Teoria dos Conjuntos

Projeto Open Logic

Idealizador

Richard Zach, *University of Calgary*

Conselho Editorial

Aldo Antonelli,† *University of California, Davis*
Andrew Arana, *Université de Lorraine*
Jeremy Avigad, *Carnegie Mellon University*
Tim Button, *University College London*
Walter Dean, *University of Warwick*
Benedict Eastaugh, *University of Warwick*
Gillian Russell, *Australian National University*
Nicole Wyatt, *University of Calgary*
Audrey Yap, *University of Victoria*

Colaboradores

Samara Burns, *Columbia University*
Dana Hägg, *University of Calgary*
Zesen Qian, *Carnegie Mellon University*

Teoria dos Conjuntos

Uma Introdução Aberta

Adaptado por Tim Button
Tradução e adaptação de Pedro Lima

Outono 2021

O Projeto Open Logic agradece o generoso apoio do Taylor Institute of Teaching and Learning da University of Calgary, e da Alberta Open Educational Resources (ABOER) Initiative, que se torna possível por meio de investimento do governo de Alberta.

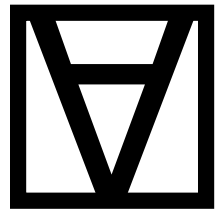
UNIVERSITY OF CALGARY
Taylor Institute for Teaching and Learning

Ilustrações de capa por Matthew Leadbeater, usadas sob uma Licença Creative Commons Atribuição-NãoComercial 4.0 Internacional.

Composição tipográfica em Baskervald X e Nimbus Sans com LaTeX.

Esta versão de *Teoria dos Conjuntos* é revisão 51c2271 (2026-07-12), com conteúdo gerado a partir da revisão 51c2271 (2026-07-12) de *Open Logic Text*. Download gratuito em:

`https://st.openlogicproject.org/pt/`



Sumário

Sobre este Livro

Este livro é um Recurso Educacional Aberto. Ele foi escrito para estudantes com algum conhecimento de lógica e de matemática do ensino médio. Seu objetivo é oferecer uma primeira visão da filosofia da teoria dos conjuntos. Ao final deste livro, os estudantes que o lerem podem ter uma noção:

1. de por que a teoria dos conjuntos surgiu;
2. de como incorporar grandes partes da matemática em teoria dos conjuntos + aritmética;
3. de como incorporar a própria aritmética na teoria dos conjuntos;
4. do que envolve a concepção cumulativa iterativa de conjunto;
5. de como se pode tentar justificar os axiomas de ZFC.

O livro surgiu de um curso curto que eu ministrei no Departamento de Filosofia de Cambridge. Antes de mim, ele foi ministrado por Luca Incurvati e Michael Potter. Ao escrever este livro—e, de forma mais geral, o curso—fiquei profundamente em dívida com Luca e Michael. Espero que isso fique claro no próprio texto do livro; mas também quero registrar aqui meus sinceros agradecimentos a ambos.

A maior parte deste livro foi originalmente publicada como *Open Set Theory*; este livro é seu sucessor. Contribuí com todo o material deste livro para o *Open Logic Project*, onde ele está disponível gratuitamente. (Capítulos 2 a 5 foram elaborados, com apenas pequenas alterações, a partir de material já existente no OLP; essas alterações agora também fazem parte do OLP.) Veja openlogicproject.org para mais informações.

Tim Button
University College London
Outubro de 2021

PARTE I

Prelúdio

CAPÍTULO 1

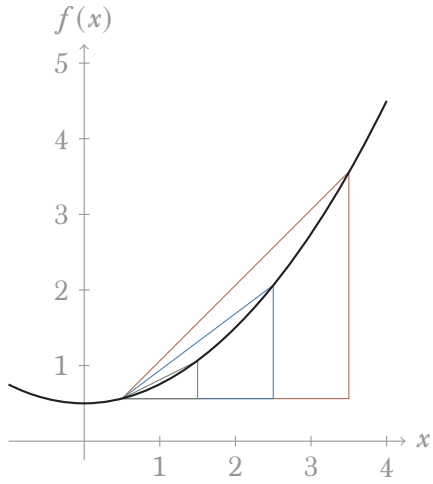
História e Mitologia

Para compreender o significado filosófico da teoria dos conjuntos, ajuda ter alguma ideia de por que ela surgiu. Para isso, ajuda pensar um pouco na história e na mitologia da matemática. Assim, antes mesmo de começarmos a discutir a teoria dos conjuntos, começaremos com uma "história" muito breve. Mas a colocamos entre aspas, porque é muito breve, extremamente seletiva e algo contestável.

1.1 Infinitesimais e Diferenciação

Newton e Leibniz descobriram o cálculo (independentemente) no fim do século XVII. Uma aplicação particularmente importante do cálculo era a *diferenciação*. Grosso modo, a diferenciação pretende dar uma noção da "taxa de variação", ou gradiente, de uma função num ponto.

Eis uma forma vívida de ilustrar a ideia. Considere a função $f(x) = x^2/4 + 1/2$, representada a preto abaixo:



Suponhamos que queremos encontrar o gradiente da função em $c = 1/2$. Começamos por desenhar um triângulo cuja hipotenusa aproxima o gradiente nesse ponto, talvez o triângulo vermelho acima. Quando β é o comprimento da base do nosso triângulo, a sua altura é $f(1/2 + \beta) - f(1/2)$, de modo que o gradiente da hipotenusa é:

$$\frac{f(1/2 + \beta) - f(1/2)}{\beta}.$$

Assim, o gradiente do nosso triângulo vermelho, com base de comprimento 3, é exatamente 1. A hipotenusa de um triângulo menor, o triângulo azul com base de comprimento 2, dá uma melhor aproximação; o seu gradiente é $3/4$. Um triângulo ainda menor, o verde com base de comprimento 1, dá uma aproximação ainda melhor, com gradiente $1/2$.

Triângulos cada vez menores dão-nos aproximações cada vez melhores. Por isso poderíamos dizer algo deste gênero: a hipotenusa de um triângulo com comprimento de base *infinitesimal* nos dá o próprio gradiente em $c = 1/2$. Deste modo, obteríamos uma fórmula para a (primeira) derivada da função f no ponto c:

$$f'(c) = \frac{f(c + \beta) - f(c)}{\beta} \text{ onde } \beta \text{ é infinitesimal.}$$

E, grosso modo, foi isto que Newton e Leibniz disseram.

Contudo, uma vez que o disseram, devemos perguntar-lhes: o que é um *infinitesimal*? Surge um dilema sério. Se $\beta = 0$, então f' fica mal definida, pois envolve dividir por 0. Mas se $\beta > 0$, obtemos apenas uma *aproximação* ao gradiente, e não o próprio gradiente.

Esta não é uma preocupação anacrônica. Eis Berkeley, a criticar os seguidores de Newton:

> Admito que os signos possam ser feitos para denotar qualquer coisa ou nada: e consequentemente que, na notação original $c + \beta$, β poderia ter significado um incremento ou nada. Mas então, seja qual for o significado atribuído a β, é preciso argumentar de modo consistente com essa significação, e não proceder com duplo sentido: o que seria um sofismo manifesto. (Berkeley 1734, §XIII, variáveis alteradas para coincidir com o texto precedente)

Para defender o cálculo infinitesimal face a Berkeley, poder-se-ia responder que a linguagem dos "infinitesimais" é meramente figurada. Poder-se-ia dizer que, desde que tomemos um triângulo *bem pequeno*, obteremos uma aproximação *suficientemente boa* à tangente. Berkeley tinha também uma resposta para isto: embora isso pudesse bastar para a engenharia, minava o *estatuto* da matemática, pois

> dizem-nos que *in rebus mathematicis errores quàm minimi non sunt contemnendi.* [No caso da matemática, os menores erros não devem ser desprezados.] (Berkeley, 1734, §IX)

A passagem em itálico é uma citação quase literal do próprio Newton na *Quadrature of Curves* (1704).

As objeções filosóficas de Berkeley são profundamente incisivas. Não obstante, o cálculo foi uma empresa enormemente bem-sucedida, e os matemáticos continuaram a usá-lo sem incorrer em erro.

1.2 Definição Rigorosa de Limites

Hoje em dia, a solução padrão ao problema anterior é livrar-se dos infinitesimais. Eis como.

Já vimos que, à medida que β fica menor, obtemos melhores aproximações do gradiente. De fato, à medida que β se aproxima arbitrariamente de 0, o valor de $f'(c)$ "tende sem limite" para o gradiente que queremos. Assim, em vez de considerar o que acontece *em* $\beta = 0$, precisamos apenas considerar a *tendência* de $f'(c)$ quando β se aproxima de 0.

Posto deste modo, o desafio geral é dar sentido a afirmações desta forma:

À medida que x se aproxima de c, $g(x)$ tende sem limite para ℓ.

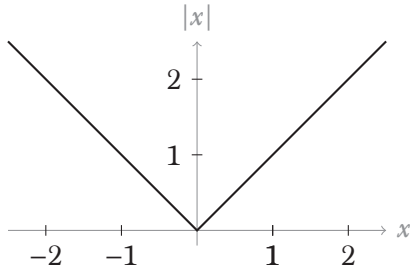
que podemos escrever de forma mais compacta assim:

$$\lim_{x \to c} g(x) = \ell.$$

No século XIX, sobre trabalho anterior de Cauchy, Weierstrass deu uma definição perfeitamente rigorosa desta expressão. A ideia é, de fato, que podemos tornar $g(x)$ tão próxima quanto queiramos de ℓ, tornando x adequadamente próximo de c. Mais precisamente, estipulamos que $\lim_{x \to c} g(x) = \ell$ significará:

$$(\forall \varepsilon > 0)(\exists \delta > 0) \forall x \left(|x - c| < \delta \to |g(x) - \ell| < \varepsilon \right).$$

As barras verticais indicam aqui valor absoluto. Isto é, $|x| = x$ quando $x \geq 0$, e $|x| = -x$ quando $x < 0$; pode representar-se *essa* função como se segue:



A definição diz portanto, grosso modo, o seguinte: pode tornar o seu "erro" inferior a ε (isto é, $|g(x) - \ell| < \varepsilon$) escolhendo argumentos que distem no máximo δ de c (isto é, $|x - c| < \delta$).

Tendo definido a noção de limite, podemos usá-la para evitar por completo os infinitesimais, estipulando que o gradiente de f em c é dado por:

$$f'(c) = \lim_{x \to 0} \left(\frac{f(c+x) - f(c)}{x} \right) \text{ quando um limite existe.}$$

É importante, no entanto, compreender por que a nossa definição precisa da ressalva "quando um limite existe". Para um exemplo simples, considere $f(x) = |x|$, cujo gráfico acabamos de ver. Evidentemente, $f'(0)$ está mal definida: se nos aproximamos de 0 "pela direita", o gradiente é sempre 1; se nos aproximamos "pela esquerda", o gradiente é sempre -1; logo o limite não está definido. Como tal, podemos acrescentar que uma função f é *diferenciável* em x se e só se tal limite existe.

Já vimos como tratar a diferenciação usando a noção de *limite*. Podemos usar a mesma noção para definir a ideia de função *contínua*. (Bolzano tinha, com efeito, percebido isto já por 1817.) O tratamento Cauchy–Weierstrass da continuidade é o seguinte. Grosso modo: uma função f é contínua (num ponto) desde que, ao exigir um certo grau de precisão relativamente à saída da função, se possa garantir essa precisão insistindo num certo grau de precisão relativamente à entrada da função. Mais precisamente: f é contínua em c desde que, à medida que x tende para zero, a diferença entre $f(c+x)$ e $f(c)$ tenda para 0. De outro modo: f é *contínua* em c se e só se $f(c) = \lim_{x \to c} f(x)$.

Ir mais longe levar-nos-ia apenas para a análise real, quando o nosso tema é a teoria dos conjuntos. Por isso, agora devemos fazer uma pausa e enunciar a moral. Durante o século XIX, os matemáticos aprenderam a passar sem infinitesimais, invocando uma noção rigorosamente definida de *limite*.

1.3 Patologias

Contudo, a definição de um *limite* acabou por permitir construções bastante "patológicas".

Por volta da década de 1830, Bolzano descobriu uma função que era *contínua em todo o lado*, mas *não diferenciável em parte nenhuma.* (Infelizmente, Bolzano nunca publicou isto; a ideia foi pela primeira vez encontrada pelos matemáticos em 1872, graças à descoberta independente de Weierstrass da mesma ideia.)[1] Isto era, no mínimo, bastante surpreendente. É fácil encontrar funções, como $|x|$, que são contínuas em todo o lado mas não diferenciáveis num dado ponto. Mas uma função contínua em todo o lado e *não diferenciável em parte nenhuma* é uma criatura muito diferente. Considere, por um momento, como tentaria desenhar tal função. Para garantir que é contínua, deve consegui-la desenhar sem nunca levantar a caneta da página; mas para garantir que não é diferenciável em nenhum ponto, teria de mudar abruptamente a direção da caneta, constantemente.

Seguiram-se mais "patologias". A 5 de janeiro de 1874, Cantor escreveu uma carta a Dedekind, colocando o problema:

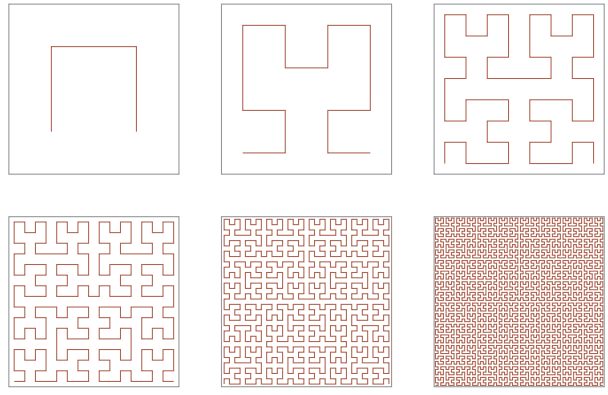
> Poderá uma superfície (digamos um quadrado incluindo o seu contorno) estar correlacionada biunivocamente com uma linha (digamos uma reta incluindo as extremidades) de modo que a cada ponto da superfície corresponda um ponto da linha e, reciprocamente, a cada ponto da linha corresponda um ponto da superfície?
>
> Ainda me parece, de momento, que a resposta a esta questão é muito difícil—embora aqui também se esteja tão impelido a dizer *não* que se queira considerar quase supérflua a prova. [Citado em Gouvêa 2011]

Mas, em 1877, Cantor provou que estivera errado. De fato, uma linha e um quadrado têm exatamente o mesmo número de pontos. Escreveu a 29 de junho de 1877 a Dedekind "*je le vois, mais*

[1] A história está documentada em notas de rodapé extremamente minuciosas do artigo da Wikipédia sobre a função de Weierstrass.

je ne le crois pas"; isto é, "eu vejo, mas não acredito". Na "história recebida" da matemática, isto é frequentemente tomado como indicador de quão *literalmente incríveis* eram estes novos resultados para os matemáticos da época. (A correspondência é apresentada em Gouvêa (2011), e voltaremos a ela em Seção 1.4. A prova de Cantor está esboçada em Seção 5.10.)

Inspirado pelo resultado de Cantor, Peano começou a considerar se seria possível mapear suavemente uma linha sobre um plano. Isto seria uma *curva que preenche o espaço*. Em 1890, Peano construiu precisamente tal curva. Isto é verdadeiramente contraintuitivo: Euclides tinha definido uma linha como "comprimento sem largura" (Livro I, Definição 2), mas Peano mostrara que, enrolando adequadamente uma linha, o seu comprimento pode transformar-se em largura. Em 1891, Hilbert descreveu uma curva preenchedora de espaço ligeiramente mais intuitiva, juntamente com algumas figuras que a ilustram. A curva é construída em sequência, e eis as primeiras seis fases da construção:



No limite—uma noção que, entretanto, recebera definição rigorosa—todo o quadrado fica inteiramente preenchido na cor vermelho. E, de passagem, a curva de Hilbert é contínua em todo o lado mas não diferenciável em parte nenhuma; intuitivamente porque, no limite infinito, a função muda abruptamente de direção a cada instante. (Esboçaremos a construção de Hilbert com mais pormenor na Seção 5.11.)

Para o bem ou para o mal, estas construções geométricas "patológicas" foram tratadas como motivo para duvidar de apelos à intuição geométrica. Tornaram-se algo próximo de *propaganda* por uma nova forma de fazer matemática, que culminaria na teoria dos conjuntos. Na construção posterior do mito em torno do assunto, repetia-se, frequentemente, que estes resultados eram ao mesmo tempo perfeitamente rigorosos e perfeitamente chocantes. Serviam portanto um duplo propósito: como aviso contra confiar na intuição geométrica, e como uma demonstração da fertilidade de novas formas de pensar.

1.4 Mais Mito do que História?

Com mais de um século de perspectiva, ao olhar para estes acontecimentos, devemos ter cuidado para não tomar estes veredictos por adquiridos. Os resultados eram certamente novos, entusiasmantes e surpreendentes. Mas quão verdadeiramente chocantes eram? E demonstraram realmente que não devemos confiar na intuição geométrica?

Quanto ao choque, Gouvêa (2011) observa que a célebre nota de Cantor a Dedekind, "*je le vois, mais je ne le crois pas*" é tirada um tanto fora de contexto. Eis mais desse contexto (citado de Gouvêa):

> Peço desculpa pelo meu zelo pelo assunto se lhe faço tantas exigências à sua bondade e paciência; as comunicações que ultimamente lhe enviei são mesmo para mim tão inesperadas, tão novas, que não posso ter paz de espírito até obter de si, amigo honrado, uma decisão sobre a sua correcção. Enquanto não concordar comigo, só posso dizer: *je le vois, mais je ne le crois pas.*

Cantor sabia que o seu resultado era "tão inesperado, tão novo". Mas é duvidoso que alguma vez achasse o seu resultado *inacreditável.* Como Gouvêa observa, estava simplesmente a pedir a Dedekind que verificasse a prova que oferecera.

Quanto à intuição geométrica: Peano publicou a sua curva preenchedora de espaço sem incluir quaisquer diagramas. Mas quando Hilbert publicou a sua curva, explicou o seu propósito: daria aos leitores um modo claro de compreender o resultado de Peano, se "se servirem da seguinte intuição geométrica"; após o que incluiu uma série de *diagramas* tal como os fornecidos em Seção 1.3.

De modo mais geral: embora os diagramas tenham caído um tanto em desuso nas provas publicadas, não há como escapar ao fato de que os matemáticos *com frequência* usam diagramas ao provar resultados. (Grosso modo: bons matemáticos sabem quando podem confiar na intuição geométrica.)

Em suma: não acredite na *exaltação*; ou, pelo menos, não a aceite por adquirido. Para mais sobre isto, pode ler Giaquinto (2007).

1.5 Roteiro

Parte da lição da secção anterior é que a história da matemática foi em grande medida escrita pelos vencedores. Tinham machados para afiar; machados filosóficos e matemáticos. O estudo sério da história da matemática é seriamente difícil (e gratificante), e a Coruja de Minerva só levanta voo ao crepúsculo.

Apesar disso, é incontestável que os resultados "patológicos" envolveram o desenvolvimento de novas ferramentas matemáticas fascinantes e uma reconsideração dos padrões de rigor matemático. Por exemplo, exigiram pensar o contínuo (a "reta real") de determinado modo, e pensar as funções como mapas ponto a ponto. E, no fim, o desenvolvimento completo de todas essas ferramentas exigiu o desenvolvimento rigoroso da teoria dos conjuntos. O restante deste livro explicará algo desse desenvolvimento.

Parte II apresentará uma versão da teoria *ingênua* dos conjuntos, facilmente suficiente para desenvolver toda a matemática acabada de descrever. Isto levará algum tempo. Mas, ao fim de

Parte II, estaremos em condições de compreender como tratar números reais como certos conjuntos, e como tratar funções sobre eles—incluindo curvas preenchedoras de espaço—como *outros* conjuntos.

Mas a *ingenuidade* dessa teoria dos conjuntos emergirá na Parte III, quando encontrarmos paradoxos da teoria dos conjuntos e a necessidade sentida de descrever as coisas com muito mais precisão. Nesse ponto, precisaremos desenvolver um tratamento axiomático dos conjuntos, com o qual possamos recuperar todos os nossos resultados ingênuos, evitando (esperemos) paradoxos. (A Coruja do rigor matemático também só levanta voo ao crepúsculo.)

PARTE II

Teoria Ingênua dos Conjuntos

Introdução à Parte II

Na Parte II, consideraremos os conjuntos de forma ingênua e informal. Capítulo 2 apresentará a ideia básica, a de que os conjuntos são coleções consideradas extensionalmente, e introduzirá algumas operações muito elementares. Em seguida, Capítulos 3 a 4 explicarão como a teoria dos conjuntos nos permite falar de relações e, portanto, de funções.

Capítulos 5 a 7 tratarão então de algumas das primeiras realizações da teoria ingênua dos conjuntos. Em Capítulo 5, exploramos como comparar conjuntos quanto ao seu tamanho. No Capítulo 6, exploramos como se podem reduzir os inteiros, os racionais e os reais à teoria dos conjuntos juntamente com aritmética elementar. No Capítulo 7, consideramos como se pode implementar aritmética elementar no seio da teoria dos conjuntos.

Repetindo, tudo isto será feito de forma *ingênua*. Mas tudo o que fizermos na Parte II pode ser feito com rigor perfeito, na teoria formal dos conjuntos que introduzimos na Parte III.

CAPÍTULO 2

Começando

2.1 Extensionalidade

Um *conjunto* é uma coleção de objetos, considerada como um único objeto. Os objetos que compõem o conjunto chamam-se *elementos* ou *membros* do conjunto. Se x é um membro de um conjunto A, escrevemos $x \in A$; caso contrário, escrevemos $x \notin A$. O conjunto que não tem membros chama-se o *conjunto vazio* e denota-se "$\emptyset$".

Não importa como *especificamos* o conjunto, ou como *ordenamos* os seus membros, ou mesmo quantas *vezes* contamos os seus membros. Só importa quais são os seus membros. Codificamos isto no princípio seguinte.

Definição 2.1 (Extensionalidade). Se A e B são conjuntos, então $A = B$ se e só se todo membro de A é também um membro de B, e vice-versa.

A extensionalidade autoriza certa notação. Em geral, quando temos objetos $a_1, \ldots, a_n$, então $\{a_1, \ldots, a_n\}$ é *o* conjunto cujos membros são $a_1, \ldots, a_n$. Enfatizamos a palavra "*o*", pois a extensionalidade nos diz que só pode haver *um* tal conjunto. De fato, a extensionalidade também autoriza o seguinte:

$$\{a, a, b\} = \{a, b\} = \{b, a\}.$$

Isto traduz a ideia de que, quando consideramos conjuntos, não nos importa a ordem dos seus membros, nem quantas vezes são especificados.

Exemplo 2.2. Sempre que temos vários objetos, podemos reuni-los num conjunto. O conjunto dos irmãos de Richard, por exemplo, é um conjunto que contém uma pessoa, e podemos escrevê-lo como $S = \{\text{Ruth}\}$. O conjunto dos inteiros positivos menores que 4 é $\{1,2,3\}$, mas também se pode escrever como $\{3,2,1\}$ ou mesmo como $\{1,2,1,2,3\}$. São todos o mesmo conjunto, pela extensionalidade. De fato, todo membro de $\{1,2,3\}$ é também um membro de $\{3,2,1\}$ (e de $\{1,2,1,2,3\}$), e vice-versa.

Com frequência especificaremos um conjunto por alguma propriedade que os seus membros compartilham. Usaremos a seguinte notação abreviada para isso: $\{x : \varphi(x)\}$, onde $\varphi(x)$ representa a propriedade que x tem de satisfazer para ser contado entre os membros do conjunto.

Exemplo 2.3. No nosso exemplo, também poderíamos ter especificado S como

$$S = \{x : x \text{ é irmão de Richard}\}.$$

Exemplo 2.4. Diz-se que um número é *perfeito* se e só se for igual à soma dos seus divisores próprios (isto é, números que o dividem sem resto mas não são iguais ao próprio número). Por exemplo, 6 é perfeito porque os seus divisores próprios são 1, 2, e 3, e $6 = 1+2+3$. De fato, 6 é o único inteiro positivo menor que 10 que é perfeito. Assim, usando a extensionalidade, podemos dizer:

$$\{6\} = \{x : x \text{ é perfeito e } 0 \leq x \leq 10\}$$

Lemos a notação à direita como "o conjunto dos x tais que x é perfeito e $0 \leq x \leq 10$". A identidade aqui confirma que, quando consideramos conjuntos, não nos importa como são especificados. E, mais em geral, a extensionalidade garante que há sempre apenas um conjunto dos x tais que $\varphi(x)$. Logo, a extensionalidade justifica chamar a $\{x : \varphi(x)\}$ *o* conjunto dos x tais que $\varphi(x)$.

A extensionalidade nos dá um modo de mostrar que conjuntos são idênticos: para mostrar que $A = B$, mostre que sempre que $x \in A$ então também $x \in B$, e sempre que $y \in B$ então também $y \in A$.

2.2 Subconjuntos e conjuntos potência

Com frequência quereremos comparar conjuntos. E um tipo óbvio de comparação que se pode fazer é o seguinte: *tudo o que está num conjunto está também no outro.* Esta situação é suficientemente importante para introduzirmos nova notação.

Definição 2.5 (Subconjunto). Se todo membro de um conjunto A é também um membro de B, então dizemos que A é um *subconjunto* de B, e escrevemos $A \subseteq B$. Se A não é subconjunto de B, escrevemos $A \nsubseteq B$. Se $A \subseteq B$ mas $A \neq B$, escrevemos $A \subsetneq B$ e dizemos que A é um *subconjunto próprio* de B.

Exemplo 2.6. Todo conjunto é subconjunto de si próprio, e $\emptyset$ é subconjunto de todo conjunto. O conjunto dos números naturais pares é subconjunto do conjunto dos números naturais. Também $\{a, b\} \subseteq \{a, b, c\}$. Mas $\{a, b, e\}$ não é subconjunto de $\{a, b, c\}$.

Exemplo 2.7. O número 2 é um membro do conjunto dos inteiros, ao passo que o conjunto dos números pares é subconjunto do conjunto dos inteiros. Contudo, um conjunto pode tanto ser um membro quanto subconjunto de algum outro conjunto, por exemplo, $\{0\} \in \{0, \{0\}\}$ e também $\{0\} \subseteq \{0, \{0\}\}$.

A extensionalidade dá um critério de identidade para conjuntos: $A = B$ se e só se todo membro de A é também um membro de B e vice-versa. A definição de "subconjunto" define $A \subseteq B$ precisamente como a primeira metade deste critério: todo membro de A é também um membro de B. Claro que a definição também se aplica se trocarmos A e B: isto é, $B \subseteq A$ se e só se todo membro de B é também um membro de A. E isso, por sua vez, é

exatamente a parte "vice-versa" da extensionalidade. Em outras palavras, a extensionalidade implica que conjuntos são iguais se e só se são subconjuntos um do outro.

Proposição 2.8. *$A = B$ se e só se $A \subseteq B$ e $B \subseteq A$.*

Agora é também boa ocasião para introduzir mais alguma notação útil. Ao definir quando A é subconjunto de B dissemos que "todo membro de A é ...," e preenchemos os "..." com "um membro de B". Mas esta é uma *forma* tão comum de expressão que convém introduzir notação formal para ela.

Definição 2.9. $(\forall x \in A)\varphi$ abrevia $\forall x(x \in A \to \varphi)$. Do mesmo modo, $(\exists x \in A)\varphi$ abrevia $\exists x(x \in A \wedge \varphi)$.

Com esta notação, podemos dizer que $A \subseteq B$ se e só se $(\forall x \in A)x \in B$.

Passamos agora a considerar um certo tipo de conjunto: o conjunto de todos os subconjuntos de um dado conjunto.

Definição 2.10 (Conjunto potência). O conjunto constituído por todos os subconjuntos de um conjunto A chama-se o *conjunto potência* de A, e denota-se $\wp(A)$.

$$\wp(A) = \{B : B \subseteq A\}$$

Exemplo 2.11. Quais são todos os subconjuntos possíveis de $\{a,b,c\}$? São: $\emptyset$, $\{a\}$, $\{b\}$, $\{c\}$, $\{a,b\}$, $\{a,c\}$, $\{b,c\}$, $\{a,b,c\}$. O conjunto de todos estes subconjuntos é $\wp(\{a,b,c\})$:

$$\wp(\{a,b,c\}) = \{\emptyset,\{a\},\{b\},\{c\},\{a,b\},\{b,c\},\{a,c\},\{a,b,c\}\}$$

2.3 Alguns Conjuntos Importantes

Exemplo 2.12. Trabalharemos sobretudo com conjuntos cujos membros são objetos matemáticos. Quatro desses conjuntos são

tão importantes que têm nomes específicos:

$$\mathbb{N} = \{0,1,2,3,\ldots\}$$
o conjunto dos números naturais
$$\mathbb{Z} = \{\ldots,-2,-1,0,1,2,\ldots\}$$
o conjunto dos números inteiros
$$\mathbb{Q} = \{m/n : m,n \in \mathbb{Z} \text{ e } n \neq 0\}$$
o conjunto dos números racionais
$$\mathbb{R} = (-\infty,\infty)$$
o conjunto dos números reais (o contínuo)

São todos conjuntos *infinitos*, isto é, cada um tem infinitos membros.

À medida que percorremos estes conjuntos, vamos acrescentando *mais* números ao nosso repertório. De fato, deve ficar claro que $\mathbb{N} \subseteq \mathbb{Z} \subseteq \mathbb{Q} \subseteq \mathbb{R}$: afinal, todo número natural é um inteiro; todo inteiro é racional; e todo racional é real. Igualmente, deve ficar claro que $\mathbb{N} \subsetneq \mathbb{Z} \subsetneq \mathbb{Q}$, pois -1 é inteiro mas não natural, e $1/2$ é racional mas não inteiro. É menos óbvio que $\mathbb{Q} \subsetneq \mathbb{R}$, isto é, que há números reais que não são racionais.

Também usaremos por vezes o conjunto dos inteiros positivos $\mathbb{Z}^+ = \{1,2,3,\ldots\}$ e o conjunto que contém apenas os dois primeiros naturais $\mathbb{B} = \{0,1\}$.

Exemplo 2.13 (Cadeias). Outro exemplo interessante é o conjunto A^* das *cadeias finitas* sobre um alfabeto A: qualquer sequência finita de elementos de A é uma cadeia sobre A. Incluímos a *cadeia vazia* Λ entre as cadeias sobre A, para todo alfabeto A. Por exemplo,

$$\mathbb{B}^* = \{\Lambda,0,1,00,01,10,11,\\ 000,001,010,011,100,101,110,111,0000,\ldots\}.$$

Se $x = x_1 \ldots x_n \in A^*$ é uma cadeia constituída por n "letras" de A, então dizemos que o *comprimento* da cadeia é n e escrevemos $\text{len}(x) = n$.

Exemplo 2.14 (Sequências infinitas). Para qualquer conjunto A podemos também considerar o conjunto A^ω das sequências infinitas de membros de A. Uma sequência infinita $a_1 a_2 a_3 a_4 \ldots$ consiste numa lista infinita numa só direção de objetos, cada um dos quais é um membro de A.

2.4 Uniões e Interseções

Na Seção 2.1, introduzimos definições de conjuntos por abstração, isto é, definições da forma $\{x : \varphi(x)\}$. Aqui invocamos alguma propriedade φ, e essa propriedade pode mencionar conjuntos que já definimos. Assim, por exemplo, se A e B são conjuntos, o conjunto $\{x : x \in A \vee x \in B\}$ consiste em todos aqueles objetos que são membros de A ou de B, isto é, é o conjunto que reúne os membros de A e B. Podemos visualizar isto como na Figura 2.1, onde a área realçada indica os membros dos dois conjuntos A e B juntos.

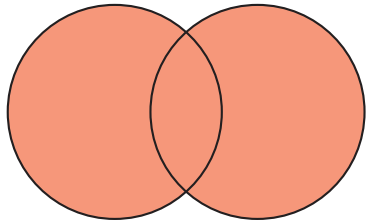


Figura 2.1: A união $A \cup B$ de dois conjuntos é o conjunto dos membros de A juntamente com os de B.

Esta operação sobre conjuntos—combiná-los—é muito útil e comum, e por isso damos-lhe um nome formal e um símbolo.

Definição 2.15 (União). A *união* de dois conjuntos A e B, escrita $A \cup B$, é o conjunto de todas as coisas que são membros de A, de B, ou de ambos.

$$A \cup B = \{x : x \in A \vee x \in B\}$$

Exemplo 2.16. Como a multiplicidade de membros não importa, a união de dois conjuntos que têm um membro em comum contém esse membro apenas uma vez, por exemplo, $\{a,b,c\} \cup \{a,0,1\} = \{a,b,c,0,1\}$.

A união de um conjunto com um dos seus subconjuntos é simplesmente o conjunto maior: $\{a,b,c\} \cup \{a\} = \{a,b,c\}$.

A união de um conjunto com o conjunto vazio é idêntica ao conjunto: $\{a,b,c\} \cup \emptyset = \{a,b,c\}$.

Também podemos considerar uma operação "dual" à união. É a operação que forma o conjunto de todos os membros que são membros de A e são também membros de B. Esta operação chama-se *interseção*, e pode representar-se como na Figura 2.2.

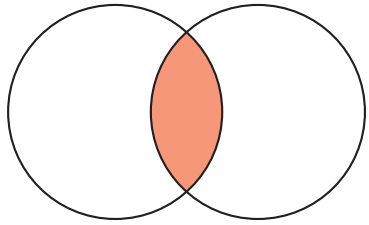


Figura 2.2: A interseção $A \cap B$ de dois conjuntos é o conjunto dos membros que têm em comum.

Definição 2.17 (Interseção). A *interseção* de dois conjuntos A e B, escrita $A \cap B$, é o conjunto de todas as coisas que são membros de ambos A e B.

$$A \cap B = \{x : x \in A \wedge x \in B\}$$

Dois conjuntos dizem-se *disjuntos* se a sua interseção é vazia. Isso significa que não têm membros em comum.

Exemplo 2.18. Se dois conjuntos não têm membros em comum, a sua interseção é vazia: $\{a,b,c\} \cap \{0,1\} = \emptyset$.

Se dois conjuntos têm membros em comum, a sua interseção é o conjunto de todos esses membros: $\{a,b,c\} \cap \{a,b,d\} = \{a,b\}$.

A interseção de um conjunto com um dos seus subconjuntos é simplesmente o conjunto menor: $\{a,b,c\} \cap \{a,b\} = \{a,b\}$.

A interseção de qualquer conjunto com o conjunto vazio é vazia: $\{a,b,c\} \cap \emptyset = \emptyset$.

Também podemos formar a união ou a interseção de mais de dois conjuntos. Uma forma elegante de tratar disso em geral é a seguinte: suponhamos que reunimos num único conjunto todos os conjuntos dos quais queremos formar a união (ou a interseção). Então podemos definir a união de todos os nossos conjuntos originais como o conjunto de todos os objetos que pertencem a pelo menos um membro desse conjunto, e a interseção como o conjunto de todos os objetos que pertencem a todo membro desse conjunto.

Definição 2.19. Se A é um conjunto de conjuntos, então $\bigcup A$ é o conjunto dos membros dos membros de A:

$$\begin{aligned}\bigcup A &= \{x : x \text{ pertence a um membro de } A\}, \text{ isto é,}\\ &= \{x : \text{existe um } B \in A \text{ tal que } x \in B\}\end{aligned}$$

Definição 2.20. Se A é um conjunto de conjuntos, então $\bigcap A$ é o conjunto dos objetos que todos os elementos de A têm em comum:

$$\begin{aligned}\bigcap A &= \{x : x \text{ pertence a todo membro de } A\}, \text{ isto é,}\\ &= \{x : \text{para todo } B \in A, x \in B\}\end{aligned}$$

Exemplo 2.21. Suponha $A = \{\{a,b\},\{a,d,e\},\{a,d\}\}$. Então $\bigcup A = \{a,b,d,e\}$ e $\bigcap A = \{a\}$.

Também poderíamos fazer o mesmo para uma sequência de conjuntos A_1, A_2, ...

$$\bigcup_i A_i = \{x : x \text{ pertence a um dos } A_i\}$$

$$\bigcap_i A_i = \{x : x \text{ pertence a todos os } A_i\}.$$

Quando temos um *índice* de conjuntos, isto é, algum conjunto I tal que consideramos A_i para cada $i \in I$, podemos também usar estas abreviaturas:

$$\bigcup_{i \in I} A_i = \bigcup \{A_i : i \in I\}$$
$$\bigcap_{i \in I} A_i = \bigcap \{A_i : i \in I\}$$

Finalmente, podemos querer pensar no conjunto de todos os membros em A que não estão em B. Podemos representá-lo como na Figura 2.3.

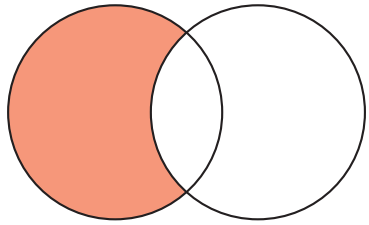


Figura 2.3: A diferença $A \setminus B$ de dois conjuntos é o conjunto dos membros de A que não são também membros de B.

Definição 2.22 (Diferença). A *diferença* de conjuntos $A \setminus B$ é o conjunto de todos os membros de A que não são também membros de B, isto é,

$$A \setminus B = \{x : x \in A \text{ e } x \notin B\}.$$

2.5 Pares, Tuplas, Produtos Cartesianos

Da extensionalidade segue-se que os conjuntos não têm ordem entre os seus elementos. Por isso, se queremos representar ordem, usamos *pares ordenados* $\langle x, y \rangle$. Num par não ordenado $\{x, y\}$,

a ordem não importa: $\{x,y\} = \{y,x\}$. Num par ordenado, importa: se $x \neq y$, então $\langle x,y\rangle \neq \langle y,x\rangle$.

Como devemos pensar os pares ordenados na teoria dos conjuntos? O essencial é preservar a ideia de que pares ordenados são idênticos se e só se compartilham o mesmo primeiro elemento e o mesmo segundo elemento, isto é:

$$\langle a,b\rangle = \langle c,d\rangle \text{ se e só se } a = c \text{ e } b = d.$$

Podemos definir pares ordenados na teoria dos conjuntos usando a definição de Wiener–Kuratowski.

Definição 2.23 (Par ordenado). $\langle a,b\rangle = \{\{a\},\{a,b\}\}$.

Tendo fixado uma definição de par ordenado, podemos usá-la para definir outros conjuntos. Por exemplo, por vezes queremos sequências ordenadas de mais de dois objetos, p.ex., *triplos* $\langle x,y,z\rangle$, *quádruplos* $\langle x,y,z,u\rangle$, e assim por diante. Podemos pensar nos triplos como pares ordenados especiais, em que o primeiro elemento é ele próprio um par ordenado: $\langle x,y,z\rangle$ é $\langle\langle x,y\rangle,z\rangle$. O mesmo vale para quádruplos: $\langle x,y,z,u\rangle$ é $\langle\langle\langle x,y\rangle,z\rangle,u\rangle$, e assim por diante. Em geral, falamos de *n-tuplas ordenadas* $\langle x_1,\ldots,x_n\rangle$.

Certos conjuntos de pares ordenados, ou outras n-tuplas ordenadas, serão úteis.

Definição 2.24 (Produto cartesiano). Dados conjuntos A e B, o seu *produto cartesiano* $A \times B$ é definido por

$$A \times B = \{\langle x,y\rangle : x \in A \text{ e } y \in B\}.$$

Exemplo 2.25. Se $A = \{0,1\}$, e $B = \{1,a,b\}$, então o seu produto é

$$A \times B = \{\langle 0,1\rangle,\langle 0,a\rangle,\langle 0,b\rangle,\langle 1,1\rangle,\langle 1,a\rangle,\langle 1,b\rangle\}.$$

Exemplo 2.26. Se A é um conjunto, o produto de A consigo próprio, $A \times A$, também se escreve A^2. É o conjunto de *todos* os pares $\langle x, y \rangle$ com $x, y \in A$. O conjunto de todos os triplos $\langle x, y, z \rangle$ é A^3, e assim por diante. Podemos dar uma definição recursiva:

$$A^1 = A$$
$$A^{k+1} = A^k \times A$$

Proposição 2.27. *Se A tem n membros e B tem m membros, então $A \times B$ tem $n \cdot m$ elementos.*

Prova. Para todo membro x em A, há m membros da forma $\langle x, y \rangle \in A \times B$. Seja $B_x = \{\langle x, y \rangle : y \in B\}$. Como sempre que $x_1 \neq x_2$, $\langle x_1, y \rangle \neq \langle x_2, y \rangle$, tem-se $B_{x_1} \cap B_{x_2} = \emptyset$. Mas se $A = \{x_1, \ldots, x_n\}$, então $A \times B = B_{x_1} \cup \cdots \cup B_{x_n}$, e portanto tem $n \cdot m$ membros.

Para visualizar isto, dispomos os membros de $A \times B$ numa grade:

$$\begin{array}{rccccc} B_{x_1} = & \{\langle x_1, y_1 \rangle & \langle x_1, y_2 \rangle & \ldots & \langle x_1, y_m \rangle\} \\ B_{x_2} = & \{\langle x_2, y_1 \rangle & \langle x_2, y_2 \rangle & \ldots & \langle x_2, y_m \rangle\} \\ \vdots & & \vdots & & \\ B_{x_n} = & \{\langle x_n, y_1 \rangle & \langle x_n, y_2 \rangle & \ldots & \langle x_n, y_m \rangle\} \end{array}$$

Como os x_i são todos distintos, e os y_j são todos distintos, quaisquer dois pares nesta grade são distintos, e há $n \cdot m$ deles. □

Exemplo 2.28. Se A é um conjunto, uma *palavra* sobre A é qualquer sequência de membros de A. Uma sequência pode ser vista como uma n-tupla de membros de A. Por exemplo, se $A = \{a, b, c\}$, então a sequência "bac" pode ser vista como o triplo $\langle b, a, c \rangle$. Palavras, isto é, sequências de símbolos, são de importância crucial na ciência da computação. Por convenção, contamos os membros de A como sequências de comprimento 1, e $\emptyset$ como a sequência de comprimento 0. O conjunto de *todas* as palavras sobre A é então

$$A^* = \{\emptyset\} \cup A \cup A^2 \cup A^3 \cup \ldots$$

2.6 Paradoxo de Russell

A extensionalidade autoriza a notação $\{x : \varphi(x)\}$, para *o* conjunto dos x tais que $\varphi(x)$. Contudo, tudo o que a extensionalidade *realmente* autoriza é o seguinte pensamento. *Se* há um conjunto cujos membros são exatamente os que satisfazem φ, *então* há apenas um tal conjunto. Dito de outro modo: fixada alguma φ, o conjunto $\{x : \varphi(x)\}$ é único, *se existir.*

Mas este condicional é importante! Crucialmente, nem toda propriedade se presta à *compreensão.* Isto é, algumas propriedades *não* definem conjuntos. Se todas definissem, cairíamos em contradições flagrantes. O exemplo mais famoso disso é o Paradoxo de Russell.

Os conjuntos podem ser membros de outros conjuntos—por exemplo, o conjunto potência de um conjunto A é constituído por conjuntos. Por isso faz sentido perguntar ou investigar se um conjunto é um membro de outro conjunto. Pode um conjunto ser membro de si próprio? Nada na ideia de conjunto parece excluir isso. Por exemplo, se *todos* os conjuntos formam uma coleção de objetos, alguém poderia pensar que podem reunir-se num único conjunto—o conjunto de todos os conjuntos. E ele, sendo conjunto, seria um membro do conjunto de todos os conjuntos.

O Paradoxo de Russell surge quando consideramos a propriedade de não ter a si próprio como um membro, de ser *não auto-membro.* E se supusermos que há um conjunto de todos os conjuntos que não têm a si próprios como um membro? Será que

$$R = \{x : x \notin x\}$$

existe? Verifica-se que podemos provar que não.

Teorema 2.29 (Paradoxo de Russell). *Não há conjunto* $R = \{x : x \notin x\}$.

Prova. Se $R = \{x : x \notin x\}$ existisse, então $R \in R$ se e só se $R \notin R$, o que é uma contradição. □

Vamos percorrer esta prova mais devagar. Se R existe, faz sentido perguntar se $R \in R$ ou não. Suponhamos que de fato $R \in R$. Agora, R foi definido como o conjunto de todos os conjuntos que não são membros de si próprios. Logo, se $R \in R$, então R não tem a propriedade definidora de R. Mas só os conjuntos que têm essa propriedade estão em R, logo R não pode ser um membro de R, isto é, $R \notin R$. Mas R não pode ao mesmo tempo ser e não ser um membro de R, logo temos uma contradição.

Como a suposição de que $R \in R$ leva a uma contradição, temos $R \notin R$. Mas isto também leva a uma contradição! Pois se $R \notin R$, então R tem a propriedade definidora de R, e assim R seria um membro de R, tal como todos os outros conjuntos não auto-membros. E de novo, R não pode ao mesmo tempo não ser e ser um membro de R.

Como montar uma teoria dos conjuntos que evite cair no Paradoxo de Russell, isto é, que evite afirmar de modo *inconsistente* que $R = \{x : x \notin x\}$ existe? Teríamos de estabelecer axiomas que nos dêem condições muito precisas para dizer quando os conjuntos existem (e quando não existem).

A teoria dos conjuntos esboçada neste capítulo não faz isso. É *genuinamente ingênua.* Diz apenas que os conjuntos obedecem à extensionalidade e que, se temos alguns conjuntos, podemos formar a sua união, interseção, etc. É possível desenvolver a teoria dos conjuntos com mais rigor do que isto.

Problemas

Problema 2.1. Prove que há no máximo um conjunto vazio, isto é, mostre que se A e B são conjuntos sem membros, então $A = B$.

Problema 2.2. Liste todos os subconjuntos de $\{a, b, c, d\}$.

Problema 2.3. Mostre que se A tem n membros, então $\wp(A)$ tem 2^n membros.

Problema 2.4. Prove que se $A \subseteq B$, então $A \cup B = B$.

Problema 2.5. Prove rigorosamente que se $A \subseteq B$, então $A \cap B = A$.

Problema 2.6. Mostre que se A é um conjunto e $A \in B$, então $A \subseteq \bigcup B$.

Problema 2.7. Prove que se $A \subsetneq B$, então $B \setminus A \neq \emptyset$.

Problema 2.8. Usando Definição 2.23, prove que $\langle a, b\rangle = \langle c, d\rangle$ se e só se $a = c$ e $b = d$.

Problema 2.9. Liste todos os membros de $\{1, 2, 3\}^3$.

Problema 2.10. Mostre, por indução em k, que para todo $k \geq 1$, se A tem n membros, então A^k tem n^k membros.

