\prod_{i\in I} A_i|$.

Há um abuso de notação na definição acima: o mesmo símbolo é utilizado tanto para o produto infinito de cardinais quanto o produto cartesiano indexado em I.

Teorema 8.27 *[Lema de König] Se κ e λ são cardinais infinitos tais que $cf(\kappa) = \lambda$, então $\kappa^\lambda > \kappa$.*

Demonstração: Seja $f : \lambda \longrightarrow \kappa$ uma função crescente e ilimitada. Mostremos que uma função $F : \kappa \longrightarrow {}^\lambda\kappa$ não pode ser sobrejetora, onde ${}^\lambda\kappa$ é o conjunto das funções de λ em κ. Defina $h : \lambda \longrightarrow \kappa$ como

$$h(\alpha) = min(\kappa \setminus \{F(\beta)(\alpha) : \beta < f(\alpha)\}).$$

Notemos que o conjunto da direita é não vazio pois $|f(\alpha)| < \kappa$.

Temos que $h(\alpha) \neq F(\beta)(\alpha)$, para todo $\beta < f(\alpha)$. Como f é ilimitada, se $F(\beta) = h$, existe $\alpha \in \lambda$ tal que $\beta < f(\alpha)$, contradizendo a definição de F. Logo, F não é sobrejetora. ■

Corolário 8.28 *Se κ é um cardinal infinito, $cf(2^\kappa) > \kappa$.*

Demonstração: Basta observar que $2^\kappa \leq \kappa^\kappa \leq (2^\kappa)^\kappa = 2^{\kappa\times\kappa} = 2^\kappa$. Aplique o teorema 8.27 para 2^κ no lugar de κ. ■

Exercícios

1. Prove que todo cardinal infinito é um ordinal limite.

2. Prove que, para todo ordinal α, existe $\beta > \alpha$ tal que $\aleph_\beta$ é um cardinal singular.

3. Podemos ter $2^{\aleph_0} = \aleph_\omega$? Justifique.

4. Prove que $\aleph_\alpha$ é um cardinal sucessor se, e somente se, α é um ordinal sucessor.

Bibliografia

[1] Aragona, J. *Números Reais*. Editora Livraria da Física, São Paulo, 2010.

[2] Barker, S, F. *Filosofia da Matemática*, 2ª ed. Zahar Editores, Rio de Janeiro, 1976.

[3] Gödel, K. *On Formally Undecidable Propositions of Principia Mathematica and Related Systems*. Dover Publications, Nova York, 1992.

[4] Halmos, P. R. *Teoria Ingênua dos Conjuntos*. Editora Polígono, São Paulo, 1973.

[5] Hrbacek, Jech, T. *Introduction to set theory.*CRC Press, Nova York, 1999.

[6] Jech, T. J. *The Axiom of Choice*. Dover Publications, Nova York, 2008.

[7] Kunen, K. *Set Theory. An Introduction to Independence Proofs*. North Holland, 1980.

[8] Miraglia, F. *Teoria dos Conjuntos: um Mínimo*. EDUSP, São Paulo, 1992.

[9] Stoll, R. R. *Set Theory and Logic*. Dover Publications, Nova York, 1979.