CAPÍTULO 3

Relações

3.1 Relações como Conjuntos

Na Seção 2.3, mencionamos alguns conjuntos importantes: $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$. Sem dúvida que se lembram de algumas relações interessantes entre os membros de alguns destes conjuntos. Por exemplo, cada um destes conjuntos tem uma *relação de ordem* completamente padrão sobre si. Há também a relação *ser idêntico a*, segundo a qual cada objeto só se relaciona consigo próprio e com nenhum outro. Há muitas outras relações interessantes que encontraremos, e ainda mais relações possíveis. Antes de as revermos, contudo, começaremos por observar que podemos ver relações como um tipo especial de conjunto.

Para isso, recordemos duas coisas da Seção 2.5. Primeiro, recordemos a noção de *par ordenado*: dados a e b, podemos formar $\langle a, b\rangle$. É importante notar que a ordem dos elementos *importa* aqui. Assim, se $a \neq b$ então $\langle a, b\rangle \neq \langle b, a\rangle$. (Contrastemos isto com pares não ordenados, isto é, conjuntos de 2 elementos, onde $\{a, b\} = \{b, a\}$.) Segundo, recordemos a noção de *produto cartesiano*: se A e B são conjuntos, então podemos formar $A \times B$, o conjunto de todos os pares $\langle x, y\rangle$ com $x \in A$ e $y \in B$. Em particular, $A^2 = A \times A$ é o conjunto de todos os pares ordenados de A.

Consideremos agora uma relação particular num conjunto: a relação $<$ no conjunto $\mathbb{N}$ dos números naturais. Seja o conjunto

de todos os pares de números $\langle n, m\rangle$ onde $n < m$, isto é,

$$R = \{\langle n, m\rangle : n, m \in \mathbb{N} \text{ e } n < m\}.$$

Há uma ligação estreita entre n ser menor que m, e o par $\langle n, m\rangle$ ser membro de R, nomeadamente:

$$n < m \text{ se e só se } \langle n, m\rangle \in R.$$

De fato, sem qualquer perda de informação, podemos considerar o conjunto R como *sendo* a relação $<$ em $\mathbb{N}$.

Do mesmo modo podemos construir um subconjunto de $\mathbb{N}^2$ para qualquer relação entre números. Reciprocamente, dado qualquer conjunto de pares de números $S \subseteq \mathbb{N}^2$, há uma relação correspondente entre números, nomeadamente, a relação segundo a qual n se liga a m se e só se $\langle n, m\rangle \in S$. Isto justifica a definição seguinte:

Definição 3.1 (Relação binária). Uma *relação binária* num conjunto A é um subconjunto de A^2. Se $R \subseteq A^2$ é uma relação binária em A e $x, y \in A$, por vezes escrevemos Rxy (ou xRy) para $\langle x, y\rangle \in R$.

Exemplo 3.2. O conjunto $\mathbb{N}^2$ dos pares de números naturais pode ser listado numa matriz bidimensional assim:

$$\begin{array}{ccccc}
\langle \mathbf{0},\mathbf{0}\rangle & \langle 0,1\rangle & \langle 0,2\rangle & \langle 0,3\rangle & \dots \\
\langle 1,0\rangle & \langle \mathbf{1},\mathbf{1}\rangle & \langle 1,2\rangle & \langle 1,3\rangle & \dots \\
\langle 2,0\rangle & \langle 2,1\rangle & \langle \mathbf{2},\mathbf{2}\rangle & \langle 2,3\rangle & \dots \\
\langle 3,0\rangle & \langle 3,1\rangle & \langle 3,2\rangle & \langle \mathbf{3},\mathbf{3}\rangle & \dots \\
\vdots & \vdots & \vdots & \vdots & \ddots
\end{array}$$

Colocamos a diagonal, aqui, em negrito, pois o subconjunto de $\mathbb{N}^2$ constituído pelos pares que estão na diagonal, isto é,

$$\{\langle 0,0\rangle, \langle 1,1\rangle, \langle 2,2\rangle, \dots\},$$

é a *relação identidade em* $\mathbb{N}$. (Como a relação identidade é popular, definamos $\mathrm{Id}_A = \{\langle x, x\rangle : x \in A\}$ para qualquer conjunto A.) O subconjunto de todos os pares acima da diagonal, isto é,

$$L = \{\langle 0,1\rangle, \langle 0,2\rangle, \ldots, \langle 1,2\rangle, \langle 1,3\rangle, \ldots, \langle 2,3\rangle, \langle 2,4\rangle, \ldots\},$$

é a relação *menor que*, isto é, Lnm se e só se $n < m$. O subconjunto dos pares abaixo da diagonal, isto é,

$$G = \{\langle 1,0\rangle, \langle 2,0\rangle, \langle 2,1\rangle, \langle 3,0\rangle, \langle 3,1\rangle, \langle 3,2\rangle, \ldots\},$$

é a relação *maior que*, isto é, Gnm se e só se $n > m$. A união de L com I, que podemos chamar $K = L \cup I$, é a relação *menor ou igual a*: Knm se e só se $n \leq m$. Do mesmo modo, $H = G \cup I$ é a *relação maior ou igual a*. Estas relações L, G, K, e H são tipos especiais de relações chamadas *ordens*. L e G gozam da propriedade de que nenhum número está em relação L ou G consigo próprio (isto é, para todo n, não se verifica nem Lnn nem Gnn). Relações com esta propriedade chamam-se *irreflexivas*, e, se também forem ordens, chamam-se *ordens estritas*.

Embora ordens e identidade sejam relações importantes e naturais, convém enfatizar que, segundo a nossa definição, *qualquer* subconjunto de A^2 é uma relação em A, por mais artificial ou forçado que pareça. Em particular, $\emptyset$ é uma relação em qualquer conjunto (a *relação vazia*, que nenhum par de elementos satisfaz), e A^2 em si é também uma relação em A (uma que todo par de elementos satisfaz), chamada a *relação universal*. Mas também algo como $E = \{\langle n, m\rangle : n > 5 \text{ ou } m \times n \geq 34\}$ conta como relação.

3.2 Reflexões Filosóficas

Na Seção 3.1, definimos relações como certos conjuntos. Devemos fazer uma pausa e colocar uma questão filosófica rápida: o que é que tal definição *faz*? É extremamente duvidoso que devamos querer dizer que *descobrimos* fatos metafísicos de identidade; que, por exemplo, a relação de ordem em $\mathbb{N}$ *acabou por*

ser o conjunto $R = \{\langle n, m\rangle : n, m \in \mathbb{N} \text{ e } n < m\}$ que definimos na Seção 3.1. Eis três razões para isso.

Primeiro: na Definição 2.23, definimos $\langle a, b\rangle = \{\{a\}, \{a, b\}\}$. Consideremos em vez disso a definição $\|a, b\| = \{\{b\}, \{a, b\}\} = \langle b, a\rangle$. Quando $a \neq b$, temos que $\langle a, b\rangle \neq \|a, b\|$. Mas poderíamos com a mesma facilidade ter tomado $\|a, b\|$ como a nossa definição de par ordenado, em vez de $\langle a, b\rangle$. Ambas as definições teriam funcionado igualmente bem. Assim, temos dois candidatos igualmente bons a "ser" a relação de ordem nos números naturais, nomeadamente:

$$R = \{\langle n, m\rangle : n, m \in \mathbb{N} \text{ e } n < m\}$$
$$S = \{\|n, m\| : n, m \in \mathbb{N} \text{ e } n < m\}.$$

Como $R \neq S$, pela extensionalidade, é claro que não podem *ambos* ser idênticos à relação de ordem em $\mathbb{N}$. Mas seria meramente arbitrário, e por isso um pouco embaraçoso, afirmar que R em vez de S (ou vice-versa) *é* a relação de ordenação, como questão de fato. (Isto é um caso muito simples de um argumento contra o reducionismo em teoria dos conjuntos que Benacerraf tornou famoso em 1965. Voltaremos a ele várias vezes.)

Segundo: se pensamos que *toda* a relação deve ser identificada com um conjunto, então a própria relação de pertinência a conjuntos, $\in$, deveria ser um conjunto particular. De fato, teria de ser o conjunto $\{\langle x, y\rangle : x \in y\}$. Mas será que este conjunto existe? Dado o Paradoxo de Russell, é uma afirmação não trivial que tal conjunto exista. De fato, é possível desenvolver a teoria dos conjuntos de modo rigoroso como teoria axiomática, e essa teoria de fato nega a existência deste conjunto. Logo, mesmo que algumas relações possam ser tratadas como conjuntos, a relação de pertinência a conjuntos terá de ser um caso especial.

Terceiro: quando "identificamos" relações com conjuntos, dissemos que nos permitiríamos escrever Rxy para $\langle x, y\rangle \in R$. Isto está bem, desde que a relação de pertinência, "$\in$", seja tratada *como* um predicado. Mas se pensamos que "$\in$" denota um certo tipo de conjunto, então a expressão "$\langle x, y\rangle \in R$" consiste apenas

em três termos singulares que denotam conjuntos: "$\langle x, y\rangle$", "$\in$", e "R". E tal lista de nomes não é mais capaz de exprimir uma proposição do que a sequência sem sentido: "o copo porta-canetas a mesa". De novo, mesmo que algumas relações possam ser tratadas como conjuntos, a relação de pertinência a conjuntos tem de ser um caso especial. (Isto reúne num só argumento uma versão simples do paradoxo do conceito *cavalo* de Frege, e uma objeção célebre que Wittgenstein dirigiu a Russell.)

Então onde nos deixa isto? Bem, não há nada de *errado* em dizermos que as relações nos números são conjuntos. Só temos de compreender o espírito em que tal observação é feita. Não enunciamos um fato metafísico de identidade. Limitamo-nos a notar que, em certos contextos, podemos (e iremos) *tratar* (certas) relações como certos conjuntos.

3.3 Propriedades Especiais das Relações

Alguns tipos de relações revelam-se tão comuns que receberam nomes especiais. Por exemplo, $\leq$ e $\subseteq$ relacionam os seus respectivos domínios (digamos, $\mathbb{N}$ no caso de $\leq$ e $\wp(A)$ no caso de $\subseteq$) de modos semelhantes. Para precisar exatamente como estas relações são semelhantes, e como diferem, as categorizamos segundo algumas propriedades especiais que as relações podem ter. Verifica-se que (combinações de) algumas destas propriedades especiais são especialmente importantes: ordens e relações de equivalência.

Definição 3.3 (Reflexividade). Uma relação $R \subseteq A^2$ é *reflexiva* se e só se, para todo $x \in A$, Rxx.

Definição 3.4 (Transitividade). Uma relação $R \subseteq A^2$ é *transitiva* se e só se, sempre que Rxy e Ryz, então também Rxz.

Definição 3.5 (Simetria). Uma relação $R \subseteq A^2$ é *simétrica* se e só se, sempre que Rxy, então também Ryx.

Definição 3.6 (Antissimetria). Uma relação $R \subseteq A^2$ é *antissimétrica* se e só se, sempre que tanto Rxy quanto Ryx, então $x = y$ (ou, noutras palavras: se $x \neq y$ então ou $\neg Rxy$ ou $\neg Ryx$).

Numa relação simétrica, Rxy e Ryx verificam-se sempre em conjunto, ou nenhuma se verifica. Numa relação antissimétrica, a única forma de Rxy e Ryx se verificarem em conjunto é que $x = y$. Note-se que isto não *exige* que Rxy e Ryx se verifiquem quando $x = y$, apenas que isso não está excluído. Assim, uma relação antissimétrica pode ser reflexiva, mas não é o caso que toda relação antissimétrica seja reflexiva. Note-se também que ser antissimétrica e ser meramente não simétrica são condições diferentes. De fato, uma relação pode ser ao mesmo tempo simétrica e antissimétrica (por exemplo, a relação identidade o é).

Definição 3.7 (Conexidade). Uma relação $R \subseteq A^2$ é *conexa* se para todos $x, y \in A$, se $x \neq y$, então ou Rxy ou Ryx.

Definição 3.8 (Irreflexividade). Uma relação $R \subseteq A^2$ diz-se *irreflexiva* se, para todo $x \in A$, não Rxx.

Definição 3.9 (Assimetria). Uma relação $R \subseteq A^2$ diz-se *assimétrica* se para nenhum par $x, y \in A$ se tem tanto Rxy quanto Ryx.

Note-se que se $A \neq \emptyset$, então nenhuma relação irreflexiva em A é reflexiva e toda relação assimétrica em A é também antissimétrica. Contudo, há $R \subseteq A^2$ que não são reflexivas nem irreflexivas, e há relações antissimétricas que não são assimétricas.

3.4 Relações de Equivalência

A relação identidade num conjunto é reflexiva, simétrica e transitiva. Relações R que têm todas as três destas propriedades são muito comuns.

Definição 3.10 (Relação de equivalência). Uma relação $R \subseteq A^2$ que é reflexiva, simétrica e transitiva chama-se uma *relação de equivalência.* Membros x e y de A dizem-se *R-equivalentes* se Rxy.

As relações de equivalência dão origem à noção de *classe de equivalência.* Uma relação de equivalência "reparte" o domínio em partições distintas. Dentro de cada partição, todos os objetos estão relacionados uns com os outros; e objetos de partições diferentes não se relacionam entre si. Por vezes, é útil falar destas partições *diretamente.* Para esse fim, introduzimos uma definição:

Definição 3.11. Seja $R \subseteq A^2$ uma relação de equivalência. Para cada $x \in A$, a *classe de equivalência* de x em A é o conjunto $[x]_R = \{y \in A : Rxy\}$. O *quociente* de A por R é $A/_R = \{[x]_R : x \in A\}$, isto é, o conjunto destas classes de equivalência.

O resultado seguinte justifica a definição de classe de equivalência, provando que as classes de equivalência são de fato as partições de A:

Proposição 3.12. *Se $R \subseteq A^2$ é uma relação de equivalência, então Rxy se e só se $[x]_R = [y]_R$.*

Prova. Para o sentido da esquerda para a direita, suponhamos Rxy, e seja $z \in [x]_R$. Por definição, então, Rxz. Como R é uma relação de equivalência, Ryz. (Explicando melhor: como Rxy e R é simétrica temos Ryx, e como Rxz e R é transitiva temos Ryz.) Logo $z \in [y]_R$. Generalizando, $[x]_R \subseteq [y]_R$. Mas de modo perfeitamente análogo, $[y]_R \subseteq [x]_R$. Logo $[x]_R = [y]_R$, pela extensionalidade.

Para o sentido da direita para a esquerda, suponhamos $[x]_R = [y]_R$. Como R é reflexiva, Ryy, logo $y \in [y]_R$. Assim também $y \in [x]_R$ pela suposição de que $[x]_R = [y]_R$. Logo Rxy. □

Exemplo 3.13. Um bom exemplo de relações de equivalência vem da aritmética modular. Para quaisquer a, b, e $n \in \mathbb{Z}^+$, digamos que $a \equiv_n b$ se e só se dividir a por n dá o mesmo resto que dividir b por n. (De modo um pouco mais simbólico: $a \equiv_n b$ se e só se, para algum $k \in \mathbb{Z}$, $a - b = kn$.) Agora, $\equiv_n$ é uma relação de equivalência, para qualquer n. E há exatamente n classes de equivalência distintas geradas por $\equiv_n$; isto é, $\mathbb{N}/_{\equiv_n}$ tem n membros. São: o conjunto dos números divisíveis por n sem resto, isto é, $[0]_{\equiv_n}$; o conjunto dos números divisíveis por n com resto 1, isto é, $[1]_{\equiv_n}$; ...; e o conjunto dos números divisíveis por n com resto $n - 1$, isto é, $[n-1]_{\equiv_n}$.

3.5 Ordens

Muitas das nossas comparações envolvem descrever certos objetos como sendo "menores que", "iguais a" ou "maiores que" outros objetos, sob algum aspecto. Isso envolve relações de *ordem*. Mas há vários tipos de relação de ordem. Por exemplo, algumas exigem que quaisquer dois objetos sejam comparáveis, outras não. Algumas incluem a identidade (como $\leq$) e outras a excluem (como $<$). Ter aqui uma taxonomia ajuda-nos.

Definição 3.14 (Pré-ordem). Uma relação que é ao mesmo tempo reflexiva e transitiva chama-se uma *pré-ordem*.

Definição 3.15 (Ordem parcial). Uma pré-ordem que também é antissimétrica chama-se uma *ordem parcial*.

Definição 3.16 (Ordem linear). Uma ordem parcial que também é conexa chama-se uma *ordem total* ou *ordem linear*.

Exemplo 3.17. Toda ordem linear é também uma ordem parcial, e toda ordem parcial é também uma pré-ordem, mas as implicações recíprocas não valem. A relação universal sobre A é uma pré-ordem, pois é reflexiva e transitiva. Mas, se A tem mais de um membro, a relação universal não é antissimétrica e, portanto, não é uma ordem parcial.

Exemplo 3.18. Considere a relação *não mais longa que* $\preccurlyeq$ sobre $\mathbb{B}^*$: $x \preccurlyeq y$ sse $\text{len}(x) \leq \text{len}(y)$. Trata-se de uma pré-ordem (reflexiva e transitiva), e até conexa, mas não é uma ordem parcial, pois não é antissimétrica. Por exemplo, $01 \preccurlyeq 10$ e $10 \preccurlyeq 01$, mas $01 \neq 10$.

Exemplo 3.19. Uma ordem parcial importante é a relação $\subseteq$ sobre um conjunto de conjuntos. Em geral não é uma ordem linear: se $a \neq b$ e consideramos $\wp(\{a, b\}) = \{\emptyset, \{a\}, \{b\}, \{a, b\}\}$, vemos que $\{a\} \not\subseteq \{b\}$ e $\{a\} \neq \{b\}$ e $\{b\} \not\subseteq \{a\}$.

Exemplo 3.20. A relação de *divisibilidade sem resto* nos dá uma ordem parcial que não é linear. Para inteiros n e m, escrevemos $n \mid m$ para significar que n divide m (exatamente), isto é, sse existe algum inteiro k tal que $m = kn$. Sobre $\mathbb{N}$, isso é uma ordem parcial, mas não linear: por exemplo, $2 \nmid 3$ e também $3 \nmid 2$. Vista como relação sobre $\mathbb{Z}$, a divisibilidade é apenas uma pré-ordem, pois não é antissimétrica: $1 \mid -1$ e $-1 \mid 1$ mas $1 \neq -1$.

Exemplo 3.21. A relação de *extensão* sobre um conjunto de sequências A^* é a seguinte: $s \sqsubseteq s'$ sse $s = \Lambda$ (a sequência vazia), $s = s'$, ou $s = \langle s_1, \ldots, s_n \rangle$ e $s' = \langle s_1, \ldots, s_n, s_{n+1}, \ldots, s_m \rangle$. Se $s \sqsubseteq s'$ dizemos também que s é um *segmento inicial* de s'. A relação de extensão sobre A^* é uma ordem parcial, mas não linear, p.ex., se $a \neq b$, então $ab \not\sqsubseteq ba$ e $ba \not\sqsubseteq ab$.

Definição 3.22 (Ordem estrita). Uma *ordem estrita* é uma relação irreflexiva, assimétrica e transitiva.

Definição 3.23 (Ordem linear estrita). Uma ordem estrita que também é conexa chama-se *ordem total estrita* ou *ordem linear estrita.*

Exemplo 3.24. $\leq$ é a ordem linear correspondente à ordem linear estrita $<$. $\subseteq$ é a ordem parcial correspondente à ordem estrita $\subsetneq$.

Qualquer ordem estrita R sobre A pode transformar-se numa ordem parcial adicionando a diagonal Id_A, isto é, adicionando todos os pares $\langle x, x\rangle$. (Isto chama-se o *fecho reflexivo* de R.) Reciprocamente, a partir de uma ordem parcial obtém-se uma ordem estrita removendo Id_A. Os dois resultados seguintes tornam isto preciso.

Proposição 3.25. *Se R é uma ordem estrita sobre A, então $R^+ = R \cup \mathrm{Id}_A$ é uma ordem parcial. Além disso, se R é uma ordem linear estrita, então R^+ é uma ordem linear.*

Prova. Suponhamos que R é uma ordem estrita, isto é, $R \subseteq A^2$ e R é irreflexiva, assimétrica e transitiva. Seja $R^+ = R \cup \mathrm{Id}_A$. Temos de mostrar que R^+ é reflexiva, antissimétrica e transitiva.

R^+ é claramente reflexiva, pois $\langle x, x\rangle \in \mathrm{Id}_A \subseteq R^+$ para todo $x \in A$.

Para mostrar que R^+ é antissimétrica, suponhamos, para redução ao absurdo, que R^+xy e R^+yx mas $x \neq y$. Como $\langle x, y\rangle \in R \cup \mathrm{Id}_A$, mas $\langle x, y\rangle \notin \mathrm{Id}_A$, temos de ter $\langle x, y\rangle \in R$, isto é, Rxy. Do mesmo modo, Ryx. Mas isto contradiz a hipótese de que R é assimétrica.

Para estabelecer a transitividade, suponhamos que R^+xy e R^+yz. Se tanto $\langle x, y\rangle \in R$ como $\langle y, z\rangle \in R$, então $\langle x, z\rangle \in R$, pois R é transitiva. Caso contrário, ou $\langle x, y\rangle \in \mathrm{Id}_A$, isto é, $x = y$, ou

$\langle y, z\rangle \in \mathrm{Id}_A$, isto é, $y = z$. No primeiro caso, temos R^+yz por hipótese, $x = y$, logo R^+xz. Analogamente no segundo caso. Em qualquer caso, R^+xz; assim, R^+ é também transitiva.

Quanto à cláusula "além disso", suponhamos que R é também conexa. Então, para todo $x \neq y$, ou Rxy ou Ryx, isto é, ou $\langle x, y\rangle \in R$ ou $\langle y, x\rangle \in R$. Como $R \subseteq R^+$, isto continua a valer para R^+, logo R^+ é também conexa. □

Proposição 3.26. *Se R é uma ordem parcial sobre A, então $R^- = R \setminus \mathrm{Id}_A$ é uma ordem estrita. Além disso, se R é uma ordem linear, então R^- é uma ordem linear estrita.*

Prova. Deixamos como exercício. □

O resultado simples a seguir estabelece que ordens lineares estritas satisfazem uma propriedade no estilo da extensionalidade:

Proposição 3.27. *Se $<$ é uma ordem linear estrita sobre A, então:*

$$(\forall a, b \in A)((\forall x \in A)(x < a \leftrightarrow x < b) \rightarrow a = b).$$

Prova. Suponhamos $(\forall x \in A)(x < a \leftrightarrow x < b)$. Se $a < b$, então $a < a$, o que contradiz o fato de $<$ ser irreflexiva; logo $a \not< b$. De forma totalmente análoga, $b \not< a$. Logo $a = b$, pois $<$ é conexa.□

3.6 Operações sobre Relações

É frequentemente útil modificar ou combinar relações. Na Proposição 3.25, consideramos a *união* de relações, que é apenas a união de duas relações vistas como conjuntos de pares. Do mesmo modo, na Proposição 3.26, consideramos a diferença relativa de relações. Eis algumas outras operações que podemos efetuar sobre relações.

Definição 3.28. Sejam R e S relações, e A um conjunto qualquer.

A *inversa* de R é $R^{-1} = \{\langle y,x\rangle : \langle x,y\rangle \in R\}$.

O *produto relativo* de R e S é $(R \mid S) = \{\langle x,z\rangle : \exists y(Rxy \wedge Syz)\}$.

A *restrição* de R a A é $R\restriction_A = R \cap A^2$.

A *aplicação* de R a A é $R[A] = \{y : (\exists x \in A)Rxy\}$

Exemplo 3.29. Seja $S \subseteq \mathbb{Z}^2$ a relação sucessora sobre $\mathbb{Z}$, isto é, $S = \{\langle x,y\rangle \in \mathbb{Z}^2 : x+1 = y\}$, de modo que Sxy sse $x+1 = y$.

S^{-1} é a relação antecessora sobre $\mathbb{Z}$, isto é, $\{\langle x,y\rangle \in \mathbb{Z}^2 : x-1 = y\}$.

$S \mid S$ é $\{\langle x,y\rangle \in \mathbb{Z}^2 : x+2 = y\}$

$S\restriction_{\mathbb{N}}$ é a relação sucessora sobre $\mathbb{N}$.

$S[\{1,2,3\}]$ é $\{2,3,4\}$.

Definição 3.30 (Fecho transitivo). Seja $R \subseteq A^2$ uma relação binária.

O *fecho transitivo* de R é $R^+ = \bigcup_{0<n\in\mathbb{N}} R^n$, onde definimos recursivamente $R^1 = R$ e $R^{n+1} = R^n \mid R$.

O *fecho reflexivo transitivo* de R é $R^* = R^+ \cup \mathrm{Id}_A$.

Exemplo 3.31. Tome a relação sucessora $S \subseteq \mathbb{Z}^2$. S^2xy sse $x+2 = y$, S^3xy sse $x+3 = y$, etc. Logo S^+xy sse $x+n = y$ para algum $n \geq 1$. Em outras palavras, S^+xy sse $x < y$, e S^*xy sse $x \leq y$.

Problemas

Problema 3.1. Liste os membros da relação $\subseteq$ no conjunto $\wp(\{a,b,c\})$.

Problema 3.2. Dê exemplos de relações que sejam (a) reflexivas e simétricas mas não transitivas, (b) reflexivas e antissimétricas, (c) antissimétricas, transitivas, mas não reflexivas, e (d) reflexivas, simétricas e transitivas. Não use relações sobre números ou conjuntos.

Problema 3.3. Mostre que $\equiv_n$ é uma relação de equivalência, para qualquer $n \in \mathbb{Z}^+$, e que $\mathbb{N}/_{\equiv_n}$ tem exatamente n membros.

Problema 3.4. Apresente uma prova da Proposição 3.26.

Problema 3.5. Mostre que o fecho transitivo de R é de fato transitivo.

CAPÍTULO 4

Funções

4.1 Noções Básicas

Uma *função* é um mapa que envia cada membro de um dado conjunto para um membro específico em algum (outro) conjunto dado. Por exemplo, a operação de somar 1 define uma função: cada número n é mapeado para um único número $n + 1$.

Mais em geral, funções podem tomar pares, triplos, etc., como entradas e devolver algum tipo de saída. Muitas funções nos são familiares da aritmética básica. Por exemplo, a adição e a multiplicação são funções. Tomam dois números e devolvem um terceiro.

Neste sentido matemático e abstrato, uma função é uma *caixa preta*: só importa que saída fica associada a que entrada, não o método de calcular a saída.

Definição 4.1 (Função). Uma *função* $f : A \to B$ é um mapeamento de cada membro de A para um membro de B.

Chamamos A de *domínio* de f e B de *contradomínio* de f. Os membros de A chamam-se entradas ou *argumentos* de f, e o membro de B que fica associado a um argumento x por f chama-se o *valor de* f para o argumento x, escrevendo-se $f(x)$.

A *imagem* $\mathrm{Im}(f)$ de f é o subconjunto do contradomínio constituído pelos valores de f para algum argumento; $\mathrm{Im}(f) =$

$\{f(x) : x \in A\}$.

O diagrama na Figura 4.1 pode ajudar a pensar em funções. A elipse à esquerda representa o *domínio* da função; a elipse à direita representa o *contradomínio*; e uma seta aponta de um *argumento* no domínio para o *valor* correspondente no contradomínio.

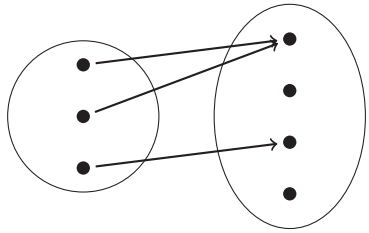


Figura 4.1: Uma função é um mapeamento de cada membro de um conjunto para um membro de outro. Uma seta vai de um argumento no domínio para o valor correspondente no contradomínio.

Exemplo 4.2. A multiplicação toma pares de números naturais como entradas e mapeia-os para números naturais como saídas, logo vai de $\mathbb{N} \times \mathbb{N}$ (o domínio) para $\mathbb{N}$ (o contradomínio). Como se vê, a imagem é também $\mathbb{N}$, pois todo $n \in \mathbb{N}$ é $n \times 1$.

Exemplo 4.3. A multiplicação é uma função porque associa cada entrada—cada par de números naturais—a uma única saída: $\times : \mathbb{N}^2 \to \mathbb{N}$. Em contraste, mapear um número natural n para um real x tal que $x^2 = n$ não é funcional, pois cada inteiro positivo n tem duas raízes quadradas: $\sqrt{n}$ e $-\sqrt{n}$. Podemos torná-la funcional devolvendo só a raiz quadrada positiva: $\sqrt{\ } : \mathbb{N} \to \mathbb{R}$.

Exemplo 4.4. A relação que associa cada aluno de uma turma à sua nota final é uma função—nenhum aluno pode ter duas notas finais diferentes na mesma turma. A relação que associa cada aluno de uma turma aos seus progenitores não é uma função: os alunos podem ter zero, ou dois, ou mais progenitores.

Podemos definir funções especificando de modo preciso qual é o valor da função para cada argumento possível. Formas de o

fazer são dar uma fórmula, descrever um método para calcular o valor, ou listar os valores para cada argumento. Como quer que as funções sejam definidas, temos de garantir que, para cada argumento, especificamos um, e apenas um, valor.

Exemplo 4.5. Seja $f\colon \mathbb{N} \to \mathbb{N}$ definida por $f(x) = x + 1$. Esta é uma definição que especifica f como uma função que toma números naturais e devolve números naturais. Diz-nos que, dado um número natural x, f devolverá o seu sucessor $x + 1$. Neste caso, o contradomínio $\mathbb{N}$ não é a imagem de f, pois o número natural 0 não é sucessor de nenhum número natural. A imagem de f é o conjunto de todos os inteiros positivos, $\mathbb{Z}^+$.

Exemplo 4.6. Seja $g\colon \mathbb{N} \to \mathbb{N}$ definida por $g(x) = x + 2 - 1$. Isto nos diz que g é uma função que toma números naturais e devolve números naturais. Dado um número natural n, g devolverá o predecessor do sucessor do sucessor de n, isto é, $n + 1$.

Acabamos de considerar duas funções, f e g, com *definições* diferentes. Contudo, são a *mesma função.* Afinal, para qualquer número natural n, temos $f(n) = n + 1 = n + 2 - 1 = g(n)$. Em outras palavras: as nossas definições de f e g especificam o mesmo mapeamento por meio de equações diferentes. Implicitamente, pois, estamos nos apoiando num princípio de extensionalidade para funções,

$$\text{se } \forall x\, f(x) = g(x), \text{ então } f = g$$

desde que f e g compartilhem o mesmo domínio e contradomínio.

Exemplo 4.7. Também podemos definir funções por casos. Por exemplo, poderíamos definir $h\colon \mathbb{N} \to \mathbb{N}$ por

$$h(x) = \begin{cases} \frac{x}{2} & \text{se } x \text{ é par} \\ \frac{x+1}{2} & \text{se } x \text{ é ímpar.} \end{cases}$$

Como todo número natural é par ou ímpar, a saída desta função será sempre um número natural. Basta lembrar que, se se define uma função por casos, cada entrada possível tem de cair em exatamente um caso. Em alguns casos, isto exigirá uma prova de que os casos são exaustivos e exclusivos.

4.2 Tipos de Funções

Será útil introduzir uma espécie de taxonomia para alguns dos tipos de funções que encontramos com mais frequência.

Para começar, podemos querer considerar funções com a propriedade de que todo *membro* do contradomínio é um valor da função. Tais funções chamam-se sobrejetivas, e podem ser esquematizadas como na Figura 4.2.

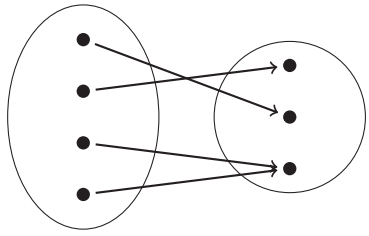


Figura 4.2: Uma função sobrejetiva tem todo membro do contradomínio como valor.

Definição 4.8 (Função Sobrejetiva). Uma função $f : A \to B$ é *sobrejetiva* se e só se B é também a imagem de f, isto é, para todo $y \in B$ existe pelo menos um $x \in A$ tal que $f(x) = y$, ou em símbolos:

$$(\forall y \in B)(\exists x \in A) f(x) = y.$$

Chamamos a tal função uma sobrejeção de A para B.

Se quiser mostrar que f é uma sobrejeção, tem de mostrar que todo objeto no contradomínio de f é o valor de $f(x)$ para alguma entrada x.

Note-se que qualquer função *induz* uma sobrejeção. Afinal, dada uma função $f : A \to B$, seja $f' : A \to \text{Im}(f)$ definida por $f'(x) = f(x)$. Como $\text{Im}(f)$ é *definida* como $\{f(x) \in B : x \in A\}$, esta função f' é garantidamente uma sobrejeção

Agora, qualquer função mapeia cada entrada possível para uma única saída. Mas há também funções que nunca mapeiam entradas diferentes para a mesma saída. Tais funções chamam-se injetivas, e podem ser esquematizadas como na Figura 4.3.

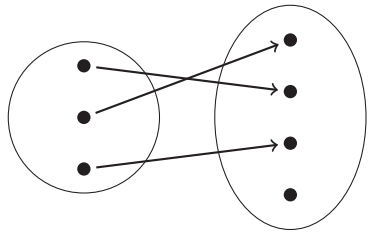


Figura 4.3: Uma função injetiva nunca mapeia dois argumentos diferentes para o mesmo valor.

Definição 4.9 (Função Injetiva). Uma função $f : A \to B$ é *injetiva* se e só se, para cada $y \in B$, há no máximo um $x \in A$ tal que $f(x) = y$. Chamamos a tal função uma injeção de A para B.

Se quiser mostrar que f é uma injeção, tem de mostrar que, para quaisquer membros x e y do domínio de f, se $f(x) = f(y)$, então $x = y$.

Exemplo 4.10. A função constante $f : \mathbb{N} \to \mathbb{N}$ dada por $f(x) = 1$ não é nem injetiva, nem sobrejetiva.

A função identidade $f : \mathbb{N} \to \mathbb{N}$ dada por $f(x) = x$ é tanto injetiva quanto sobrejetiva.

A função sucessora $f : \mathbb{N} \to \mathbb{N}$ dada por $f(x) = x+1$ é injetiva mas não sobrejetiva.

A função $f : \mathbb{N} \to \mathbb{N}$ definida por:

$$f(x) = \begin{cases} \frac{x}{2} & \text{se } x \text{ é par} \\ \frac{x+1}{2} & \text{se } x \text{ é ímpar.} \end{cases}$$

é sobrejetiva, mas não injetiva.

Muitas vezes, queremos considerar funções que são tanto injetivas como sobrejetivas. Chamamos a tais funções bijetivas. Parecem-se com a função esquematizada na Figura 4.4. Bijeções são também por vezes chamadas *correspondências biunívocas*, pois associam de modo único elementos do contradomínio a elementos do domínio.

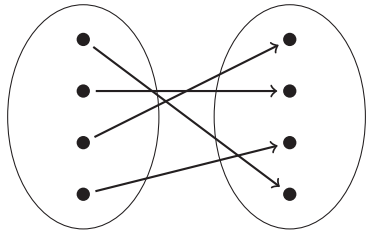


Figura 4.4: Uma função bijetiva associa de modo único os elementos do contradomínio aos do domínio.

Definição 4.11 (Bijeção). Uma função $f : A \to B$ é *bijetiva* se e só se é tanto sobrejetiva como injetiva. Chamamos a tal função uma bijeção de A para B (ou entre A e B).

4.3 Funções como Relações

Uma função que mapeia membros de A para membros de B define obviamente uma relação entre A e B, nomeadamente a relação que se verifica entre x e y se e só se $f(x) = y$. De fato, podemos até—se nos interessa reduzir os blocos de construção da matemática, por exemplo—*identificar* a função f com esta relação, isto é, com um conjunto de pares. Isto levanta então a questão: que relações definem funções deste modo?

Definição 4.12 (Gráfico de uma função). Seja $f : A \to B$ uma função. O *gráfico* de f é a relação $R_f \subseteq A \times B$ definida

por

$$R_f = \{\langle x, y\rangle : f(x) = y\}.$$

O gráfico de uma função fica unicamente determinado, pela extensionalidade. Além disso, a extensionalidade (sobre conjuntos) justifica de imediato o princípio implícito de extensionalidade para funções, segundo o qual, se f e g compartilham domínio e contradomínio, então são idênticas se coincidem em todos os valores.

Do mesmo modo, se uma relação é "funcional", então é o gráfico de uma função.

Proposição 4.13. *Seja $R \subseteq A \times B$ tal que:*

1. *Se Rxy e Rxz, então $y = z$; e*
2. *para todo $x \in A$ existe algum $y \in B$ tal que $\langle x, y\rangle \in R$.*

Então R é o gráfico da função $f : A \to B$ definida por $f(x) = y$ se e só se Rxy.

Prova. Suponhamos que existe y tal que Rxy. Se existisse outro $z \neq y$ tal que Rxz, a condição sobre R seria violada. Logo, se existe y tal que Rxy, esse y é único, e assim f está bem definida. Obviamente, $R_f = R$. □

Toda função $f : A \to B$ tem um gráfico, isto é, uma relação em $A \times B$ definida por $f(x) = y$. Por outro lado, toda relação $R \subseteq A \times B$ com as propriedades dadas na Proposição 4.13 é o gráfico de uma função $f : A \to B$. Por causa desta ligação estreita entre funções e os seus gráficos, podemos pensar numa função simplesmente como o seu gráfico. Em outras palavras, funções podem identificar-se com certas relações, isto é, com certos conjuntos de tuplos. Note, contudo, que o espírito desta "identificação" é como na Seção 3.2: não é uma afirmação sobre a metafísica das funções, mas a observação de que é conveniente *tratar* funções como certos conjuntos. Uma razão pela

qual isto é tão conveniente é que podemos agora considerar fazer operações semelhantes sobre funções às que fizemos sobre relações (ver Seção 3.6). Em particular:

Definição 4.14. Seja $f : A \to B$ uma função com $C \subseteq A$.

A *restrição* de f a C é a função $f\restriction_C : C \to B$ definida por $(f\restriction_C)(x) = f(x)$ para todo $x \in C$. Em outras palavras, $f\restriction_C = \{\langle x, y\rangle \in R_f : x \in C\}$.

A *aplicação* de f a C é $f[C] = \{f(x) : x \in C\}$. Chamamos-lhe também a *imagem* de C por f.

Segue-se destas definições que $\mathrm{Im}(f) = f[\mathrm{dom}(f)]$, para qualquer função f. Estas noções são exatamente as que se esperaria, dadas as definições na Seção 3.6 e a nossa identificação de funções com relações. Mas outras duas operações—inversas e produtos relativos—exigem um pouco mais de detalhe. Trataremos disso na Seção 4.4 e na Seção 4.5.

4.4 Inversas de Funções

Pensamos nas funções como mapas. Uma pergunta natural sobre funções é, pois, se o mapeamento pode ser "invertido". Por exemplo, a função sucessora $f(x) = x + 1$ pode ser invertida, no sentido em que a função $g(y) = y - 1$ "desfaz" o que f faz.

Mas é preciso ter cuidado. Embora a definição de g defina uma função $\mathbb{Z} \to \mathbb{Z}$, não define uma *função* $\mathbb{N} \to \mathbb{N}$, pois $g(0) \notin \mathbb{N}$. Assim, mesmo em casos simples, não é óbvio que uma função possa ser invertida; pode depender do domínio e do contradomínio.

Isto torna-se mais preciso com a noção de inversa de uma função.

Definição 4.15. Uma função $g : B \to A$ é uma *inversa* de uma

função $f: A \to B$ se $f(g(y)) = y$ e $g(f(x)) = x$ para todo $x \in A$ e $y \in B$.

Se f tem uma inversa g, escrevemos muitas vezes f^{-1} em vez de g.

Agora determinaremos quando as funções têm inversas. Um bom candidato a inversa de $f: A \to B$ é $g: B \to A$ "definida por"

$$g(y) = \text{"o" } x \text{ tal que } f(x) = y.$$

Mas as aspas em "definida por" (e em "o") sugerem que isto não é uma definição. Pelo menos, nem sempre funcionará com total generalidade. Pois, para que esta definição especifique uma função, tem de haver um e apenas um x tal que $f(x) = y$—a saída de g tem de ficar unicamente determinada. Além disso, tem de estar especificada para todo $y \in B$. Se existirem x_1 e $x_2 \in A$ com $x_1 \neq x_2$ mas $f(x_1) = f(x_2)$, então $g(y)$ não ficaria unicamente determinada para $y = f(x_1) = f(x_2)$. E se não existir x algum tal que $f(x) = y$, então $g(y)$ nem sequer fica especificada. Em outras palavras, para g estar definida, f tem de ser tanto injetiva quanto sobrejetiva.

Vamos com calma. Dividiremos a questão em duas: dada uma função $f: A \to B$, quando existe uma função $g: B \to A$ tal que $g(f(x)) = x$? Tal g "desfaz" o que f faz, e chama-se *inversa à esquerda* de f. Em segundo lugar, quando existe uma função $h: B \to A$ tal que $f(h(y)) = y$? Tal h chama-se *inversa à direita* de f—f "desfaz" o que h faz.

Proposição 4.16. *Se $f: A \to B$ é injetiva, então existe uma* inversa à esquerda *$g: B \to A$ de f tal que $g(f(x)) = x$ para todo $x \in A$.*

Prova. Suponhamos que $f: A \to B$ é injetiva. Consideremos $y \in B$. Se $y \in \mathrm{Im}(f)$, existe $x \in A$ tal que $f(x) = y$. Como f é injetiva, há apenas um tal $x \in A$. Então podemos definir: $g(y) = x$, isto é, $g(y)$ é "o" $x \in A$ tal que $f(x) = y$. Se $y \notin \mathrm{Im}(f)$, podemos mapeá-lo para qualquer $a \in A$. Assim, podemos escolher $a \in A$

e definir $g\colon B \to A$ por:

$$g(y) = \begin{cases} x & \text{se } f(x) = y \\ a & \text{se } y \notin \mathrm{Im}(f). \end{cases}$$

Está definida para todo $y \in B$, pois para cada tal $y \in \mathrm{Im}(f)$ há exatamente um $x \in A$ tal que $f(x) = y$. Por definição, se $y = f(x)$, então $g(y) = x$, isto é, $g(f(x)) = x$. □

Proposição 4.17. *Se $f\colon A \to B$ é sobrejetiva, então existe uma* inversa à direita *$h\colon B \to A$ de f tal que $f(h(y)) = y$ para todo $y \in B$.*

Prova. Suponhamos que $f\colon A \to B$ é sobrejetiva. Consideremos $y \in B$. Como f é sobrejetiva, existe $x_y \in A$ com $f(x_y) = y$. Então podemos definir: $h(y) = x_y$, isto é, para cada $y \in B$ escolhemos algum $x \in A$ tal que $f(x) = y$; como f é sobrejetiva há sempre pelo menos um de onde escolher.[1] Por definição, se $x = h(y)$, então $f(x) = y$, isto é, para todo $y \in B$, $f(h(y)) = y$. □

Combinando as ideias da prova anterior, obtemos agora que toda bijeção tem uma inversa, isto é, existe uma única função que é tanto inversa à esquerda quanto à direita de f.

Proposição 4.18. *Se $f\colon A \to B$ é bijetiva, existe uma função $f^{-1}\colon B \to A$ tal que, para todo $x \in A$, $f^{-1}(f(x)) = x$ e, para todo $y \in B$, $f(f^{-1}(y)) = y$.*

Prova. Exercício. □

[1] Como f é sobrejetiva, para todo $y \in B$ o conjunto $\{x : f(x) = y\}$ é não vazio. A nossa definição de h exige que escolhamos um único x em cada um destes conjuntos. Que isto seja sempre possível não é de todo óbvio—a possibilidade de fazer estas escolhas é simplesmente assumida como axioma. Em outras palavras, esta proposição pressupõe o chamado Axioma da Escolha, questão a que voltaremos no Capítulo 16. Contudo, em muitos casos concretos, p.ex., quando $A = \mathbb{N}$ ou é finito, ou quando f é bijetiva, o Axioma da Escolha não é necessário. (No caso em que f é bijetiva, para cada $y \in B$ o conjunto $\{x : f(x) = y\}$ tem exatamente um membro, logo não há escolha a fazer.)

Há uma forma ligeiramente mais geral de extrair inversas. Vimos na Seção 4.2 que toda função f induz uma sobrejeção $f' \colon A \to \mathrm{Im}(f)$ pondo $f'(x) = f(x)$ para todo $x \in A$. Claramente, se f é injetiva, então f' é bijetiva, logo tem uma inversa única pela Proposição 4.18. Por um pequeno abuso de notação, chamamos por vezes à inversa de f' simplesmente "a inversa de f."

Proposição 4.19. *Mostre que, se $f \colon A \to B$ tem uma inversa à esquerda g e uma inversa à direita h, então $h = g$.*

Prova. Exercício. □

Proposição 4.20. *Toda função f tem no máximo uma inversa.*

Prova. Suponhamos que g e h são ambas inversas de f. Então, em particular, g é uma inversa à esquerda de f e h é uma inversa à direita. Pela Proposição 4.19, $g = h$. □

4.5 Composição de Funções

Vimos na Seção 4.4 que a inversa f^{-1} de uma bijeção f é ela própria uma função. Outra operação sobre funções é a composição: podemos definir uma nova função compondo duas funções, f e g, isto é, aplicando primeiro f e depois g. Claro que isto só é possível se as imagens e os domínios forem compatíveis, isto é, a imagem de f tem de ser um subconjunto do domínio de g. Esta operação sobre funções é o análogo da operação de produto relativo sobre relações na Seção 3.6.

Um diagrama pode ajudar a explicar a ideia de composição. Na Figura 4.5, representamos duas funções $f \colon A \to B$ e $g \colon B \to C$ e a sua composição $(g \circ f)$. A função $(g \circ f) \colon A \to C$ associa cada membro de A a um membro de C. Especificamos a que membro de C fica associado um membro de A do seguinte modo: dada uma entrada $x \in A$, aplicamos primeiro a função f a x,

obtendo algum $f(x) = y \in B$, depois aplicamos a função g a y, obtendo algum $g(f(x)) = g(y) = z \in C$.

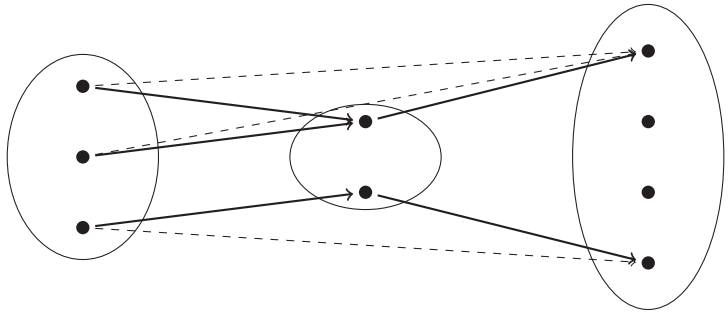


Figura 4.5: A composição $g \circ f$ de duas funções f e g.

Definição 4.21 (Composição). Sejam $f : A \to B$ e $g : B \to C$ funções. A *composição* de f com g é $g \circ f : A \to C$, onde $(g \circ f)(x) = g(f(x))$.

Exemplo 4.22. Consideremos as funções $f(x) = x + 1$, e $g(x) = 2x$. Como $(g \circ f)(x) = g(f(x))$, para cada entrada x deve primeiro tomar-se o sucessor, depois multiplicar o resultado por dois. Logo a composição é dada por $(g \circ f)(x) = 2(x + 1)$.

Problemas

Problema 4.1. Mostre que, se $f : A \to B$ tem uma inversa à esquerda g, então f é injetiva.

Problema 4.2. Mostre que, se $f : A \to B$ tem uma inversa à direita h, então f é sobrejetiva.

Problema 4.3. Prove Proposição 4.18. Tem de definir f^{-1}, mostrar que é uma função, e mostrar que é uma inversa de f, isto é, $f^{-1}(f(x)) = x$ e $f(f^{-1}(y)) = y$ para todo $x \in A$ e $y \in B$.

Problema 4.4. Prove Proposição 4.19.

Problema 4.5. Mostre que, se $f : A \to B$ e $g : B \to C$ são ambas injetiva, então $g \circ f : A \to C$ é injetiva.

Problema 4.6. Mostre que, se $f : A \to B$ e $g : B \to C$ são ambas sobrejetiva, então $g \circ f : A \to C$ é sobrejetiva.

Problema 4.7. Suponhamos $f : A \to B$ e $g : B \to C$. Mostre que o gráfico de $g \circ f$ é $R_f \mid R_g$.

CAPÍTULO 5

O Tamanho dos Conjuntos

5.1 Introdução

Quando Georg Cantor desenvolveu a teoria dos conjuntos na década de 1870, um dos seus objetivos era tornar aceitável a ideia de uma coleção infinita—um infinito *atual*, como diriam os medievais. Parte central disto foi o seu tratamento do *tamanho* de conjuntos diferentes. Se a, b e c são todos distintos, então o conjunto $\{a, b, c\}$ é intuitivamente *maior* que $\{a, b\}$. Mas e os conjuntos infinitos? Serão todos do mesmo tamanho? Acontece que não.

A primeira ideia importante aqui é a de *enumeração.* Podemos listar qualquer conjunto finito enumerando todos os seus membros. Para alguns conjuntos infinitos, também podemos listar todos os seus membros se permitirmos que a própria lista seja infinita. Tais conjuntos chamam-se enumeráveis. O resultado surpreendente de Cantor, que compreenderemos completamente no fim deste capítulo, foi que alguns conjuntos infinitos *não* são enumeráveis.

5.2 Enumerações e Conjuntos Enumeráveis

Podemos especificar um conjunto finito simplesmente enumerando os seus membros. Fazemo-lo quando definimos um conjunto assim:

$$A = \{a_1, a_2, \ldots, a_n\}.$$

Supondo que os membros a_1, ..., a_n são todos distintos, obtemos uma bijeção entre A e os primeiros n números naturais 0, ..., $n-1$. Reciprocamente, como todo o conjunto finito tem apenas finitos membros, todo o conjunto finito pode ser posto nessa correspondência. Em outras palavras, se A é finito, existe uma bijeção entre A e $\{0, \ldots, n-1\}$, onde n é o número de membros de A.

Se admitirmos certos tipos de conjuntos infinitos, admitiremos também que alguns conjuntos infinitos possam ser enumerados. Podemos precisar isto dizendo que um conjunto infinito é enumerado por uma bijeção entre ele e todo o $\mathbb{N}$.

Definição 5.1 (Enumeração, no sentido conjuntista). Uma *enumeração* de um conjunto A é uma bijeção cujo contradomínio é A e cujo domínio é ou um conjunto inicial de números naturais $\{0, 1, \ldots, n\}$ *ou* todo o conjunto dos números naturais $\mathbb{N}$.

Há uma base intuitiva para este uso da palavra *enumeração*. Dizer que enumeramos um conjunto A é dizer que existe uma bijeção f que nos permite *contar* os elementos do conjunto A. O elemento de índice 0 é $f(0)$, o de índice 1 é $f(1)$, ... o de índice n é $f(n)$....[1] O raciocínio fica ainda mais claro com o seguinte:

Definição 5.2. Um conjunto A é enumerável se e só se $A = \emptyset$ ou

[1] Sim, contamos a partir de 0. Claro que também poderíamos começar em 1. Não faria grande diferença. Bastaria substituir $\mathbb{N}$ por $\mathbb{Z}^+$.

existe uma enumeração de A. Dizemos que A é não enumerável se e só se A *não* é enumerável.

Logo um conjunto é enumerável se e só se é vazio ou se podemos usar uma enumeração para contar os seus membros.

Exemplo 5.3. Uma função que enumera os números naturais é simplesmente a função identidade $\mathrm{Id}_{\mathbb{N}} \colon \mathbb{N} \to \mathbb{N}$ dada por $\mathrm{Id}_{\mathbb{N}}(n) = n$. Uma função que enumera os números naturais *positivos*, $\mathbb{N}^+ = \mathbb{N} \setminus \{0\}$, é a função $g(n) = n + 1$, isto é, a função sucessora.

Exemplo 5.4. As funções $f \colon \mathbb{N} \to \mathbb{N}$ e $g \colon \mathbb{N} \to \mathbb{N}$ dadas por

$$f(n) = 2n \text{ e}$$
$$g(n) = 2n + 1$$

enumeram, respectivamente, os números naturais pares e os ímpares. Mas nenhuma é sobrejetiva, logo nenhuma é uma enumeração de $\mathbb{N}$.

Exemplo 5.5. Seja $\lceil x \rceil$ a função *teto*, que arredonda x por excesso para o inteiro mais próximo. Então a função $f \colon \mathbb{N} \to \mathbb{Z}$ dada por:

$$f(n) = (-1)^n \left\lceil \tfrac{n}{2} \right\rceil$$

enumera o conjunto dos inteiros $\mathbb{Z}$ do seguinte modo:

$$\begin{array}{cccccccc} f(0) & f(1) & f(2) & f(3) & f(4) & f(5) & f(6) & \dots \\ \left\lceil \frac{0}{2} \right\rceil & -\left\lceil \frac{1}{2} \right\rceil & \left\lceil \frac{2}{2} \right\rceil & -\left\lceil \frac{3}{2} \right\rceil & \left\lceil \frac{4}{2} \right\rceil & -\left\lceil \frac{5}{2} \right\rceil & \left\lceil \frac{6}{2} \right\rceil & \dots \\ 0 & -1 & 1 & -2 & 2 & -3 & 3 & \dots \end{array}$$

Repare como f gera os valores de $\mathbb{Z}$ "saltitando" entre inteiros positivos e negativos. Também se pode ver f definida por casos assim:

$$f(n) = \begin{cases} \frac{n}{2} & \text{se } n \text{ é par} \\ -\frac{n+1}{2} & \text{se } n \text{ é ímpar} \end{cases}$$

5.3 Método do Zigue-Zague de Cantor

Já consideramos algumas enumerações "fáceis". Vejamos algo um pouco mais difícil. Considere-se o conjunto dos pares de números naturais, que definimos na Seção 2.5 assim:

$$\mathbb{N} \times \mathbb{N} = \{\langle n, m\rangle : n, m \in \mathbb{N}\}$$

Podemos organizar estes pares ordenados numa *matriz*, assim:

	0	**1**	**2**	**3**	...
0	$\langle 0,0\rangle$	$\langle 0,1\rangle$	$\langle 0,2\rangle$	$\langle 0,3\rangle$	...
1	$\langle 1,0\rangle$	$\langle 1,1\rangle$	$\langle 1,2\rangle$	$\langle 1,3\rangle$	...
2	$\langle 2,0\rangle$	$\langle 2,1\rangle$	$\langle 2,2\rangle$	$\langle 2,3\rangle$	...
3	$\langle 3,0\rangle$	$\langle 3,1\rangle$	$\langle 3,2\rangle$	$\langle 3,3\rangle$	...
⋮	⋮	⋮	⋮	⋮	⋱

Claramente, cada par ordenado em $\mathbb{N} \times \mathbb{N}$ aparece exatamente uma vez na matriz. Em particular, $\langle n, m\rangle$ aparece na n-ésima linha e na m-ésima coluna. Mas como organizar os elementos de tal matriz numa lista "unidimensional"? O padrão na matriz abaixo demonstra uma forma de o fazer (embora haja, claro, muitas outras):

	0	**1**	**2**	**3**	**4**	...
0	0	1	3	6	10	...
1	2	4	7	11	...	...
2	5	8	12	...	...	...
3	9	13	...	...	...	...
4	14	...	...	...	...	...
⋮	⋮	⋮	⋮	⋮	⋮	⋱

Este padrão chama-se *método do zigue-zague de Cantor*. Enumera $\mathbb{N} \times \mathbb{N}$ do seguinte modo:

$$\langle 0,0\rangle, \langle 0,1\rangle, \langle 1,0\rangle, \langle 0,2\rangle, \langle 1,1\rangle, \langle 2,0\rangle, \langle 0,3\rangle, \langle 1,2\rangle, \langle 2,1\rangle, \langle 3,0\rangle, \ldots$$

Isto estabelece o seguinte:

Proposição 5.6. $\mathbb{N} \times \mathbb{N}$ *é enumerável.*

Prova. Seja $f : \mathbb{N} \to \mathbb{N} \times \mathbb{N}$ a função que a cada $k \in \mathbb{N}$ associa o tuplo $\langle n, m \rangle \in \mathbb{N} \times \mathbb{N}$ tal que k é o valor na n-ésima linha e m-ésima coluna da matriz do zigue-zague de Cantor. □

Esta técnica generaliza-se bem. Por exemplo, podemos usá-la para enumerar o conjunto dos triplos ordenados de números naturais, isto é:

$$\mathbb{N} \times \mathbb{N} \times \mathbb{N} = \{\langle n, m, k \rangle : n, m, k \in \mathbb{N}\}$$

Pensamos em $\mathbb{N} \times \mathbb{N} \times \mathbb{N}$ como o produto cartesiano de $\mathbb{N} \times \mathbb{N}$ com $\mathbb{N}$, isto é,

$$\mathbb{N}^3 = (\mathbb{N} \times \mathbb{N}) \times \mathbb{N} = \{\langle \langle n, m \rangle, k \rangle : n, m, k \in \mathbb{N}\}$$

e assim podemos enumerar $\mathbb{N}^3$ com uma matriz rotulando um eixo com a enumeração de $\mathbb{N}$ e o outro com a enumeração de $\mathbb{N}^2$:

	0	**1**	**2**	**3**	...
$\langle \mathbf{0}, \mathbf{0} \rangle$	$\langle 0,0,0 \rangle$	$\langle 0,0,1 \rangle$	$\langle 0,0,2 \rangle$	$\langle 0,0,3 \rangle$	...
$\langle \mathbf{0}, \mathbf{1} \rangle$	$\langle 0,1,0 \rangle$	$\langle 0,1,1 \rangle$	$\langle 0,1,2 \rangle$	$\langle 0,1,3 \rangle$	...
$\langle \mathbf{1}, \mathbf{0} \rangle$	$\langle 1,0,0 \rangle$	$\langle 1,0,1 \rangle$	$\langle 1,0,2 \rangle$	$\langle 1,0,3 \rangle$	...
$\langle \mathbf{0}, \mathbf{2} \rangle$	$\langle 0,2,0 \rangle$	$\langle 0,2,1 \rangle$	$\langle 0,2,2 \rangle$	$\langle 0,2,3 \rangle$	...
⋮	⋮	⋮	⋮	⋮	⋱

Assim, com um método como o zigue-zague de Cantor, obtemos analogamente uma enumeração de $\mathbb{N}^3$. E podemos continuar, obtendo enumerações de $\mathbb{N}^n$ para qualquer número natural n. Logo:

Proposição 5.7. $\mathbb{N}^n$ *é enumerável, para todo* $n \in \mathbb{N}$.

5.4 Funções de Pareamento e Códigos

O método do zigue-zague de Cantor torna visualmente evidente a enumerabilidade de $\mathbb{N}^n$. Mas vamos concentrar-nos na nossa matriz que representa $\mathbb{N}^2$. Seguindo a linha em zigue-zague na matriz e contando as posições, podemos verificar que $\langle 1,2\rangle$ fica associado ao número 7. Seria útil, contudo, poder calcular isto de modo mais direto. Ou seja, seria útil dispor da *inversa* da enumeração em zigue-zague, $g\colon \mathbb{N}^2 \to \mathbb{N}$, tal que

$$g(\langle 0,0\rangle) = 0,\ g(\langle 0,1\rangle) = 1,\ g(\langle 1,0\rangle) = 2,\ \ldots, g(\langle 1,2\rangle) = 7,\ \ldots$$

Isto permitir-nos-ia calcular exatamente em que posição da enumeração cai $\langle n,m\rangle$.

De fato, podemos definir g diretamente com base em duas observações. Primeiro: se a n-ésima linha e a m-ésima coluna contêm o valor v, então a $(n+1)$-ésima linha e a $(m-1)$-ésima coluna contêm o valor $v+1$. Segundo: a primeira linha da nossa enumeração consiste nos números triangulares, começando por 0, 1, 3, 6, etc. O k-ésimo número triangular é a soma dos números naturais $< k$, que se pode calcular como $k(k+1)/2$. Juntando estas duas observações, consideremos a função:

$$g(n,m) = \frac{(n+m+1)(n+m)}{2} + n$$

Escreve-se frequentemente $g(n,m)$ em vez de $g(\langle n,m\rangle)$, já que é mais fácil para os olhos. Isto diz-nos que se determine primeiro o $(n+m)$-ésimo número triangular e depois que lhe somemos n. A função preenche a matriz exatamente da maneira que desejamos. Em particular, o par $\langle 1,2\rangle$ é enviado para $\frac{4\times 3}{2} + 1 = 7$.

Esta função g é a *inversa* de uma enumeração de um conjunto de pares. Tais funções chamam-se *funções de emparelhamento.*

Definição 5.8 (Função de emparelhamento). Uma função $f\colon A \times B \to \mathbb{N}$ é uma *função de emparelhamento* aritmética se f for injetiva. Dizemos também que f *codifica* $A \times B$, e que $f(x,y)$

é o *código* de $\langle x, y \rangle$.

Podemos usar funções de emparelhamento para codificar, por exemplo, pares de números naturais; ou, em outras palavras, podemos representar cada *par* de elementos por um *único* número. Usando a inversa da função de emparelhamento, podemos *descodificar* o número, isto é, determinar que par representa.

5.5 Conjuntos Não Enumeráveis

O conjunto $\mathbb{N}$ dos números naturais é infinito. Também é trivialmente enumerável. O fato notável, porém, é que existem conjuntos *não enumeráveis*, isto é, conjuntos que não são enumeráveis (ver Definição 5.2).

Isto pode surpreender. Afinal, dizer que A é não enumerável é dizer que *não* existe bijeção $f : \mathbb{N} \to A$; isto é, nenhuma função que mapeie os infinitos membros de $\mathbb{N}$ para A esgota todo A. Logo, se A é não enumerável, há "mais" membros em A do que números naturais.

Para provar que um conjunto é não enumerável, é preciso mostrar que não pode existir uma bijeção adequada. A melhor forma é mostrar que qualquer tentativa de enumerar membros de A tem de deixar pelo menos um membro de fora; isto mostra que nenhuma função $f : \mathbb{N} \to A$ é sobrejetiva. Uma estratégia geral para o estabelecer é usar o *método diagonal* de Cantor. Dada uma lista de membros de A, digamos x_1, x_2, ..., construímos outro membro de A que, pela sua construção, não pode estar nessa lista.

Tudo isto compreende-se melhor com um exemplo. O nosso primeiro exemplo é o conjunto $\mathbb{B}^\omega$ de todas as cadeias infinitas de 0s e 1s. (O símbolo '$\mathbb{B}$' significa binário, e podemos pensar nele como o conjunto de dois elementos $\{0,1\}$.)

Teorema 5.9. *$\mathbb{B}^\omega$ é não enumerável.*

Prova. Considere-se qualquer enumeração de um subconjunto de $\mathbb{B}^\omega$. Temos então uma lista s_0, s_1, s_2, ... em que cada s_n é uma cadeia infinita de 0s e 1s. Seja $s_n(m)$ o m-ésimo dígito da cadeia s_n. Podemos pois pensar na lista como uma matriz, colocando $s_n(m)$ na n-ésima linha e m-ésima coluna:

	0	1	2	3	...
0	$\mathbf{s_0(0)}$	$s_0(1)$	$s_0(2)$	$s_0(3)$	...
1	$s_1(0)$	$\mathbf{s_1(1)}$	$s_1(2)$	$s_1(3)$	...
2	$s_2(0)$	$s_2(1)$	$\mathbf{s_2(2)}$	$s_2(3)$	...
3	$s_3(0)$	$s_3(1)$	$s_3(2)$	$\mathbf{s_3(3)}$	...
⋮	⋮	⋮	⋮	⋮	⋱

Construiremos agora uma cadeia infinita, d, de 0s e 1s que não está nesta lista, especificando cada uma das suas entradas, isto é, $d(n)$ para todo $n \in \mathbb{N}$. Intuitivamente, lemos a diagonal da matriz acima (daí o nome "método diagonal") e trocamos cada 0 por 1 e cada 1 por 0. De modo mais abstrato, definimos $d(n)$ como 0 ou 1 consoante o n-ésimo membro da diagonal, $s_n(n)$, seja 1 ou 0, isto é:

$$d(n) = \begin{cases} 1 & \text{se } s_n(n) = 0 \\ 0 & \text{se } s_n(n) = 1 \end{cases}$$

Claramente $d \in \mathbb{B}^\omega$, pois é uma cadeia infinita de 0s e 1s. Mas construímos d de modo que $d(n) \neq s_n(n)$ para qualquer $n \in \mathbb{N}$, isto é, d difere de s_n na sua n-ésima entrada. Logo $d \neq s_n$ para qualquer $n \in \mathbb{N}$, pelo que d não pode estar na lista s_0, s_1, s_2, ...

Mostramos assim que, dada uma enumeração arbitrária de algum subconjunto de $\mathbb{B}^\omega$, ela omite algum membro de $\mathbb{B}^\omega$. Logo não existe enumeração do conjunto $\mathbb{B}^\omega$, isto é, $\mathbb{B}^\omega$ é não enumerável. □

Este método de prova chama-se "diagonalização" porque usa a diagonal da matriz para definir d. Contudo, a diagonalização não exige necessariamente uma matriz. De fato, podemos mostrar que um conjunto é não enumerável por uma ideia semelhante

mesmo quando não há matriz nem diagonal literal. O resultado seguinte ilustra como.

Teorema 5.10. *$\wp(\mathbb{N})$ não é enumerável.*

Prova. Procedemos da mesma forma, mostrando que qualquer lista de subconjuntos de $\mathbb{N}$ omite algum subconjunto de $\mathbb{N}$. Suponhamos, pois, que temos uma lista $N_0, N_1, N_2, \ldots$ de subconjuntos de $\mathbb{N}$. Definimos um conjunto D assim: $n \in D$ sse $n \notin N_n$:

$$D = \{n \in \mathbb{N} : n \notin N_n\}$$

Claramente $D \subseteq \mathbb{N}$. Mas D não pode estar na lista: por construção, $n \in D$ sse $n \notin N_n$, logo $D \neq N_n$ para qualquer $n \in \mathbb{N}$. □

A prova precedente não mencionou uma diagonal. Ainda assim, pode pensar-se nela como envolvendo uma diagonal, imaginando os conjuntos N_0, N_1, ... dispostos numa matriz em que escrevemos N_n na n-ésima linha colocando m na m-ésima coluna sse $m \in N_n$. Por exemplo, se os primeiros quatro conjuntos da lista são $\{0,1,2,\ldots\}$, $\{1,3,5,\ldots\}$, $\{0,1,4\}$ e $\{2,3,4,\ldots\}$, a matriz começaria com

$$\begin{array}{lccccccc}
N_0 = \{\mathbf{0}, & 1, & 2, & & & & \ldots\} \\
N_1 = \{ & \mathbf{1}, & & 3, & & 5, & \ldots\} \\
N_2 = \{0, & 1, & & & 4 & & \} \\
N_3 = \{ & & 2, & \mathbf{3}, & 4, & & \ldots\} \\
\vdots & & & & & & \ddots
\end{array}$$

Então D é o conjunto obtido percorrendo a diagonal, pondo $n \in D$ sse *n não* está na diagonal. No caso acima, excluiríamos 0 e 1, incluiríamos 2, excluiríamos 3, etc.

5.6 Redução

Provamos que $\mathbb{B}^\omega$ é não enumerável por um argumento de diagonalização. Usamos um argumento semelhante para mostrar que

$\wp(\mathbb{N})$ é não enumerável. Há outra forma de provar que $\wp(\mathbb{N})$ é não enumerável: mostrar que *se $\wp(\mathbb{N})$ é enumerável então $\mathbb{B}^\omega$ também é enumerável.* Como sabemos que $\mathbb{B}^\omega$ é não enumerável, segue que $\wp(\mathbb{N})$ também o é.

Isto chama-se *reduzir* um problema a outro. Neste caso, reduzimos o problema de enumerar $\mathbb{B}^\omega$ ao problema de enumerar $\wp(\mathbb{N})$. Uma solução para este último—uma enumeração de $\wp(\mathbb{N})$—daria uma solução para o primeiro—uma enumeração de $\mathbb{B}^\omega$.

Para reduzir o problema de enumerar um conjunto B ao de enumerar um conjunto A, indicamos uma forma de transformar uma enumeração de A numa enumeração de B. A forma mais simples é definir uma sobrejeção $f\colon A \to B$. Se $x_1, x_2, \ldots$ enumeram A, então $f(x_1), f(x_2), \ldots$ enumerariam B. No nosso caso, procuramos uma sobrejeção $f\colon \wp(\mathbb{N}) \to \mathbb{B}^\omega$.

Prova do Teorema 5.10 por redução. Para a redução, suponhamos que $\wp(\mathbb{N})$ é enumerável e portanto que existe uma enumeração $N_1, N_2, N_3, \ldots$

Definimos a função $f\colon \wp(\mathbb{N}) \to \mathbb{B}^\omega$ pondo $f(N)$ igual à cadeia s_k tal que $s_k(n) = 1$ sse $n \in N$, e $s_k(n) = 0$ caso contrário.

Isto define claramente uma função, pois sempre que $N \subseteq \mathbb{N}$, qualquer $n \in \mathbb{N}$ ou é um membro de N ou não é. Por exemplo, o conjunto $2\mathbb{N} = \{2n : n \in \mathbb{N}\} = \{0, 2, 4, 6, \ldots\}$ dos naturais pares é mapeado para a cadeia $1010101\ldots$; $\emptyset$ é mapeado para $0000\ldots$; $\mathbb{N}$ é mapeado para $1111\ldots$.

Também é sobrejetiva: toda a cadeia de 0s e 1s corresponde a algum conjunto de números naturais, nomeadamente aquele que tem como membros os naturais nas posições em que a cadeia tem um 1. Com mais precisão, se $s \in \mathbb{B}^\omega$, defina-se $N \subseteq \mathbb{N}$ por:

$$N = \{n \in \mathbb{N} : s(n) = 1\}$$

Então $f(N) = s$, como se verifica consultando a definição de f.

Consideremos agora a lista

$$f(N_1), f(N_2), f(N_3), \ldots$$

Como f é sobrejetiva, todo o membro de $\mathbb{B}^\omega$ tem de aparecer como valor de f para algum argumento, e portanto na lista. Esta lista teria de enumerar todo $\mathbb{B}^\omega$.

Logo, se $\wp(\mathbb{N})$ fosse enumerável, $\mathbb{B}^\omega$ seria enumerável. Mas $\mathbb{B}^\omega$ é não enumerável (Teorema 5.9). Logo $\wp(\mathbb{N})$ é não enumerável. □

5.7 Equinumerosidade

Temos uma noção intuitiva do "tamanho" dos conjuntos, que funciona bem para conjuntos finitos. E quanto aos infinitos? Se quisermos uma forma formal de comparar os tamanhos de dois conjuntos de *qualquer* cardinalidade, é boa ideia começar por definir quando dois conjuntos têm o mesmo tamanho. Eis Frege:

> Se um garçom quer ter certeza de que pôs exatamente tantas facas quantos pratos sobre a mesa, não precisa contar nem uns nem outros, se simplesmente puser uma faca à direita de cada prato, de modo que cada faca na mesa fique à direita de algum prato. Os pratos e as facas ficam assim correlacionados de modo único, e de fato por essa mesma relação espacial. (Frege, 1884, § 70)

A intuição central desta passagem traduz-se numa definição formal:

Definição 5.11. A é *equinumeroso* a B, escrevendo-se $A \approx B$, sse existir uma bijeção $f : A \to B$.

Proposição 5.12. *A relação de equinumerosidade é uma relação de equivalência.*

Prova. É preciso mostrar que a equinumerosidade é reflexiva, simétrica e transitiva. Sejam A, B e C conjuntos.

Reflexividade. A aplicação identidade $\mathrm{Id}_A \colon A \to A$, com $\mathrm{Id}_A(x) = x$ para todo $x \in A$, é uma bijeção. Logo $A \approx A$.

Simetria. Suponhamos $A \approx B$, isto é, que existe uma bijeção $f \colon A \to B$. Como f é bijetiva, a inversa f^{-1} existe e também é bijetiva. Logo $f^{-1} \colon B \to A$ é uma bijeção, pelo que $B \approx A$.

Transitividade. Suponhamos $A \approx B$ e $B \approx C$, isto é, que existem bijeções $f \colon A \to B$ e $g \colon B \to C$. Então a composição $g \circ f \colon A \to C$ é bijetiva, logo $A \approx C$. □

Proposição 5.13. *Se $A \approx B$, então A é enumerável se e só se B o é.*

Prova. Suponhamos $A \approx B$, logo existe alguma bijeção $f \colon A \to B$, e suponhamos que A é enumerável. Então ou $A = \emptyset$ ou existe uma bijeção g cujo contradomínio é A e cujo domínio é $\mathbb{N}$ ou um segmento inicial de números naturais. Se $A = \emptyset$, então $B = \emptyset$ também (caso contrário existiria algum $y \in B$ sem $x \in A$ tal que $g(x) = y$). Suponhamos então que temos a bijeção g. Então $f \circ g$ é uma bijeção com contradomínio B e domínio o mesmo que o de g (isto é, $\mathbb{N}$ ou um segmento inicial), logo B é enumerável.

Se B é enumerável, obtemos que A é enumerável repetindo o argumento com a bijeção $f^{-1} \colon B \to A$ em lugar de f. □

5.8 Conjuntos de Tamanhos Diferentes e o Teorema de Cantor

Apresentamos uma formulação precisa da ideia de que dois conjuntos têm o mesmo tamanho. Também podemos precisar a ideia de que um conjunto é menor que outro. A nossa definição de "é menor do que (ou equinumeroso)" exigirá, em vez de uma bijeção entre os conjuntos, uma injeção do primeiro para o segundo. Se tal função existir, o tamanho do primeiro conjunto é menor ou igual ao do segundo. Intuitivamente, uma injeção de um conjunto em outro garante que a imagem da função tem pelo menos tantos membros quanto o domínio, pois não há dois membros do domínio que vão parar ao mesmo membro da imagem.

Definição 5.14. *A não é maior do que B*, escrevendo-se $A \preceq B$, sse existir uma injeção $f : A \to B$.

É claro que isto é uma relação reflexiva e transitiva, mas não simétrica (fica como exercício). Podemos ainda introduzir uma noção que diz que um conjunto é (estritamente) menor que outro.

Definição 5.15. *A é menor do que B*, escrevendo-se $A \prec B$, sse existir uma injeção $f : A \to B$ mas não existir bijeção $g : A \to B$, isto é, $A \preceq B$ e $A \not\approx B$.

É claro que esta relação é irreflexiva e transitiva. (Também fica como exercício.) Com esta notação, um conjunto A é enumerável sse $A \preceq \mathbb{N}$, e A é não enumerável sse $\mathbb{N} \prec A$. Isto permite reformular o Teorema 5.10 como a observação de que $\mathbb{N} \prec \wp(\mathbb{N})$. De fato, Cantor (1892) provou que este último ponto é *completamente geral*:

Teorema 5.16 (Cantor). *$A \prec \wp(A)$, para qualquer conjunto A.*

Prova. A aplicação $f(x) = \{x\}$ é uma injeção $f : A \to \wp(A)$, pois se $x \neq y$, então também $\{x\} \neq \{y\}$ por extensionalidade, logo $f(x) \neq f(y)$. Temos pois $A \preceq \wp(A)$.

Resta mostrar que $A \not\approx \wp(A)$. Por *redução ao absurdo*, suponhamos $A \approx \wp(A)$, isto é, que existe alguma bijeção $g : A \to \wp(A)$. Consideremos:

$$D = \{x \in A : x \notin g(x)\}$$

Note-se que $D \subseteq A$, logo $D \in \wp(A)$. Como g é uma bijeção, existe algum $y \in A$ tal que $g(y) = D$. Mas então:

$$y \in g(y) \text{ sse } y \in D \text{ sse } y \notin g(y).$$

Isto é uma contradição; logo $A \not\approx \wp(A)$. □

É instrutivo comparar a prova do Teorema 5.16 com a do Teorema 5.10. Ali mostramos que, para qualquer lista N_0, N_1, N_2, ..., de subconjuntos de $\mathbb{N}$, podemos construir um conjunto D de números garantidamente fora da lista. Estava garantido que não estava na lista porque $n \in N_n$ sse $n \notin D$, para todo $n \in \mathbb{N}$. Seguimos a mesma ideia aqui, exceto que os índices n são agora membros de A em vez de $\mathbb{N}$. O conjunto D está definido de modo a diferir de $g(x)$ para cada $x \in A$, porque $x \in g(x)$ sse $x \notin D$.

A prova vale também a pena comparar com a do Paradoxo de Russell, Teorema 2.29. De fato, o Teorema de Cantor foi a inspiração para o paradoxo de Russell.

5.9 A Noção de Tamanho e Schröder-Bernstein

Eis um pensamento intuitivo: se A não é maior do que B e B não é maior do que A, então A e B são equinumerosos. Sejamos honestos: se isto estivesse *errado*, dificilmente poderíamos justificar que a noção definida de equinumerosidade tem a ver com comparações de "tamanhos" entre conjuntos. Felizmente, o pensamento intuitivo está certo. Isto justifica-se pelo Teorema de Schröder-Bernstein.

Teorema 5.17 (Schröder-Bernstein). *Se $A \preceq B$ e $B \preceq A$, então $A \approx B$.*

Em outras palavras, se existir uma injeção de A para B e uma injeção de B para A, então existe uma bijeção de A para B.

Este resultado é, contudo, realmente bastante *difícil* de provar. De fato, embora Cantor o tivesse enunciado, outros provaram-no.[2] Por agora, pode (e deve) aceitá-lo como verdadeiro.

Felizmente, Schröder-Bernstein é *correto* e legitima pensarmos nas relações que definimos, isto é, $A \approx B$ e $A \preceq B$, como tendo a

[2] Para mais sobre a história, ver por exemplo Potter (2004, pp. 165–6).

ver com "tamanho". Além disso, Schröder-Bernstein é muito *útil*. Pode ser difícil imaginar uma bijeção entre dois conjuntos equinumerosos. O Teorema de Schröder-Bernstein permite decompor a comparação em casos em que só precisamos de uma injeção do primeiro para o segundo e vice-versa.

5.10 Cantor na Reta e no Plano

Algumas circunstâncias em torno da prova de Schröder-Bernstein relacionam-se com a história que discutimos em Seção 1.3. Recordemos que, em 1877, Cantor provou que há tantos pontos num quadrado quanto num dos seus lados. Aqui apresentaremos a sua prova (primeira tentativa).

Seja L o intervalo unitário, isto é, o conjunto de pontos $[0,1]$. Seja S o quadrado unitário, isto é, o conjunto de pontos $\mathrm{L} \times \mathrm{L}$. Nestes termos, Cantor provou que $\mathrm{L} \approx \mathrm{S}$. Escreveu uma nota a Dedekind, que continha essencialmente o seguinte argumento.

Teorema 5.18. *$L \approx S$*

Prova: primeira parte.. Fixemos $a, b \in \mathrm{L}$. Escrevamos a e b em notação binária, de modo que temos sequências infinitas de 0s e 1s, a_1, a_2, ..., e b_1, b_2, ..., tais que:

$$a = 0.a_1a_2a_3a_4\ldots$$
$$b = 0.b_1b_2b_3b_4\ldots$$

Consideremos a função $f\colon \mathrm{S} \to \mathrm{L}$ dada por

$$f(a,b) = 0.a_1b_1a_2b_2a_3b_3a_4b_4\ldots$$

A função f é uma injeção, pois se $f(a,b) = f(c,d)$, então $a_n = c_n$ e $b_n = d_n$ para todo $n \in \mathbb{N}$, logo $a = c$ e $b = d$. □

Infelizmente, como Dedekind apontou a Cantor, isto não responde à questão original. Considere-se $0.\dot{1}\dot{0} = 0.1010101010\ldots$.

Para que $f(a,b) = 0.\dot{1}\dot{0}$, seria preciso:

$$a = 0.\dot{1}\dot{1} = 0.111111\ldots$$
$$b = 0$$

Mas $a = 0.\dot{1}\dot{1} = 1$. Assim, quando dizemos "escrever a e b em notação binária", temos de escolher *qual* notação usar; e, como f deve ser uma *função*, só podemos usar *uma* das duas notações possíveis. Mas se, por exemplo, usarmos a notação simples e escrevermos a como "$1.000\ldots$", então não existe par $\langle a,b\rangle$ tal que $f(a,b) = 0.\dot{1}\dot{0}$.

Em resumo: Dedekind observou que, dada a possibilidade de certas expansões binárias periódicas, a função f de Cantor é uma injeção mas *não* é uma sobrejeção. Cantor tinha mostrado apenas que $\mathrm{S} \preceq \mathrm{L}$ e *não* que $\mathrm{S} \approx \mathrm{L}$.

Cantor respondeu a Dedekind quase de imediato, sugerindo essencialmente que a prova poderia completar-se assim:

Prova: completada.. Mostramos que $\mathrm{S} \preceq \mathrm{L}$. Mas é óbvio que existe uma injeção de L para S: basta "deitar" a reta sobre um lado do quadrado. Logo $\mathrm{L} \preceq \mathrm{S}$ e $\mathrm{S} \preceq \mathrm{L}$. Pelo teorema de Schröder–Bernstein (Teorema 5.17), $\mathrm{L} \approx \mathrm{S}$. □

Claro que Cantor não podia justificar a última linha nestes termos, pois o teorema de Schröder-Bernstein ainda não estava provado. De fato, embora Cantor viesse a formulá-lo como conjectura geral, só viria a ser provado de modo satisfatório em 1897. (E assim, mais tarde em 1877, Cantor deu outra prova de Teorema 5.18, que não passava por Schröder–Bernstein.)

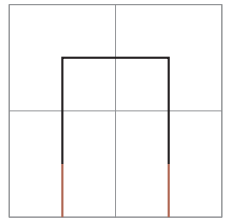
5.11 Apêndice: Curvas Preenchedoras de Espaço de Hilbert

No capítulo Seção 1.3, mencionamos que a prova de Cantor de que uma reta e um quadrado têm exatamente o mesmo número de pontos (Teorema 5.18) levou Peano a perguntar se poderia

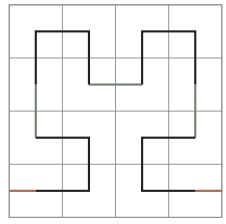
existir uma *curva* que preenchesse o espaço. Obteve resposta afirmativa em 1890. Nesta secção explicamos (de modo informal) como construir a curva preenchedora de espaço de Hilbert (com um pequeno ajuste).[3]

Temos de definir uma função, h, como limite de uma sequência de funções h_1, h_2, h_3, ... Primeiro descrevemos a construção. Depois mostramos que preenche o espaço. Depois mostramos que é uma curva.

Tomaremos a imagem de h como sendo o quadrado unitário, S. Eis a nossa primeira aproximação a h, isto é, h_1:



Para orientação, impusemos uma grelha 2×2 ao quadrado. Podemos pensar na curva como começando no quadrante inferior esquerdo, indo para o superior esquerdo, depois para o superior direito e finalmente para o inferior direito. Eis a segunda etapa da construção, isto é, h_2:

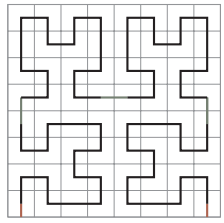


As cores diferentes ajudam a explicar como h_2 foi construída. Colocamos primeiro cópias reduzidas do troço não-vermelho de h_1

[3]Para uma explicação mais rigorosa, ver Rose (2010). O ajuste consiste em incluir as partes vermelhas das curvas abaixo. Isto torna ligeiramente mais fácil verificar que a curva é contínua.

no inferior esquerdo, superior esquerdo, superior direito e inferior direito do quadrado (traçadas a preto). Depois ligamos estas quatro figuras (com linhas verde). Por fim, ligamos a figura ao contorno do quadrado (com linhas vermelho).

Passamos a h_3. Tal como h_2 foi feita a partir de quatro cópias ligadas e reduzidas do troço não vermelho de h_1, h_3 compõe-se de quatro cópias reduzidas do troço não vermelho de h_2 (a preto), que depois se unem (com linhas verde) e se ligam ao contorno do quadrado (com linhas vermelho).



Agora vê-se o padrão geral para definir h_{n+1} a partir de h_n. Por fim definimos a curva *h em si* considerando o limite ponto a ponto destas funções sucessivas h_1, h_2, ... Isto é, para cada $x \in$ S:

$$h(x) = \lim_{n\to\infty} h_n(x)$$

Mostramos agora que esta curva preenche o espaço. Ao traçar h_n, impomos uma grelha $2^n \times 2^n$ sobre S. Pelo teorema de Pitágoras, a diagonal de cada célula da grelha tem comprimento:

$$\sqrt{(1/2^n)^2 + (1/2^n)^2} = 2^{(\frac{1}{2}-n)}$$

e é evidente que h_n passa por todas as células. Logo cada ponto de S está *a uma distância no máximo* $2^{(\frac{1}{2}-n)}$ de algum ponto de h_n. Mas h é definida como o limite das funções h_1, h_2, h_3, ... Assim, a distância máxima de qualquer ponto a h é dada por:

$$\lim_{n\to\infty} 2^{(\frac{1}{2}-n)} = 0.$$

Ou seja: todo o ponto de S está à distância 0 de h. Em outras palavras, todo o ponto de S situa-se *sobre* a curva. Logo h preenche o espaço!

Resta mostrar que h é, de fato, uma *curva.* Para isso é preciso definir o conceito. A definição moderna baseia-se na de Jordan, de 1887 (isto é, poucos anos antes da primeira curva preenchedora de espaço):

Definição 5.19. Uma curva é uma aplicação contínua de L para $\mathbb{R}^2$.

Isto é bastante intuitivo: uma curva é, intuitivamente, uma aplicação "suave" que leva uma reta canônica para o plano $\mathbb{R}^2$. A nossa função h é de fato uma aplicação de L para $\mathbb{R}^2$. Logo só precisamos de mostrar que h é contínua. Definimos continuidade em Seção 1.2 com a notação ε/δ. Em linguagem corrente, queremos estabelecer o seguinte: *Se especificar um ponto p em S, juntamente com qualquer nível de precisão ε, podemos encontrar um intervalo aberto de L tal que, dado qualquer x nesse intervalo, $h(x)$ fica a uma distância inferior a ε de p.*

Assim: suponhamos que p e ε estão fixos. Isto corresponde, em efeito, a traçar um círculo com centro p e raio ε sobre S. (O círculo pode sair para fora de S, mas não importa.) Recordemos que, ao descrever h_n, traçamos uma grelha $2^n \times 2^n$ sobre S. É óbvio que, por pequeno que seja ε, existe algum n tal que alguma célula individual da grelha $2^n \times 2^n$ sobre S fica inteiramente dentro do círculo com centro p e raio ε.

Tomemos esse n, e seja I a maior parte aberta de L que h_n mapeia inteiramente para a célula relevante. (É claro que I existe, pois já notamos que h_n passa por todas as células da grelha $2^n \times 2^n$.) Basta agora mostrar que, sempre que $x \in I$, o ponto $h(x)$ pertence à mesma célula. E, para *isso*, basta mostrar que $h_m(x)$ pertence à mesma célula, para qualquer $m > n$. Mas isto é óbvio. Se considerarmos o que acontece com h_m para $m > n$, vemos que exatamente a "mesma parte" do intervalo unitário é

mapeada para a mesma célula; apenas a descrevemos de modo cada vez mais esticado e sinuoso.

Problemas

Problema 5.1. Mostre que um conjunto A é enumerável se e só se $A = \emptyset$ ou existe uma sobrejeção $f : \mathbb{N} \to A$. Mostre que A é enumerável se e só se existe uma injeção $g : A \to \mathbb{N}$.

Problema 5.2. Defina uma enumeração dos quadrados perfeitos 1, 4, 9, 16, ...

Problema 5.3. Mostre que se A e B são enumeráveis, então $A \cup B$ também é.

Problema 5.4. Mostre por indução em n que se A_1, A_2, ..., A_n são todos enumeráveis, então $A_1 \cup \cdots \cup A_n$ também é.

Problema 5.5. Mostre que $(\mathbb{Z}^+)^n$ é enumerável, para todo $n \in \mathbb{N}$.

Problema 5.6. Mostre que $(\mathbb{Z}^+)^*$ é enumerável. Pode assumir PROBLEMA 5.5.

Problema 5.7. Dê uma enumeração do conjunto de todos os números racionais não negativos.

Problema 5.8. Mostre que $\mathbb{Q}$ é enumerável. Recordemos que qualquer número racional se pode escrever como uma fração z/m com $z \in \mathbb{Z}$, $m \in \mathbb{N}^+$.

Problema 5.9. Defina uma enumeração de $\mathbb{B}^*$.

Problema 5.10. Recordando o curso introdutório de lógica, cada tabela de verdade possível exprime uma função de verdade. Em outras palavras, as funções de verdade são todas as funções de $\mathbb{B}^k \to \mathbb{B}$ para algum k. Prove que o conjunto de todas as funções de verdade é enumerável.

Problema 5.11. Mostre que o conjunto de todos os subconjuntos finitos de um conjunto infinito enumerável arbitrário é enumerável.

Problema 5.12. Um subconjunto de $\mathbb{N}$ diz-se *confinito* sse for o complemento de um conjunto finito; isto é, $A \subseteq \mathbb{N}$ é confinito sse $\mathbb{N} \setminus A$ for finito. Seja I o conjunto cujos membros são exatamente os subconjuntos finitos e confinitos de $\mathbb{N}$. Mostre que I é enumerável.

Problema 5.13. Mostre que a união enumerável de conjuntos enumeráveis é enumerável. Isto é, sempre que A_1, A_2, ... são conjuntos e cada A_i é enumerável, então a união $\bigcup_{i=1}^{\infty} A_i$ de todos eles também é enumerável. [Nota: isto é difícil!]

Problema 5.14. Seja $f : A \times B \to \mathbb{N}$ uma função de emparelhamento arbitrária. Mostre que a inversa de f é uma enumeração de $A \times B$.

Problema 5.15. Especifique uma função que codifique $\mathbb{N}^3$.

Problema 5.16. Mostre que o conjunto de todas as funções $f : \mathbb{N} \to \mathbb{N}$ é não enumerável por um argumento diagonal explícito. Isto é, mostre que, se f_1, f_2, ... é uma lista de funções e cada $f_i : \mathbb{N} \to \mathbb{N}$, então existe alguma $g : \mathbb{N} \to \mathbb{N}$ que não está nesta lista.

Problema 5.17. Mostre que, se existir uma função injetiva $g : B \to A$, e B for não enumerável, então A também o é. Faça-o mostrando como pode usar g para transformar uma enumeração de A numa de B.

Problema 5.18. Mostre que o conjunto X de todas as funções $f : \mathbb{N} \to \mathbb{N}$ é não enumerável por um argumento de redução (Sugestão: dê uma função sobrejetora de X para $\mathbb{B}^\omega$.)

Problema 5.19. Mostre que o conjunto de todos os *conjuntos de* pares de números naturais, isto é, $\wp(\mathbb{N} \times \mathbb{N})$, é não enumerável por um argumento de redução.

Problema 5.20. Mostre que $\mathbb{N}^\omega$, o conjunto das sequências infinitas de números naturais, é não enumerável por um argumento de redução.

Problema 5.21. Seja S o conjunto de todas as sobrejeções de $\mathbb{N}$ para o conjunto $\{0,1\}$, isto é, S consiste em todas as sobrejeções $f : \mathbb{N} \to \mathbb{B}$. Mostre que S é não enumerável.

Problema 5.22. Mostre que o conjunto $\mathbb{R}$ de todos os números reais é não enumerável.

Problema 5.23. Mostre que, se $A \approx C$ e $B \approx D$, e $A \cap B = C \cap D = \emptyset$, então $A \cup B \approx C \cup D$.

Problema 5.24. Mostre que, se A é infinito e enumerável, então $A \approx \mathbb{N}$.

Problema 5.25. Mostre que não pode existir uma injeção $g\colon \wp(A) \to A$, para qualquer conjunto A. Sugestão: suponha $g\colon \wp(A) \to A$ injetiva. Considere $D = \{g(B) : B \subseteq A \text{ e } g(B) \notin B\}$. Seja $x = g(D)$. Use o fato de g ser injetiva para obter uma contradição.

CAPÍTULO 6

Aritmetização

No Capítulo 1, consideramos parte do enquadramento histórico, nomeadamente *por que* temos sequer teoria dos conjuntos. Capítulos 2 a 5 percorreram depois alguns princípios da teoria ingênua dos conjuntos. Compreendemos assim, de modo ingênuo, como construir relações e funções, comparar os tamanhos dos conjuntos e *coisas desse gênero*.

Com isto assimilado, podemos abordar algumas das primeiras conquistas da teoria dos conjuntos na redução (em certo sentido) de grandes parcelas da matemática à teoria dos conjuntos e à aritmética. Esse é o objetivo deste capítulo.

6.1 De $\mathbb{N}$ para $\mathbb{Z}$

Eis duas constatações básicas (no sentido de "ideias que retemos"):

1. Todo inteiro pode escrever-se na forma $n-m$, com $n, m \in \mathbb{N}$.
2. A informação codificada numa expressão $n-m$ pode igualmente codificar-se por um par ordenado $\langle n, m \rangle$.

Já sabemos que os pares ordenados de números naturais são os membros de $\mathbb{N}^2$; estamos a assumir, também, que compreendemos $\mathbb{N}$. Eis uma sugestão ingênua, fundada nestas duas constatações: *tratar os inteiros como pares ordenados de números naturais.*

De fato, esta sugestão é demasiado ingênua. Obviamente, queremos que seja verdade que $0-2 = 4-6$. Mas, evidentemente, $\langle 0,2\rangle \neq \langle 4,6\rangle$. Logo, não podemos simplesmente dizer que $\mathbb{N}^2$ é o conjunto dos inteiros.

Generalizando a partir do problema precedente, o que queremos é o seguinte:

$$a - b = c - d \text{ sse } a + d = c + b$$

(Deve ser óbvio que assim é que os inteiros *devem* comportar-se: basta somar b e d a ambos os lados.) A forma fácil de garantir este comportamento é definir uma relação de equivalência entre pares ordenados, $\sim$, do seguinte modo:

$$\langle a,b\rangle \sim \langle c,d\rangle \text{ sse } a + d = c + b$$

Resta-nos mostrar que isto é uma relação de equivalência.

Proposição 6.1. *$\sim$ é uma relação de equivalência.*

Prova. Há que mostrar que $\sim$ é reflexiva, simétrica e transitiva.

Reflexividade: Evidentemente $\langle a,b\rangle \sim \langle a,b\rangle$, pois $a+b = b+a$.

Simetria: Suponhamos $\langle a,b\rangle \sim \langle c,d\rangle$, logo $a+d = c+b$. Então $c + b = a + d$, pelo que $\langle c,d\rangle \sim \langle a,b\rangle$.

Transitividade: Suponhamos $\langle a,b\rangle \sim \langle c,d\rangle \sim \langle m,n\rangle$. Logo, $a + d = c + b$ e $c + n = m + d$. Assim, $a + d + c + n = c + b + m + d$, e portanto $a + n = m + b$. Logo, $\langle a,b\rangle \sim \langle m,n\rangle$. □

Podemos agora usar esta relação de equivalência para formar classes de equivalência:

Definição 6.2. Os inteiros são as classes de equivalência, sob $\sim$, dos pares ordenados de números naturais; isto é, $\mathbb{Z} = \mathbb{N}^2/_\sim$.

Agora, podem surgir várias reações *filosóficas* distintas face a esta definição estipulativa. Antes de as considerarmos, porém, vale a pena continuar com alguns pormenores técnicos.

Tendo dito o que são os inteiros, precisamos de definir funções e relações básicas sobre eles. Escrevamos $[m,n]_\sim$ para a classe de equivalência sob $\sim$ que tem $\langle m,n\rangle$ como um membro.[1] Isto é:

$$[m,n]_\sim = \{\langle a,b\rangle \in \mathbb{N}^2 : \langle a,b\rangle \sim \langle m,n\rangle\}$$

Apresentamos, pois, algumas definições:

$$\begin{aligned}
[a,b]_\sim + [c,d]_\sim &= [a+c, b+d]_\sim \\
[a,b]_\sim \times [c,d]_\sim &= [ac+bd, ad+bc]_\sim \\
[a,b]_\sim \leq [c,d]_\sim &\text{ sse } a+d \leq b+c
\end{aligned}$$

(Como é habitual, uso 'ab' no sentido de '$(a \times b)$', só para tornar os axiomas mais legíveis.) Há, ainda, que garantir que estas definições se comportam como *devem*. Explicitar o que isto significa, e verificá-lo, é trabalhoso; relegamos os pormenores a Seção 6.6. Em resumo, contudo: tudo funciona!

Resta um último ponto. Construímos os inteiros a partir dos números naturais. Isto significa que os naturais *não são, eles próprios, inteiros*. Voltaremos ao significado filosófico disto na Seção 6.5. Do ponto de vista puramente técnico, porém, precisamos de alguma forma de tratar números naturais *como* inteiros. A ideia é simples: para cada $n \in \mathbb{N}$, estipulamos $n_\mathbb{Z} = [n,0]_\sim$. Há que confirmar que esta definição é bem comportada, isto é, que para quaisquer $m, n \in \mathbb{N}$

$$\begin{aligned}
(m+n)_\mathbb{Z} &= m_\mathbb{Z} + n_\mathbb{Z} \\
(m \times n)_\mathbb{Z} &= m_\mathbb{Z} \times n_\mathbb{Z} \\
m \leq n &\leftrightarrow m_\mathbb{Z} \leq n_\mathbb{Z}
\end{aligned}$$

Mas isto é bastante direto. Por exemplo, para mostrar que a segunda condição se verifica, podemos apoiar-nos no comportamento dos naturais e raciocinar assim:

$$(m \times n)_\mathbb{Z} = [m \times n, 0]_\sim$$

[1] Nota: com a notação introduzida na Definição 3.11, escreveríamos $[\langle m,n\rangle]_\sim$ para o mesmo objeto; isso, porém, torna a leitura um pouco mais difícil.

$$
\begin{aligned}
&= [m \times n + 0 \times 0, m \times 0 + 0 \times n]_{\sim} \\
&= [m, 0]_{\sim} \times [n, 0]_{\sim} \\
&= m_{\mathbb{Z}} \times n_{\mathbb{Z}}
\end{aligned}
$$

Deixamos como exercício confirmar que as outras duas condições se verificam.

6.2 De $\mathbb{Z}$ para $\mathbb{Q}$

Acabamos de ver como construir os inteiros a partir dos naturais, usando alguma teoria ingênua dos conjuntos. Veremos agora como construir os racionais a partir dos inteiros, de modo muito semelhante. As constatações iniciais são:

1. Todo racional pode escrever-se na forma $^{i}/_{j}$, onde i e j são inteiros, mas j é diferente de zero.

2. A informação codificada numa expressão $^{i}/_{j}$ pode igualmente codificar-se num par ordenado $\langle i, j \rangle$.

A abordagem óbvia seria pensar nos racionais *como* pares ordenados de $\mathbb{Z} \times (\mathbb{Z} \setminus \{0_{\mathbb{Z}}\})$. Como antes, porém, isso seria um pouco demasiado ingênuo, pois queremos $^{3}/_{2} = {}^{6}/_{4}$, mas $\langle 3, 2 \rangle \neq \langle 6, 4 \rangle$. Mais geralmente, quereremos o seguinte:

$$^{a}/_{b} = {}^{c}/_{d} \text{ sse } a \times d = b \times c$$

Para obter isto, definimos uma *relação de equivalência* em $\mathbb{Z} \times (\mathbb{Z} \setminus \{0_{\mathbb{Z}}\})$ assim:

$$\langle a, b \rangle \backsim \langle c, d \rangle \text{ sse } a \times d = b \times c$$

Há que verificar que isto é uma relação de equivalência. É muito parecido com o caso de $\sim$, e deixamo-lo como exercício. Mas isto permite-nos dizer o seguinte:

Definição 6.3. Os racionais são as classes de equivalência, sob $\backsim$, dos pares de inteiros (cujo segundo elemento é não nulo). Isto é, $\mathbb{Q} = (\mathbb{Z} \times (\mathbb{Z} \setminus \{0_{\mathbb{Z}}\}))/_{\backsim}$.

Tal como para os inteiros, queremos também definir algumas operações básicas. Sendo $[i,j]_{\backsim}$ a classe de equivalência sob $\backsim$ com $\langle i,j\rangle$ como um membro, definimos:

$$[a,b]_{\backsim} + [c,d]_{\backsim} = [ad + bc, bd]_{\backsim}$$
$$[a,b]_{\backsim} \times [c,d]_{\backsim} = [ac, bd]_{\backsim}.$$

Para definir $r \leq s$ nestes racionais, usamos o fato de que $r \leq s$ sse $s - r$ não é negativo, isto é, $r - s$ pode escrever-se como i/j com i não negativo e j positivo:

$$[a,b]_{\backsim} \leq [c,d]_{\backsim} \text{ sse } [c,d]_{\backsim} - [a,b]_{\backsim} = [i_{\mathbb{Z}}, j_{\mathbb{Z}}]_{\backsim}$$

para algum $i \in \mathbb{N}$ e $0 \neq j \in \mathbb{N}$.

Depois há que verificar que estas definições se comportam como *devem*; relegamo-lo a Seção 6.6. Mas de fato comportam-se! Por fim, queremos alguma forma de tratar inteiros *como* racionais; logo, para cada $i \in \mathbb{Z}$, estipulamos $i_{\mathbb{Q}} = [i, 1_{\mathbb{Z}}]_{\backsim}$. De novo, verificamos na Seção 6.6 que tudo isto se comporta corretamente.

6.3 A Reta Real

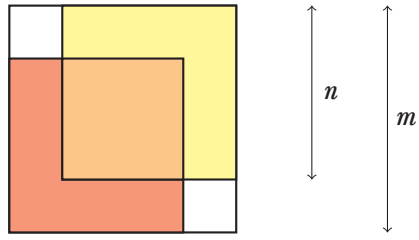
O passo seguinte é mostrar como construir os reais a partir dos racionais. Antes disso, há que perceber o que há de *distintivo* nos reais.

Os reais comportam-se de modo muito parecido com os racionais. (Tecnicamente, ambos são exemplos de *corpos ordenados*; para a definição, ver Definição 6.9.) Agora, se resolveu os exercícios de Capítulo 5, saberá que há estritamente mais reais do que racionais, isto é, que $\mathbb{Q} \prec \mathbb{R}$. Cantor foi o primeiro a provar isto. Há, contudo, cerca de dois milênios e meio que se conhecem números irracionais, isto é, reais que não são racionais. De fato:

Teorema 6.4. *$\sqrt{2}$ não é racional, isto é, $\sqrt{2} \notin \mathbb{Q}$*

Prova. Suponhamos, por *reductio*, que $\sqrt{2}$ é racional. Então $\sqrt{2} = m/n$ para alguns números naturais m e n. De fato, podemos escolher m e n de modo a que a fração não possa ser mais reduzida. Reorganizando, $m^2 = 2n^2$. A partir daqui, completamos a prova de duas formas:

Primeiro, geometricamente (seguindo Tennenbaum).[2] Consideremos estes quadrados:



□

Como $m^2 = 2n^2$, a região em que os dois quadrados de lado n se sobrepõem tem a mesma área que a região que nenhum dos dois quadrados cobre; isto é, a área do quadrado cor de laranja é igual à soma das áreas dos dois quadrados não sombreados. Logo, se o quadrado cor de laranja tem lado p e cada quadrado não sombreado tem lado q, então $p^2 = 2q^2$. Mas então $\sqrt{2} = p/q$, com $p < m$, $q < n$ e $p, q \in \mathbb{N}$. Isto contradiz o fato de m e n terem sido escolhidos tão pequenos quanto possível.

Segundo, formalmente. Como $m^2 = 2n^2$, segue-se que m é par. (É fácil mostrar que, se x é ímpar, então x^2 é ímpar.) Logo $m = 2r$, para algum $r \in \mathbb{N}$. Rearranjando, $2r^2 = n^2$, logo n também é par. Assim m e n são ambos pares, e portanto a fração m/n *pode* ser ainda reduzida. Contradição!

De passagem, esta prova diagramática permite relembrar o material de Seção 1.4. Tennenbaum (1927–2006) era um matemático inequivocamente moderno; mas a prova é inegavelmente bela, completamente rigorosa e apela à intuição geométrica!

[2] Esta prova é relatada por Conway (2006).

Em todo o caso: os reais são "mais extensos" do que os racionais. Nalgum sentido, há "lacunas" nos racionais, e os reais preenchem-nas. Weierstrass apercebeu-se de que isto descreve uma única propriedade dos números reais que os distingue dos racionais, nomeadamente a *propriedade de completude*: *Todo conjunto não vazio de números reais com majorante tem um supremo (menor majorante).*

É fácil ver que os racionais não têm a propriedade de completude. Por exemplo, considere-se o conjunto dos racionais menores que $\sqrt{2}$, isto é:

$$\{p \in \mathbb{Q} : p^2 < 2 \text{ ou } p < 0\}$$

Este conjunto tem um majorante nos racionais; os seus membros são todos menores que 3, por exemplo. Mas qual é o seu supremo? Queremos dizer "$\sqrt{2}$"; contudo, acabamos de ver que $\sqrt{2}$ *não* é racional. E não existe racional mínimo entre os que são estritamente maiores que $\sqrt{2}$. Logo, o conjunto tem majorante, mas não tem supremo. Assim, os racionais não satisfazem a propriedade de completude.

Em contraste, o contínuo "moralmente deveria" ter a propriedade de completude. Não queremos apenas que $\sqrt{2}$ seja um número real; queremos preencher todas as "lacunas" da reta racional. Queremos, ainda, que o próprio contínuo não tenha "lacunas". É exatamente isso que obteremos via completude.

6.4 De $\mathbb{Q}$ para $\mathbb{R}$

Em essência, a propriedade de completude mostra que qualquer ponto α da reta real divide essa reta em duas metades de modo perfeito: numa, α é o supremo; noutra, α é o ínfimo. Para *construir* os números reais a partir dos racionais, Dedekind sugeriu que pensássemos nos reais simplesmente como os *cortes* que particionam os racionais. Isto é, identificamos $\sqrt{2}$ com o *corte* que separa os racionais $< \sqrt{2}$ dos racionais $> \sqrt{2}$.

Vamos pôr isto em ordem. Se cortarmos os racionais em duas metades, podemos identificar de modo único a partição que fizemos considerando apenas a metade *inferior*. Sendo precisos, apresentamos a seguinte definição:

Definição 6.5 (Corte). Um *corte* α é qualquer segmento inicial próprio não vazio dos racionais sem máximo. Isto é, α é um corte sse:

1. *não vazio, próprio*: $\emptyset \neq \alpha \subsetneq \mathbb{Q}$
2. *inicial*: para todos $p, q \in \mathbb{Q}$: se $p < q \in \alpha$ então $p \in \alpha$
3. *sem máximo*: para todo $p \in \alpha$ existe $q \in \alpha$ tal que $p < q$

Então $\mathbb{R}$ é o conjunto dos cortes.

Podemos pois dizer que $\sqrt{2} = \{p \in \mathbb{Q} : p^2 < 2 \text{ ou } p < 0\}$. Claro que há que verificar que isto *é* de fato um corte, mas relegamo-lo a Seção 6.6.

Como antes, tendo definido certas entidades, precisamos a seguir de definir funções e relações básicas sobre elas. Começamos por uma fácil:

$$\alpha \leq \beta \text{ sse } \alpha \subseteq \beta$$

Esta definição de ordem permite *enunciar* o resultado central, a saber que o conjunto dos cortes tem a propriedade de completude. Por extenso, o enunciado tem esta forma. Se S é um conjunto não vazio de cortes com majorante, então S tem supremo. Com mais pormenor: existe um corte, λ, que é majorante de S, isto é, $(\forall \alpha \in S)\alpha \subseteq \lambda$, e λ é o menor tal corte, isto é, $(\forall \beta \in \mathbb{R})((\forall \alpha \in S)\alpha \subseteq \beta \rightarrow \lambda \subseteq \beta)$. Segue-se a prova do resultado:

Teorema 6.6. *O conjunto dos cortes tem a propriedade de completude.*

Prova. Seja S um conjunto não vazio de cortes com majorante. Seja $\lambda = \bigcup S$. Começamos por alegar que λ é um corte:

1. Como S tem majorante, pelo menos um corte pertence a S, logo $\emptyset \neq \lambda$. Como S é um conjunto de cortes, $\lambda \subseteq \mathbb{Q}$. Como S tem majorante, algum $p \in \mathbb{Q}$ está ausente de todo o corte $\alpha \in S$. Logo $p \notin \lambda$, e portanto $\lambda \subsetneq \mathbb{Q}$.

2. Suponhamos $p < q \in \lambda$. Então existe algum $\alpha \in S$ tal que $q \in \alpha$. Como α é um corte, $p \in \alpha$. Logo $p \in \lambda$.

3. Suponhamos $p \in \lambda$. Então existe algum $\alpha \in S$ tal que $p \in \alpha$. Como α é um corte, existe algum $q \in \alpha$ tal que $p < q$. Logo $q \in \lambda$.

Isto prova a alegação. Além disso, claramente $(\forall \alpha \in S)\alpha \subseteq \bigcup S = \lambda$, isto é, λ é majorante de S. Suponhamos agora que $\beta \in \mathbb{R}$ também é majorante, isto é, $(\forall \alpha \in S)\alpha \subseteq \beta$. Para qualquer $p \in \mathbb{Q}$, se $p \in \lambda$, então existe $\alpha \in S$ tal que $p \in \alpha$, logo $p \in \beta$. Generalizando, $\lambda \subseteq \beta$. Logo λ é o supremo de S. □

Temos pois um conjunto de entidades que satisfaz a propriedade de completude. Uma forma de o dizer é: não há "lacunas" nos nossos cortes. (Assim: tomar novos "cortes" de reais, em vez de racionais, não produziria objetos novos interessantes.)

Segue-se a definição de algumas operações sobre os reais. Começamos por embutir os racionais nos reais estipulando $p_{\mathbb{R}} = \{q \in \mathbb{Q} : q < p\}$ para cada $p \in \mathbb{Q}$. Definimos então:

$$\alpha + \beta = \{p + q : p \in \alpha \wedge q \in \beta\}$$
$$\alpha \times \beta = \{p \times q : 0 \leq p \in \alpha \wedge 0 \leq q \in \beta\} \cup 0^{\mathbb{R}} \quad \text{se } \alpha, \beta \geq 0_{\mathbb{R}}$$

Para tratar os outros casos da multiplicação, definamos primeiro:

$$-\alpha = \{p - q : p < 0 \wedge q \notin \alpha\}$$

e estipulemos:

$$\alpha \times \beta = \begin{cases} -\alpha \times -\beta & \text{se } \alpha < 0_{\mathbb{R}} \text{ e } \beta < 0_{\mathbb{R}} \\ -(-\alpha \times \beta) & \text{se } \alpha < 0_{\mathbb{R}} \text{ e } \beta > 0_{\mathbb{R}} \\ -(\alpha \times -\beta) & \text{se } \alpha > 0_{\mathbb{R}} \text{ e } \beta < 0_{\mathbb{R}} \end{cases}$$

Depois há que verificar que cada uma destas definições produz sempre um corte. E, por fim, há que fazer uma demonstração fácil (mas longa) de que os cortes, assim definidos, se comportam exatamente como devem. Mas relegamos tudo isto a Seção 6.6.

6.5 Algumas Reflexões Filosóficas

Quanto aos pormenores técnicos, por ora basta. Mas o que foi que alcançaram?

Bem, de forma pouco contestável, deram-nos uma matemática pura *belíssima*. Além disso, houve conquistas conceptuais profundas. Foi um insight profundo perceber que a propriedade de completude exprime a diferença crucial entre os reais e os racionais. Além disso, a construção explícita dos reais, como cortes de Dedekind, põe o objeto de estudo da análise em bases sólidas. Sabemos que a noção de *corpo ordenado completo* é coerente, pois os cortes formam precisamente um tal corpo.

Todavia, importa também levantar algumas reservas quanto a essas conquistas.

Primeiro, não é claro que pensar os reais em termos de cortes seja *mais* rigoroso do que pensá-los nas suas expansões decimais familiares (eventualmente infinitas). Esta última "construção" dos reais tem alguma semelhança com a construção via sequências de Cauchy; mas, de fato, era, essencialmente, conhecida dos matemáticos desde princípios do século XVII (ver Seção 6.7). O verdadeiro aumento de rigor veio da tomada de consciência de que os reais têm a propriedade de completude; a possibilidade de construir números reais como conjuntos particulares talvez não seja, por si só, tão interessante.

É ainda menos claro que a aritmetização (muito mais fácil) dos inteiros ou dos racionais aumente o rigor nessas áreas. Aqui, vale a pena uma observação simples. Tendo *construído* os inteiros como classes de equivalência de pares ordenados de naturais, e depois construído os racionais como classes de equivalência de

pares ordenados de inteiros, e depois construído os reais como conjuntos de racionais, *nos esquecemos de imediato* das construções. Em particular: ninguém quereria *invocar* estas construções no decurso de uma prova matemática (exceto, claro, numa prova de que as construções se comportam como deviam). É muito mais fácil falar de um real diretamente do que falar de algum conjunto de conjuntos de conjuntos de conjuntos de conjuntos de conjuntos de conjuntos de naturais.

O mais duvidoso, porém, é que estas definições nos digam o que os inteiros, racionais ou reais *são*, *do ponto de vista metafísico*. Isto é, é duvidoso que os reais (digamos) *sejam* certos conjuntos (de conjuntos de conjuntos...). A principal barreira a tal visão é que a construção poderia ter sido feita de muitas maneiras diferentes. No caso dos reais, há construções genuinamente diferentes e interessantes (ver Seção 6.7). Mas eis uma forma trivial de obter construções diferentes: como em Seção 3.2, poderíamos ter definido pares ordenados de modo ligeiramente diferente; se tivéssemos usado essa noção alternativa de par ordenado, as nossas construções teriam funcionado exatamente tão bem como funcionaram, mas teríamos acabado com objetos diferentes. Assim, há muitas construções rivais, em teoria dos conjuntos, dos inteiros, racionais e reais. Seria, ademais, arbitrário (e embaraçoso) afirmar que os inteiros (digamos) são *estes* conjuntos, e não *aqueles*. (Como na Seção 3.2, isto é um caso de um argumento tornado célebre por Benacerraf 1965.)

Um outro ponto merece ser levantado: há algo de bastante *estranho* nas nossas construções. Começamos pelos números naturais. Depois, construímos os inteiros e construímos "o 0 dos inteiros", isto é, $[0,0]_\sim$. Mas $0 \neq [0,0]_\sim$. De fato, dadas as nossas construções, *nenhum* natural é um inteiro. Isto parece fortemente contraintuitivo. Na Seção 2.3, afirmamos, sem grande argumento, que $\mathbb{N} \subseteq \mathbb{Q}$. Agora, se as construções nos dissessem exatamente *o que* são os números, esta afirmação seria trivialmente falsa.

Recuando um pouco, onde chegamos? Trabalhando numa teoria ingênua dos conjuntos e tomando os naturais como dados,

podemos *tratar* inteiros, racionais e reais como certos conjuntos. Nesse sentido, podemos *embutir* as teorias destas entidades numa teoria dos conjuntos. Mas a importância filosófica deste embutimento não é assim tão direta.

Claro que isto não é a última palavra! O ponto é apenas este: mostrar que a aritmetização dos reais *é de* profunda significação filosófica exigiria argumento filosófico adicional.

6.6 Anéis e Corpos Ordenados

Ao longo deste capítulo, afirmamos que certas definições se comportam "como devem". Neste apêndice técnico, explicitaremos o que queremos dizer com isso e (esboçaremos como) mostrar que as definições se comportam "corretamente".

Na Seção 6.1, definimos adição e multiplicação em $\mathbb{Z}$. Queremos mostrar que, assim definidas, conferem a $\mathbb{Z}$ a estrutura que "quereríamos" que tivesse. Em particular, a estrutura em causa é a de um anel comutativo.

Definição 6.7. Um *anel comutativo* é um conjunto S, equipado com elementos distintos 0 e 1 e operações + e $\times$, que satisfazem estas oito condições:

Associatividade	$a + (b + c) = (a + b) + c$
	$(a \times b) \times c = a \times (b \times c)$
Comutatividade	$a + b = b + a$
	$a \times b = b \times a$
Neutros	$a + 0 = a$
	$a \times 1 = a$
Simétrico aditivo	$(\exists b \in S) 0 = a + b$
Distributividade	$a \times (b + c) = (a \times b) + (a \times c)$

Implicitamente, todas estas fórmulas estão quantificadas univer-

salmente sobre S. Note-se que os elementos 0 e 1 aqui não têm de ser os números naturais com o mesmo nome.

Logo, para verificar que os inteiros formam um anel comutativo, basta verificar estas oito condições. Nenhuma é *difícil* de estabelecer, mas o trabalho é um pouco laborioso. Por exemplo, eis como provar *associatividade*, no caso da adição:

Prova. Fixemos $i, j, k \in \mathbb{Z}$. Então existem $a_1, b_1, a_2, b_2, a_3, b_3 \in \mathbb{N}$ tais que $i = [a_1, b_1]$ e $j = [a_2, b_2]$ e $k = [a_3, b_3]$. (Por legibilidade, escrevemos "$[x, y]$" em vez de "$[x, y]_\sim$"; faremos o mesmo ao longo desta secção.) Então:

$$\begin{aligned} i + (j + k) &= [a_1, b_1] + ([a_2, b_2] + [a_3, b_3]) \\ &= [a_1, b_1] + [a_2 + a_3, b_2 + b_3] \\ &= [a_1 + (a_2 + a_3), b_1 + (b_2 + b_3)] \\ &= [(a_1 + a_2) + a_3, (b_1 + b_2) + b_3] \\ &= [a_1 + a_2, b_1 + b_2] + [a_3, b_3] \\ &= ([a_1, b_1] + [a_2, b_2]) + [a_3, b_3] \\ &= (i + j) + k \end{aligned}$$

apoiando-nos livremente no comportamento da adição em $\mathbb{N}$. □

Do mesmo modo, eis uma prova do *simétrico aditivo*:

Prova. Fixemos $i \in \mathbb{Z}$, de modo que $i = [a, b]$ para alguns $a, b \in \mathbb{N}$. Seja $j = [b, a] \in \mathbb{Z}$. Apoiando-nos no comportamento dos naturais, $(a + b) + 0 = 0 + (a + b)$, logo $\langle a + b, b + a \rangle \sim_{\mathbb{Z}} \langle 0, 0 \rangle$ por definição, e portanto $[a + b, b + a] = [0, 0] = 0_{\mathbb{Z}}$. Assim $i + j = [a, b] + [b, a] = [a + b, b + a] = [0, 0] = 0_{\mathbb{Z}}$. □

E eis uma prova da *distributividade*:

Prova. Como acima, fixemos $i = [a_1, b_1]$ e $j = [a_2, b_2]$ e $k = [a_3, b_3]$. Então:

$$i \times (j + k) = [a_1, b_1] \times ([a_2, b_2] + [a_3, b_3])$$

$$
\begin{aligned}
&= [a_1, b_1] \times [a_2 + a_3, b_2 + b_3] \\
&= [a_1(a_2 + a_3) + b_1(b_2 + b_3), a_1(b_2 + b_3) + b_1(a_2 + a_3)] \\
&= [a_1a_2 + a_1a_3 + b_1b_2 + b_1b_3, a_1b_2 + a_1b_3 + a_2b_1 + a_3b_1] \\
&= [a_1a_2 + b_1b_2, a_1b_2 + a_2b_1] + [a_1a_3 + b_1b_3, a_1b_3 + a_3b_1] \\
&= ([a_1, b_1] \times [a_2, b_2]) + ([a_1, b_1] \times [a_3, b_3]) \\
&= (i \times j) + (i \times k) \qquad \square
\end{aligned}
$$

Deixamos como exercício provar as cinco condições restantes. Tendo feito isso, mostramos que $\mathbb{Z}$ constitui um anel comutativo, isto é, que a adição e a multiplicação (assim definidas) se comportam como devem.

Mas a tarefa não termina. Para além de definir adição e multiplicação em $\mathbb{Z}$, definimos uma relação de ordem, $\leq$, e há que verificar que esta se comporta como deve. Com mais pormenor, há que mostrar que $\mathbb{Z}$ constitui um *anel ordenado*.[3]

Definição 6.8. Um *anel ordenado* é um anel comutativo que está também equipado com uma relação de ordem total, $\leq$, tal que:

$$
\begin{aligned}
a \leq b &\rightarrow a + c \leq b + c \\
(a \leq b \wedge 0 \leq c) &\rightarrow a \times c \leq b \times c
\end{aligned}
$$

Como antes, é laborioso mas rotineiro mostrar que $\mathbb{Z}$, tal como construído, é um anel ordenado. Deixamo-lo ao leitor.

Isto trata dos inteiros. Mas agora precisamos de mostrar coisas muito parecidas para os racionais. Em particular, há que mostrar que os racionais formam um *corpo ordenado*, com as nossas definições de $+$, $\times$ e $\leq$:

[3]Como na Definição 3.16, uma ordem total é uma relação reflexiva, transitiva, antissimétrica e conexa. No contexto de relações de ordem, a conexividade é por vezes chamada *tricotomia*, pois, para quaisquer a e b, temos $a \leq b \vee a = b \vee a \geq b$.

Definição 6.9. Um *corpo ordenado* é um anel ordenado que também satisfaz:

$$\textit{Inverso multiplicativo} \qquad (\forall a \in S \setminus \{0\})(\exists b \in S)\, a \times b = 1$$

Uma vez mostrado que $\mathbb{Z}$ constitui um anel ordenado, é fácil mas laborioso mostrar que $\mathbb{Q}$ constitui um corpo ordenado.

Tendo tratado dos inteiros e dos racionais, só restam os reais. Em particular, há que mostrar que $\mathbb{R}$ constitui um corpo ordenado *completo*, isto é, um corpo ordenado com a propriedade de completude. Agora, Teorema 6.6 estabeleceu que $\mathbb{R}$ tem a propriedade de completude. Contudo, cumpre ainda percorrer a tediosa verificação de que $\mathbb{R}$ é um corpo ordenado.

Antes de nos lançarmos a *esse* exercício laborioso, há que confirmar algumas coisas mais "imediatas". Por exemplo, precisamos de garantia de que $\alpha + \beta$, tal como definida, é de fato um *corte*, para quaisquer cortes α e β. Segue-se uma prova desse fato:

Prova. Como α e β são ambos cortes, $\alpha + \beta = \{p + q : p \in \alpha \wedge q \in \beta\}$ é um subconjunto próprio não vazio de $\mathbb{Q}$. Suponhamos agora $x < p+q$ para alguns $p \in \alpha$ e $q \in \beta$. Então $x - p < q$, logo $x - p \in \beta$, e $x = p + (x - p) \in \alpha + \beta$. Logo $\alpha + \beta$ é um segmento inicial de $\mathbb{Q}$. Finalmente, para qualquer $p + q \in \alpha + \beta$, como α e β são cortes, existem $p_1 \in \alpha$ e $q_1 \in \beta$ tais que $p < p_1$ e $q < q_1$; logo $p + q < p_1 + q_1 \in \alpha + \beta$; logo $\alpha + \beta$ não tem máximo. □

Esforços semelhantes permitem verificar que $\alpha - \beta$ e $\alpha \times \beta$ e $\alpha \div \beta$ são cortes (no último caso, ignorando o caso em que β é o corte zero). De novo, deixamo-lo ao leitor.

Mas eis um pequeno fio solto a atar. Em Seção 6.4, sugerimos que podemos tomar $\sqrt{2} = \{p \in \mathbb{Q} : p < 0 \text{ ou } p^2 < 2\}$. Mas há que mostrar que este conjunto é de fato um *corte*. Segue-se uma prova desse fato:

Prova. Claramente isto é um segmento inicial próprio não vazio dos racionais; basta pois mostrar que não tem máximo. Em

particular, basta mostrar que, sendo p um racional positivo com $p^2 < 2$ e $q = \frac{2p+2}{p+2}$, se tem $p < q$ e $q^2 < 2$. Para ver que $p < q$, basta notar:

$$\begin{aligned} p^2 &< 2 \\ p^2 + 2p &< 2 + 2p \\ p(p+2) &< 2 + 2p \\ p &< \tfrac{2+2p}{p+2} = q \end{aligned}$$

Para ver que $q^2 < 2$, basta notar:

$$\begin{aligned} p^2 &< 2 \\ 2p^2 + 4p + 2 &< p^2 + 4p + 4 \\ 4p^2 + 8p + 4 &< 2(p^2 + 4p + 4) \\ (2p+2)^2 &< 2(p+2)^2 \\ \tfrac{(2p+2)^2}{(p+2)^2} &< 2 \\ q^2 &< 2 \end{aligned}$$

□

6.7 Apêndice: os Reais como Sequências de Cauchy

Na Seção 6.4, construímos os reais como cortes de Dedekind. Nesta secção, explicamos uma construção alternativa. Apoia-se na definição, devida a Cauchy, do que hoje chamamos *sequência de Cauchy*; mas o uso desta definição para *construir* os reais deve-se a outros autores do século XIX, nomeadamente Weierstrass, Heine, Méray e Cantor. (Para uma boa história, ver O'Connor and Robertson 2005.)

Antes de chegarmos ao século XIX, vale a pena considerar Simon Stevin (1548–1620). Em resumo, Stevin apercebeu-se de que podemos pensar cada real em termos da sua expansão decimal. Assim, mesmo um número irracional, como $\sqrt{2}$, tem uma

expansão decimal agradável, que começa:

$$1.41421356237\ldots$$

É muito fácil modelar expansões decimais na teoria dos conjuntos: basta considerá-las como funções $d\colon \mathbb{N} \to \mathbb{N}$, onde $d(n)$ é o n-ésimo algarismo decimal que nos interessa. Depois precisamos de um pequeno ajuste, para tratar a parte do número real que vem antes da vírgula decimal (aqui, apenas 1). Precisaremos também de outro ajuste (uma relação de equivalência) para garantir que, por exemplo, $0.999\ldots = 1$. Mas não é difícil apresentar uma construção perfeitamente rigorosa dos números reais, à maneira de Stevin, dentro da teoria dos conjuntos.

Stevin não é o nosso foco. (Para mais sobre Stevin, ver Katz and Katz 2012.) Mas eis um pensamento estreitamente relacionado. Em vez de tratar diretamente a expansão decimal de $\sqrt{2}$, podemos antes considerar uma *sequência* de aproximações racionais cada vez mais precisas a $\sqrt{2}$, olhando para as expansões cada vez mais precisas:

$$1, 1.4, 1.414, 1.4142, 1.41421, \ldots$$

A ideia de que os reais podem ser considerados via "aproximações cada vez melhores" nos dá a base para outra sequência de insights, semelhante às constatações que retivemos ao construir $\mathbb{Q}$ a partir de $\mathbb{Z}$, ou $\mathbb{Z}$ a partir de $\mathbb{N}$. Os insights básicos são estes:

1. Todo o real pode escrever-se como uma expansão decimal (eventualmente infinita).

2. A informação codificada por uma expansão decimal (eventualmente infinita) pode igualmente codificar-se por uma sequência de números racionais.

3. Uma sequência de números racionais pode pensar-se como uma função de $\mathbb{N}$ para $\mathbb{Q}$; basta definir $f(n)$ como o n-ésimo racional na sequência.

Claro que nem *qualquer* função de $\mathbb{N}$ para $\mathbb{Q}$ nos dará um número real. Por exemplo, considere-se esta função:

$$f(n) = \begin{cases} 1 & \text{se } n \text{ é ímpar} \\ 0 & \text{se } n \text{ é par} \end{cases}$$

Essencialmente, a preocupação é que a sequência $0,1,0,1,0,1,0,\ldots$ não parece "convergir" para nenhum real. Logo: para garantir que consideramos sequências que de fato convergem para algum real, precisamos de restringir a atenção a sequências que tenham algum *limite*.

Já encontramos a ideia de limite em Seção 1.2. Mas não podemos usar *exatamente* a mesma definição que ali. A expressão "$(\forall \varepsilon > 0)$" envolvia tacitamente quantificação sobre os números reais; e estávamos a considerar limites de funções nos reais; logo, invocar essa definição seria valer-nos dos números reais — e eram precisamente eles que estávamos a tentar *construir*. Felizmente, podemos trabalhar com uma ideia estreitamente relacionada com a de limite.

Definição 6.10. Uma função $f : \mathbb{N} \to \mathbb{Q}$ é uma *sequência de Cauchy* sse, para qualquer $\varepsilon \in \mathbb{Q}$ positivo, se tem $(\exists \ell \in \mathbb{N})(\forall m, n > \ell)|f(m) - f(n)| < \varepsilon$.

A ideia geral de limite é a mesma de antes: se se quer um certo nível de precisão (medido por ε), há uma "região" onde olhar (qualquer entrada maior que ℓ). E é fácil ver que a nossa sequência 1, 1.4, 1.414, 1.4142, 1.41421... tem limite: se se quer aproximar $\sqrt{2}$ com erro inferior a $1/10^n$, basta olhar para qualquer entrada depois da n-ésima.

O pensamento óbvio seria então dizer que um número real *é* simplesmente qualquer sequência de Cauchy. Mas, como nas construções de $\mathbb{Z}$ e $\mathbb{Q}$, isso seria demasiado ingênuo: para um dado número real, várias sequências de Cauchy diferentes "indicam" esse número. Uma forma simples de ver isto é a seguinte. Dada uma sequência de Cauchy f, defina-se g como exatamente

a mesma função que f, exceto que $g(0) \neq f(0)$. Como as duas sequências coincidem em todo o lado depois do primeiro número, quereremos, em última análise, dizer que têm o mesmo limite, no sentido de Definição 6.10, e portanto devem ser pensadas como "definindo" o mesmo real. Logo, devemos pensar nestas sequências de Cauchy como o mesmo número real.

Consequentemente, voltamos a precisar de definir uma relação de equivalência sobre as sequências de Cauchy, e identificar os números reais com *classes* de equivalência. Precisamos primeiro da ideia de uma função que tende para 0 no limite. Para qualquer função $h : \mathbb{N} \to \mathbb{Q}$, digamos que h *tende para* 0 sse, para qualquer $\varepsilon \in \mathbb{Q}$ positivo, se tem que $(\exists \ell \in \mathbb{N})(\forall n > \ell)|h(n)| < \varepsilon$.[4] Além disso, sendo f e g funções $\mathbb{N} \to \mathbb{Q}$, definamos $(f - g)(n) = f(n) - g(n)$. Definimos agora:

$$f \eqsim g \text{ sse } (f - g) \text{ tende para } 0.$$

Há que verificar que $\eqsim$ é uma relação de equivalência; e de fato é. Podemos então, se quisermos, definir os reais como as classes de equivalência, sob $\eqsim$, de todas as sequências de Cauchy de $\mathbb{N} \to \mathbb{Q}$.

Feito isto, como de costume escrevemos $[f]_{\eqsim}$ para a classe de equivalência que tem f como um membro. Porém, para manter a legibilidade, no que se segue omitimos o subscrito e escrevemos apenas $[f]$. Estipulamos também que, para cada $q \in \mathbb{Q}$, se tem $q_{\mathbb{R}} = [c_q]$, onde c_q é a função constante $c_q(n) = q$ para todo $n \in \mathbb{N}$. Definimos depois relações e operações básicas sobre os reais, por exemplo:

$$[f] + [g] = [(f + g)]$$
$$[f] \times [g] = [(f \times g)]$$

onde $(f + g)(n) = f(n) + g(n)$ e $(f \times g)(n) = f(n) \times g(n)$. Claro que também há que verificar que cada uma de $(f + g)$, $(f - g)$ e $(f \times g)$ é sequência de Cauchy quando f e g o são; mas o são, e deixamo-lo ao leitor.

[4] Compare-se com a definição de $\lim_{x \to \infty} f(x) = 0$ em Seção 1.2.

Por fim, definimos uma noção de ordem. Dizemos que $[f]$ é *positivo* sse tanto $[f] \neq 0_{\mathbb{Q}}$ como $(\exists \ell \in \mathbb{N})(\forall n > \ell) 0 < f(n)$. Depois dizemos $[f] < [g]$ sse $[(g - f)]$ é positivo. Há que verificar que isto está bem definido (isto é, que não depende da escolha da função "representante" na classe de equivalência). Mas tendo feito isto, é bastante fácil mostrar que obtemos as propriedades algébricas certas; isto é:

Teorema 6.11. *As sequências de Cauchy constituem um corpo ordenado.*

Prova. Exercício. □

É mais difícil provar que os reais, assim construídos, têm a propriedade de completude, pelo que daremos a prova.

Teorema 6.12. *Todo o conjunto não vazio de sequências de Cauchy com majorante tem supremo (menor majorante).*

Esboço de prova. Seja S um conjunto não vazio de sequências de Cauchy com majorante. Então existe algum $p \in \mathbb{Q}$ tal que $p_{\mathbb{R}}$ é majorante de S. Seja $r \in S$; então existe algum $q \in \mathbb{Q}$ tal que $q_{\mathbb{R}} < r$. Logo, se existe supremo de S, este está entre $q_{\mathbb{R}}$ e $p_{\mathbb{R}}$ (inclusive).

Nos aproximaremos do supremo, abordando-o simultaneamente por baixo e por cima. Em particular, definimos duas funções, $f, g \colon \mathbb{N} \to \mathbb{Q}$, com a ideia de que f se irá aproximar do supremo por cima e g por baixo. Começamos por definir:

$$f(0) = p$$
$$g(0) = q$$

Então, sendo $a_n = \frac{f(n)+g(n)}{2}$, definamos:[5]

$$f(n+1) = \begin{cases} a_n & \text{se } (\forall h \in S)[h] \leq (a_n)_{\mathbb{R}} \\ f(n) & \text{caso contrário} \end{cases}$$

[5]Isto é uma definição recursiva. Mas ainda *não* demos qualquer razão para crer que definições recursivas são legítimas.

$$g(n+1) = \begin{cases} a_n & \text{se } (\exists h \in S)[h] \geq (a_n)_{\mathbb{R}} \\ g(n) & \text{caso contrário} \end{cases}$$

Tanto f como g são sequências de Cauchy. (Isto verifica-se com relativa facilidade, mas deixamo-lo como exercício.) Note-se que a função $(f-g)$ tende para 0, pois a diferença entre f e g reduz-se à metade a cada passo. Logo $[f] = [g]$.

Mostramos primeiro que $[f]$ é majorante de S, isto é, que $(\forall h \in S)[h] \leq [f]$. (Invocaremos Teorema 6.11 ao longo do argumento.) Seja $h \in S$ e suponhamos, por *reductio*, que $[f] < [h]$, de modo que $0_{\mathbb{R}} < [(h-f)]$. Como f é uma sequência de Cauchy monotonamente decrescente, existe algum $n \in \mathbb{N}$ tal que $[(c_{f(n)} - f)] < [(h-f)]$. Logo:

$$(f(n))_{\mathbb{R}} = [c_{f(n)}] < [f] + [(h-f)] = [h],$$

o que contradiz o fato de que, por construção, $[h] \leq (f(n))_{\mathbb{R}}$.

Mostramos a seguir que $[f] = [g]$ é o supremo de S. Seja j qualquer sequência de Cauchy e suponhamos $[j] < [g]$. Raciocinando como acima (usando o fato de g ser *crescente*), existe $n \in \mathbb{N}$ tal que $[j] < (g(n))_{\mathbb{R}}$. Mas, por construção, existe $h \in S$ tal que $(g(n))_{\mathbb{R}} \leq [h]$, logo $[j] < [h]$ e portanto $[j]$ não é majorante de S. □

Problemas

Problema 6.1. Mostre que $(m+n)_{\mathbb{Z}} = m_{\mathbb{Z}} + n_{\mathbb{Z}}$ e $m \leq n \leftrightarrow m_{\mathbb{Z}} \leq n_{\mathbb{Z}}$, para quaisquer $m, n \in \mathbb{N}$.

Problema 6.2. Mostre que $\backsim$ é uma relação de equivalência.

Problema 6.3. Mostre que $(i+j)_{\mathbb{Q}} = i_{\mathbb{Q}} + j_{\mathbb{Q}}$ e $(i \times j)_{\mathbb{Q}} = i_{\mathbb{Q}} \times j_{\mathbb{Q}}$ e $i \leq j \leftrightarrow i_{\mathbb{Q}} \leq j_{\mathbb{Q}}$, para quaisquer $i, j \in \mathbb{Z}$.

Problema 6.4. Mostre que $\mathbb{Z}$ é um anel comutativo.

Problema 6.5. Mostre que $\mathbb{Z}$ é um anel ordenado.

Problema 6.6. Mostre que $\mathbb{Q}$ é um corpo ordenado.

Problema 6.7. Mostre que $\mathbb{R}$ é um corpo ordenado.

Problema 6.8. Seja $f(n) = 0$ para todo n. Seja $g(n) = \frac{1}{(n+1)^2}$. Mostre que ambas são sequências de Cauchy e que o limite de ambas as funções é 0, de modo que também $f \sim_{\mathbb{R}} g$.

Problema 6.9. Mostre que as sequências de Cauchy constituem um corpo ordenado.

CAPÍTULO 7

Conjuntos Infinitos

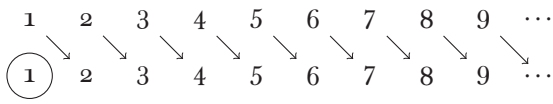
No capítulo anterior, mostramos como construir, de certo modo, várias coisas—inteiros, racionais e reais—assumindo uma teoria ingênua dos conjuntos e os números naturais. A pergunta deste capítulo é: podemos construir o próprio conjunto dos números naturais *em si* usando a teoria dos conjuntos?

7.1 Hotel de Hilbert

O conjunto dos números naturais é obviamente infinito. Logo, se não quisermos *valer-nos* dos números naturais, o nosso primeiro passo é caracterizar um conjunto infinito em termos que não exijam mencionar os próprios números naturais. Eis uma abordagem agradável, apresentada por Hilbert numa conferência de 1924. Ele nos pede que imaginemos

> [...] um hotel com um número finito de quartos. Todos esses quartos devem estar ocupados por exatamente um hóspede. Se os hóspedes trocam de quarto de algum modo, [mas] de tal forma que cada quarto continua a conter no máximo uma pessoa, então nenhum quarto ficará livre, e o dono do hotel não pode,

> deste modo, criar um novo lugar para um hóspede que acaba de chegar [...¶...]
>
> Agora estipulamos que o hotel tem infinitos quartos numerados 1, 2, 3, 4, 5, ..., cada um ocupado por exatamente um hóspede. Logo que chega um novo hóspede, o dono só precisa de mudar cada um dos hóspedes antigos para o quarto associado ao número imediatamente superior, e o quarto 1 ficará livre para o recém-chegado.



> (publicado em Hilbert 2013, 730; a nossa tradução)

O ponto crucial é que o Hotel de Hilbert tem infinitos quartos; e podemos tomar a sua explicação como definição do que significa dizer isto. De fato, foi esta a abordagem de Dedekind (apresentada aqui, claro, com enorme anacronismo; a definição de Dedekind é de 1888):

Definição 7.1. Um conjunto A é *infinito de Dedekind* sse existe uma injeção de A para um subconjunto próprio de A. Isto é, existe algum $o \in A$ e uma injeção $f : A \to A$ tal que $o \notin \mathrm{Im}(f)$.

7.2 Álgebras de Dedekind

Não queremos apenas que os números naturais sejam infinitos; queremos que tenham certas propriedades (algébricas): precisam de comportar-se bem sob adição, multiplicação, e assim por diante.

A ideia de Dedekind foi tomar como primitiva a noção de *função sucessor* e depois caracterizar os números como aqueles que têm as seguintes propriedades:

1. Existe um número, 0, que não é sucessor de nenhum número
   isto é, $0 \notin \mathrm{Im}(s)$
   isto é, $\forall x\ s(x) \neq 0$

2. Números distintos têm sucessores distintos
   isto é, s é uma injeção
   isto é, $\forall x \forall y (s(x) = s(y) \rightarrow x = y)$

3. Todo o número se obtém de 0 por aplicações repetidas da função sucessor.

As duas primeiras condições são fáceis de lidar com, usando lógica de primeira ordem (ver acima). Mas não podemos tratar (3) apenas com lógica de primeira ordem. O avanço de Dedekind foi reformular a condição (3), em termos de teoria dos conjuntos, assim:

3′. Os números naturais são o menor conjunto que é *fechado para a função sucessor*: isto é, se aplicarmos s a qualquer membro do conjunto, obtemos outro membro do conjunto.

Mas precisamos de explicitar isto com calma.

Definição 7.2. Para qualquer função f, o conjunto X é *f-fechado* sse $(\forall x \in X) f(x) \in X$. Definimos agora, para qualquer o:

$$\mathrm{clo}_f(o) = \bigcap \{X : o \in X \text{ e } X \text{ é } f\text{-fechado}\}$$

Logo $\mathrm{clo}_f(o)$ é a interseção de todos os conjuntos f-fechados que têm o como um membro. Intuitivamente, então, $\mathrm{clo}_f(o)$ é o conjunto f-fechado *menor* que tem o como um membro. O resultado seguinte torna preciso esse pensamento intuitivo;

Lema 7.3. *Para qualquer função f e qualquer $o \in A$:*

1. $o \in \mathrm{clo}_f(o)$; *e*

2. $\mathrm{clo}_f(o)$ *é* f*-fechado; e*

3. *se* X *é* f*-fechado e* $o \in X$*, então* $\mathrm{clo}_f(o) \subseteq X$

Prova. Note-se que existe pelo menos um conjunto f-fechado com o como um membro, nomeadamente $\mathrm{Im}(f) \cup \{o\}$. Logo $\mathrm{clo}_f(o)$, a interseção de *todos* os tais conjuntos, existe. Cumpre agora verificar (1)–(3).

Quanto a (1): $o \in \mathrm{clo}_f(o)$ pois é uma interseção de conjuntos que todos têm o como um membro.

Quanto a (2): suponhamos $x \in \mathrm{clo}_f(o)$. Se $o \in X$ e X é f-fechado, então $x \in X$, e agora $f(x) \in X$ pois X é f-fechado. Logo $f(x) \in \mathrm{clo}_f(o)$.

Quanto a (3): de modo geral, se $X \in C$ então $\bigcap C \subseteq X$. □

Com isto, podemos dizer:

Definição 7.4. Uma *álgebra de Dedekind* é um conjunto A juntamente com uma função $f : A \to A$ e algum $o \in A$ tal que:

1. $o \notin \mathrm{Im}(f)$

2. f é uma injeção

3. $A = \mathrm{clo}_f(o)$

Como $A = \mathrm{clo}_f(o)$, o resultado anterior nos diz que A é o menor conjunto f-fechado com o como um membro. Claramente uma álgebra de Dedekind é infinita de Dedekind; basta olhar para as cláusulas (1) e (2) da definição. Mas o fato mais relevante é que qualquer conjunto infinito de Dedekind pode ser transformado numa álgebra de Dedekind.

Teorema 7.5. *Se existe um conjunto infinito de Dedekind, então existe uma álgebra de Dedekind.*

Prova. Seja D infinito de Dedekind. Logo existe uma injeção $g\colon D \to D$ e um elemento $o \in D \setminus \mathrm{Im}(g)$. Seja agora $A = \mathrm{clo}_g(o)$; por Lema 7.3, A existe e $o \in A$. Seja $f = g\restriction_A$. Mostraremos que A, f, o formam uma álgebra de Dedekind.

Quanto a (1): $o \notin \mathrm{Im}(g)$ e $\mathrm{Im}(f) \subseteq \mathrm{Im}(g)$, logo $o \notin \mathrm{Im}(f)$.

Quanto a (2): g é injetiva em D; logo $f \subseteq g$ tem de ser injetiva.

Quanto a (3): por Lema 7.3, A é g-fechado; *a fortiori*, A é f-fechado. Logo $\mathrm{clo}_f(o) \subseteq A$ por Lema 7.3. Como também $\mathrm{clo}_f(o)$ é f-fechado e $f = g\restriction_A$, segue-se que $\mathrm{clo}_f(o)$ é g-fechado. Logo $A \subseteq \mathrm{clo}_f(o)$ por Lema 7.3. □

7.3 Álgebras de Dedekind e Indução Aritmética

Crucialmente, agora, uma álgebra de Dedekind—de fato, *qualquer* álgebra de Dedekind—servirá como substituto dos números naturais. Isto deve-se à seguinte consequência trivial:

Teorema 7.6 (Indução aritmética). *Sejam N, s, o uma álgebra de Dedekind. Então, para qualquer conjunto X:*

$$\textit{se } o \in X \textit{ e } (\forall n \in N \cap X)s(n) \in X\textit{, então } N \subseteq X.$$

Prova. Pela definição de álgebra de Dedekind, $N = \mathrm{clo}_s(o)$. Ora, se tanto $o \in X$ como $(\forall n \in N)(n \in X \to s(n) \in X)$, então $N = \mathrm{clo}_s(o) \subseteq X$. □

Como a indução é característica dos números naturais, o ponto é este. Dado qualquer conjunto infinito de Dedekind, podemos formar uma álgebra de Dedekind e usar essa álgebra como substituto dos números naturais.

Admitidamente, Teorema 7.6 formula a indução em termos *de teoria dos conjuntos.* Mas podemos facilmente pôr o princípio em termos mais familiares:

Corolário 7.7. *Sejam N, s, o uma álgebra de Dedekind. Então, para qualquer fórmula $\varphi(x)$, que pode ter parâmetros:*

$$\text{se } \varphi(o) \text{ e } (\forall n \in N)(\varphi(n) \to \varphi(s(n))), \text{ então } (\forall n \in N)\varphi(n)$$

Prova. Seja $X = \{n \in N : \varphi(n)\}$, e use-se agora Teorema 7.6 □

Neste resultado, falamos de uma fórmula "ter parâmetros". O que isto significa, em traços grosseiros, é que, para quaisquer objetos $c_1, \ldots, c_k$, podemos trabalhar com $\varphi(x, c_1, \ldots, c_k)$. Com mais precisão, podemos enunciar o resultado sem mencionar "parâmetros" do seguinte modo. Para qualquer fórmula $\varphi(x, v_1, \ldots, v_k)$, cujas variáveis livres são todas as exibidas, temos:

$$\begin{aligned}\forall v_1 \ldots \forall v_k((&\varphi(o, v_1, \ldots, v_k) \wedge \\ &(\forall x \in N)(\varphi(x, v_1, \ldots, v_k) \to \varphi(s(x), v_1, \ldots, v_k))) \to \\ &(\forall x \in N)\varphi(x, v_1, \ldots, v_k))\end{aligned}$$

Evidentemente, falar em "ter parâmetros" pode tornar a leitura muito mais fácil. (Na Parte III, usaremos este expediente com frequência.)

Voltando às álgebras de Dedekind: dada qualquer álgebra de Dedekind, podemos também definir as habituais funções aritméticas de adição, multiplicação e exponenciação. Isto não é trivial, porém, e envolve a técnica de *definição recursiva.* É uma técnica que introduziremos e justificaremos muito mais tarde, e num contexto bem mais geral. (Os entusiastas podem querer voltar a isto depois de Capítulo 13, ou talvez ler um tratamento alternativo, como Potter 2004, pp. 95–8.) Mas, onde N, s, o formam uma álgebra de Dedekind, acabaremos por poder estipular o seguinte:

$$\begin{array}{ccc} a + o = a & a \times o = o & a^o = s(o) \\ a + s(b) = s(a + b) & a \times s(b) = (a \times b) + a & a^{s(b)} = a^b \times a \end{array}$$

e mostrar que estas definições se comportam como se espera.

7.4 "Prova" de Dedekind da existência de um conjunto infinito

Neste capítulo, apresentamos um tratamento em teoria dos conjuntos dos números naturais, em termos de álgebras de Dedekind. Em Seção 6.5, refletimos sobre o significado filosófico da aritmetização da análise (entre outras coisas). Cumpre agora refletir sobre o significado do que aqui alcançamos.

Ao longo de Capítulo 6, tomamos os números naturais como dados e os usamos para construir os inteiros, os racionais e os reais, de modo explícito. Neste capítulo, não demos uma construção explícita dos números naturais. Mostramos apenas que, *dado qualquer conjunto infinito de Dedekind*, podemos definir um conjunto que se comportará como queremos que $\mathbb{N}$ se comporte.

Obviamente, pois, não podemos pretender ter respondido a uma questão metafísica, como *quais os objetos que são os números naturais.* Mas isso é bom. Afinal, na Seção 6.5, sublinhamos que estaríamos errados se pensássemos na definição de $\mathbb{R}$ como conjunto de cortes de Dedekind como uma *descoberta*, em vez de uma estipulação conveniente. A observação crucial é que os cortes de Dedekind exemplificam as propriedades matemáticas centrais dos números reais. Assim também aqui: a observação crucial é que *qualquer* álgebra de Dedekind exemplifica as propriedades matemáticas centrais dos números naturais. (De fato, Dedekind insistiu neste ponto ao provar que todas as álgebras de Dedekind são *isomorfas* (1888, Theorems 132–3). Não surpreende, então, que muitos "estruturalistas" contemporâneos cite Dedekind como precursor.)

Além disso, mostramos como embutir a teoria dos números naturais numa teoria ingênua simples dos conjuntos, que ainda permanece bastante informal, mas que (aparentemente) não assume os números naturais como dados. Assim, podemos estar no caminho de realizar o próprio projeto ambicioso de Dedekind, que ele explicou assim:

> Na ciência nada que seja provável deve ser acredi-

> tado sem prova. Embora esta exigência pareça razoável, não posso considerá-la satisfeita sequer nos métodos mais recentes de lançar os fundamentos da ciência mais simples; a saber, aquela parte da lógica que trata da teoria dos números. Ao falar de aritmética (álgebra, análise) como mera parte da lógica, quero insinuar que considero o conceito de número totalmente independente das noções ou intuições de espaço e tempo—que o considero antes um produto imediato das leis puras do pensamento. (Dedekind, 1888, preface)

A ideia audaciosa de Dedekind é esta. Acabamos de mostrar como construir os números naturais usando só (ingenuamente) a teoria dos conjuntos. Em Capítulo 6, vimos como construir os reais dados os números naturais e alguma teoria dos conjuntos. Logo, talvez, "a aritmética (álgebra, análise)" acabe por ser "meramente parte da lógica" (no sentido alargado que Dedekind dá à palavra "lógica").

Essa é a ideia. Mas esperemos um momento. A nossa construção de uma álgebra de Dedekind (o nosso substituto dos números naturais) é condicional à existência de um conjunto infinito de Dedekind. (Basta olhar para trás, para Teorema 7.5.) A menos que a existência de um conjunto infinito de Dedekind possa ser estabelecida via "lógica" ou "as leis puras do pensamento", o projeto estagna.

Então, *pode* a existência de um conjunto infinito de Dedekind ser estabelecida pelas "leis puras do pensamento"? Eis o esforço de Dedekind:

> O meu próprio reino de pensamentos, isto é, a totalidade S de todas as coisas que podem ser objetos do meu pensamento, é infinita. Pois, se s designa um elemento de S, então o pensamento s' de que s pode ser objeto do meu pensamento é ele próprio um elemento de S. Se considerarmos isto como uma

> imagem $\varphi(s)$ do elemento s, então ... S é infinito de Dedekind, como queríamos ter provado. (Dedekind, 1888, §66)

Isto é algo de espantoso de encontrar no meio de um livro que em grande parte consiste em provas matemáticas altamente rigorosas. Duas observações valem a pena.

Primeiro: esta "prova" mal tem o que hoje reconheceríamos como caráter "matemático". Fala de objetos psicológicos (pensamentos), e apenas de objetos *possíveis*.

Segundo: pelo menos tal como apresentamos as álgebras de Dedekind, esta "prova" tem uma lacuna técnica direta. Se o argumento de Dedekind é bem-sucedido, estabelece só que há infinitas coisas (nomeadamente, infinitos pensamentos). Mas Dedekind também precisa de nos dar uma razão para ver S como um único *conjunto*, com infinitos membros, em vez de pensar em S como *algumas coisas* (no plural).

O fato de Dedekind não ver aqui uma lacuna pode sugerir que o seu uso da palavra "totalidade" não corresponde exatamente ao *nosso* uso da palavra "conjunto".[1] Mas isto não seria demasiado surpreendente. O projeto que perseguimos nos últimos dois capítulos—uma "construção" dos naturais e, a partir deles, uma "construção" dos inteiros, reais e racionais—foi todo ele feito de modo ingênuo. Nos apoiamos neste conjunto, ou naquele conjunto, sempre que precisamos, sem enunciar muitos princípios gerais sobre exatamente que conjuntos existem, e quando. Mas sabemos que precisamos de *alguns* princípios gerais, pois caso contrário caímos no Paradoxo de Russell.

Chegou a hora de superarmos a nossa ingenuidade.

[1] De fato, temos outras razões para pensar que não correspondia; ver Potter (2004, p. 23).

7.5 Apêndice: Provando Schröder-Bernstein

Antes de deixarmos a teoria ingênua dos conjuntos, há ainda uma última prova ingênua (mas sofisticada!) a considerar. Trata-se de uma prova de Schröder-Bernstein (Teorema 5.17): se $A \preceq B$ e $B \preceq A$ então $A \approx B$; isto é, dados injeções $f : A \to B$ e $g : B \to A$, existe uma bijeção $h : A \to B$.

Neste capítulo, seguimos a noção dedekindiana de *fechos*. De fato, Dedekind deu uma prova muito elegante de Schröder-Bernstein com esta noção, e aqui a apresentamos. A prova segue de perto Potter (2004, pp. 157–8), se quiser um tratamento ligeiramente diferente mas essencialmente semelhante. Uma breve pesquisa na Internet também convencerá o leitor de que este é um teorema—tal como a irracionalidade de $\sqrt{2}$—para o qual existem *muitas* provas interessantes e distintas.

Usando notação semelhante à de Definição 7.2, seja

$$\mathrm{Clo}_f(B) = \bigcap \{X : B \subseteq X \text{ e } X \text{ é } f\text{-fechado}\}$$

para cada conjunto B e função f. Assim definido, $\mathrm{Clo}_f(B)$ é o menor conjunto f-fechado que contém B, no sentido de que:

Lema 7.8. *Para qualquer função f, e qualquer B:*

1. $B \subseteq \mathrm{Clo}_f(B)$*; e*
2. $\mathrm{Clo}_f(B)$ *é f-fechado; e*
3. *se X é f-fechado e $B \subseteq X$, então* $\mathrm{Clo}_f(B) \subseteq X$.

Prova. Exatamente como no Lema 7.3. □

Precisamos de mais um fato para chegar a Schröder-Bernstein:

Proposição 7.9. *Se $A \subseteq B \subseteq C$ e $A \approx C$, então $A \approx B \approx C$.*

Prova. Dada uma bijeção $f\colon C \to A$, seja $F = \mathrm{Clo}_f(C \setminus B)$ e defina-se uma função g com domínio C do seguinte modo:

$$g(x) = \begin{cases} f(x) & \text{se } x \in F \\ x & \text{caso contrário} \end{cases}$$

Mostraremos que g é uma bijeção de C para B, donde seguirá que $g \circ f^{-1}\colon A \to B$ é uma bijeção, completando a prova.

Primeiro alegamos que, se $x \in F$ mas $y \notin F$, então $g(x) \neq g(y)$. Por reductio, suponhamos o contrário, de modo que $y = g(y) = g(x) = f(x)$. Como $x \in F$ e F é f-fechado por Lema 7.8, temos $y = f(x) \in F$, contradição.

Suponhamos agora $g(x) = g(y)$. Então, pelo que precede, $x \in F$ sse $y \in F$. Se $x, y \in F$, então $f(x) = g(x) = g(y) = f(y)$, logo $x = y$ pois f é uma bijeção. Se $x, y \notin F$, então $x = g(x) = g(y) = y$. Logo g é uma injeção.

Resta mostrar que $\mathrm{Im}(g) = B$. Fixemos $x \in B \subseteq C$. Se $x \notin F$, então $g(x) = x$. Se $x \in F$, então $x = f(y)$ para algum $y \in F$, pois caso contrário $F \setminus \{x\}$ seria f-fechado e estenderia $C \setminus B$, o que é impossível por Lema 7.8; agora $g(y) = f(y) = x$. □

Finalmente, eis a prova do resultado principal. Recordemos que, dada uma função h e um conjunto D, definimos $h[D] = \{h(x) : x \in D\}$.

Prova de Schröder-Bernstein. Sejam $f\colon A \to B$ e $g\colon B \to A$ injeções. Como $f[A] \subseteq B$, temos $g[f[A]] \subseteq g[B] \subseteq A$. Além disso, $g \circ f\colon A \to g[f[A]]$ é uma injeção, pois tanto g como f o são; e de fato $g \circ f$ é uma bijeção, apenas pela forma como definimos o seu contradomínio. Logo $g[f[A]] \approx A$, e portanto, por Proposição 7.9, existe uma bijeção $h\colon A \to g[B]$. Além disso, g^{-1} é uma bijeção $g[B] \to B$. Logo $g^{-1} \circ h\colon A \to B$ é uma bijeção. □

PARTE III

A Concepção Iterativa

Introdução à Parte III

Parte II tratou os conjuntos de modo ingênuo e informal. É tempo de apertar o rigor e apresentar uma teoria formal dos conjuntos. Esse é o objetivo da Parte III.

A nossa teoria formal é uma teoria de primeira ordem com um único predicado binário, $\in$. Enunciaremos vários axiomas que regem o comportamento da relação de pertença. Contudo, introduziremos estes axiomas só à medida que forem necessários, e consideraremos como poderíamos justificá-los quando os encontrarmos. Como resultado, introduziremos os nossos axiomas *ao longo* dos capítulos seguintes.

Pode, no entanto, ser útil ao leitor ter uma lista de todos os axiomas num só lugar. Eis *todos* os axiomas que consideraremos na Parte III. Como na Parte II, a escolha de minúsculas e maiúsculas nas letras orienta-se apenas pela legibilidade:

Extensionalidade.

$$\forall A \forall B (\forall x (x \in A \leftrightarrow x \in B) \rightarrow A = B)$$

União.

$$\forall A \exists U \forall x (x \in U \leftrightarrow (\exists b \in A) x \in b),$$

isto é, $\bigcup A$ existe para qualquer conjunto A.

Pares.

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

isto é, $\{a, b\}$ existe para quaisquer a e b.

Conjuntos potência.

$$\forall A \exists P \forall x (x \in P \leftrightarrow (\forall z \in x) z \in A),$$

isto é, $\wp(A)$ existe para qualquer conjunto A.

Infinito.

$$\exists I((\exists o \in I)\forall x(x \notin o) \wedge (\forall x \in I)(\exists s \in I)\forall z(z \in s \leftrightarrow (z \in x \vee z = x))),$$

isto é, existe um conjunto que tem $\emptyset$ como um membro e que é fechado para $x \mapsto x \cup \{x\}$.

Fundamentação.

$$\forall A(\forall x\, x \notin A \vee (\exists b \in A)(\forall x \in A)x \notin b),$$

isto é, $A = \emptyset$ ou $(\exists b \in A)A \cap b = \emptyset$.

Boa ordenação. Para todo o conjunto A, existe uma relação que bem ordena A. (Escrever este axioma em lógica de primeira ordem seria demasiado penoso para valer a pena aqui.)

Esquema de separação. Para qualquer fórmula $\varphi(x)$ que não contenha "S":

$$\forall A\exists S\forall x(x \in S \leftrightarrow (\varphi(x) \wedge x \in A)),$$

isto é, $\{x \in A : \varphi(x)\}$ existe para qualquer conjunto A.

Esquema de substituição. Para qualquer fórmula $\varphi(x,y)$ que não contenha "B":

$$\forall A((\forall x \in A)\exists! y\, \varphi(x,y) \rightarrow \exists B\forall y(y \in B \leftrightarrow (\exists x \in A)\varphi(x,y))),$$

isto é, $\{y : (\exists x \in A)\varphi(x,y)\}$ existe para qualquer A, se φ for "funcional".

Em ambos os esquemas, as fórmulas podem conter parâmetros. De fato, ao longo de Parte III, seguimos a convenção de que qualquer fórmula pode conter parâmetros. (Ver Seção 7.3 para recordar o que significa dizer que uma fórmula pode conter parâmetros.)

Z$^-$ é Extensionalidade, União, Pares, Conjuntos potência, Infinito, Separação.

Z é **Z**$^-$ mais Fundamentação.

ZF$^-$ é **Z** mais Substituição.

ZF é **ZF**$^-$ mais Fundamentação.

ZFC é **ZF** mais Boa ordenação.

CAPÍTULO 8

A Concepção Iterativa

8.1 Extensionalidade

O primeiro ponto a notar é que os conjuntos são individuados pelos seus membros. Mais precisamente:

Extensionalidade (Extensionalidade). Se os conjuntos A e B têm os mesmos membros, então A e B são o mesmo conjunto.

$$\forall A \forall B \, (\forall x \, (x \in A \leftrightarrow x \in B) \rightarrow A = B)$$

Assumimos isto ao longo de Parte II. Mas vale a pena repetir. O axioma de extensionalidade exprime a ideia fundamental de que um conjunto é determinado pelos seus membros. (Assim, os conjuntos podem contrastar-se com os *conceitos*, nos quais precisamente os mesmos objetos podem cair sob muitos conceitos diferentes.)

Por que aceitar este princípio? Ora, é plausível dizer que qualquer recusa da extensionalidade é uma decisão de abandonar tudo o que possa ser chamado *teoria dos conjuntos*. A teoria dos conjuntos não é mais nem menos do que a teoria de coleções extensionais.

O verdadeiro desafio na Parte III, contudo, é estabelecer princípios que nos digam *que conjuntos existem.* E acaba por se verificar que a única resposta verdadeiramente "óbvia" a esta questão é provadamente errada.

8.2 Paradoxo de Russell (de novo)

Na Parte II, trabalhamos com uma teoria ingênua dos conjuntos. Mas, segundo uma concepção *muito* ingênua, os conjuntos são apenas as extensões dos predicados. Este pensamento ingênuo imporia o seguinte princípio:

Compreensão ingênua. $\{x : \varphi(x)\}$ existe para qualquer fórmula φ.

Por tentador que seja este princípio, é provadamente inconsistente. Vimo-lo na Seção 2.6, mas o resultado é tão importante, e tão direto, que vale a pena repetir. Ipsis verbis.

Teorema 8.1 (Paradoxo de Russell). *Não existe conjunto* $R = \{x : x \notin x\}$

Prova. Se $R = \{x : x \notin x\}$ existisse, então $R \in R$ sse $R \notin R$, o que é uma contradição. □

Russell descobriu este resultado em junho de 1901. (Não colocou, porém, o paradoxo exatamente na forma em que o acabamos de apresentar, pois estava a considerar a teoria dos conjuntos de Frege, tal como delineada em *Grundgesetze.* Voltaremos a isto na Seção 8.6.) Russell escreveu a Frege a 16 de junho de 1902, explicando a inconsistência no sistema de Frege. Para a correspondência, e um pouco de contexto, ver Heijenoort (1967, pp. 124–8).

Vale a pena enfatizar que esta prova de duas linhas é um resultado de *lógica pura.* De fato, usamos implicitamente um axioma (não lógico?) de extensionalidade na nossa notação $\{x : x \notin x\}$; pois $\{x : \varphi(x)\}$ deve ser *o único* (pela extensionalidade) conjunto

dos φs, se algum existir. Mas podemos evitar até o vestígio da extensionalidade, bastando enunciar o resultado assim: *não existe conjunto cujos membros sejam exatamente os conjuntos que não são membros de si próprios.* E isto pouco tem a ver com conjuntos. Como o próprio Russell observou, raciocínio rigorosamente análogo levanos a concluir: *nenhum homem barbeia exatamente os homens que não se barbeiam a si próprios.* Ou: *nenhum pug cheira exatamente os pugs que não se cheiram a si próprios.* E assim por diante. Esquematicamente, a forma do resultado é apenas:

$$\neg\exists x\forall z(Rzx \leftrightarrow \neg Rzz).$$

E isto não passa de um teorema (esquema) da lógica de primeira ordem. Consequentemente, não podemos evitar o paradoxo de Russell apenas mexendo na nossa teoria dos conjuntos; ele surge antes sequer de *chegarmos* à teoria dos conjuntos. Se formos usar lógica (clássica) de primeira ordem, temos simplesmente de *aceitar* que não existe conjunto $R = \{x : x \notin x\}$.

A consequência é esta. Se quisermos aceitar a compreensão ingênua enquanto *evitamos* a inconsistência, não basta mexer na *teoria dos conjuntos.* Teríamos, em vez disso, de reformular a nossa *lógica.*

É claro que já foram apresentadas teorias dos conjuntos com lógicas não clássicas. Mas são—no mínimo—não standard. A abordagem usual ao paradoxo de Russell é tratá-lo como uma prova direta de não existência e depois tentar aprender a viver com ela. É essa a abordagem que seguiremos.

8.3 Predicativo e Impredicativo

O conjunto de Russell, R, foi definido via $\{x : x \notin x\}$. Dito com mais detalhe, R seria o conjunto que contém todos e apenas aqueles conjuntos que não são membros de si próprios. Assim, ao definir R, quantificamos sobre o domínio que conteria R (se este existisse).

Isto é uma definição *impredicativa*. Mais em geral, podemos dizer que uma definição é impredicativa sse quantifica sobre um domínio que contém o objeto que está a ser definido.

Na sequência dos paradoxos, Whitehead, Russell, Poincaré e Weyl rejeitaram tais definições impredicativas como "viciosamente circulares":

> Uma análise dos paradoxos a evitar mostra que todos eles resultam de uma espécie de círculo vicioso. Os círculos viciosos em questão surgem ao supor que uma coleção de objetos pode conter membros que só podem ser definidos por meio da coleção como um todo[.... ¶]
>
> O princípio que nos permite evitar totalidades ilegítimas pode enunciar-se do seguinte modo: "Tudo o que envolve *todos* os membros de uma coleção não pode ser um dos membros da coleção"; ou, inversamente: "Se, supondo que uma certa coleção tivesse um todo, ela teria membros só definíveis em termos desse todo, então a dita coleção não tem todo." Chamaremos a isto o "princípio do círculo vicioso", porque nos permite evitar os círculos viciosos envolvidos na suposição de totalidades ilegítimas. (Whitehead and Russell, 1910, p. 37)

Se os acompanharmos na rejeição das definições impredicativas motivada pelo *princípio do círculo vicioso*, então podemos tentar substituir o desastroso esquema de compreensão ingênua (de Seção 8.2) por algo deste gênero:

Compreensão predicativa. Para toda a fórmula φ que quantifica apenas sobre conjuntos: existe o conjunto′ $\{x : \varphi(x)\}$.

Enquanto os conjuntos′ não forem conjuntos, não se seguirá contradição.

Infelizmente, a compreensão predicativa não é muito *abrangente.* Afinal, introduz-nos novas entidades, conjuntos′. Logo teremos de considerar fórmulas que quantificam sobre conjuntos′. Se estas derem sempre um conjunto′, então o paradoxo de Russell surgirá de novo, bastando considerar o conjunto′ de todos os conjuntos′ que não são membros de si próprios. Assim, a seguir o mesmo fio de ideias, temos de dizer que uma fórmula que quantifica sobre conjuntos′ produz um conjunto′′ correspondente. E então precisaremos de conjuntos′′′, conjuntos′′′′, etc. Para evitar uma enxurrada de primos, será mais fácil pensar nestes como conjuntos$_0$, conjuntos$_1$, conjuntos$_2$, conjuntos$_3$, conjuntos$_4$,.... E isto dar-nos-ia entrada na teoria (simples) dos tipos.

Há algumas objeções óbvias a tal teoria (embora não seja óbvio que sejam objeções *esmagadoras*). Em resumo: a teoria resultante é pouco manejável; é prolixa ao postular diferentes tipos de objetos; e não está claro, no fim, que as definições impredicativas sejam sequer *assim tão más.*

Para ilustrar o último ponto, consideremos esta observação de Ramsey:

> podemos referir-nos a um homem como o mais alto de um grupo, identificando-o por meio de uma totalidade da qual ele próprio é membro sem que haja círculo vicioso. (Ramsey, 1925)

O ponto de Ramsey é que "o homem mais alto do grupo" *é* uma definição impredicativa; mas é obviamente perfeitamente legítima.

Alguém poderá responder que, neste caso, poderíamos identificar a pessoa mais alta por meios *predicativos.* Por exemplo, talvez pudéssemos simplesmente apontar para o homem em questão. A objeção às definições impredicativas, então, teria claramente de se limitar a entidades que *só* possam ser identificadas impredicativamente. Mas mesmo assim, seria preciso ouvir mais, sobre por que tal "impredicatividade essencial" seria tão grave.[1]

[1] Para mais, ver Linnebo (2010).

Admitidamente, as definições impredicativas são péssimas notícias, se quisermos que as definições nos forneçam algo como uma receita para *criar* um objeto. Pois, dada uma definição impredicativa, estar-se-ia genuinamente preso a um círculo vicioso: para criar o objeto especificado impredicativamente, seria preciso *primeiro* criar todos os objetos (incluindo o objeto especificado impredicativamente), uma vez que este é especificado em termos de todos os objetos; logo seria preciso criar o objeto especificado impredicativamente antes de o ter criado. Mas, de novo, isto só é uma objeção séria a conjuntos especificados de modo "essencialmente impredicativo", se pensarmos nos conjuntos como coisas que *criamos*. E nós (provavelmente) não os criamos.

Assim—para o bem ou para o mal—a abordagem que se tornou comum não envolve tomar uma posição rígida sobre (im)predicatividade. Envolve antes o que hoje se considera a abordagem cumulativa-iterativa. No fim, isto permitir-nos-á estratificar os conjuntos em "estágios"—um *pouco* como a abordagem predicativa estratifica entidades em conjuntos$_0$, conjuntos$_1$, conjuntos$_2$, ...—mas não postularemos diferença de natureza entre eles.

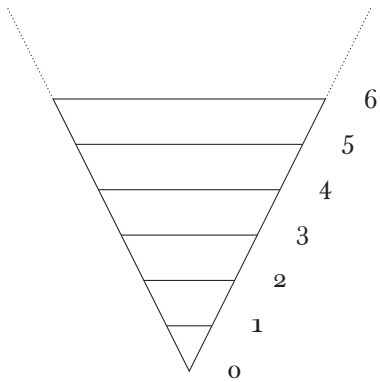
8.4 A Abordagem Cumulativa-Iterativa

Eis um enunciado um pouco mais completo de como estratificaremos os conjuntos em estágios:

> Os conjuntos formam-se em *estágios*. Para cada estágio S, há certos estágios que são *anteriores* a S. No estágio S, cada coleção constituída por conjuntos formados em estágios anteriores a S é formada num conjunto. Não há conjuntos além dos que são formados nos estágios. (Shoenfield, 1977, p. 323)

Isto é um esboço da *concepção cumulativa-iterativa do conjunto*. Servirá de base à teoria formal dos conjuntos que apresentamos em Parte III.

Exploremos isto com um pouco mais de detalhe. Como Shoenfield descreve o processo, em cada estágio formamos novos conjuntos a partir dos conjuntos que tínhamos disponíveis nos estágios anteriores. Assim, na imagem de Shoenfield, no estágio inicial, estágio 0, não há estágios *anteriores*, e portanto *a fortiori* não há conjuntos disponíveis a partir de estágios anteriores.[2] Logo formamos apenas um conjunto: o conjunto sem membros $\emptyset$. No estágio 1, está exatamente um conjunto disponível a partir de estágios anteriores, logo apenas um conjunto novo é $\{\emptyset\}$. No estágio 2, dois conjuntos estão disponíveis a partir de estágios anteriores, e formamos dois conjuntos novos $\{\{\emptyset\}\}$ e $\{\emptyset, \{\emptyset\}\}$. No estágio 3, quatro conjuntos estão disponíveis a partir de estágios anteriores, logo formamos doze conjuntos novos.... Assim, a imagem cumulativa-iterativa dos conjuntos ficará algo assim (com números a indicar estágios):



Então: por que deveríamos aceitar esta narrativa?

Uma razão é que é uma narrativa agradável e tratável. Dado o fim da narrativa mais óbvia, isto é, a compreensão ingênua, precisamos de algo de substância.

Mas a narrativa não é *apenas* agradável. Temos boa razão para acreditar que qualquer teoria dos conjuntos baseada nesta narrativa será *consistente*. Eis a razão.

[2]Por que deveríamos supor que *existe* um primeiro estágio? Ver a nota de rodapé a *Estágios-são-ordenados* em Seção 9.1.

Dada a concepção cumulativa-iterativa do conjunto, formamos conjuntos em estágios; e os seus membros têm de ser objetos que já estavam disponíveis. Assim, para qualquer estágio S, podemos formar o conjunto

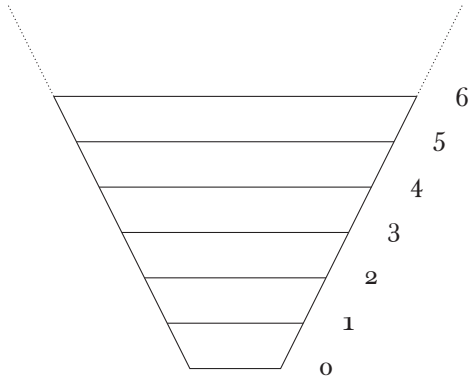
$$R_S = \{x : x \notin x \text{ e } x \text{ já estava disponível antes de } S\}$$

O raciocínio envolvido na prova do paradoxo de Russell estabelece agora que o próprio R_S não está disponível antes do estágio S. E isso não é uma contradição. Além disso, se aceitarmos a concepção cumulativa-iterativa do conjunto, nem sequer deveríamos ter *esperado* poder formar o próprio conjunto de Russell. Pois esse seria o conjunto de todos os conjuntos que não são membros de si próprios que "alguma vez estarão disponíveis". Em resumo: o fato de, provadamente, não podermos formar o conjunto de Russell não é *surpreendente*, dada a narrativa cumulativa-iterativa; é antes *previsível*.

8.5 Urelementos ou Não?

Nos próximos capítulos, tentaremos extrair axiomas da concepção cumulativa-iterativa do conjunto. Mas, antes de avançar mais, precisamos de dizer algo mais sobre os *urelementos*.

A imagem de Seção 8.4 só nos permitia formar novos conjuntos a partir de *conjuntos* antigos. Contudo, podemos querer admitir que certos *não-conjuntos*—vacas, porcos, grãos de areia, ou o que seja—possam ser membros de conjuntos. Nesse caso, começaríamos com certos elementos básicos, *urelementos*, e diríamos que em cada estágio S formaríamos "todos os possíveis" conjuntos constituídos por urelementos ou conjuntos formados em estágios anteriores a S (em qualquer combinação). A imagem resultante parecer-se-ia mais com isto:



Temos pois uma decisão a tomar: *devemos admitir urelementos?*

Do ponto de vista filosófico, faz sentido incluir urelementos na nossa teorização. A principal razão é tornar a teoria dos conjuntos *aplicável*. Para ilustrar, recordemos de Capítulo 5 que dizemos que dois conjuntos A e B têm o mesmo tamanho, isto é, $A \approx B$, sse existe uma bijeção entre eles. Agora, se as vacas no campo e os porcos no chiqueiro constituem ambos conjuntos, podemos oferecer um tratamento conjunto-teórico da afirmação "há tantas vacas como porcos". Mas se proibirmos urelementos, de modo que as vacas e os porcos *não* constituem conjuntos, esse tratamento conjunto-teórico deixará de estar disponível. De fato, não teremos capacidade direta de aplicar a teoria dos conjuntos a mais nada além dos próprios conjuntos. (Para mais razões para incluir urelementos, ver Potter 2004, pp. vi, 24, 50–1.)

Matematicamente, contudo, é bastante raro admitir urelementos. Em parte, porque é *ligeiramente* mais fácil formular a teoria dos conjuntos sem urelementos. Mas, ocasionalmente, encontram-se justificações mais interessantes para excluir urelementos da teoria dos conjuntos:

> De acordo com a convicção de que a teoria dos conjuntos é o fundamento da matemática, devemos ser capazes de captar toda a matemática apenas falando de conjuntos, pelo que as nossas variáveis não devem

> quantificar sobre objetos como vacas e porcos. (Kunen, 1980, p. 8)

Assim: um foco na aplicabilidade sugeriria *incluir* urelementos; um foco num objetivo fundacional redutivo (reduzir a matemática à teoria pura dos conjuntos) poderia sugerir *excluí-los*. Uma ligeira preguiça também aponta no sentido de excluir urelementos.

Seguiremos o caminho mais preguiçoso. Em parte, há também uma justificação pedagógica. O nosso objetivo é introduzir os elementos da teoria dos conjuntos de que precisa para começar na filosofia da teoria dos conjuntos. E a maior parte dessa literatura filosófica discute teorias dos conjuntos formuladas *sem* urelementos. Logo este livro será, talvez, mais útil se se mantiver bastante próximo dessa literatura.

8.6 Apêndice: Lei Básica V de Frege

Na Seção 8.2, explicamos que Russell formulou o paradoxo como um problema para o sistema que Frege delineou nos seus *Grundgesetze*. O sistema de Frege não incluía uma formulação direta da compreensão ingênua. Logo, neste apêndice, explicaremos muito sumariamente o que o sistema de Frege *incluía*, como se relaciona com a compreensão ingênua e como se relaciona com o paradoxo de Russell.

O sistema de Frege é de *segunda ordem* e foi concebido para formular a noção de *extensão de um conceito*.[3] Usando notação inspirada em Frege, escreveremos $\epsilon x\, F(x)$ para *a extensão do conceito* F. Este é um dispositivo que toma um *predicado*, "F", e transforma-o num *termo* (de primeira ordem), "$\epsilon x\, F(x)$". Com este dispositivo, Frege propôs a seguinte *definição* de pertença:

$$a \in b =_{\text{df}} \exists G(b = \epsilon x\, G(x) \wedge Ga)$$

[3]Estritamente falando, Frege tenta formalizar uma noção mais geral: o "valor de curso" de uma função. As extensões de conceitos são um caso especial da noção mais geral. Ver Heck (2012, pp. 8–9) para pormenores.

grosso modo: $a \in b$ sse a cai sob um conceito cuja extensão é b. (Note que o quantificador "$\exists G$" é de segunda ordem.) Frege também sustentava o seguinte princípio, conhecido como *Lei básica V*:

$$\epsilon x\, F(x) = \epsilon x\, G(x) \leftrightarrow \forall x(Fx \leftrightarrow Gx)$$

grosso modo: conceitos têm extensões idênticas sse são coextensivos. (De novo, tanto "F" como "G" estão em posição de predicado.) Um princípio simples liga agora a pertença à satisfação de propriedades:

Lema 8.2 (em *Grundgesetze*). $\forall F \forall a(a \in \epsilon x\, F(x) \leftrightarrow Fa)$

Prova. Fixe F e a. Ora $a \in \epsilon x\, F(x)$ sse $\exists G(\epsilon x\, F(x) = \epsilon x\, G(x) \wedge Ga)$ (pela definição de pertença) sse $\exists G(\forall x(Fx \leftrightarrow Gx) \wedge Ga)$ (pela lei básica V) sse Fa (por lógica elementar de segunda ordem).□

E disto resulta quase de imediato a compreensão ingênua:

Lema 8.3 (em *Grundgesetze.*). $\forall F \exists s \forall a(a \in s \leftrightarrow Fa)$

Prova. Fixe F; agora Lema 8.2 dá $\forall a(a \in \epsilon x\, F(x) \leftrightarrow Fa)$; logo $\exists s \forall a(a \in s \leftrightarrow Fa)$ por generalização existencial. O resultado segue-se, pois F era arbitrário. □

O paradoxo de Russell segue tomando F dado por $\forall x(Fx \leftrightarrow x \notin x)$.

CAPÍTULO 9

Passos rumo a Z

No capítulo anterior, consideramos a concepção iterativa de conjunto. Neste capítulo, tentaremos extrair a maior parte dos axiomas da teoria de conjuntos de Zermelo, isto é, **Z**. A abordagem é *inteiramente* inspirada em Boolos (1971), Scott (1974) e Shoenfield (1977).

9.1 A história com mais pormenor

Na Seção 8.4, citamos a descrição de Shoenfield do processo de formação de conjuntos. Queremos agora escrever mais alguns princípios, para tornar esta história um pouco mais precisa. Eis-os:

> *Estágios-são-chave.* Cada conjunto forma-se por algum estágio.
>
> *Estágios-são-ordenados.* Os estágios estão ordenados: alguns vêm *antes* de outros.[1]

[1] Assumiremos de fato—tacitamente—que os estágios estão *bem ordenados.* O que isto significa explica-se no Capítulo 10. É um pressuposto substancial. De fato, usando uma técnica muito engenhosa devida a Scott (1974), este pres-

> *Estágios-acumulam.* Para qualquer estágio S, e para quaisquer conjuntos que se formaram *antes* do estágio S: forma-se, no estágio S, um conjunto cujos membros são exatamente esses conjuntos. Mais nada se forma no estágio S.

Estes são princípios informais, mas poderemos usá-los para justificar vários dos axiomas da teoria de conjuntos de Zermelo.

(Convém uma palavra de cautela. Embora vamos apresentar alguns axiomas completamente standard, com nomes completamente standard, os princípios em itálico que acabamos de apresentar não têm nomes particulares na literatura. São apenas alcunhas que esperamos serem úteis.)

9.2 Separação

Começamos com um princípio que substitui a compreensão ingênua:

Esquema de Separação (Esquema de Separação). Para cada fórmula $\varphi(x)$, isto é um axioma: para qualquer A, o conjunto $\{x \in A : \varphi(x)\}$ existe.

Note que isto não é um único axioma. É um *esquema* de axiomas. Há *infinitamente muitos* axiomas de Separação; um para cada fórmula $\varphi(x)$. O esquema pode igualmente ser (e normalmente é) escrito do seguinte modo:

Para qualquer fórmula $\varphi(x)$ que não contenha "S", isto é um axioma:

$$\forall A \exists S \forall x (x \in S \leftrightarrow (\varphi(x) \wedge x \in A)).$$

---

suposto pode ser *evitado* e depois *deduzido*. (Isto também explicará por que devemos pensar que existe um estágio inicial.) Não podemos aprofundar isso aqui; para mais, veja Button (2021).

Em conformidade com a convenção assinalada no início da Parte III, as fórmulas φ nos axiomas de Separação podem ter parâmetros.[2]

A Separação justifica-se de imediato pela concepção cumulativa iterativa de conjuntos que vamos expondo. Para ver por que, seja A um conjunto. Então A forma-se por algum estágio S (por *Estágios-são-chave*). Uma vez que A se formou no estágio S, todos os membros de A formaram-se antes do estágio S (por *Estágios-acumulam*). Consideremos em particular todos os conjuntos que são membros de A e que também satisfazem φ; claramente todos estes conjuntos também se formaram antes do estágio S. Logo formam-se num conjunto $\{x \in A : \varphi(x)\}$ também no estágio S (por *Estágios-acumulam*).

Ao contrário da compreensão ingênua, isto evita o paradoxo de Russell. Pois não podemos simplesmente afirmar a existência do conjunto $\{x : x \notin x\}$. Pelo contrário, *dado* algum conjunto A, podemos afirmar a existência do conjunto $R_A = \{x \in A : x \notin x\}$. Mas isto só prova que $R_A \notin R_A$ e $R_A \notin A$, o que não é preocupante.

Contudo, a Separação tem uma consequência imediata e marcante:

Teorema 9.1. *Não existe conjunto* universal*, isto é,* $\{x : x = x\}$ *não existe.*

Prova. Por redução ao absurdo, suponhamos que V é um conjunto universal. Então, por Separação, $R = \{x \in V : x \notin x\} = \{x : x \notin x\}$ existe, o que contradiz o paradoxo de Russell. □

A ausência de um conjunto universal—de fato, a abertura da hierarquia de conjuntos—é uma das ideias mais fundamentais por detrás da concepção cumulativa iterativa. Por isso vale a pena ver que, intuitivamente, poderíamos chegar lá por outro caminho. Um conjunto universal teria de ser um membro de

[2]Para uma explicação do que isto significa, veja a discussão imediatamente a seguir a Corolário 7.7.

si próprio. Mas, na nossa concepção cumulativa iterativa, cada conjunto aparece (pela primeira vez) na hierarquia no primeiro estágio imediatamente a seguir a todos os seus membros. Isto implica que *nenhum* conjunto pertence a si próprio. Pois qualquer conjunto que pertencesse a si próprio teria de ocorrer pela primeira vez imediatamente depois do estágio em que ocorreu pela primeira vez, o que é absurdo. (Veremos na Definição 11.15 como tornar esta explicação mais rigorosa, usando a noção de "posto" de um conjunto. Contudo, precisaremos de mais alguns axiomas para o fazer.)

Eis mais algumas consequências da Separação e da Extensionalidade.

Proposição 9.2. *Se existe algum conjunto, então $\emptyset$ existe.*

Prova. Se A é um conjunto, $\emptyset = \{x \in A : x \neq x\}$ existe por Separação. □

Proposição 9.3. *$A \setminus B$ existe quaisquer que sejam os conjuntos A e B*

Prova. $A \setminus B = \{x \in A : x \notin B\}$ existe por Separação. □

Também se verifica que existem (quase) interseções arbitrárias:

Proposição 9.4. *Se $A \neq \emptyset$, então $\bigcap A = \{x : (\forall y \in A) x \in y\}$ existe.*

Prova. Seja $A \neq \emptyset$, logo existe algum $c \in A$. Então $\bigcap A = \{x : (\forall y \in A) x \in y\} = \{x \in c : (\forall y \in A) x \in y\}$, que existe por Separação. □

Note, porém, a condição de que $A \neq \emptyset$; pois $\bigcap \emptyset$ seria o conjunto universal, vacuamente, o que contradiz Teorema 9.1.

9.3 União

Proposição 9.4 deu-nos interseções. Mas, se quisermos que existam uniões arbitrárias, precisamos de enunciar outro axioma:

União (União). Para qualquer conjunto A, o conjunto $\bigcup A = \{x : (\exists b \in A) x \in b\}$ existe.

$$\forall A \exists U \forall x (x \in U \leftrightarrow (\exists b \in A) x \in b)$$

Este axioma também se justifica pela concepção cumulativa iterativa. Seja A um conjunto, logo A forma-se por algum estágio S (por *Estágios-são-chave*). Cada membro de A formou-se *antes* de S (por *Estágios-acumulam*); logo, raciocinando de modo análogo, cada membro de cada membro de A formou-se antes de S. Assim, todos *esses* conjuntos estão disponíveis antes de S, para serem reunidos num conjunto em S. E esse conjunto é precisamente $\bigcup A$.

9.4 Pares

O próximo axioma a considerar é o seguinte:

Pares (Pares). Quaisquer que sejam os conjuntos a, b, o conjunto $\{a, b\}$ existe.

$$\forall a \forall b \exists P \forall x (x \in P \leftrightarrow (x = a \vee x = b))$$

Eis como justificar este axioma à luz da concepção iterativa. Suponhamos que a está disponível no estágio S, e b no estágio T. Seja M o que quer que seja dos estágios S e T que ocorre mais tarde. Então, uma vez que a e b estão ambos disponíveis no estágio M, o conjunto $\{a, b\}$ é uma coleção possível disponível em qualquer estágio depois de M (o que quer que seja o posterior).

Mas espera! Porquê assumir que *existem* estágios depois de M? Se não existir nenhum, a nossa justificação falha. Por isso,

para justificar Pares, teremos de acrescentar outro princípio à história que contamos na Seção 9.1, nomeadamente:

Estágios-prosseguem. Não há último estágio.

Este princípio está justificado? Nada na história de Shoenfield afirmava *explicitamente* que não há último estágio. Ainda assim, mesmo que seja (estritamente falando) um acréscimo à nossa história, encaixa bem na ideia básica de que os conjuntos se formam em estágios. Aceitá-lo-emos simplesmente no que se segue. E, assim, aceitaremos também o Axioma dos Pares.

Armados com este novo axioma, podemos provar a existência de muitos mais conjuntos. Por exemplo:

Proposição 9.5. *Quaisquer que sejam os conjuntos a e b, existem os seguintes conjuntos:*

1. $\{a\}$
2. $a \cup b$
3. $\langle a, b\rangle$

Prova. (1). Por Pares, $\{a, a\}$ existe, o que é $\{a\}$ por Extensionalidade.

(2). Por Pares, $\{a, b\}$ existe. Ora, $a \cup b = \bigcup\{a, b\}$ existe por União.

(3). Por (1), $\{a\}$ existe. Por Pares, $\{a, b\}$ existe. Ora, $\{\{a\}, \{a, b\}\} = \langle a, b\rangle$ existe, de novo por Pares. □

9.5 Conjuntos potência

Prosseguiremos com outro axioma:

Conjuntos potência (Conjuntos potência). Para qualquer

conjunto A, existe o conjunto $\wp(A) = \{x : x \subseteq A\}$.

$$\forall A \exists P \forall x (x \in P \leftrightarrow (\forall z \in x) z \in A)$$

A nossa justificação é bastante direta. Suponhamos que A se forma no estágio S. Então todos os membros de A estavam disponíveis antes de S (por *Estágios-acumulam*). Logo, raciocinando como na justificação da Separação, cada subconjunto de A forma-se até ao estágio S. Estão todos disponíveis, para serem reunidos num único conjunto, em qualquer estágio depois de S. E sabemos que existe algum tal estágio, pois S não é o último estágio (por *Estágios-prosseguem*). Logo $\wp(A)$ existe.

Eis uma boa consequência dos conjuntos potência:

Proposição 9.6. *Dados quaisquer conjuntos A, B, o seu produto cartesiano $A \times B$ existe.*

Prova. O conjunto $\wp(\wp(A \cup B))$ existe por Conjuntos potência e Proposição 9.5. Logo, por Separação, existe este conjunto:

$$C = \{z \in \wp(\wp(A \cup B)) : (\exists x \in A)(\exists y \in B) z = \langle x, y \rangle\}.$$

Agora, para quaisquer $x \in A$ e $y \in B$, o conjunto $\langle x, y \rangle$ existe por Proposição 9.5. Além disso, uma vez que $x, y \in A \cup B$, temos $\{x\}, \{x, y\} \in \wp(A \cup B)$, e $\langle x, y \rangle \in \wp(\wp(A \cup B))$. Logo $A \times B = C$.□

Nesta prova, o axioma dos conjuntos potência interage com a Separação. E isso não surpreende. Sem Separação, o princípio dos conjuntos potência não seria muito *poderoso*. Afinal, a Separação nos diz que subconjuntos de um conjunto existem e, por conseguinte, determina quão "gordo" é cada conjunto potência.

9.6 Infinito

Já temos axiomas suficientes para garantir que há infinitamente muitos conjuntos (se é que existe algum). De fato, suponhamos que algum conjunto existe; então $\emptyset$ existe (por Proposição 9.2).

Agora, para qualquer conjunto x, o conjunto $x \cup \{x\}$ existe por Proposição 9.5. Logo, aplicando isto algumas vezes, obtemos conjuntos como se segue:

0. $\emptyset$
1. $\{\emptyset\}$
2. $\{\emptyset, \{\emptyset\}\}$
3. $\{\emptyset, \{\emptyset\}, \{\emptyset, \{\emptyset\}\}\}$
4. $\{\emptyset, \{\emptyset\}, \{\emptyset, \{\emptyset\}\}, \{\emptyset, \{\emptyset\}, \{\emptyset, \{\emptyset\}\}\}\}$

e podemos verificar que cada um destes conjuntos é distinto.

Começamos a numeração em 0, por várias razões. Mas uma delas é esta. Não é difícil verificar que o conjunto que rotulamos "n" tem exatamente n membros e que (intuitivamente) se forma no n-ésimo estágio.

Contudo. Isto nos dá *infinitamente muitos* conjuntos, mas não garante que exista um conjunto *infinito*, isto é, um conjunto com infinitamente muitos membros. E isto importa deveras: a menos que encontremos um conjunto infinito de Dedekind, não podemos construir uma álgebra de Dedekind. Mas queremos uma álgebra de Dedekind, para a podermos tratar como o conjunto dos números naturais. (Compare-se com Seção 7.4.)

Importa que os axiomas que estabelecemos até aqui *não* garantem a existência de nenhum conjunto infinito. Por isso temos de enunciar um novo axioma:

Infinito (Infinito). Existe um conjunto, I, tal que $\emptyset \in I$ e $x \cup \{x\} \in I$ sempre que $x \in I$.

$$\exists I((\exists o \in I)\forall x\ x \notin o \wedge \\ (\forall x \in I)(\exists s \in I)\forall z(z \in s \leftrightarrow (z \in x \vee z = x)))$$

É fácil ver que o conjunto I que nos é dado pelo Axioma do Infinito é infinito de Dedekind. O seu elemento distinguido é $\emptyset$,

e a injeção em I é dada por $s(x) = x \cup \{x\}$. Agora, Teorema 7.5 mostrou como extrair uma álgebra de Dedekind de um conjunto infinito de Dedekind; e trataremos isso como o nosso conjunto de números naturais. Mais precisamente:

Definição 9.7. Seja I qualquer conjunto que nos seja dado pelo Axioma do Infinito. Seja s a função $s(x) = x \cup \{x\}$. Seja $\omega = \mathrm{clo}_s(\emptyset)$. Chamamos *números naturais* aos membros de ω, e dizemos que n é o resultado de n aplicações de s a $\emptyset$.

Pode agora voltar atrás e verificar que o conjunto rotulado "n", há uns parágrafos, será tratado *como* o número n.

Discutiremos o significado desta estipulação na Seção 9.8. Por ora, permite-nos provar um resultado intuitivo:

Proposição 9.8. *Nenhum número natural é infinito de Dedekind.*

Prova. A prova é por indução, isto é, Teorema 7.6. Claramente $0 = \emptyset$ não é infinito de Dedekind. Para o passo indutivo, estabeleceremos a contrapositiva: se (absurdamente) $s(n)$ for infinito de Dedekind, então n é infinito de Dedekind.

Suponhamos, pois, que $s(n)$ é infinito de Dedekind, isto é, que existe alguma injeção f com $\mathrm{Im}(f) \subsetneq \mathrm{dom}(f) = s(n) = n \cup \{n\}$. Há dois casos a considerar.

Caso 1: $n \notin \mathrm{Im}(f)$. Então $\mathrm{Im}(f) \subseteq n$, e $f(n) \in n$. Seja $g = f \restriction_n$; agora $\mathrm{Im}(g) = \mathrm{Im}(f) \setminus \{f(n)\} \subsetneq n = \mathrm{dom}(g)$. Logo n é infinito de Dedekind.

Caso 2: $n \in \mathrm{Im}(f)$. Fixemos $m \in \mathrm{dom}(f) \setminus \mathrm{Im}(f)$, e definamos uma função h com domínio $s(n) = n \cup \{n\}$:

$$h(x) = \begin{cases} f(x) & \text{se } f(x) \neq n \\ m & \text{se } f(x) = n \end{cases}$$

Assim h e f coincidem em toda a parte, exceto que $h(f^{-1}(n)) = m \neq n = f(f^{-1}(n))$. Uma vez que f é uma injeção, $n \notin \mathrm{Im}(h)$; e $\mathrm{Im}(h) \subsetneq \mathrm{dom}(h) = s(n)$. Logo n é infinito de Dedekind, pelo argumento do Caso 1. □

A questão que permanece é como poderíamos *justificar* o Axioma do Infinito. A resposta curta é que precisaremos de acrescentar outro princípio à história que vamos contando. Esse princípio é o seguinte:

> *Estágios-atingem-o-infinito.* Há um estágio infinito. Isto é, há um estágio que (a) não é o primeiro estágio, e que (b) tem alguns estágios antes de si, mas que (c) não tem predecessor imediato.

O Axioma do Infinito segue-se diretamente deste princípio. Sabemos que o número natural n se forma no estágio n. Logo o conjunto ω forma-se no primeiro estágio infinito. E o próprio ω testemunha o Axioma do Infinito.

Isto, porém, apenas nos leva de volta à questão de como justificar *Estágios-atingem-o-infinito.* Tal como *Estágios-prosseguem*, não fazia parte explícita da história que contamos sobre a hierarquia cumulativa iterativa. Mas mais do que isso: nada na própria ideia de uma hierarquia iterativa, em que os conjuntos se formam estágio a estágio, nos obriga a pensar que o processo envolve um estágio *infinito.* Parece perfeitamente coerente pensar que os estágios estão ordenados como os números naturais.

Isto, contudo, levanta um problema óbvio. Na Seção 7.4, consideramos a "prova" de Dedekind de que existe um conjunto infinito de Dedekind (de pensamentos). Isto poderá não lhe ter parecido muito satisfatório. Mas se *Estágios-atingem-o-infinito* não nos é "imposto" pela concepção iterativa de conjunto (nem pelas "leis do pensamento"), continuamos sem uma justificação intrínseca para a afirmação de que existe um conjunto infinito de Dedekind.

Há muito mais a dizer aqui, naturalmente. Mas esperamos que esteja agora num ponto em que possa começar a refletir sobre o que poderá *custar* justificar um axioma (ou um princípio). No que se segue tomaremos simplesmente *Estágios-atingem-o-infinito* como concedido.

9.7 $\mathbf{Z}^-$: um marco

Voltaremos a *Estágios-atingem-o-infinito* na parte seguinte. Contudo, com o Axioma do Infinito, atingimos um marco importante. Temos agora todos os axiomas necessários para a teoria $\mathbf{Z}^-$. Em detalhe:

Definição 9.9. A teoria $\mathbf{Z}^-$ tem estes axiomas: Extensionalidade, União, Pares, Conjuntos potência, Infinito, e todas as instâncias do esquema de Separação.

O nome abrevia *Zermelo* (*menos* algo que abordaremos mais tarde). Zermelo merece a honra, pois formulou essencialmente esta teoria no seu trabalho de 1908a.[3]

Esta teoria é bastante poderosa para permitir-nos fazer uma quantidade enorme de matemática. Em particular, *deverá* voltar atrás na Parte II e convencer-se de que tudo o que fizemos, ingenuamente, pode ser feito de modo mais formal dentro de $\mathbf{Z}^-$. (Depois de o fazer durante algum tempo, poderá querer avançar e ler a Seção 9.9.) Por isso, daqui em diante, e sem mais comentários, assumiremos estar a trabalhar em $\mathbf{Z}^-$ (pelo menos).

9.8 A escolher os números naturais

Na Definição 9.7, definimos explicitamente a expressão "números naturais". Como deve entender esta estipulação? Não é uma afirmação metafísica, mas apenas uma decisão de *tratar* certos conjuntos como os números naturais. Já abordamos razões para pensar assim na Seção 3.2, na Seção 6.5 e na Seção 7.4. Mas podemos tornar essas razões ainda mais nítidas.

O nosso Axioma do Infinito segue von Neumann (1925). Mas eis outro axioma, que poderíamos ter adotado em alternativa:

[3] Para comentários interessantes sobre a história e os pormenores técnicos, veja Potter (2004, Apêndice A).

Axioma do Infinito de Zermelo 1908a. Existe um conjunto A tal que $\emptyset \in A$ e $(\forall x \in A)\{x\} \in A$.

Se tivéssemos usado o axioma de Zermelo em vez do nosso Axioma do Infinito (inspirado em von Neumann), ter-nos-iam sido dado igualmente um conjunto infinito de Dedekind e, portanto, uma álgebra de Dedekind. Na abordagem de Zermelo, o elemento distinguido da nossa álgebra seria de novo $\emptyset$ (o nosso sucedâneo de 0), mas a injeção seria dada pela aplicação $x \mapsto \{x\}$, em vez de $x \mapsto x \cup \{x\}$. A consequência mais simples é que Zermelo trata 2 como $\{\{\emptyset\}\}$, ao passo que nós (com von Neumann) tratamos 2 como $\{\emptyset, \{\emptyset\}\}$.

Porquê escolher um axioma do Infinito em vez do outro? A razão prática principal é que a abordagem de von Neumann "escala" bastante bem para números transfinitos. Exploraremos isto a partir do Capítulo 10 em diante. Porém, na simples perspectiva de *fazer aritmética*, ambas as abordagens serviriam igualmente bem. Por isso, se alguém lhe disser que os números naturais *são* conjuntos, a pergunta óbvia é: *Que conjuntos são eles?*

Esta pergunta precisa foi tornada célebre por Benacerraf (1965). Mas vale a pena sublinhar que é apenas o exemplo mais famoso de um fenómeno que já encontramos muitas vezes. O ponto essencial é este. A teoria dos conjuntos nos dá uma maneira de *simular* vários tipos de entidades "intuitivas": os reais, os racionais, os inteiros e os naturais, sim; mas também pares ordenados, funções e relações. Contudo, a teoria dos conjuntos nunca nos fornece uma escolha *única* de simulação. Há *sempre* alternativas que—de modo direto—nos teriam servido tão bem.

9.9 Apêndice: Fechamento, compreensão e interseção

Na Seção 9.7, sugerimos que voltasse atrás no trabalho ingênuo da Parte II e verificasse que este pode ser realizado em $\mathbf{Z}^-$. Se

seguiu esse conselho, *um* ponto poderá tê-lo feito tropeçar: o uso da *interseção* no tratamento por Dedekind dos *fechamentos*.

Recordemos, de Definição 7.2, que

$$\mathrm{clo}_f(o) = \bigcap \{X : o \in X \text{ e } X \text{ é } f\text{-fechado}\}.$$

A forma geral disto é uma definição da forma:

$$C = \bigcap \{X : \varphi(X)\}.$$

Mas isto deve fazer soar o alarme: uma vez que a compreensão ingênua falha, não há garantia de que $\{X : \varphi(X)\}$ exista. Parece, então, perigosamente, que tais definições são *truques*.

Felizmente, não são truques; ou melhor, se *o forem* tal como estão, então podemos empenhar-nos honestamente para as tornar legítimas. Esse trabalho honesto foi antecipado na Proposição 9.4, quando explicamos por que $\bigcap A$ existe para qualquer $A \neq \emptyset$. Mas vamos explicitá-lo.

Dada a Extensionalidade, se tentarmos definir C como $\bigcap\{X : \varphi(X)\}$, tudo o que realmente pedimos é um objeto C que satisfaça o seguinte:

$$\forall x(x \in C \leftrightarrow \forall X(\varphi(X) \rightarrow x \in X)) \qquad (*)$$

Suponhamos, agora, que existe *algum* conjunto, S, tal que $\varphi(S)$. Então, para satisfazer Eq. (*), podemos simplesmente definir C por *Separação*, do seguinte modo:

$$C = \{x \in S : \forall X(\varphi(X) \rightarrow x \in X)\}.$$

Deixamos como exercício verificar que esta definição produz Eq. (*), como desejado. E esta estratégia geral permitir-nos-á contornar qualquer uso aparente da compreensão ingênua na definição de interseções. No caso particular que nos levou a esta linha de raciocínio, nomeadamente o de $\mathrm{clo}_f(o)$, eis como isso funcionaria. Começamos a prova de Lema 7.3 ao notar que $o \in \mathrm{Im}(f) \cup \{o\}$ e que $\mathrm{Im}(f) \cup \{o\}$ é f-fechado. Assim, podemos definir o que queremos assim:

$$\mathrm{clo}_f(o) = \{x \in \mathrm{Im}(f) \cup \{o\} : (\forall X \ni o)(X \text{ é } f\text{-fechado} \rightarrow x \in X)\}.$$

Problemas

Problema 9.1. Mostre que, quaisquer que sejam os conjuntos a, b, c, o conjunto $\{a, b, c\}$ existe.

Problema 9.2. Mostre que, quaisquer que sejam os conjuntos $a_1, \ldots, a_n$, o conjunto $\{a_1, \ldots, a_n\}$ existe.

Problema 9.3. Mostre que, quaisquer que sejam os conjuntos A, B: (i) existe o conjunto de todas as relações com domínio A e contradomínio B; e (ii) existe o conjunto de todas as funções de A para B.

Problema 9.4. Seja A um conjunto, e seja $\sim$ uma relação de equivalência em A. Prove que existe o conjunto das classes de equivalência sob $\sim$ em A, isto é, $A/_\sim$.

CAPÍTULO 10

Ordinais

10.1 Introdução

No Capítulo 9, postulamos que existe um estágio infinito da hierarquia, na forma de *Estágios-atingem-o-infinito* (ver também o nosso axioma do Infinito). Contudo, dado *Estágios-prosseguem*, não podemos parar no estágio infinito; temos de continuar. Assim: no estágio seguinte ao primeiro estágio infinito, formamos todas as coleções possíveis de conjuntos que estavam disponíveis no primeiro estágio infinito; e repetimos; e repetimos; e repetimos; ...

Implicitamente, o que aqui aconteceu foi começarmos a invocar uma noção "intuitiva" de número, segundo a qual podem existir números *depois* de todos os números naturais. Em particular, a noção em causa é a de *ordinal transfinito*. O objetivo deste capítulo é tornar esta ideia mais rigorosa. Exploraremos a noção geral de ordinal e depois definiremos explicitamente certos conjuntos para serem os nossos ordinais.

10.2 A Ideia Geral de um Ordinal

Considere os números naturais, na sua ordem habitual:

$$0 < 1 < 2 < 3 < 4 < 5 < \cdots$$

Chamamos a isto, na gíria, uma ω-sucessão. E, de fato, esta ordenação geral espelha-se na nossa construção inicial dos estágios da hierarquia de conjuntos. Mas suponha agora que movemos 0 para o fim desta sucessão, de modo que fica depois de todos os outros números:

$$1 < 2 < 3 < 4 < 5 < \cdots < 0$$

Temos aqui as mesmas entidades, mas ordenadas de modo fundamentalmente diferente: a nossa primeira ordenação não tinha último elemento; a nova tem. De fato, a nova ordenação consiste numa ω-sucessão de entidades $(1, 2, 3, 4, 5, \ldots)$, seguida de outra entidade. Será uma sucessão do tipo $\omega + 1$.

Podemos gerar ainda mais tipos de ordenação, usando apenas estas entidades. Por exemplo, considere todos os números pares (na sua ordem natural) seguidos de todos os ímpares (na sua ordem natural):

$$0 < 2 < 4 < \cdots < 1 < 3 < \cdots$$

Isto é uma ω-sucessão seguida de outra ω-sucessão; uma sucessão do tipo $\omega + \omega$.

Pois bem, podemos continuar. Mas o que gostaríamos é de ter uma forma geral de entender esta linguagem sobre *ordenações*.

10.3 Boas Ordenações

A noção fundamental é a seguinte:

Definição 10.1. A relação $<$ *bem ordena* A sse cumpre estas duas condições:

1. $<$ é conexa, isto é, para todos $a, b \in A$, ou $a < b$ ou $a = b$ ou $b < a$;
2. todo o subconjunto não vazio de A tem um membro $<$-minimal, isto é, se $\emptyset \neq X \subseteq A$ então $(\exists m \in X)(\forall z \in X) z \not< $

m

É fácil ver que os três exemplos que acabamos de considerar eram de fato relações de boa ordenação.

Eis algumas observações elementares mas extremamente importantes sobre boa ordenação.

Proposição 10.2. *Se $<$ bem ordena A, então todo o subconjunto não vazio de A tem um único membro $<$-mínimo, e $<$ é irreflexiva, assimétrica e transitiva.*

Prova. Se X é um subconjunto não vazio de A, tem um membro $<$-minimal m, isto é, $(\forall z \in X) z \not< m$. Como $<$ é conexa, $(\forall z \in X) m \leq z$. Logo m é o membro $<$-mínimo de X.

Para a irreflexividade, fixe $a \in A$; o membro $<$-mínimo de $\{a\}$ é a, logo $a \not< a$. Para a transitividade, se $a < b < c$, então como $\{a, b, c\}$ tem um membro $<$-mínimo, $a < c$. A assimetria segue da irreflexividade e da transitividade □

Proposição 10.3. *Se $<$ bem ordena A, então para qualquer fórmula $\varphi(x)$:*

$$\textit{se } (\forall a \in A)((\forall b < a)\varphi(b) \to \varphi(a)), \textit{ então } (\forall a \in A)\varphi(a).$$

Prova. Provaremos a contrapositiva. Suponha $\neg(\forall a \in A)\varphi(a)$, isto é, que $X = \{x \in A : \neg\varphi(x)\} \neq \emptyset$. Então X tem um membro $<$-minimal, a. Logo $(\forall b < a)\varphi(b)$ mas $\neg\varphi(a)$. □

Esta última propriedade deve recordar-lhe o princípio de indução forte nos naturais, isto é: se $(\forall n \in \omega)((\forall m < n)\varphi(m) \to \varphi(n))$, então $(\forall n \in \omega)\varphi(n)$. E esta propriedade torna a boa ordenação numa noção muito *robusta*.[1]

[1] Lembrete: todas as fórmulas podem ter parâmetros (salvo indicação em contrário).

10.4 Isomorfismos de Ordem

Para explicar *quão* robusta é a boa ordenação, começaremos por introduzir um método para comparar boas ordenações.

Definição 10.4. Uma *boa ordenação* é um par $\langle A, < \rangle$, tal que $<$ bem ordena A. As boas ordenações $\langle A, < \rangle$ e $\langle B, \lessdot \rangle$ são *isomorfas por ordem* sse existe uma bijeção $f : A \to B$ tal que: $x < y$ sse $f(x) \lessdot f(y)$. Neste caso, escrevemos $\langle A, < \rangle \cong \langle B, \lessdot \rangle$, e dizemos que f é um *isomorfismo de ordem.*

Daqui em diante, por brevidade, falaremos em "isomorfismos" em vez de "isomorfismos de ordem". Intuitivamente, os isomorfismos são bijeções que preservam estrutura. Eis alguns fatos simples sobre isomorfismos.

Lema 10.5. *Composições de isomorfismos são isomorfismos, isto é: se $f : A \to B$ e $g : B \to C$ são isomorfismos, então $(g \circ f) : A \to C$ é um isomorfismo.*

Prova. Fica como exercício. □

Corolário 10.6. *$X \cong Y$ é uma relação de equivalência.*

Proposição 10.7. *Se $\langle A, < \rangle$ e $\langle B, \lessdot \rangle$ são boas ordenações isomorfas, então o isomorfismo entre elas é único.*

Prova. Sejam f e g isomorfismos $A \to B$. Provaremos o resultado por indução, isto é, usando Proposição 10.3. Fixe $a \in A$, e suponha (para indução) que $(\forall b < a) f(b) = g(b)$. Fixe $x \in B$.

Se $x \lessdot f(a)$, então $f^{-1}(x) < a$, logo $g(f^{-1}(x)) \lessdot g(a)$, invocando o fato de f e g serem isomorfismos. Mas como $f^{-1}(x) < a$, pela suposição $x = f(f^{-1}(x)) = g(f^{-1}(x))$. Logo $x \lessdot g(a)$. Analogamente, se $x \lessdot g(a)$ então $x \lessdot f(a)$.

Generalizando, $(\forall x \in B)(x \preccurlyeq f(a) \leftrightarrow x \preccurlyeq g(a))$. Segue que $f(a) = g(a)$ por Proposição 3.27. Logo $(\forall a \in A) f(a) = g(a)$ por Proposição 10.3. □

Isto dá alguma ideia de que as boas ordenações são robustas. Mas para continuar a explicar isto, convém introduzir mais notação.

Definição 10.8. Quando $\langle A, < \rangle$ é uma boa ordenação com $a \in A$, seja $A_a = \{x \in A : x < a\}$. Dizemos que A_a é um *segmento inicial* próprio de A (e permitimos que o próprio A seja um segmento inicial impróprio de A). Seja $<_a$ a restrição de $<$ ao segmento inicial, isto é, $< \restriction_{A_a^2}$.

Com esta notação, podemos enunciar e provar que nenhuma boa ordenação é isomorfa a nenhum dos seus segmentos iniciais próprios.

Lema 10.9. *Se $\langle A, < \rangle$ é uma boa ordenação com $a \in A$, então $\langle A, < \rangle \ncong \langle A_a, <_a \rangle$*

Prova. Por redução ao absurdo, suponha $f : A \to A_a$ um isomorfismo. Como f é uma bijeção e $A_a \subsetneq A$, usando Proposição 10.2 seja $b \in A$ o membro $<$-mínimo de A tal que $b \neq f(b)$. Mostraremos que $(\forall x \in A)(x < b \leftrightarrow x < f(b))$, donde seguirá, por Proposição 3.27, que $b = f(b)$, completando a redução ao absurdo.

Suponha $x < b$. Logo $x = f(x)$, pela escolha de b. E $f(x) < f(b)$, por f ser um isomorfismo. Logo $x < f(b)$.

Suponha $x < f(b)$. Logo $f^{-1}(x) < b$, por f ser um isomorfismo, e portanto $f^{-1}(x) = x$ pela escolha de b. Logo $x < b$. □

O nosso resultado seguinte mostra, grosso modo, que um "segmento inicial" de um isomorfismo é um isomorfismo:

Lema 10.10. *Sejam $\langle A, <\rangle$ e $\langle B, \lessdot\rangle$ boas ordenações. Se $f : A \to B$ é um isomorfismo e $a \in A$, então $f \restriction_{A_a} : A_a \to B_{f(a)}$ é um isomorfismo.*

Prova. Como f é um isomorfismo:

$$\begin{aligned} f[A_a] &= f[\{x \in A : x < a\}] \\ &= f[\{f^{-1}(y) \in A : f^{-1}(y) < a\}] \\ &= \{y \in B : y \lessdot f(a)\} \\ &= B_{f(a)} \end{aligned}$$

E $f \restriction_{A_a}$ preserva a ordem porque f preserva. □

Os nossos dois resultados seguintes estabelecem que as boas ordenações são sempre comparáveis:

Lema 10.11. *Sejam $\langle A, <\rangle$ e $\langle B, \lessdot\rangle$ boas ordenações. Se $\langle A_{a_1}, <_{a_1}\rangle \cong \langle B_{b_1}, \lessdot_{b_1}\rangle$ e $\langle A_{a_2}, <_{a_2}\rangle \cong \langle B_{b_2}, \lessdot_{b_2}\rangle$, então $a_1 < a_2$ sse $b_1 \lessdot b_2$*

Prova. Provaremos da *esquerda para a direita*; a outra direção é análoga. Suponha ambos $\langle A_{a_1}, <_{a_1}\rangle \cong \langle B_{b_1}, \lessdot_{b_1}\rangle$ e $\langle A_{a_2}, <_{a_2}\rangle \cong \langle B_{b_2}, \lessdot_{b_2}\rangle$, com $f : A_{a_2} \to B_{b_2}$ o nosso isomorfismo. Seja $a_1 < a_2$; então $\langle A_{a_1}, <_{a_1}\rangle \cong \langle B_{f(a_1)}, \lessdot_{f(a_1)}\rangle$ por Lema 10.10. Logo $\langle B_{b_1}, \lessdot_{b_1}\rangle \cong \langle B_{f(a_1)}, \lessdot_{f(a_1)}\rangle$, e portanto $b_1 = f(a_1)$ por Lema 10.9. Ora, $b_1 \lessdot b_2$ pois o domínio de f é B_{b_2}. □

Teorema 10.12. *Dadas quaisquer duas boas ordenações, uma é isomorfa a um segmento inicial (não necessariamente próprio) da outra.*

Prova. Sejam $\langle A, <\rangle$ e $\langle B, \lessdot\rangle$ boas ordenações. Usando a Separação, seja

$$f = \{\langle a, b\rangle \in A \times B : \langle A_a, <_a\rangle \cong \langle B_b, \lessdot_b\rangle\}.$$

Por Lema 10.11, $a_1 < a_2$ sse $b_1 \lessdot b_2$ para todos $\langle a_1, b_1\rangle, \langle a_2, b_2\rangle \in f$. Logo $f : \mathrm{dom}(f) \to \mathrm{Im}(f)$ é um isomorfismo.

Se $a_2 \in \mathrm{dom}(f)$ e $a_1 < a_2$, então $a_1 \in \mathrm{dom}(f)$ por Lema 10.10; logo $\mathrm{dom}(f)$ é um segmento inicial de A. Analogamente, $\mathrm{Im}(f)$ é um segmento inicial de B. Por redução ao absurdo, suponha que ambos são segmentos iniciais *próprios*. Então seja a o membro $<$-mínimo de $A \setminus \mathrm{dom}(f)$, de modo que $\mathrm{dom}(f) = A_a$, e seja b o membro $\lessdot$-mínimo de $B \setminus \mathrm{Im}(f)$, de modo que $\mathrm{Im}(f) = B_b$. Logo $f : A_a \to B_b$ é um isomorfismo, e portanto $\langle a, b\rangle \in f$, contradição. □

10.5 Construção de Von Neumann dos Ordinais

Teorema 10.12 sugere uma ideia. Poderíamos introduzir certos objetos, chamados *tipos de ordem*, para servirem de representantes das boas ordenações. Escrevendo $\mathrm{ord}(A, <)$ para o tipo de ordem da boa ordenação $\langle A, <\rangle$, gostaríamos de garantir os dois princípios seguintes:

$\mathrm{ord}(A, <) = \mathrm{ord}(B, \lessdot)$ sse $\langle A, <\rangle \cong \langle B, \lessdot\rangle$

$\mathrm{ord}(A, <) < \mathrm{ord}(B, \lessdot)$ sse $\langle A, <\rangle \cong \langle B_b, \lessdot_b\rangle$ para algum $b \in B$

Além disso, poderíamos esperar introduzir tipos de ordem *como certos conjuntos*, tal como podemos introduzir os números naturais como certos conjuntos.

A forma mais comum de o fazer—e a abordagem que seguiremos—é definir estes tipos de ordem via certos conjuntos bem ordenados *canónicos*. Estes conjuntos canónicos foram introduzidos pela primeira vez por von Neumann:

Definição 10.13. O conjunto A é *transitivo* sse $(\forall x \in A) x \subseteq A$. Então A é um *ordinal* sse A é transitivo e bem ordenado por $\in$.

Daqui em diante, usaremos letras gregas para ordinais. Segue-se imediatamente da definição que, se α é um ordinal, então $\langle \alpha, \in_\alpha\rangle$ é uma boa ordenação, onde $\in_\alpha = \{\langle x, y\rangle \in \alpha^2 : x \in y\}$. Assim,

abusando um pouco da notação, podemos dizer que o próprio α *é* uma boa ordenação.

Eis os primeiros ordinais:

$$\emptyset, \{\emptyset\}, \{\emptyset, \{\emptyset\}\}, \{\emptyset, \{\emptyset\}, \{\emptyset, \{\emptyset\}\}\}, \ldots$$

Repare que estes são precisamente os primeiros ordinais que encontramos no nosso Axioma do Infinito, isto é, na definição de von Neumann de ω (ver Seção 9.6). Isto não é coincidência. A definição de von Neumann dos ordinais trata os números naturais como ordinais, mas permite também ordinais transfinitos.

Como sempre, podemos agora perguntar: *são* estes os ordinais? Ou von Neumann deu-nos apenas alguns conjuntos que podemos *tratar* como os ordinais? O tipo de discussões que se podem ter sobre esta questão é semelhante às que tivemos na Seção 3.2, Seção 6.5, Seção 7.4, e Seção 9.8, pelo que não insistiremos no ponto. Em vez disso, daqui em diante usaremos "os ordinais" para falar dos "ordinais de von Neumann".

10.6 Propriedades Básicas dos Ordinais

Observamos que os primeiros ordinais são os números naturais. A razão principal para desenvolver uma teoria dos ordinais é alargar o princípio de indução que vale nos números naturais. Iremos construir isso através duma sucessão de resultados elementares.

Lema 10.14. *Todo o membro de um ordinal é um ordinal.*

Prova. Seja α um ordinal com $b \in \alpha$. Como α é transitivo, $b \subseteq \alpha$. Logo $\in$ bem ordena b pois $\in$ bem ordena α.

Para ver que b é transitivo, suponha $x \in c \in b$. Logo $c \in \alpha$ pois $b \subseteq \alpha$. Outra vez, como α é transitivo, $c \subseteq \alpha$, de modo que $x \in \alpha$. Logo $x, c, b \in \alpha$. Mas $\in$ bem ordena α, logo $\in$ é uma relação transitiva em α por Proposição 10.2. Logo, como $x \in c \in b$, temos $x \in b$. Generalizando, $c \subseteq b$ □

Corolário 10.15. $\alpha = \{\beta \in \alpha : \beta$ *é um ordinal*$\}$, *para qualquer ordinal* α

Prova. Imediato a partir de Lema 10.14. □

A ideia grosseira dos dois resultados principais seguintes, Teorema 10.16 e Teorema 10.17, é que os ordinais eles próprios são bem ordenados por pertença:

Teorema 10.16 (Indução Transfinita). *Para qualquer fórmula* $\varphi(x)$*:*

$$\textit{se } \exists\alpha\varphi(\alpha), \textit{ então } \exists\alpha(\varphi(\alpha) \wedge (\forall\beta \in \alpha)\neg\varphi(\beta))$$

onde os quantificadores exibidos ficam implicitamente restritos a ordinais.

Prova. Suponha $\varphi(\alpha)$, para algum ordinal α. Se $(\forall\beta \in \alpha)\neg\varphi(\beta)$, terminamos. Caso contrário, como α é um ordinal, tem um membro $\in$-mínimo que verifica φ, e este é um ordinal por Lema 10.14.□

Repare que podemos exprimir igualmente Teorema 10.16 como o esquema:

$$\text{se } \forall\alpha((\forall\beta \in \alpha)\varphi(\beta) \rightarrow \varphi(\alpha)), \text{ então } \forall\alpha\varphi(\alpha)$$

tomando simplesmente $\neg\varphi(\alpha)$ no Teorema 10.16, e depois fazendo manipulações lógicas elementares.

Teorema 10.17 (Tricotomia). $\alpha \in \beta \vee \alpha = \beta \vee \beta \in \alpha$, *para quaisquer ordinais* α *e* β.

Prova. A prova é por dupla indução, isto é, usando Teorema 10.16 duas vezes. Diga-se que x é *comparável* com y sse $x \in y \vee x = y \vee y \in x$.

Para a indução, suponha que todo o ordinal em α é comparável com *todo* o ordinal. Para uma segunda indução, suponha que α é comparável com todo o ordinal em β. Mostraremos que

α é comparável com β. Por indução em β, seguirá que α é comparável com todo o ordinal; e por indução em α, *todo* o ordinal é comparável com *todo* o ordinal, como pretendido. Basta supor que $\alpha \notin \beta$ e $\beta \notin \alpha$, e mostrar que $\alpha = \beta$.

Para mostrar que $\alpha \subseteq \beta$, fixe $\gamma \in \alpha$; este é um ordinal por Lema 10.14. Logo, pela primeira hipótese de indução, γ é comparável com β. Mas se $\gamma = \beta$ ou $\beta \in \gamma$ então $\beta \in \alpha$ (invocando a transitividade de α se necessário), contra a suposição; logo $\gamma \in \beta$. Generalizando, $\alpha \subseteq \beta$.

Raciocínio exatamente análogo, usando a segunda hipótese de indução, mostra que $\beta \subseteq \alpha$. Logo $\alpha = \beta$. □

Assim, por vezes escreveremos $\alpha < \beta$ em vez de $\alpha \in \beta$, já que $\in$ se comporta como relação de ordem. Não há razões profundas para além da familiaridade, e porque é mais fácil escrever $\alpha \leq \beta$ do que $\alpha \in \beta \vee \alpha = \beta$.[2]

Eis duas consequências rápidas dos últimos resultados, a primeira das quais põe a nova notação em prática:

Corolário 10.18. *Se* $\exists\alpha\varphi(\alpha)$*, então* $\exists\alpha(\varphi(\alpha) \wedge \forall\beta(\varphi(\beta) \to \alpha \leq \beta))$*. Além disso, para quaisquer ordinais* α, β, γ*, tanto* $\alpha \notin \alpha$ *como* $\alpha \in \beta \in \gamma \to \alpha \in \gamma$*.*

Prova. Tal como Proposição 10.2. □

Corolário 10.19. *A é um ordinal sse A é um conjunto transitivo de ordinais.*

Prova. Da esquerda para a direita. Por Lema 10.14. *Da direita para a esquerda.* Se A é um conjunto transitivo de ordinais, então $\in$ bem ordena A por Teorema 10.16 e Teorema 10.17. □

Agora, descrevemos Teorema 10.16 e Teorema 10.17 como dizendo que $\in$ bem ordena os ordinais. Contudo, temos de ser *muito cautelosos* com este tipo de afirmação, graças ao resultado seguinte:

[2] Poderíamos escrever $\alpha \subseteq \beta$; mas isso seria totalmente não convencional.

Teorema 10.20 (Paradoxo de Burali-Forti). *Não existe conjunto de todos os ordinais*

Prova. Por redução ao absurdo, suponha O o conjunto de todos os ordinais. Se $\alpha \in \beta \in O$, então α é um ordinal, por Lema 10.14, logo $\alpha \in O$. Logo O é transitivo, e portanto O é um ordinal por Corolário 10.19. Daí $O \in O$, contradizendo Corolário 10.18. □

Este resultado deve o nome a Burali-Forti. Mas foi Cantor em 1899—numa carta a Dedekind—quem viu pela primeira vez com clareza a *contradição* em supor que existe um conjunto de todos os ordinais. Como van Heijenoort explica:

> O próprio Burali-Forti considerou a contradição como estabelecendo, por *reductio ad absurdum*, o resultado de que a ordenação natural dos ordinais é apenas uma ordenação parcial. (Heijenoort, 1967, p. 105)

Deixando de lado o erro de Burali-Forti, podemos resumir o precedente como segue. Os ordinais são conjuntos que individualmente são bem ordenados por pertença, e coletivamente são bem ordenados por pertença (sem coletivamente constituírem um conjunto).

Para concluir, eis mais propriedades básicas dos ordinais que seguem de Teorema 10.16 e Teorema 10.17.

Proposição 10.21. *Toda a sucessão estritamente descendente de ordinais é finita.*

Prova. Toda a sucessão infinita estritamente descendente de ordinais $\alpha_0 > \alpha_1 > \alpha_2 > \ldots$ não tem membro $<$-mínimo, contradizendo Teorema 10.16. □

Proposição 10.22. *$\alpha \subseteq \beta \vee \beta \subseteq \alpha$, para quaisquer ordinais α, β.*

Prova. Se $\alpha \in \beta$, então $\alpha \subseteq \beta$ pois β é transitivo. Analogamente, se $\beta \in \alpha$, então $\beta \subseteq \alpha$. E se $\alpha = \beta$, então $\alpha \subseteq \beta$ e $\beta \subseteq \alpha$. Logo, por Teorema 10.17 terminamos. □

Proposição 10.23. *$\alpha = \beta$ sse $\alpha \cong \beta$, para quaisquer ordinais α, β.*

Prova. Os ordinais são boas ordenações; logo isto é imediato da Tricotomia (Teorema 10.17) e de Lema 10.9. □

10.7 Substituição

Na Seção 10.5, motivamos a introdução dos ordinais ao sugerir que os poderíamos tratar como tipos de ordem, isto é, como representantes canónicos de boas ordenações. Para que isso funcione, seria preciso provar que *toda a boa ordenação é isomorfa a algum ordinal.* Isso permitiria definir $\mathrm{ord}(A, <)$ como o ordinal α tal que $\langle A, < \rangle \cong \alpha$.

Infelizmente, *não* podemos provar o resultado desejado apenas com os axiomas introduzidos até aqui. (Veremos por que em Seção 12.2, mas por agora basta: não podemos.) Precisamos de uma ideia nova, e eis-a:

Esquema de Substituição (Esquema de Substituição). Para qualquer fórmula $\varphi(x, y)$, o seguinte é um axioma:

> para qualquer A, se $(\forall x \in A) \exists! y \, \varphi(x, y)$, então $\{y : (\exists x \in A) \varphi(x, y)\}$ existe.

Como na Separação, isto é um esquema: produz infinitamente muitos axiomas, um para cada uma das infinitas φ distintas. E pode igualmente ser (e normalmente é) escrito assim:

Para qualquer fórmula $\varphi(x,y)$ que não contenha "B", o seguinte é um axioma:

$$\forall A[(\forall x \in A)\exists! y\, \varphi(x,y) \to \exists B \forall y(y \in B \leftrightarrow (\exists x \in A)\varphi(x,y))]$$

À primeira vista, contudo, esta é uma fórmula bastante intrincada. A consequência imediata seguinte da Substituição dá provavelmente uma expressão *mais clara* à ideia intuitiva com que trabalhamos:[3]

Corolário 10.24. *Para qualquer termo $\tau(x)$, e qualquer conjunto A, este conjunto existe:*

$$\{\tau(x) : x \in A\} = \{y : (\exists x \in A) y = \tau(x)\}.$$

Prova. Como τ é um *termo*, $\forall x \exists! y\, \tau(x) = y$. *A fortiori*, $(\forall x \in A)\exists! y\, \tau(x) = y$. Logo $\{y : (\exists x \in A)\tau(x) = y\}$ existe pela Substituição. □

Isto sugere que "Substituição" é um bom nome para o axioma: dado um conjunto A, pode formar um novo conjunto, $\{\tau(x) : x \in A\}$, substituindo cada membro de A pela sua imagem sob τ. De fato, seguindo a notação para a imagem de um conjunto sob uma função, poderíamos escrever $\tau[A]$ para $\{\tau(x) : x \in A\}$.

O essencial, porém, é que τ é um *termo*. Não tem de ser (um nome de) uma *função*, no sentido de Seção 4.3, isto é, um certo conjunto de pares ordenados. Afinal, se f é uma função (nesse sentido), então o conjunto $f[A] = \{f(x) : x \in A\}$ é apenas um subconjunto particular de $\mathrm{Im}(f)$, e isso já está garantido a existir apenas com os axiomas de $\mathbf{Z}^-$.[4] A Substituição, em contraste,

[3] Um termo é uma expressão que designa exatamente um objeto (em qualquer instanciação das suas variáveis livres). Por exemplo, "$\emptyset$" é um termo que designa o conjunto vazio; "$\{x\}$" é um termo que designa o singular de x (seja qual for x); "$x \cup y$" é um termo que designa a união de x e de y (sejam quais forem).

[4] Baste considerar $\{y \in \bigcup\bigcup f : (\exists x \in A) y = f(x)\}$.

é um acréscimo *poderoso* aos nossos axiomas, como veremos no Capítulo 12.

10.8 $\mathbf{ZF}^-$: um marco

A questão de como justificar o Esquema de Substituição (se é que o justificamos) não é linear. Como tal, a reservamos para o Capítulo 12. No entanto, com a adição da Substituição, atingimos outro marco importante. Dispomos agora de todos os axiomas necessários para a teoria $\mathbf{ZF}^-$. Em detalhe:

Definição 10.25. A teoria $\mathbf{ZF}^-$ tem estes axiomas: Extensionalidade, União, Pares, Conjuntos das Partes, Infinito, e todas as instâncias dos esquemas de Separação e Substituição. Por outras palavras, $\mathbf{ZF}^-$ acrescenta a Substituição a $\mathbf{Z}^-$.

Isto designa a teoria de conjuntos de *Zermelo–Fraenkel* (*menos* algo que veremos mais adiante). Fraenkel recebe a honra, uma vez que lhe é atribuída a formulação da Substituição em 1922, embora a primeira formulação precisa seja devida a Skolem (1922).

10.9 Ordinais como Tipos de Ordem

Armados com a Substituição, e portanto trabalhando agora em $\mathbf{ZF}^-$, podemos finalmente provar o resultado a que vimos aspirando:

Teorema 10.26. *Toda a boa ordenação é isomorfa a um único ordinal.*

Prova. Seja $\langle A, < \rangle$ uma boa ordenação. Por Proposição 10.23, é isomorfa no máximo a um ordinal. Logo, por redução ao absurdo, suponha que $\langle A, < \rangle$ não é isomorfa a *nenhum* ordinal. Começaremos por "tornar $\langle A, < \rangle$ o mais pequena possível". Em detalhe: se algum segmento inicial próprio $\langle A_a, <_a \rangle$ não for isomorfo a nenhum ordinal, existe um $a \in A$ mínimo com essa

propriedade; então seja $B = A_a$ e $\lhd = <_a$. Caso contrário, seja $B = A$ e $\lhd = <$.

Por definição, todo o segmento inicial próprio de B é isomorfo a algum ordinal, que é único como acima. Logo, pela Substituição, o conjunto seguinte existe, e é uma função:

$$f = \{\langle \beta, b\rangle : b \in B \text{ e } \beta \cong \langle B_b, \lhd_b\rangle\}$$

Para completar a redução ao absurdo, mostraremos que f é um isomorfismo $\alpha \to B$, para algum ordinal α.

É óbvio que $\mathrm{Im}(f) = B$. E por Lema 10.11, f preserva a ordenação, isto é, $\gamma \in \beta$ sse $f(\gamma) \lhd f(\beta)$. Para mostrar que $\mathrm{dom}(f)$ é um ordinal, por Corolário 10.19 basta mostrar que $\mathrm{dom}(f)$ é transitivo. Fixe então $\beta \in \mathrm{dom}(f)$, isto é, $\beta \cong \langle B_b, \lhd_b\rangle$ para algum b. Se $\gamma \in \beta$, então $\gamma \in \mathrm{dom}(f)$ por Lema 10.10; generalizando, $\beta \subseteq \mathrm{dom}(f)$. □

Este resultado autoriza a definição seguinte, que queríamos apresentar desde Seção 10.5:

Definição 10.27. Se $\langle A, <\rangle$ é uma boa ordenação, então o seu tipo de ordem, $\mathrm{ord}(A, <)$, é o único ordinal α tal que $\langle A, <\rangle \cong \alpha$.

Além disso, esta definição autoriza dois princípios convenientes:

Corolário 10.28. *Sendo $\langle A, <\rangle$ e $\langle B, \lhd\rangle$ boas ordenações:*

$\mathrm{ord}(A, <) = \mathrm{ord}(B, \lhd)$ *sse* $\langle A, <\rangle \cong \langle B, \lhd\rangle$
$\mathrm{ord}(A, <) \in \mathrm{ord}(B, \lhd)$ *sse* $\langle A, <\rangle \cong \langle B_b, \lhd_b\rangle$ *para algum* $b \in B$

Prova. A identidade vale por Proposição 10.23. Para provar a segunda afirmação, seja $\mathrm{ord}(A, <) = \alpha$ e $\mathrm{ord}(B, \lhd) = \beta$, e seja $f : \beta \to \langle B, \lhd\rangle$ o nosso isomorfismo. Então:

$$\begin{aligned} \alpha \in \beta \text{ sse } & f\restriction_\alpha : \alpha \to B_{f(\alpha)} \text{ é um isomorfismo} \\ \text{sse } & \langle A, <\rangle \cong \langle B_{f(\alpha)}, \lhd_{f(\alpha)}\rangle \end{aligned}$$

$$\text{sse } \langle A, <\rangle \cong \langle B_b, <_b\rangle \text{ para algum } b \in B$$

por Proposição 10.7, Lema 10.10, e Corolário 10.15. □

10.10 Ordinais Sucessores e Limites

Nos próximos capítulos, usaremos muito os ordinais. Por isso convém introduzir algumas noções simples.

Definição 10.29. Para qualquer ordinal α, o seu *sucessor* é $\alpha^+ = \alpha \cup \{\alpha\}$. Dizemos que α é um ordinal *sucessor* se $\beta^+ = \alpha$ para algum ordinal β. Dizemos que α é um ordinal *limite* sse α não é nem vazio nem um ordinal sucessor.

O resultado seguinte mostra que esta é a noção certa de *sucessor*:

Proposição 10.30. *Para qualquer ordinal α:*

1. $\alpha \in \alpha^+$;
2. α^+ *é um ordinal;*
3. *não existe ordinal β tal que $\alpha \in \beta \in \alpha^+$.*

Prova. Trivialmente, $\alpha \in \alpha \cup \{\alpha\} = \alpha^+$. Igualmente, α^+ é um conjunto transitivo de ordinais, e portanto um ordinal por Corolário 10.19. E é impossível que $\alpha \in \beta \in \alpha^+$, pois então ou $\beta \in \alpha$ ou $\beta = \alpha$, contradizendo Corolário 10.18. □

Isto autoriza também uma variante de prova por indução transfinita:

Teorema 10.31 (Indução Transfinita Simples). *Seja $\varphi(x)$ uma fórmula tal que:*

1. $\varphi(\emptyset)$; *e*
2. *para qualquer ordinal α, se $\varphi(\alpha)$ então $\varphi(\alpha^+)$; e*

3. se α é um ordinal limite e $(\forall \beta \in \alpha)\varphi(\beta)$, então $\varphi(\alpha)$.

Então $\forall \alpha \varphi(\alpha)$.

Prova. Provamos a contrapositiva. Logo, suponha que existe algum ordinal para o qual vale $\neg\varphi$; seja γ o menor tal ordinal. Então ou $\gamma = \emptyset$, ou $\gamma = \alpha^+$ para algum α tal que $\varphi(\alpha)$; ou γ é um ordinal limite e $(\forall \beta \in \gamma)\varphi(\beta)$. □

Um último bocado de notação será útil mais adiante:

Definição 10.32. Se X é um conjunto de ordinais, então $\operatorname{lsub}(X) = \bigcup_{\alpha \in X} \alpha^+$.

Aqui, "lsub" abrevia "least strict upper bound" (majorante superior estrito mínimo).[5] O resultado seguinte explica isto:

Proposição 10.33. *Se X é um conjunto de ordinais,* $\operatorname{lsub}(X)$ *é o menor ordinal estritamente maior do que todo o ordinal em X.*

Prova. Seja $Y = \{\alpha^+ : \alpha \in X\}$, de modo que $\operatorname{lsub}(X) = \bigcup Y$. Como os ordinais são transitivos e todo o membro de um ordinal é um ordinal, $\operatorname{lsub}(X)$ é um conjunto transitivo de ordinais, e portanto um ordinal por Corolário 10.19.

Se $\alpha \in X$, então $\alpha^+ \in Y$, logo $\alpha^+ \subseteq \bigcup Y = \operatorname{lsub}(X)$, e portanto $\alpha \in \operatorname{lsub}(X)$. Logo $\operatorname{lsub}(X)$ é estritamente maior do que todo o ordinal em X.

Inversamente, se $\alpha \in \operatorname{lsub}(X)$, então $\alpha \in \beta^+ \in Y$ para algum $\beta \in X$, de modo que $\alpha \leq \beta \in X$. Logo $\operatorname{lsub}(X)$ é o majorante superior estrito *mínimo* de X. □

[5] Alguns livros usam "sup(X)" para isto. Mas outros usam "sup(X)" para o majorante superior *não estrito* mínimo, isto é, simplesmente $\bigcup X$. Se X tem um maior elemento, α, estas noções divergem: o majorante superior estrito mínimo é α^+, enquanto o majorante superior não estrito mínimo é apenas α.

Problemas

Problema 10.1. Seção 10.2 apresentou três ordenações exemplo sobre os números naturais. Verifique que cada uma é uma boa ordenação.

Problema 10.2. Prove Lema 10.5.

Problema 10.3. Complete o raciocínio "exatamente análogo" na prova de Teorema 10.17.

Problema 10.4. Prove que, se cada membro de X é um ordinal, então $\bigcup X$ é um ordinal.

CAPÍTULO 11

Estágios e postos

11.1 Definindo os estágios como os V_α

No Capítulo 10, definimos boas ordenações e os ordinais (de von Neumann). Neste capítulo, usaremos isto para caracterizar a própria hierarquia de conjuntos. Para tal, recordemos que na Seção 10.10 definimos a noção de ordinal sucessor e ordinal limite. Usamos estas ideias na definição seguinte:

Definição 11.1.

$$\begin{aligned} V_\emptyset &= \emptyset \\ V_{\alpha^+} &= \wp(V_\alpha) && \text{para qualquer ordinal } \alpha \\ V_\alpha &= \bigcup_{\gamma<\alpha} V_\gamma && \text{quando } \alpha \text{ é um ordinal limite} \end{aligned}$$

Será uma definição por *recursão transfinita* nos ordinais. A este respeito, convém comparar com definições recursivas de funções nos números naturais.[1] Como ao tratar dos números naturais,

[1] Cf. as definições de adição, multiplicação e exponenciação em Seção 7.2.

define-se um caso base e casos sucessor; mas ao tratar de ordinais, também precisamos de descrever o comportamento nos casos *limite*.

Esta definição dos V_α será um marco importante. Motivamos informalmente a nossa hierarquia de conjuntos como formando conjuntos por *estágios*. Os V_α são, em efeito, precisamente esses estágios. Importa, contudo, que se trata de uma caracterização *interna* dos estágios. Em vez de sugerir um possível *modelo* da teoria, teremos definido os estágios *dentro* da nossa teoria dos conjuntos.

11.2 Teorema(s) de recursão transfinita

O primeiro passo é, contudo, confirmar que Definição 11.1 é uma definição bem sucedida. De modo mais geral, precisamos de provar que qualquer tentativa de definir por recursão transfinita (transfinita) terá êxito. Esse é o objetivo deste texto.

Aviso: este material é delicado. A lição geral, no entanto, é bem simples: a indução transfinita mais o Esquema de Substituição garantem a legitimidade de (várias versões da) recursão transfinita.[2]

Definição 11.2. Seja $\tau(x)$ um termo; seja f uma função; seja α um ordinal. Dizemos que f é uma *aproximação de ordem* α para τ sse $\mathrm{dom}(f) = \alpha$ e $(\forall \beta \in \alpha) f(\beta) = \tau(f \restriction_\beta)$.

Lema 11.3 (Recursão limitada). *Para qualquer termo $\tau(x)$ e qualquer ordinal α, existe uma única aproximação de ordem α para τ.*

Prova. Mostraremos que, para qualquer $\gamma \leq \alpha$, existe uma única aproximação de ordem γ.

[2] Lembrete: todas as fórmulas e termos podem ter parâmetros (salvo indicação em contrário).

Começamos por estabelecer a unicidade. Sejam g e h (respetivamente) aproximações de ordem γ e δ. Uma indução transfinita nos seus argumentos mostra que $g(\beta) = h(\beta)$ para qualquer $\beta \in \mathrm{dom}(g) \cap \mathrm{dom}(h) = \gamma \cap \delta = \min(\gamma, \delta)$. Logo as aproximações são únicas (se existem) e coincidem em todos os valores.

Para estabelecer a existência, usamos agora uma indução transfinita simples (Teorema 10.31) nos ordinais $\delta \leq \alpha$.

A função vazia é trivialmente uma aproximação de ordem $\emptyset$.

Se g é uma aproximação de ordem γ, então $g \cup \{\langle \gamma, \tau(g)\rangle\}$ é uma aproximação de ordem γ^+.

Se γ é um ordinal limite e g_δ é uma aproximação de ordem δ para todo o $\delta < \gamma$, seja $g = \bigcup_{\delta \in \gamma} g_\delta$. Isto é uma função, pois as várias g_δ coincidem em todos os valores. E se $\delta \in \gamma$ então $g(\delta) = g_{\delta^+}(\delta) = \tau(g_{\delta^+} \restriction_\delta) = \tau(g \restriction_\delta)$.

Isto completa a prova por indução transfinita. □

Se nos permitirmos definir um *termo* em vez de uma função, podemos retirar a cota α do resultado anterior. No enunciado e na prova do resultado seguinte, quando σ é um termo, definimos $\sigma \restriction_\alpha = \{\langle \beta, \sigma(\beta)\rangle : \beta \in \alpha\}$.

Teorema 11.4 (Recursão geral). *Para qualquer termo $\tau(x)$, podemos definir explicitamente um termo $\sigma(x)$ tal que $\sigma(\alpha) = \tau(\sigma \restriction_\alpha)$ para qualquer ordinal α.*

Prova. Para cada α, por Lema 11.3 existe uma única aproximação de ordem α, f_α, para τ. Definimos $\sigma(\alpha)$ como $f_{\alpha^+}(\alpha)$. Ora:

$$\begin{aligned}
\sigma(\alpha) &= f_{\alpha^+}(\alpha) \\
&= \tau(f_{\alpha^+} \restriction_\alpha) \\
&= \tau(\{\langle \beta, f_{\alpha^+}(\beta)\rangle : \beta \in \alpha\}) \\
&= \tau(\{\langle \beta, f_{\beta^+}(\beta)\rangle : \beta \in \alpha\}) \\
&= \tau(\sigma \restriction_\alpha)
\end{aligned}$$

notando que $f_{\beta^+}(\beta) = f_{\alpha^+}(\beta)$ para todo o $\beta < \alpha$, como no Lema 11.3. □

Notemos que Teorema 11.4 é um *esquema*. Crucialmente, não podemos esperar que σ defina uma função, isto é, um certo tipo de *conjunto*, pois então $\mathrm{dom}(\sigma)$ seria o conjunto de todos os ordinais, contradizendo o paradoxo de Burali-Forti (Teorema 10.20).

Ainda falta mostrar, contudo, que Teorema 11.4 justifica a nossa definição dos V_α. Isto pode não ser óbvio de imediato; mas tornar-se-á claro com uma última versão simples da recursão transfinita.

Teorema 11.5 (Recursão simples). *Para quaisquer termos $\tau(x)$ e $\theta(x)$ e qualquer conjunto A, podemos definir explicitamente um termo $\sigma(x)$ tal que:*

$$\begin{aligned} \sigma(\emptyset) &= A \\ \sigma(\alpha^+) &= \tau(\sigma(\alpha)) && \textit{para qualquer ordinal } \alpha \\ \sigma(\alpha) &= \theta(\mathrm{Im}(\sigma\restriction_\alpha)) && \textit{quando } \alpha \textit{ é um ordinal limite} \end{aligned}$$

Prova. Começamos por definir um termo, $\xi(x)$, do seguinte modo:

$$\xi(x) = \begin{cases} A & \text{se } x \text{ não é uma função cujo} \\ & \quad \text{domínio é um ordinal; caso contrário:} \\ \tau(x(\alpha)) & \text{se } \mathrm{dom}(x) = \alpha^+ \\ \theta(\mathrm{Im}(x)) & \text{se } \mathrm{dom}(x) \text{ é um ordinal limite} \end{cases}$$

Por Teorema 11.4, existe um termo $\sigma(x)$ tal que $\sigma(\alpha) = \xi(\sigma\restriction_\alpha)$ para todo o ordinal α; além disso, $\sigma\restriction_\alpha$ é uma função com domínio α. Mostramos que σ tem as propriedades pretendidas, por indução transfinita simples (Teorema 10.31).

Primeiro, $\sigma(\emptyset) = \xi(\emptyset) = A$.

Segundo, $\sigma(\alpha^+) = \xi(\sigma\restriction_{\alpha^+}) = \tau(\sigma\restriction_{\alpha^+}(\alpha)) = \tau(\sigma(\alpha))$.

Por fim, $\sigma(\alpha) = \xi(\sigma\restriction_\alpha) = \theta(\mathrm{Im}(\sigma\restriction_\alpha))$, quando α é limite. □

Ora, para justificar Definição 11.1, basta tomar $A = \emptyset$, $\tau(x) = \wp(x)$ e $\theta(x) = \bigcup x$. Por fim, isto justifica a definição dos V_α!

11.3 Propriedades básicas dos estágios

Para evidenciar a importância fundacional da definição dos V_α, apresentaremos alguns resultados básicos sobre eles. Começamos com uma definição:[3]

Definição 11.6. O conjunto A é *potente* sse $\forall x((\exists y \in A)x \subseteq y \to x \in A)$.

Lema 11.7. *Para cada ordinal α:*

1. *Cada V_α é transitivo.*
2. *Cada V_α é potente.*
3. *Se $\gamma \in \alpha$, então $V_\gamma \in V_\alpha$ (e portanto também $V_\gamma \subseteq V_\alpha$, por (1))*

Prova. Provamos isto por indução transfinita (simultânea). Para a indução, supomos que (1)–(3) valem para cada ordinal $\beta < \alpha$.

O caso $\alpha = \emptyset$ é trivial.

Suponhamos $\alpha = \beta^+$. Para mostrar (3), se $\gamma \in \alpha$ então $V_\gamma \subseteq V_\beta$ por hipótese, logo $V_\gamma \in \wp(V_\beta) = V_\alpha$. Para mostrar (2), suponhamos $A \subseteq B \in V_\alpha$, isto é, $A \subseteq B \subseteq V_\beta$; então $A \subseteq V_\beta$, logo $A \in V_\alpha$. Para mostrar (1), notemos que se $x \in A \in V_\alpha$ temos $A \subseteq V_\beta$, logo $x \in V_\beta$, logo $x \subseteq V_\beta$ pois V_β é transitivo por hipótese, e assim $x \in V_\alpha$.

Suponhamos que α é um ordinal limite. Para mostrar (3), se $\gamma \in \alpha$ então $\gamma \in \gamma^+ \in \alpha$, de modo que $V_\gamma \in V_{\gamma^+}$ por hipótese, logo $V_\gamma \in \bigcup_{\beta\in\alpha} V_\beta = V_\alpha$. Para mostrar (1) e (2), basta observar que uma união de conjuntos transitivos (respetivamente, potentes) é transitiva (respetivamente, potente). □

[3] Não há terminologia padrão para «potente»; este é o nome usado por Button (2021).

Lema 11.8. *Para cada ordinal α, $V_\alpha \notin V_\alpha$.*

Prova. Por indução transfinita. Evidentemente $V_\emptyset \notin V_\emptyset$.

Se $V_{\alpha^+} \in V_{\alpha^+} = \wp(V_\alpha)$, então $V_{\alpha^+} \subseteq V_\alpha$; e como $V_\alpha \in V_{\alpha^+}$ por Lema 11.7, temos $V_\alpha \in V_\alpha$. Reciprocamente: se $V_\alpha \notin V_\alpha$ então $V_{\alpha^+} \notin V_{\alpha^+}$

Se α é limite e $V_\alpha \in V_\alpha = \bigcup_{\beta\in\alpha} V_\beta$, então $V_\alpha \in V_\beta$ para algum $\beta \in \alpha$; mas então também $V_\beta \in V_\alpha$, logo $V_\beta \in V_\beta$ por Lema 11.7 (duas vezes). Reciprocamente, se $V_\beta \notin V_\beta$ para todo o $\beta \in \alpha$, então $V_\alpha \notin V_\alpha$. □

Corolário 11.9. *Para quaisquer ordinais α, β: $\alpha \in \beta$ sse $V_\alpha \in V_\beta$*

Prova. Lema 11.7 dá um sentido. Reciprocamente, suponhamos $V_\alpha \in V_\beta$. Então $\alpha \neq \beta$ por Lema 11.8; e $\beta \notin \alpha$, pois caso contrário teríamos $V_\beta \in V_\alpha$ e logo $V_\beta \in V_\beta$ por Lema 11.7 (duas vezes), contradizendo Lema 11.8. Logo $\alpha \in \beta$ por tricotomia. □

Tudo isto permite pensar em cada V_α como o α-ésimo estágio da hierarquia. Eis a razão.

Certamente os nossos V_α podem ser vistos como formados num processo *iterativo*, pois o uso de ordinais acompanha a noção de iteração. Além disso, se um estágio é formado antes do outro, isto é, $V_\beta \in V_\alpha$, isto é, $\beta \in \alpha$, então o nosso processo de formação é *cumulativo*, pois $V_\beta \subseteq V_\alpha$. Finalmente, estamos de fato a formar *todas* as coleções possíveis de conjuntos que estavam disponíveis em qualquer estágio anterior, pois qualquer estágio sucessor V_{α^+} é o conjunto potência do seu predecessor V_α.

Em resumo: com $\mathbf{ZF}^-$, estamos *quase* a terminar de articular a nossa visão da hierarquia cumulativa e iterativa de conjuntos. (Embora, naturalmente, ainda precisemos de justificar o Esquema de Substituição.)

11.4 Fundação

Estamos apenas *quase* concluídos—e não *totalmente* concluídos—porque nada em $\mathbf{ZF}^-$ garante que *todo* o conjunto está em algum V_α, isto é, que todo o conjunto é formado em algum estágio.

Há, contudo, um sentido (matemático) bastante direto em que não *importa* se existem conjuntos fora da hierarquia. (Se os houver, podemos simplesmente ignorá-los.) Mas motivamos o nosso *conceito* de conjunto com a ideia de que todo o conjunto é formado em algum estágio (ver *Estágios-são-chave* na Seção 9.1). Logo queremos excluir a possibilidade de conjuntos que caem fora da hierarquia. Por conseguinte, devemos acrescentar um novo axioma, que assegura que todo o conjunto surge em algum lugar da hierarquia.

Uma vez que os V_α são os nossos estágios, poderíamos simplesmente considerar acrescentar o seguinte como axioma:

Regularidade. $\forall A \exists \alpha\ A \subseteq V_\alpha$

Seria uma abordagem perfeitamente razoável. Contudo, por razões que se explicarão adiante, adotaremos antes um axioma alternativo:

Fundação (Fundação). $(\forall A \neq \emptyset)(\exists B \in A) A \cap B = \emptyset$.

Com algum trabalho, podemos mostrar (em $\mathbf{ZF}^-$) que a Fundação implica a Regularidade:

Definição 11.10. Para cada conjunto A, definimos:

$$\begin{aligned} \mathrm{cl}_0(A) &= A, \\ \mathrm{cl}_{n+1}(A) &= \bigcup \mathrm{cl}_n(A), \\ \mathrm{trcl}(A) &= \bigcup_{n<\omega} \mathrm{cl}_n(A). \end{aligned}$$

Chamamos a $\operatorname{trcl}(A)$ o *fecho transitivo* de A.

O nome «fecho transitivo» é adequado:

Proposição 11.11. *$A \subseteq \operatorname{trcl}(A)$ e $\operatorname{trcl}(A)$ é um conjunto transitivo.*

Prova. Evidentemente $A = \operatorname{cl}_0(A) \subseteq \operatorname{trcl}(A)$. E se $x \in b \in \operatorname{trcl}(A)$, então $b \in \operatorname{cl}_n(A)$ para algum n, logo $x \in \operatorname{cl}_{n+1}(A) \subseteq \operatorname{trcl}(A)$. □

Lema 11.12. *Se A é um conjunto transitivo, então existe algum α tal que $A \subseteq V_\alpha$.*

Prova. Recordando a definição de «$\operatorname{lsub}(X)$» em Definição 10.32, definimos dois conjuntos:

$$D = \{x \in A : \forall \delta\ x \not\subseteq V_\delta\}$$
$$\alpha = \operatorname{lsub}\{\delta : (\exists x \in A)(x \subseteq V_\delta \wedge (\forall \gamma \in \delta) x \not\subseteq V_\gamma)\}$$

Suponhamos $D = \emptyset$. Se $x \in A$, então existe algum δ tal que $x \subseteq V_\delta$ e, pela boa ordenação dos ordinais, $(\forall \gamma \in \delta) x \not\subseteq V_\gamma$; logo $\delta \in \alpha$ e assim $x \in V_\alpha$ por Lema 11.7. Logo $A \subseteq V_\alpha$, como pretendido.

Basta pois mostrar que $D = \emptyset$. Por redução ao absurdo, suponhamos o contrário. Pela Fundação, existe algum $B \in D \subseteq A$ tal que $D \cap B = \emptyset$. Se $x \in B$ então $x \in A$, pois A é transitivo, e como $x \notin D$, segue que $\exists \delta\ x \subseteq V_\delta$. Seja agora

$$\beta = \operatorname{lsub}\{\delta : (\exists x \in B)(x \subseteq V_\delta \wedge (\forall \gamma < \delta) x \not\subseteq V_\gamma)\}.$$

Como antes, $B \subseteq V_\beta$, contradizendo a afirmação de que $B \in D$.□

Teorema 11.13. *A Regularidade vale.*

Prova. Fixemos A; agora $A \subseteq \operatorname{trcl}(A)$ por Proposição 11.11, que é transitivo. Logo existe algum α tal que $A \subseteq \operatorname{trcl}(A) \subseteq V_\alpha$ por Lema 11.12 □

Estes resultados mostram que $\mathbf{ZF}^-$ prova a implicação *Fundação* $\Rightarrow$ *Regularidade*. Em Proposição 11.22, mostraremos que $\mathbf{ZF}^-$ prova *Regularidade* $\Rightarrow$ *Fundação*. Como tal, a Fundação e a Regularidade são *equivalentes* (módulo $\mathbf{ZF}^-$). Mas isto significa que, dado $\mathbf{ZF}^-$, podemos justificar a Fundação notando que é equivalente à Regularidade. E podemos justificar a Regularidade de imediato com base em *Estágios-são-chave*.

11.5 Z e ZF: um marco

Com a Fundação, atingimos outro marco importante. Consideramos as teorias $\mathbf{Z}^-$ e $\mathbf{ZF}^-$, que dissemos ser certas teorias «menos» alguma coisa. Essa coisa é a Fundação. Logo:

Definição 11.14. A teoria **Z** acrescenta a Fundação a $\mathbf{Z}^-$. Os seus axiomas são Extensionalidade, União, Pares, Conjuntos das Partes, Infinito, Fundação, e todas as instâncias do esquema de Separação.

A teoria **ZF** acrescenta a Fundação a $\mathbf{ZF}^-$. Por outras palavras, **ZF** acrescenta todas as instâncias de Substituição a **Z**.

Ainda assim, uma pergunta pode ter surgido. Se a Regularidade é equivalente sobre $\mathbf{ZF}^-$ à Fundação, e a justificação da Regularidade é clara, por que dar voltas e tomar a Fundação como axioma básico, em vez da Regularidade?

Deixando de lado razões históricas (quem formulou o quê e quando), a razão essencial é que a Fundação pode ser apresentada sem recorrer à definição dos V_α. Essa definição dependeu de todo o trabalho na Seção 11.2: foi preciso provar a recursão transfinita, para mostrar que era legítima. Mas a nossa prova da recursão transfinita usou o *Esquema de Substituição*. Logo, embora a Fundação e a Regularidade sejam equivalentes módulo $\mathbf{ZF}^-$, não são equivalentes módulo $\mathbf{Z}^-$.

De fato, a questão é mais grave do que esta observação simples sugere. Embora ultrapasse o âmbito deste livro, verifica-se

que tanto $\mathbf{Z}^-$ como $\mathbf{Z}$ são demasiado fracas para definir os V_α. Logo, se se trabalha apenas em $\mathbf{Z}$, então a Regularidade (tal como a formulamos) nem sequer faz *sentido.* É por isso que o nosso axioma oficial é a Fundação, e não a Regularidade.

Daqui em diante, trabalharemos em **ZF** (salvo indicação em contrário), sem mais comentários.

11.6 Posto

Tendo definido os estágios como os V_α e sabendo que todo o conjunto é subconjunto de algum estágio, podemos definir o *posto* de um conjunto. Intuitivamente, o posto de A é o primeiro momento em que A é formado. Com mais precisão:

Definição 11.15. Para cada conjunto A, $\text{posto}(A)$ é o menor ordinal α tal que $A \subseteq V_\alpha$.

Proposição 11.16. $\text{posto}(A)$ *existe, para qualquer* A.

Prova. Fica como exercício. □

A boa ordenação dos postos permite provar alguns resultados importantes:

Proposição 11.17. *Para qualquer ordinal* α, $V_\alpha = \{x : \text{posto}(x) \in \alpha\}$.

Prova. Se $\text{posto}(x) \in \alpha$ então $x \subseteq V_{\text{posto}(x)} \in V_\alpha$, logo $x \in V_\alpha$ pois V_α é potente (invocando Lema 11.7 várias vezes). Reciprocamente, se $x \in V_\alpha$ então $x \subseteq V_\alpha$, logo $\text{posto}(x) \leq \alpha$; agora uma indução transfinita simples mostra que $x \notin V_\alpha$. □

Proposição 11.18. *Se $B \in A$, então* $\mathrm{posto}(B) \in \mathrm{posto}(A)$.

Prova. $A \subseteq V_{\mathrm{posto}(A)} = \{x : \mathrm{posto}(x) \in \mathrm{posto}(A)\}$ por Proposição 11.17. □

Com este fato, podemos estabelecer um resultado que nos permite provar afirmações sobre *todos os conjuntos* por uma forma de indução:

Teorema 11.19 (Esquema de indução-$\in$). *Para qualquer fórmula φ:*

$$\forall A((\forall x \in A)\varphi(x) \to \varphi(A)) \to \forall A \varphi(A).$$

Prova. Provaremos a contrapositiva. Suponhamos, pois, $\neg\forall A\varphi(A)$. Pela indução transfinita (Teorema 10.16), existe algum não-φ de posto mínimo possível; isto é, algum A tal que $\neg\varphi(A)$ e $\forall x(\mathrm{posto}(x) \in \mathrm{posto}(A) \to \varphi(x))$. Ora, se $x \in A$ então $\mathrm{posto}(x) \in \mathrm{posto}(A)$, por Proposição 11.18, logo $\varphi(x)$; isto é, $(\forall x \in A)\varphi(x) \wedge \neg\varphi(A)$. □

Eis uma forma informal de interpretar este resultado poderoso. Digamos que φ é *hereditária* sse sempre que todo o membro de um conjunto é φ, o próprio conjunto é φ. Então a indução-$\in$ diz o seguinte: se φ é hereditária, todo o conjunto é φ.

Para concluir a discussão dos postos (por agora), provaremos algumas afirmações que já antecipamos várias vezes.

Proposição 11.20. $\mathrm{posto}(A) = \mathrm{lsub}_{x \in A} \mathrm{posto}(x)$.

Prova. Seja $\alpha = \mathrm{lsub}_{x \in A} \mathrm{posto}(x)$. Por Proposição 11.18, $\alpha \leq \mathrm{posto}(A)$. Mas se $x \in A$ então $\mathrm{posto}(x) \in \alpha$, logo $x \in V_\alpha$ por Proposição 11.17, e portanto $A \subseteq V_\alpha$, isto é, $\mathrm{posto}(A) \leq \alpha$. Logo $\mathrm{posto}(A) = \alpha$. □

Corolário 11.21. *Para qualquer ordinal* α, $\text{posto}(\alpha) = \alpha$.

Prova. Suponhamos, por indução transfinita, que $\text{posto}(\beta) = \beta$ para todo o $\beta \in \alpha$. Ora, $\text{posto}(\alpha) = \text{lsub}_{\beta\in\alpha}\,\text{posto}(\beta) = \text{lsub}_{\beta\in\alpha}\,\beta = \alpha$ por Proposição 11.20. □

Finalmente, eis uma prova rápida do resultado prometido no fim de Seção 11.4, a saber que $\mathbf{ZF}^-$ prova a implicação *Regularidade* $\Rightarrow$ *Fundação.* (Notemos que a noção de «posto» e Proposição 11.18 estão disponíveis para uso nesta prova pois—como indicamos no início deste texto—podem ser apresentadas usando $\mathbf{ZF}^-$ + Regularidade.)

Proposição 11.22 (trabalhando em $\mathbf{ZF}^-$ + Regularidade). *A Fundação vale.*

Prova. Fixemos $A \neq \emptyset$, e algum $B \in A$ de posto mínimo possível. Se $c \in B$ então $\text{posto}(c) \in \text{posto}(B)$ por Proposição 11.18, logo $c \notin A$ pela escolha de B. □

Problemas

Problema 11.1. Prove Proposição 11.16.

Problema 11.2. Complete a indução transfinita simples mencionada na Proposição 11.17.

CAPÍTULO 12

Substituição

12.1 Introdução

A Substituição é o esquema axiomático que distingue **ZF** de **Z**. Recorremos a ele ao longo de Capítulos 10 a 11. Neste capítulo, abordaremos finalmente a questão: a Substituição está justificada?

Para tornar a questão precisa, convém observar que a Substituição é realmente bastante *forte*. Teremos uma ideia de quão forte é ao longo deste capítulo (e de novo na Seção 15.5). Isto sugere que uma justificação é de fato necessária.

Discutiremos dois tipos de justificação. Em traços largos: uma justificação *extrínseca* é uma tentativa de justificar um axioma pelos seus frutos; uma justificação *intrínseca* é uma tentativa de justificar um axioma sugerindo que ele é vindicado pelos conceitos matemáticos em causa. Teremos uma ideia mais clara do que isto significa ao longo deste capítulo, mas não passa da ponta de um icebergue. Para mais pormenores, veja em particular Maddy (1988a e 1988b).

12.2 A Força da Substituição

Começamos com uma observação simples sobre a força da Substituição: a menos que ultrapassemos **Z**, não podemos provar a

existência de qualquer ordinal de von Neumann maior ou igual a $\omega + \omega$.

Eis um esboço do porquê. Trabalhando em **ZF**, consideremos o conjunto $V_{\omega+\omega}$. Este conjunto serve como domínio de um *modelo* para **Z**. Para ver isto, introduzimos alguma notação para a *relativização* de uma fórmula:

Definição 12.1. Para qualquer conjunto M, e qualquer fórmula φ, seja φ^M a fórmula que resulta de restringir todos os quantificadores de φ a M. Isto é, substituir "$\exists x$" por "$(\exists x \in M)$", e substituir "$\forall x$" por "$(\forall x \in M)$".

Pode mostrar-se que, para todo o axioma φ de **Z**, tem-se $\mathbf{ZF} \vdash \varphi^{V_{\omega+\omega}}$. Mas $\omega + \omega$ não está *em* $V_{\omega+\omega}$, por Corolário 11.21. Logo **Z** é consistente com a não existência de $\omega + \omega$.

Foi por isso que dissemos, na Seção 10.7, que Teorema 10.26 não pode ser provado sem Substituição. Pois é fácil, dentro de **Z**, definir uma boa ordenação explícita que intuitivamente *deveria* ter tipo de ordem $\omega + \omega$. De fato, demos um exemplo informal disto em Seção 10.2, quando apresentamos a ordenação nos números naturais dada por:

$$\begin{aligned} n \lessdot m \text{ sse ou } & n < m \text{ e } m - n \text{ é par,} \\ & \text{ou } n \text{ é par e } m \text{ é ímpar.} \end{aligned}$$

Mas se $\omega + \omega$ não existe, esta boa ordenação não é isomorfa a nenhum ordinal. Logo **Z** *não* prova Teorema 10.26.

Invertendo o prisma: a Substituição permite provar a existência de $\omega + \omega$, e portanto tem de permitir provar a existência de $V_{\omega+\omega}$. E não só. Para *qualquer* boa ordenação que possamos definir, Teorema 10.26 nos diz que existe algum α isomorfo com essa boa ordenação, e portanto que V_α existe. De forma direta, então, a Substituição garante que a hierarquia de conjuntos tem de ser *muito alta*.

Nas próximas partes do capítulo, e de novo em Seção 15.5, teremos uma ideia mais clara de quão *alta* a Substituição força a

hierarquia a ser. O ponto simples, por agora, é que a Substituição *precisa* mesmo de justificação!

12.3 Considerações Extrínsecas sobre a Substituição

Começamos por considerar uma tentativa *extrínseca* de justificar a Substituição. Boolos sugere uma, como segue.

> [...] a razão para adotar os axiomas da substituição é bastante simples: têm muitas consequências desejáveis e (aparentemente) nenhuma indesejável. Para além de teoremas sobre a concepção iterativa, as consequências incluem uma teoria satisfatória, embora não ideal, dos números infinitos, e um resultado altamente desejável que justifica definições indutivas em relações bem fundadas. (Boolos, 1971, 229)

A ideia central de Boolos é que devemos justificar a Substituição pelos seus frutos. E os frutos específicos que menciona são os que temos discutido nos últimos capítulos. A Substituição permitiu-nos provar que os ordinais de von Neumann eram excelentes sucedâneos da ideia de um tipo de boa ordenação (isto é a nossa "teoria satisfatória, embora não ideal, dos números infinitos"). A Substituição permitiu-nos também definir os V_α, estabelecer a noção de posto e provar a indução-$\in$ (isto corresponde aos nossos "teoremas sobre a concepção iterativa"). Por fim, a Substituição permite provar o teorema de recursão transfinita (isto são as "definições indutivas em relações bem fundadas").

São, de fato, consequências desejáveis. Mas estas consequências desejáveis bastam para *justificar* a Substituição? *Não.* Ou pelo menos, não de forma direta.

Eis um problema simples. Embora tenhamos enunciado algumas consequências desejáveis da Substituição, muitas delas poderiam ter sido obtidas por outras vias. Isto não é tão conhecido como devia, pelo que convém explicar a situação.

Existe uma teoria simples de conjuntos, a Teoria dos Níveis, ou **LT** abreviadamente.[1] Os axiomas de **LT** são apenas Extensionalidade, Separação e a afirmação de que todo o conjunto é subconjunto de algum *nível*, onde "nível" é definido de modo astuto para que os níveis se comportem como os nossos amigos, os V_α. Logo **ZF** prova **LT**; mas **LT** é *muito* mais fraca que **ZF**. De fato, **LT** não dá Pares, Conjuntos potencia, Infinito nem Substituição. Seja **Zr** o resultado de acrescentar Infinito e Conjuntos potencia a **LT**; isto fornece também Pares, logo **Zr** é pelo menos tão forte quanto **Z**. Mas, de fato, **Zr** é estritamente mais forte que **Z**, pois acrescenta a afirmação de que todo o conjunto tem posto (daí a sugestão de o chamar **Zr**). De fato, **Zr** fornece: uma teoria perfeitamente satisfatória dos ordinais; resultados que estratificam a hierarquia em estágios bem ordenados; uma prova da indução-$\in$; e uma *versão* da recursão transfinita.

Em resumo: embora Boolos não soubesse disto, todas as consequências desejáveis que menciona poderiam ter sido obtidas *sem* Substituição; bastaria usar **Zr** em vez de **Z**.

(Diante disto, por que seguimos a rota convencional, ensinando-lhe **ZF** em vez de **LT** e **Zr**? Há duas razões. Primeiro: por razões puramente históricas, começar por **LT** é pouco standard; queríamos dotá-lo de ferramentas para ler discussões mais standard de teoria de conjuntos. Segundo: quando estiver preparado para apreciar **LT** e **Zr**, pode simplesmente ler Potter 2004 e Button 2021.)

Naturalmente, como **Zr** é estritamente mais fraca que **ZF**, há resultados que **ZF** prova e que **Zr** deixa em aberto. Logo alguém poderia tentar justificar a Substituição em bases extrínsecas apontando para um desses resultados. Mas, uma vez que saiba usar **Zr**, é bastante difícil encontrar muitos exemplos de coisas que (a) ficam decididas pela Substituição mas não de outro modo, e (b) são intuitivamente verdadeiras. (Para mais sobre

[1] As primeiras versões de **LT** são apresentadas por Montague (1965) e Scott (1974); foi simplificada e recebeu tratamento de livro por Potter (2004); e Button (2021) simplificou recentemente ainda mais **LT**.

isto, veja Potter 2004, §13.2.)

A conclusão é esta. Para fornecer uma justificação extrínseca convincente da Substituição, seria preciso encontrar um resultado que *não* possa ser obtido sem Substituição. E isso não é empresa fácil.

Consideremos um problema adicional que surge para qualquer tentativa de oferecer uma justificação puramente extrínseca da Substituição. (Este problema é talvez mais fundamental que o primeiro.) Boolos não se limita a observar que a Substituição tem muitas consequências desejáveis. Afirma também que a Substituição "(aparentemente) não tem" consequências indesejáveis. Mas esta ressalva entre parênteses, "aparentemente", é seguramente absolutamente crucial.

Relembre como chegamos aqui: a Compreensão ingênua encontrou inconsistência, e respondemos a essa inconsistência abraçando a concepção cumulativa-iterativa de conjunto. Esta concepção vem equipada com uma narrativa que, esperamos, nos assegura a sua consistência. Mas se não pudermos justificar a Substituição dentro dessa narrativa, então não temos (ainda) razão para crer que **ZF** é consistente. Ou melhor: não temos razão para crer que **ZF** é consistente, para além do fato (talvez meramente contingente) de que ninguém descobriu ainda uma contradição. Exatamente nesse sentido, o comentário de Boolos parece reduzir-se a isto: "(aparentemente) **ZF** é consistente". Deveríamos exigir maior garantia de consistência do que isto.

Esta questão afecta qualquer tentativa *puramente* extrínseca de justificar a Substituição, isto é, qualquer justificação formulada unicamente em termos das consequências (conhecidas) de **ZF**. Como tal, convém procurar uma justificação *intrínseca* da Substituição, isto é, uma justificação que sugira que a narrativa que contamos sobre conjuntos de algum modo "já" nos compromete com a Substituição.

12.4 Limitação de Tamanho

Talvez a tentativa mais comum de oferecer uma justificação "intrínseca" da Substituição passe pela seguinte noção:

> *Limitação-de-tamanho.* Quaisquer coisas formam um conjunto, desde que não haja demasiadas.

Este princípio vindicará de imediato a Substituição. Afinal, qualquer conjunto formado por Substituição não pode ser maior do que o conjunto a partir do qual foi formado. Enunciado com precisão: suponha que forma um conjunto $\tau[A] = \{\tau(x) : x \in A\}$ usando Substituição; então $\tau[A] \preceq A$; logo, se os membroos de A não eram demasiado numerosos para formar um conjunto, as suas imagens não são demasiado numerosas para formar $\tau[A]$.

A dificuldade óbvia em invocar *Limitação-de-tamanho* para justificar a Substituição é que *ainda* não estabelecemos qualquer princípio como *Limitação-de-tamanho.* Além disso, quando contamos a nossa história sobre a concepção cumulativa-iterativa de conjunto em Capítulos 8 a 9, nada *sugeriu* a direção de *Limitação-de-tamanho.* Isto é, de fato, precisamente a razão pela qual Boolos escreveu num dado momento: "Talvez se possa concluir que há pelo menos dois pensamentos 'por detrás' da teoria de conjuntos" (1989, p. 19). Por um lado, as ideias em torno da concepção cumulativa-iterativa de conjunto destinam-se a vindicar **Z**. Por outro, *Limitação-de-tamanho* visa vindicar a Substituição.

Mas a questão não é apenas termos estado até aqui *em silêncio* sobre *Limitação-de-tamanho.* Antes, a questão é que *Limitação-de-tamanho* (tal como acabamos de o formular) parece assentar mal com a noção cumulativa-iterativa de conjunto. Afinal, não menciona nada sobre a ideia de conjuntos formados em *estágios.*

Isto não surpreende muito, dada a história destes "dois pensamentos" (isto é, a concepção cumulativa-iterativa de conjunto, e *Limitação-de-tamanho*). Estes "dois pensamentos" acabam por ser dois projetos bastante diferentes para bloquear os paradoxos da teoria de conjuntos. A noção cumulativa-iterativa de conjunto bloqueia o paradoxo de Russell dizendo, em traços largos: *nunca*

deveríamos ter esperado que existisse um conjunto de Russell, porque não seria "formado" em nenhum estágio. Por contraste, *Limitação-de-tamanho* visa excluir o conjunto de Russell, dizendo, em traços largos: *nunca deveríamos ter esperado que existisse um conjunto de Russell, porque seria demasiado grande.*

Posto assim, sejamos diretos: considerada como resposta aos paradoxos, *Limitação-de-tamanho* carece de *muito* mais justificação. Considere, por exemplo, esta versão do paradoxo de Russell: *nenhum pug fareja exatamente os pugs que não se cheiram a si próprios* (ver Seção 8.2). Se perguntar "por que não há tal pug?", não é boa resposta dizer-lhe que tal pug teria de cheirar demasiados pugs. Logo por que seria boa explicação intuitiva, da não existência de um conjunto de Russell, que teria de ser "demasiado grande" para existir?

Em suma, é perdoável estar um pouco perplexo quanto à motivação "intuitiva" para *Limitação-de-tamanho.*

12.5 Substituição e "Infinito Absoluto"

Deixamos agora *Limitação-de-tamanho* para trás e exploramos uma outra família de tentativas (intrínsecas) de justificar a Substituição, que levam a sério a ideia de conjuntos formados em estágios.

Quando delineamos pela primeira vez o processo iterativo, oferecemos alguns princípios que explicam o que acontece em cada estágio. Foram *Estágios-são-chave*, *Estágios-são-ordenados*, e *Estágios-acumulam.* Mais tarde, acrescentamos alguns princípios que nos dizem algo sobre o número de estágios: *Estágios-prosseguem* disse-nos que o processo de formação de conjuntos nunca termina, e *Estágios-atingem-o-infinito* disse-nos que o processo passa por um estágio infinito-ésimo.

É razoável sugerir que estes dois últimos princípios decorrem de algum princípio mais amplo, como:

> *Estágios-são-inesgotáveis.* Há absolutamente infinitamente muitos estágios; a hierarquia é tão alta quanto possível.

Obviamente isto é um princípio informal. Mas mesmo que não seja imediatamente *implicado* pela concepção cumulativa-iterativa de conjunto, certamente parece *consonante* com ela. Pelo menos, e ao contrário de *Limitação-de-tamanho*, retém a ideia de que os conjuntos são formados estágio a estágio.

A esperança, agora, é alavancar *Estágios-são-inesgotáveis* numa justificação da Substituição. Vejamos como isso poderia fazer-se.

Na Seção 10.2, vimos que é fácil construir uma boa ordenação que (moralmente) deveria ser isomorfa a $\omega+\omega$. Dito de outro modo, podemos facilmente imaginar um processo iterativo estágio a estágio, cujo tipo de ordem (moralmente) é $\omega + \omega$. Como tal, se aceitamos *Estágios-são-inesgotáveis*, então deveríamos certamente aceitar que há pelo menos um estágio $\omega + \omega$-ésimo da hierarquia, isto é, $V_{\omega+\omega}$, pois a hierarquia seguramente *poderia* continuar até aqui.

Este pensamento generaliza-se como segue: para qualquer boa ordenação, o processo de construir a hierarquia iterativa deveria correr pelo menos tão longe quanto essa boa ordenação. E poderíamos garantir isto tratando Teorema 10.26 como um *axioma*. Isto diria-nos que toda a boa ordenação é isomorfa a um ordinal de von Neumann. Como cada ordinal de von Neumann será igual ao seu próprio posto, Teorema 10.26 dir-nos-á então que, sempre que possamos descrever uma boa ordenação na nossa teoria de conjuntos, o processo iterativo de formação de conjuntos tem de ultrapassar essa boa ordenação.

Esta ideia parece certamente um corolário de *Estágios-são-inesgotáveis*. Infelizmente, se o nosso objetivo é extrair a Substituição desta ideia, enfrentamos um obstáculo simples e técnico: a Substituição é estritamente mais forte que Teorema 10.26. (Esta observação é feita por Potter (2004, §13.2); prová-lo-emos na Seção 12.8.)

A conclusão é que, se vamos entender *Estágios-são-inesgotáveis* de modo a obter a Substituição, então não pode *meramente* dizer que a hierarquia ultrapassa qualquer boa ordenação. Tem de fazer uma afirmação mais forte. Com este fim, Shoenfield (1977) propôs um reforço muito natural da ideia, como segue: a hie-

rarquia não é *cofinal* com nenhum conjunto.[2] Com um pouco mais de pormenor: se τ é um mapeamento que envia conjuntos para estágios da hierarquia, a imagem de qualquer conjunto A sob τ não esgota a hierarquia. Dito de outro modo (esquematicamente):

> *Estágios-são-supercofinais.* Se A é um conjunto e $\tau(x)$ é um estágio para todo o $x \in A$, então existe um estágio que vem depois de cada $\tau(x)$ para $x \in A$.

É óbvio que **ZF** prova uma versão adequadamente formalizada de *Estágios-são-supercofinais.* Reciprocamente, podemos argumentar informalmente que *Estágios-são-supercofinais* justifica a Substituição.[3] De fato, suponha $(\forall x \in A)\exists! y\, \varphi(x,y)$. Então, para cada $x \in A$, seja $\sigma(x)$ o y tal que $\varphi(x,y)$, e seja $\tau(x)$ o estágio em que $\sigma(x)$ é formado pela primeira vez. Por *Estágios-são-supercofinais*, existe um estágio V tal que $(\forall x \in A)\tau(x) \in V$. Ora, como cada $\tau(x) \in V$ e $\sigma(x) \subseteq \tau(x)$, por Separação podemos obter $\{y \in V : (\exists x \in A)\sigma(x) = y\} = \{y : (\exists x \in A)\varphi(x,y)\}$.

Logo *Estágios-são-supercofinais* vindica a Substituição. E é pelo menos plausível que *Estágios-são-inesgotáveis* vindique *Estágios-são-supercofinais.* De fato, suponha que *Estágios-são-supercofinais* falha. Logo a hierarquia é cofinal com algum conjunto A, isto é, temos uma aplicação τ tal que para qualquer estágio S existe algum $x \in A$ tal que $S \in \tau(x)$. Nesse caso, temos de fato uma forma de apreender o suposto "infinito absoluto" da hierarquia: ele é *esgotado* pela imagem de τ aplicada a A. E isso compromete a ideia de que a hierarquia é "absolutamente infinita". Contrapondo: *Estágios-são-inesgotáveis* implica *Estágios-são-supercofinais*, que por sua vez justifica a Substituição.

---

[2] Gödel parece ter proposto um pensamento semelhante; ver Potter (2004, p. 223). Para discussão de Gödel e Shoenfield, ver Incurvati (2020, 90–5).

[3] Seria mais difícil provar a Substituição usando alguma formalização de *Estágios-são-supercofinais*, pois **Z** por si só não é forte o suficiente para definir os estágios, logo não é claro como se formalizaria *Estágios-são-supercofinais.* Uma opção, contudo, é trabalhar nalguma extensão de **LT**, como discutido na Seção 12.3.

Isto representa uma tentativa genuinamente promissora de fornecer uma justificação intrínseca da Substituição. Mas se funciona ultimamente ou não, teremos de deixar para si decidir.

12.6 Substituição e Reflexão

A nossa última tentativa de justificar a Substituição, via *Estágios-são-inesgotáveis*, começa com um resultado profundo e elegante:[4]

Teorema 12.2 (Esquema de reflexão). *Para qualquer fórmula φ:*

$$\forall\alpha\exists\beta > \alpha(\forall x_1 \ldots, x_n \in V_\beta)(\varphi(x_1, \ldots, x_n) \leftrightarrow \varphi^{V_\beta}(x_1, \ldots, x_n))$$

Como na Definição 12.1, φ^{V_β} é o resultado de restringir todos os quantificadores de φ ao conjunto V_β. Assim, intuitivamente, a Reflexão diz o seguinte: se φ é verdadeira em toda a hierarquia, então φ é verdadeira em arbitrariamente muitos *segmentos iniciais* da hierarquia.

Montague (1961) e Lévy (1960) mostraram que (formulações adequadas da) Substituição e da Reflexão são equivalentes, módulo **Z**, de modo que acrescentar qualquer uma delas dá **ZF**. (Provamos estes resultados na Seção 12.7.) Dada esta equivalência, poder-se-á esperar justificar a Reflexão e a Substituição via *Estágios-são-inesgotáveis* do seguinte modo: dado *Estágios-são-inesgotáveis*, a hierarquia deverá ser muito, muito alta; tão alta, de fato, que nada do que podemos dizer sobre ela é suficiente para limitar a sua altura. E podemos entender isto como a ideia de que, se qualquer sentença φ é verdadeira em toda a hierarquia, então é verdadeira em arbitrariamente muitos segmentos iniciais da hierarquia. E isso é precisamente a Reflexão.

De novo, isto parece uma tentativa genuinamente promissora de fornecer uma justificação intrínseca da Substituição. Mas há

[4] Lembrete: todas as fórmulas podem ter parâmetros (salvo indicação explícita em contrário).

demasiado a dizer sobre o assunto aqui. Cabe-lhe agora decidir por si se tem êxito.[5]

12.7 Apêndice: Resultados sobre Substituição

A seguir, provaremos a Reflexão dentro de **ZF**. Provaremos também em que sentido a Reflexão é equivalente à Substituição. E provaremos uma consequência interessante de tudo isto, relativa à força da Reflexão/Substituição. *Aviso: isto é, de longe, a parte mais avançada de matemática deste manual.*

Começamos com um lema que, por brevidade, recorre ao expediente notacional do *sobrelinhado* para lidar com sequências de variáveis ou objetos. Assim: "$\overline{a}_k$" abrevia "$a_{k_1}, \ldots, a_{k_n}$", onde n é determinado pelo contexto.

Lema 12.3. *Para cada $1 \leq i \leq k$, seja $\varphi_i(\overline{v}_i, x)$ uma fórmula. Então, para cada α existe algum $\beta > \alpha$ tal que, para quaisquer $\overline{a}_1, \ldots, \overline{a}_k \in V_\beta$ e cada $1 \leq i \leq k$:*

$$\exists x \varphi_i(\overline{a}_i, x) \to (\exists x \in V_\beta)\varphi_i(\overline{a}_i, x)$$

Prova. Definimos um termo μ como segue: $\mu(\overline{a}_1, \ldots, \overline{a}_k)$ é o estágio mínimo, V, que satisfaz todos os seguintes condicionais, para $1 \leq i \leq k$:

$$\exists x \varphi_i(\overline{a}_i, x) \to (\exists x \in V)\varphi_i(\overline{a}_i, x))$$

É fácil verificar que $\mu(\overline{a}_1, \ldots, \overline{a}_k)$ existe para todos $\overline{a}_1, \ldots, \overline{a}_k$. Ora, usando a Substituição e o nosso teorema de recursão, definimos:

$$S_0 = V_{\alpha+1}$$
$$S_{n+1} = S_n \cup \bigcup \{\mu(\overline{a}_1, \ldots, \overline{a}_k) : \overline{a}_1, \ldots, \overline{a}_k \in S_n\}$$

[5] Poderá, no entanto, querer continuar lendo Incurvati (2020, 95–100).

$$S = \bigcup_{m<\omega} S_n.$$

Cada S_n, e portanto o próprio S, é um estágio depois de V_α. Fixemos $\overline{a}_1, \ldots, \overline{a}_k \in S$; logo existe algum $n < \omega$ tal que $\overline{a}_1, \ldots, \overline{a}_k \in S_n$. Fixemos algum $1 \leq i \leq k$, e suponhamos que $\exists x \varphi_i(\overline{a}_i, x)$. Então $(\exists x \in \mu(\overline{a}_1, \ldots, \overline{a}_k))\varphi_i(\overline{a}_i, x)$ por construção, logo $(\exists x \in S_{n+1})\varphi_i(\overline{a}_i, x)$ e portanto $(\exists x \in S)\varphi_i(\overline{a}_i, x)$. Logo S é o nosso V_β. □

Podemos agora provar Teorema 12.2 de forma bastante direta:

Prova. Fixemos α. Sem perda de generalidade, podemos assumir que os únicos conectivos de φ são $\exists$, $\neg$ e $\wedge$ (pois estes são expressivamente adequados). Sejam $\psi_1, \ldots, \psi_k$ uma enumeração de cada uma das subfórmulas de φ por ordem de complexidade, de modo que $\psi_k = \varphi$. Por Lema 12.3, existe um $\beta > \alpha$ tal que, para qualquer $\overline{a}_i \in V_\beta$ e cada $1 \leq i \leq k$:

$$\exists x \psi_i(\overline{a}_i, x) \rightarrow (\exists x \in V_\beta)\psi_i(\overline{a}_i, x) \qquad (*)$$

Por indução na complexidade de ψ_i, mostraremos que $\psi_i(\overline{a}_i) \leftrightarrow \psi_i^{V_\beta}(\overline{a}_i)$, para qualquer $\overline{a}_i \in V_\beta$. Se ψ_i é atómica, isto é trivial. A bicondicional estabelece também que, quando ψ_i é uma negação ou conjunção de subfórmulas que satisfazem esta propriedade, a própria ψ_i satisfaz esta propriedade. Logo o único caso interessante é o da quantificação. Fixemos $\overline{a}_i \in V_\beta$; então:

$(\exists x \psi_i(\overline{a}_i, x))^{V_\beta}$	sse $(\exists x \in V_\beta)\psi_i^{V_\beta}(\overline{a}_i, x)$	por definição
	sse $(\exists x \in V_\beta)\psi_i(\overline{a}_i, x)$	por hipótese
	sse $\exists x \psi_i(\overline{a}_i, x)$	por (*)

Isto completa a indução; o resultado segue pois $\psi_k = \varphi$. □

Provamos a Reflexão em **ZF**. A nossa prova seguiu essencialmente Montague (1961). Queremos agora provar em **Z** que a Reflexão implica a Substituição. A prova segue Lévy (1960), com uma simplificação.

Como trabalhamos em **Z**, não podemos apresentar a Reflexão exatamente na forma dada acima. Afinal, formulamos a Reflexão usando a notação "V_α", e isso não pode ser definido em **Z** (ver Seção 11.5). Em vez disso, ofereceremos uma formulação aparentemente mais fraca da Substituição, como segue:

Reflexão fraca. Para qualquer fórmula φ, existe um conjunto transitivo S tal que 0, 1 e quaisquer parâmetros de φ são membroos de S, e $(\forall \overline{x} \in S)(\varphi \leftrightarrow \varphi^S)$.

Para usar isto na prova da Substituição, seguiremos primeiro Lévy (1960, primeira parte do Teorema 2) e mostraremos que podemos "refletir" duas fórmulas de uma vez:

Lema 12.4 (em Z + Reflexão fraca.). *Para quaisquer fórmulas ψ, χ, existe um conjunto transitivo S tal que 0 e 1 (e quaisquer parâmetros das fórmulas) são membroos de S, e $(\forall \overline{x} \in S)((\psi \leftrightarrow \psi^S) \wedge (\chi \leftrightarrow \chi^S))$.*

Prova. Seja φ a fórmula $(z = 0 \wedge \psi) \vee (z = 1 \wedge \chi)$.

Aqui usamos uma abreviatura; devemos explicitar "$z = 0$" como "$\forall t\ t \notin z$" e "$z = 1$" como "$\forall s(s \in z \leftrightarrow \forall t\ t \notin s)$". Mas como $0,1 \in S$ e S é transitivo, estas fórmulas são *absolutas* para S; isto é, aplicam-se ao mesmo objeto quer restrinjamos os seus quantificadores a S.[6]

Pela Reflexão fraca, temos algum S adequado tal que:

$$
\begin{aligned}
&(\forall z, \overline{x} \in S)(\varphi \leftrightarrow \varphi^S)\\
\text{isto é } &(\forall z, \overline{x} \in S)(((z = 0 \wedge \psi) \vee (z = 1 \wedge \chi)) \leftrightarrow\\
&\qquad\qquad ((z = 0 \wedge \psi) \vee (z = 1 \wedge \chi))^S)\\
\text{isto é } &(\forall z, \overline{x} \in S)(((z = 0 \wedge \psi) \vee (z = 1 \wedge \chi)) \leftrightarrow\\
&\qquad\qquad ((z = 0 \wedge \psi^S) \vee (z = 1 \wedge \chi^S)))\\
\text{isto é } &(\forall \overline{x} \in S)((\psi \leftrightarrow \psi^S) \wedge (\chi \leftrightarrow \chi^S))
\end{aligned}
$$

[6] Mais formalmente, seja ξ qualquer uma destas fórmulas; então $\xi(z) \leftrightarrow \xi^S(z)$.

A segunda afirmação implica a terceira porque "$z = 0$" e "$z = 1$" são absolutas para S; a quarta segue pois $0 \neq 1$. □

Podemos agora obter a Substituição, seguindo e simplificando Lévy (1960, Teorema 6):

Teorema 12.5 (em Z + Reflexão fraca). *Para qualquer fórmula $\varphi(v,w)$, e qualquer A, se $(\forall x \in A)\exists!y\, \varphi(x,y)$, então $\{y : (\exists x \in A)\varphi(x,y)\}$ existe.*

Prova. Fixemos A tal que $(\forall x \in A)\exists!y\, \varphi(x,y)$, e definimos fórmulas:

$$\psi \text{ é } (\varphi(x,z) \wedge A = A)$$
$$\chi \text{ é } \exists y\, \varphi(x,y)$$

Usando Lema 12.4, como A é um parâmetro de ψ, existe um conjunto transitivo S tal que $0,1,A \in S$ (juntamente com quaisquer outros parâmetros), e tal que:

$$(\forall x,z \in S)((\psi \leftrightarrow \psi^S) \wedge (\chi \leftrightarrow \chi^S))$$

Logo, em particular:

$$(\forall x,z \in S)(\varphi(x,z) \leftrightarrow \varphi^S(x,z))$$
$$(\forall x \in S)(\exists y \varphi(x,y) \leftrightarrow (\exists y \in S)\varphi^S(x,y))$$

Combinando isto, e observando que $A \subseteq S$ pois $A \in S$ e S é transitivo:

$$(\forall x \in A)(\exists y \varphi(x,y) \leftrightarrow (\exists y \in S)\varphi(x,y))$$

Ora, $(\forall x \in A)(\exists! y \in S)\varphi(x,y)$, porque $(\forall x \in A)\exists!y\, \varphi(x,y)$. Ora, a Separação dá $\{y \in S : (\exists x \in A)\varphi(x,y)\} = \{y : (\exists x \in A)\varphi(x,y)\}$. □

12.8 Apêndice: Axiomatizabilidade finita

Terminamos este capítulo extraindo alguns resultados da Substituição. O primeiro resultado é devido a Montague (1961); note que não é uma prova *dentro* de **ZF**, mas uma prova *sobre* **ZF**:

Teorema 12.6. **ZF** *não é finitamente axiomatizável. Mais geralmente: se* $\mathbf{T}$ *é finita e* $\mathbf{T} \vdash \mathbf{ZF}$*, então* $\mathbf{T}$ *é inconsistente.*

(Aqui nos restringimos tacitamente a sentenças de primeira ordem cujo único primitivo não lógico é $\in$*, e escrevemos* $\mathbf{T} \vdash \mathbf{ZF}$ *para indicar que* $\mathbf{T} \vdash \varphi$ *para todo o* $\varphi \in \mathbf{ZF}$*.)*

Prova. Fixemos $\mathbf{T}$ finita tal que $\mathbf{T} \vdash \mathbf{ZF}$. Logo $\mathbf{T}$ prova a Reflexão, isto é, Teorema 12.2. Como $\mathbf{T}$ é finita, podemos reescrevê-la como uma única conjunção, θ. Refletindo com esta fórmula, $\mathbf{T} \vdash \exists\beta\,(\theta \leftrightarrow \theta^{V_\beta})$. Como trivialmente $\mathbf{T} \vdash \theta$, obtemos $\mathbf{T} \vdash \exists\beta\ \theta^{V_\beta}$.

Ora, seja $\psi(X)$ uma abreviatura de:

$$\theta^X \wedge X \text{ é transitivo} \wedge (\forall Y \in X)(Y \text{ é transitivo} \rightarrow \neg\theta^Y)$$

isto diz, em traços largos: X é um modelo transitivo de θ, e $\in$-minimal a esse respeito. Ora, recordando que $\mathbf{T} \vdash \exists\beta\ \theta^{V_\beta}$, por fatos básicos sobre postos dentro de **ZF** e portanto dentro de $\mathbf{T}$, temos:

$$\mathbf{T} \vdash \exists M \psi(M). \tag{*}$$

Usando o primeiro conjunto de $\psi(X)$, sempre que $\mathbf{T} \vdash \sigma$, temos $\mathbf{T} \vdash \forall X(\psi(X) \rightarrow \sigma^X)$. Logo, por (*):

$$\mathbf{T} \vdash \forall X(\psi(X) \rightarrow (\exists N \psi(N))^X)$$

Usando isto, e (*) de novo:

$$\mathbf{T} \vdash \exists M(\psi(M) \wedge (\exists N \psi(N))^M)$$

Em particular, então:

$$\mathbf{T} \vdash \exists M(\psi(M) \wedge (\exists N \in M)((N \text{ é transitivo})^N \wedge (\theta^N)^M))$$

Logo, por raciocínio elementar sobre transitividade:

$$\mathbf{T} \vdash \exists M(\psi(M) \wedge (\exists N \in M)(N \text{ é transitivo} \wedge \theta^N))$$

Logo $\mathbf{T}$ é inconsistente.[7] □

Eis um resultado semelhante, assinalado por Potter (2004, 223):

Proposição 12.7. *Seja* $\mathbf{T}$ *uma extensão de* $\mathbf{Z}$ *com um número finito de novos axiomas. Se* $\mathbf{T} \vdash \mathbf{ZF}$*, então* $\mathbf{T}$ *é inconsistente. (Aqui usamos as mesmas restrições tácitas que para Teorema 12.6.)*

Prova. Usemos θ para a conjunção de todos os axiomas de $\mathbf{T}$ *exceto* as (infinitas) instâncias de Separação. Definindo ψ a partir de θ como no Teorema 12.6, podemos mostrar que $\mathbf{T} \vdash \exists M \psi(M)$.

Como no Teorema 12.6, podemos estabelecer o esquema de que, sempre que $\mathbf{T} \vdash \sigma$, temos $\mathbf{T} \vdash \forall X(\psi(X) \rightarrow \sigma^X)$. Terminamos a prova exatamente como no Teorema 12.6.

No entanto, estabelecer o esquema envolve um pouco mais de trabalho do que no Teorema 12.6. Afinal, as instâncias de Separação estão em $\mathbf{T}$, mas não são conjuntos de θ. Contudo, podemos ultrapassar este obstáculo provando que $\mathbf{T} \vdash \forall X(X \text{ é transitivo} \rightarrow \sigma^X)$, para toda a instância de Separação σ. Deixamos isto ao leitor. □

Como observamos na Seção 12.5, isto mostra que a Substituição é estritamente mais forte que Teorema 10.26. Ou, um pouco mais precisamente: se $\mathbf{Z}$ + "toda a boa ordenação é isomorfa a um ordinal único" for consistente, então não prova alguma instância da Substituição.

[7] Este "raciocínio elementar" envolve provar certos "fatos de absolutidade" para conjuntos transitivos.

Problemas

Problema 12.1. Apresente uma formalização de *Estágios-são-supercofinais* em **ZF**.

Problema 12.2. Mostre que, para toda a instância de Separação σ, tem-se: $\mathbf{Z} \vdash \forall X (X \text{ é transitivo} \to \sigma^X)$. (Usamos este esquema na Proposição 12.7.)

Problema 12.3. Mostre que, para todo o $\varphi \in \mathbf{Z}$, tem-se $\mathbf{ZF} \vdash \varphi^{V_{\omega+\omega}}$.

Problema 12.4. Confirme os restantes resultados esquemáticos invocados nas provas de Teorema 12.6 e Proposição 12.7.

CAPÍTULO 13

Aritmética Ordinal

13.1 Introdução

No Capítulo 10, desenvolvemos uma teoria dos números ordinais. Vimos no Capítulo 11 que podemos pensar nos ordinais como uma espinha dorsal em torno da qual se constrói o restante da hierarquia. Mas esse não é o único papel dos ordinais. Há também a tarefa de efetuar a aritmética ordinal.

Já aludimos a isto, em Seção 10.2, quando falamos de ω, $\omega+1$ e $\omega + \omega$. Na altura falamos informalmente; chegou a altura de o expor com rigor. Devemos contudo mencionar que não há muita filosofia neste capítulo; apenas desenvolvimentos técnicos, juntamente com uma observação (ligeiramente) interessante de que podemos fazer a mesma coisa de duas maneiras diferentes.

13.2 Adição Ordinal

Suponha-se que se pretende somar α e β. Pode simplesmente colocar-se uma cópia de β imediatamente a seguir a uma cópia de α. (É preciso tomar *cópias*, pois sabemos por Proposição 10.22 que ou $\alpha \subseteq \beta$ ou $\beta \subseteq \alpha$.) O efeito intuitivo disto é percorrer

uma α-sucessão de passos, *e depois* percorrer uma β-sucessão. A sucessão resultante fica bem ordenada; logo, por Teorema 10.26 é isomorfa a um ordinal (único). Esse ordinal pode considerar-se a *soma* de α e β.

Esta é a ideia intuitiva por detrás da adição ordinal. Para a definir com rigor, começa-se pela ideia de tomar *cópias* de conjuntos. A ideia é usar etiquetas arbitrárias, 0 e 1, para saber de onde veio cada objeto:

Definição 13.1. A *soma disjunta* de A e B é $A \sqcup B = (A \times \{0\}) \cup (B \times \{1\})$.

Define-se em seguida uma ordem em pares de ordinais:

Definição 13.2. Para quaisquer ordinais $\alpha_1, \alpha_2, \beta_1, \beta_2$, diga-se que:

$$\langle \alpha_1, \alpha_2 \rangle \lessdot \langle \beta_1, \beta_2 \rangle \text{ sse ou } \alpha_2 \in \beta_2 \text{ ou tanto } \alpha_2 = \beta_2 \text{ como } \alpha_1 \in \beta_1$$

Trata-se de uma ordem *lexicográfica inversa*, pois ordena-se pelo segundo elemento e depois pelo primeiro. Recorde-se que se pretendia definir $\alpha + \beta$ como o tipo de ordem de uma cópia de α seguida de uma cópia de β. Para o conseguir, define-se:

Definição 13.3. Para quaisquer ordinais α, β, a sua soma é $\alpha + \beta = \mathrm{ord}(\alpha \sqcup \beta, \lessdot)$.

Note-se que aqui se abusou ligeiramente da notação; estritamente deveria escrever-se "$\{\langle x, y \rangle \in \alpha \sqcup \beta : x \lessdot y\}$" em lugar de "$\lessdot$". Por brevidade, contudo, continuaremos a abusar da notação deste modo no que se segue.

O resultado seguinte, juntamente com Teorema 10.26, confirma que a nossa definição está bem formada:

Lema 13.4. *$\langle \alpha \sqcup \beta, \lhd \rangle$ é um bom ordenamento, para quaisquer ordinais α e β.*

Prova. Obviamente $\lhd$ é conexa em $\alpha \sqcup \beta$. Para mostrar que é bem fundada, fixe-se um $X \subseteq \alpha \sqcup \beta$ não vazio. Seja Y o subconjunto de X cuja segunda coordenada é a menor possível, isto é, $Y = \{\langle \gamma, i \rangle \in X : (\forall \langle \delta, j \rangle \in X) i \leq j\}$. Escolha-se agora o elemento de Y com menor primeira coordenada. □

Temos assim uma definição explícita e clara da adição ordinal. Eis um fato não surpreendente (recorde-se que $1 = \{0\}$, por Definição 9.7):

Proposição 13.5. *$\alpha + 1 = \alpha^+$, para qualquer ordinal α.*

Prova. Considere-se o isomorfismo f de $\alpha^+ = \alpha \cup \{\alpha\}$ para $\alpha \sqcup 1 = (\alpha \times \{0\}) \sqcup (\{0\} \times \{1\})$ dado por $f(\gamma) = \langle \gamma, 0 \rangle$ para $\gamma \in \alpha$, e $f(\alpha) = \langle 0, 1 \rangle$. □

Além disso, é fácil mostrar que a adição obedece a certas condições recursivas:

Lema 13.6. *Para quaisquer ordinais α, β, tem-se:*

$$\begin{aligned} \alpha + 0 &= \alpha \\ \alpha + (\beta + 1) &= (\alpha + \beta) + 1 \\ \alpha + \beta &= \operatorname*{lsub}_{\delta < \beta} (\alpha + \delta) \qquad \textit{se } \beta \textit{ é um ordinal limite} \end{aligned}$$

Prova. Verifica-se caso a caso; primeiro:

$$\begin{aligned} \alpha + 0 &= \operatorname{ord}((\alpha \times \{0\}) \cup (0 \times \{1\}), \lhd) \\ &= \operatorname{ord}((\alpha \times \{0\}) \cup \{0\}, \lhd) \\ &= \alpha \\ \alpha + (\beta + 1) &= \operatorname{ord}((\alpha \times \{0\}) \cup (\beta^+ \times \{1\}), \lhd) \\ &= \operatorname{ord}((\alpha \times \{0\}) \cup (\beta \times \{1\}), \lhd) + 1 \end{aligned}$$

$$= (\alpha + \beta) + 1$$

Seja agora $\beta \neq \emptyset$ um limite. Se $\delta < \beta$ então também $\delta + 1 < \beta$, logo $\alpha + \delta$ é um segmento inicial próprio de $\alpha + \beta$. Logo $\alpha + \beta$ é um majorante estrito de $X = \{\alpha + \delta : \delta < \beta\}$. Além disso, se $\alpha \leq \gamma < \alpha + \beta$, então claramente $\gamma = \alpha + \delta$ para algum $\delta < \beta$. Logo $\alpha + \beta = \mathrm{lsub}_{\delta<\beta}(\alpha + \delta)$. □

Mas eis um fato marcante. Para definir a adição ordinal, poder-se-ia *em alternativa* ter usado simplesmente o teorema da recursão transfinita, e estipulado as equações recursivas, exatamente como em Lema 13.6 (embora com "β^+" em vez de "$\beta+1$").

Há, pois, duas maneiras diferentes de definir operações nos ordinais. Podem definir-se *sinteticamente*, construindo explicitamente um conjunto bem ordenado e considerando o seu tipo de ordem. Ou podem definir-se *recursivamente*, apenas estipulando as equações recursivas. Feito corretamente, porém, o resultado é idêntico. De fato, Teorema 10.26 garante que estas equações recursivas determinam ordinais *únicos*.

Em muitos aspectos, a aritmética ordinal comporta-se como a adição dos números naturais. Por exemplo, pode provar-se o seguinte:

Lema 13.7. *Se α, β, γ são ordinais, então:*

1. *se $\beta < \gamma$, então $\alpha + \beta < \alpha + \gamma$*
2. *se $\alpha + \beta = \alpha + \gamma$, então $\beta = \gamma$*
3. *$\alpha + (\beta + \gamma) = (\alpha + \beta) + \gamma$, isto é, a adição é associativa*
4. *Se $\alpha \leq \beta$, então $\alpha + \gamma \leq \beta + \gamma$*

Prova. Provamos (3), deixando o resto como exercício. A prova é por indução transfinita simples em γ, usando Lema 13.6. Quando $\gamma = 0$:

$$(\alpha + \beta) + 0 = \alpha + \beta = \alpha + (\beta + 0)$$

Quando $\gamma = \delta + 1$, suponha-se por indução que $(\alpha + \beta) + \delta = \alpha + (\beta + \delta)$; usando Lema 13.6 três vezes:

$$\begin{aligned}(\alpha + \beta) + (\delta + 1) &= ((\alpha + \beta) + \delta) + 1 \\ &= (\alpha + (\beta + \delta)) + 1 \\ &= \alpha + ((\beta + \delta) + 1) \\ &= \alpha + (\beta + (\delta + 1))\end{aligned}$$

Quando γ é um ordinal limite, suponha-se por indução que se $\delta \in \gamma$ então $(\alpha + \beta) + \delta = \alpha + (\beta + \delta)$; então:

$$\begin{aligned}(\alpha + \beta) + \gamma &= \operatorname*{lsub}_{\delta<\gamma}((\alpha + \beta) + \delta) \\ &= \operatorname*{lsub}_{\delta<\gamma}(\alpha + (\beta + \delta)) \\ &= \alpha + \operatorname*{lsub}_{\delta<\gamma}(\beta + \delta) \\ &= \alpha + (\beta + \gamma)\end{aligned}$$

□

Deste modo, a adição ordinal deve parecer muito familiar. Há contudo um ponto crucial em que a adição ordinal *não* se parece com a adição nos naturais.

Proposição 13.8. *A adição ordinal* não *é comutativa;* $1 + \omega = \omega < \omega + 1$.

Prova. Note-se que $1 + \omega = \operatorname{lsub}_{n<\omega}(1 + n) = \omega \in \omega \cup \{\omega\} = \omega^+ = \omega + 1$. □

Embora isto possa surpreender à primeira vista, *não devia.* Por um lado, ao considerar $1 + \omega$, pensa-se no tipo de ordem que se obtém colocando um elemento extra *antes* de todos os números naturais. Raciocinando como com o Hotel de Hilbert em Seção 7.1, intuitivamente esse primeiro elemento extra não altera o tipo de ordem global. Por outro lado, ao considerar $\omega + 1$, pensa-se no tipo de ordem que se obtém colocando um elemento extra *depois* de todos os números naturais. E isso é algo radicalmente diferente!

13.3 Usando Adição Ordinal

Com a adição nos ordinais, podemos calcular explicitamente os postos de vários conjuntos, no sentido de Definição 11.15:

Lema 13.9. *Se* $\text{posto}(A) = \alpha$ *e* $\text{posto}(B) = \beta$*, então:*

1. $\text{posto}(\wp(A)) = \alpha + 1$
2. $\text{posto}(\{A, B\}) = \max(\alpha, \beta) + 1$
3. $\text{posto}(A \cup B) = \max(\alpha, \beta)$
4. $\text{posto}(\langle A, B\rangle) = \max(\alpha, \beta) + 2$
5. $\text{posto}(A \times B) \leq \max(\alpha, \beta) + 2$
6. $\text{posto}(\bigcup A) = \alpha$ *quando* α *é vazio ou limite;* $\text{posto}(\bigcup A) = \gamma$ *quando* $\alpha = \gamma + 1$

Prova. Ao longo de todo o raciocínio, invocamos repetidamente Proposição 11.20.

(1). Se $x \subseteq A$ então $\text{posto}(x) \leq \text{posto}(A)$. Logo $\text{posto}(\wp(A)) \leq \alpha + 1$. Como em particular $A \in \wp(A)$, tem-se $\text{posto}(\wp(A)) = \alpha + 1$.

(2). Por Proposição 11.20

(3). Por Proposição 11.20.

(4). Por (2), duas vezes.

(5). Note-se que $A \times B \subseteq \wp(\wp(A \cup B))$, e invoque (4).

(6). Se $\alpha = \gamma + 1$, existe algum $c \in A$ com $\text{posto}(c) = \gamma$, e nenhum membro de A tem posto superior; logo $\text{posto}(\bigcup A) = \gamma$. Se α é um ordinal limite, então A tem membros com posto arbitrariamente próximo (mas estritamente inferior) de α, pelo que $\bigcup A$ também tem membros com posto arbitrariamente próximo (mas estritamente inferior) de α, e portanto $\text{posto}(\bigcup A) = \alpha$. □

Deixamos como exercício mostrar por que (5) envolve uma *des*igualdade.

Estamos também agora em condições de mostrar que várias noções razoáveis do que possa significar descrever um ordinal como "finito" ou "infinito" coincidem:

Lema 13.10. *Para qualquer ordinal α, as seguintes afirmações são equivalentes:*

1. $\alpha \notin \omega$, *isto é,* α *não é um número natural*
2. $\omega \leq \alpha$
3. $1 + \alpha = \alpha$
4. $\alpha \approx \alpha + 1$, *isto é,* α *e* $\alpha + 1$ *são equipotentes*
5. α *é infinito de Dedekind*

Temos assim cinco formas provavelmente equivalentes de precisar o que exige que um ordinal seja (in)finito.

Prova. (1) $\Rightarrow$ (2). Por tricotomia.

(2) $\Rightarrow$ (3). Fixe $\alpha \geq \omega$. Por indução transfinita, existe um ordinal γ mínimo (possivelmente 0) tal que existe um ordinal limite β com $\alpha = \beta + \gamma$. Ora:

$$1 + \alpha = 1 + (\beta + \gamma) = (1 + \beta) + \gamma = \operatorname*{lsub}_{\delta<\beta}(1 + \delta) + \gamma = \beta + \gamma = \alpha.$$

(3) $\Rightarrow$ (4). Existe claramente uma bijeção $f : (\alpha \sqcup 1) \to (1 \sqcup \alpha)$. Se $1 + \alpha = \alpha$, existe um isomorfismo $g : (1 \sqcup \alpha) \to \alpha$. Considere-se $g \circ f$.

(4) $\Rightarrow$ (5). Se $\alpha \approx \alpha + 1$, existe uma bijeção $f : (\alpha \sqcup 1) \to \alpha$. Defina $g(\gamma) = f(\gamma, 0)$ para cada $\gamma < \alpha$; esta injeção mostra que α é infinito de Dedekind, pois $f(0,1) \in \alpha \setminus \mathrm{Im}(g)$.

(5) $\Rightarrow$ (1). Isto é Proposição 9.8. □

13.4 Multiplicação Ordinal

Passamos agora à multiplicação ordinal, abordando-a de forma muito semelhante à adição ordinal. Suponha-se que se pretende

multiplicar α por β. Pode imaginar-se uma grelha retangular, com largura α e altura β; o produto de α e β é então o resultado de percorrer cada linha e depois passar à linha seguinte... até percorrer toda a grelha. Dito de outro modo, o produto de α e β obtém-se substituindo *cada* elemento de β por uma cópia de α.

Para formalizar, usa-se simplesmente a ordem lexicográfica inversa no produto cartesiano de α e β:

Definição 13.11. Para quaisquer ordinais α, β, o seu produto $\alpha \cdot \beta = \mathrm{ord}(\alpha \times \beta, \lhd)$.

É preciso de novo confirmar que esta definição está bem formada:

Lema 13.12. *$\langle \alpha \times \beta, \lhd \rangle$ é um bom ordenamento, para quaisquer ordinais α e β.*

Prova. Exatamente como para Lema 13.4. □

E não é difícil provar que a multiplicação se comporta assim:

Lema 13.13. *Para quaisquer ordinais α, β:*

$$\begin{aligned} \alpha \cdot 0 &= 0 \\ \alpha \cdot (\beta + 1) &= (\alpha \cdot \beta) + \alpha \\ \alpha \cdot \beta &= \operatorname*{lsub}_{\delta < \beta}(\alpha \cdot \delta) \qquad \textit{quando } \beta \textit{ é um ordinal limite.} \end{aligned}$$

Prova. Deixado como exercício. □

De fato, tal como no caso da adição, poderíamos ter definido a multiplicação ordinal por estas equações recursivas, em vez de dar uma definição direta. Tal como na adição, certo comportamento é familiar:

Lema 13.14. *Se α, β, γ são ordinais, então:*

1. *se $\alpha \neq 0$ e $\beta < \gamma$, então $\alpha \cdot \beta < \alpha \cdot \gamma$;*

2. *se* $\alpha \neq 0$ *e* $\alpha \cdot \beta = \alpha \cdot \gamma$, *então* $\beta = \gamma$;

3. $\alpha \cdot (\beta \cdot \gamma) = (\alpha \cdot \beta) \cdot \gamma$;

4. *Se* $\alpha \leq \beta$, *então* $\alpha \cdot \gamma \leq \beta \cdot \gamma$;

5. $\alpha \cdot (\beta + \gamma) = (\alpha \cdot \beta) + (\alpha \cdot \gamma)$.

Prova. Deixado como exercício. □

Pode provar-se (ou consultar) outros resultados, à vontade. Mas, dado Proposição 13.8, o que se segue não deve surpreender:

Proposição 13.15. *A multiplicação ordinal não é comutativa:* $2 \cdot \omega = \omega < \omega \cdot 2$

Prova. $2 \cdot \omega = \mathrm{lsub}_{n<\omega}(2 \cdot n) = \omega \in \mathrm{lsub}_{n<\omega}(\omega + n) = \omega + \omega = \omega \cdot 2$. □

De novo, a justificação intuitiva é direta. Para calcular $2 \cdot \omega$, substitui-se cada número natural por duas entidades. Obtém-se o mesmo tipo de ordem se simplesmente se inserirem todos os números "meios" nos naturais, isto é, se se considerar a ordem natural em $\{n/2 : n \in \omega\}$. Assim posto, o tipo de ordem é claramente o mesmo que o de ω em si. Mas, para calcular $\omega \cdot 2$, dispõem-se duas cópias de ω, uma a seguir à outra.

13.5 Exponenciação Ordinal

Passamos agora à exponenciação ordinal. Infelizmente, não há uma definição *simpática* sintética para a exponenciação ordinal.

Certo que *existem* definições sintéticas explícitas. Eis uma. Seja $\mathrm{finfun}(\alpha, \beta)$ o conjunto de todas as funções $f : \alpha \to \beta$ tais que $\{\gamma \in \alpha : f(\gamma) \neq 0\}$ é equipotente a algum número natural. Defina-se um bom ordenamento em $\mathrm{finfun}(\alpha, \beta)$ por $f \sqsubset g$ sse $f \neq g$ e $f(\gamma_0) < g(\gamma_0)$, onde $\gamma_0 = \max\{\gamma \in \alpha : f(\gamma) \neq g(\gamma)\}$. Então pode definir-se $\alpha^{(\beta)}$ como $\mathrm{ord}(\mathrm{finfun}(\alpha, \beta), \sqsubset)$. Potter usa esta definição explícita e explica de seguida:

> A escolha desta ordenação determina-se unicamente pelo desejo de obter uma definição da exponenciação ordinal que obedeça à condição recursiva adequada..., e é muito mais difícil de visualizar do que a soma ou o produto ordenados. (Potter, 2004, p. 199)

Exatamente. Explicamos a adição como "uma cópia de α seguida de uma cópia de β", e a multiplicação como "uma β-sucessão de cópias de α". Mas não temos nada de sucinto a dizer sobre finfun(α, γ). Por isso, apresentaremos a definição da exponenciação ordinal apenas *por* recursão transfinita, isto é:

Definição 13.16.

$$\begin{aligned} \alpha^{(0)} &= 1 \\ \alpha^{(\beta+1)} &= \alpha^{(\beta)} \cdot \alpha \\ \alpha^{(\beta)} &= \bigcup_{\delta<\beta} \alpha^{(\delta)} \qquad \text{quando } \beta \text{ é um ordinal limite} \end{aligned}$$

Se trabalhássemos *como* teoristas de conjuntos, poderíamos querer explorar algumas propriedades da exponenciação ordinal. Mas pouco mais temos a acrescentar, salvo notar o fato pouco surpreendente de que a exponenciação ordinal não comuta. Assim $2^{(\omega)} = \bigcup_{\delta<\omega} 2^{(\delta)} = \omega$, enquanto $\omega^{(2)} = \omega \cdot \omega$. Mas então não se deve *esperar* que comute, pois já não comuta nos naturais: $2^{(3)} = 8 < 9 = 3^{(2)}$.

Problemas

Problema 13.1. Prove o resto de Lema 13.7.

Problema 13.2. Exiba conjuntos A e B tais que posto$(A \times B) = \max(\text{posto}(A), \text{posto}(B))$. Exiba conjuntos A e B tais que posto$(A \times B) = \max(\text{posto}(A), \text{posto}(B)) + 2$. São possíveis outros postos?

Problema 13.3. Prove Lema 13.12, Lema 13.13, e Lema 13.14

Problema 13.4. Usando indução transfinita, prove que, se se define $\alpha^{(\beta)} = \mathrm{ord}(\mathrm{finfun}(\alpha, \beta), \sqsubset)$, obtêm-se as equações recursivas de Definição 13.16.

CAPÍTULO 14

Cardinais

14.1 Princípio de Cantor

Volte o pensamento a Seção 10.5. Estávamos a discutir conjuntos bem ordenados e sugerimos que seria desejável ter objetos que funcionem como representantes dos bons ordenamentos. Com isto em mente, introduzimos os ordinais e depois mostramos em Corolário 10.28 que estes se comportam como desejamos, isto é:

$$\mathrm{ord}(A, <) = \mathrm{ord}(B, \lessdot) \text{ sse } \langle A, < \rangle \cong \langle B, \lessdot \rangle.$$

Volte ainda mais atrás, a Seção 5.7. Ali, trabalhando ingenuamente, introduzimos a noção de "tamanho" de um conjunto. Especificamente, dissemos que dois conjuntos são equipotentes, $A \approx B$, precisamente quando existe uma bijeção $f : A \to B$. Esta noção é intrinsecamente mais simples do que a de um bom ordenamento: interessam-nos apenas as bijeções, e não (como com os isomorfismos de ordem) se as bijeções "preservam alguma estrutura".

Tudo isto sugere um pensamento óbvio. Tal como introduzimos certos objetos, *ordinais*, para calibrar bons ordenamentos, podemos introduzir certos objetos, *cardinais*, para calibrar tamanho. Esse é o objetivo deste capítulo.

Antes de dizer o que serão estes cardinais, convém enunciar um princípio que devem satisfazer. Escrevendo $|X|$ para a cardi-

nalidade do conjunto X, desejaríamos que obedecessem a:

$$|A| = |B| \text{ sse } A \approx B.$$

Chamaremos a isto o Princípio *de Cantor*, já que Cantor foi provavelmente o primeiro a tê-lo claramente presente. (Diremos mais sobre a sua relação com o Princípio de *Hume* na Seção 14.5.) O nosso objetivo é definir $|X|$, para cada X, de modo a que satisfaça o Princípio de Cantor.

14.2 Cardinais como Ordinais

De fato, a nossa teoria dos cardinais fará uso (descarado) da nossa teoria dos ordinais. Ou seja: definiremos os cardinais como certos ordinais específicos. Em particular, proporemos o seguinte:

Definição 14.1. Se A pode ser bem ordenado, então $|A|$ é o menor ordinal γ tal que $A \approx \gamma$. Para qualquer ordinal γ, diz-se que γ é um *cardinal* sse $\gamma = |\gamma|$.

Acabamos de usar a locução "A pode ser bem ordenado". Como quase sempre em matemática, a locução modal aqui é apenas um atalho para uma asserção existencial: dizer "A pode ser bem ordenado" é dizer "existe uma relação que bem ordena A".

Há contudo um problema com Definição 14.1. Gostaríamos que *todo* o conjunto tivesse um tamanho, isto é, que $|A|$ existisse para todo A. A definição que acabamos de dar, porém, começa por uma condicional: "*Se* A pode ser bem ordenado...". Se existir algum conjunto A que não possa ser bem ordenado, a nossa definição simplesmente deixa de definir um objeto $|A|$.

Logo, para usar Definição 14.1, precisamos de uma garantia de que todo o conjunto pode ser bem ordenado. Infelizmente, essa garantia não está disponível em **ZF**. Assim, se queremos usar Definição 14.1, não há alternativa senão acrescentar um novo axioma, tal como:

Bom ordenamento (Bom ordenamento). Todo o conjunto pode ser bem ordenado.

Discutiremos se o axioma do bom ordenamento é aceitável em Capítulo 16. Daqui em diante, porém, simplesmente o admitiremos. E, com ele, é bastante direto provar que os cardinais (tal como definidos na Definição 14.1) existem e comportam-se bem:

Lema 14.2. *Para todo o conjunto A:*

1. *$|A|$ existe e é único;*
2. *$|A| \approx A$;*
3. *$|A|$ é um cardinal, isto é, $|A| = ||A||$;*

Prova. Fixe A. Pelo Bom ordenamento, existe um bom ordenamento $\langle A, R\rangle$. Por Teorema 10.26, $\langle A, R\rangle$ é isomorfo a um único ordinal, β. Logo $A \approx \beta$. Por indução transfinita, existe um único ordinal mínimo, γ, tal que $A \approx \gamma$. Logo $|A| = \gamma$, o que estabelece (1) e (2). Para (3), note-se que se $\delta \in \gamma$ então $\delta \prec A$, pela escolha de γ, logo também $\delta \prec \gamma$ pois a equipotência é uma relação de equivalência (Proposição 5.12). Logo $\gamma = |\gamma|$. □

O resultado seguinte garante o Princípio de Cantor, e mais ainda. (Note-se que os cardinais herdam a ordem dos ordinais, isto é, $\mathfrak{a} < \mathfrak{b}$ sse $\mathfrak{a} \in \mathfrak{b}$. Ao formular isto, usaremos letras Fraktur para objetos que sabemos serem cardinais. Isto é bastante usual. Uma alternativa comum é usar letras gregas, já que os cardinais são ordinais, escolhendo-as do meio do alfabeto, por exemplo κ, λ.):

Lema 14.3. *Para quaisquer conjuntos A e B:*

$$A \approx B \text{ sse } |A| = |B|$$
$$A \preceq B \text{ sse } |A| \leq |B|$$

$$A \prec B \textit{ sse } |A| < |B|$$

Prova. Provaremos o sentido da esquerda para a direita da segunda afirmação (os outros casos são semelhantes e ficam como exercício). Considere-se o diagrama seguinte:

As setas duplas indicam bijeções, cuja existência é garantida por Lema 14.2. Ao supor $A \preceq B$, existe uma injeção $A \to B$. Seguindo as setas de $|A|$ para A, para B e para $|B|$, obtém-se uma injeção $|A| \to |B|$ (a seta tracejada). □

Pode ainda usar-se Lema 14.3 para voltar a provar Cantor–Bernstein. Trata-se da afirmação de que se $A \preceq B$ e $B \preceq A$ então $A \approx B$. A enunciamos como Teorema 5.17, mas a provamos primeiro—com algum esforço—em Seção 7.5. Considere-se:

Nova prova de Schröder-Bernstein. Se $A \preceq B$ e $B \preceq A$, então $|A| \leq |B|$ e $|B| \leq |A|$ por Lema 14.3. Logo $|A| = |B|$ e $A \approx B$ por tricotomia e Lema 14.3. □

Embora seja uma prova muito simples, apoia-se implicitamente na Substituição (para assegurar Teorema 10.26) e no Bom ordenamento (para garantir Lema 14.3). Em contraste, a prova na Seção 7.5 era muito mais autônoma (de fato, pode efetuar-se em $\mathbf{Z}^-$).

14.3 ZFC: um marco

Com a adição do axioma do bom ordenamento, atingimos o último marco teórico. Dispomos agora de todos os axiomas necessários para **ZFC**. Em detalhe:

Definição 14.4. A teoria **ZFC** tem estes axiomas: Extensionalidade, União, Pares, Conjuntos potência, Infinitude, Fundação, Bom ordenamento e todas as instâncias dos esquemas de Separação e Substituição. Dito de outro modo, **ZFC** acrescenta o Bom ordenamento a **ZF**.

ZFC significa teoria de conjuntos de *Zermelo–Fraenkel* com *Escolha*. Isto pode parecer ligeiramente estranho, pois o axioma que adicionamos se chama "Bom ordenamento", não "Escolha". Mas, quando mais tarde formularmos a *Escolha*, verificar-se-á que o Bom ordenamento lhe é equivalente (módulo **ZF**) (ver Teorema 16.6). Logo, qual tomar como axioma "básico" é indiferente. E o nome "**ZFC**" é inteiramente standard na literatura.

14.4 Finito, Enumerável, Não Enumerável

Agora que fomos introduzidos aos cardinais, vale a pena dedicar um pouco de tempo a variedades de cardinais; nomeadamente, cardinais finitos, enumerável e não enumerável.

Os nossos dois primeiros resultados implicam que os cardinais finitos serão exatamente os ordinais finitos, que definimos como os nossos *números naturais* em Definição 9.7:

Proposição 14.5. *Sejam $n, m \in \omega$. Então $n = m$ sse $n \approx m$.*

Prova. O sentido *da esquerda para a direita* é trivial. Para provar *da direita para a esquerda*, suponha-se $n \approx m$ embora $n \neq m$. Por tricotomia, ou $n \in m$ ou $m \in n$; suponha-se, sem perda de generalidade, $n \in m$. Então $n \subsetneq m$ e existe uma bijeção $f : m \to n$, logo m é infinito de Dedekind, o que contradiz Proposição 9.8.□

Corolário 14.6. *Se $n \in \omega$, então n é um cardinal.*

Prova. Imediato. □

Segue-se também que várias noções razoáveis do que possa significar descrever um cardinal como "finito" ou "infinito" coincidem:

Teorema 14.7. *Para qualquer conjunto A, as seguintes afirmações são equivalentes:*

1. $|A| \notin \omega$, *isto é, A não é um número natural;*
2. $\omega \leq |A|$;
3. A *é infinito de Dedekind.*

Prova. Por Lema 13.10, Lema 14.3, e Corolário 14.6. □

Isto autoriza a seguinte *definição* de algumas noções que usamos de forma informal na Parte II:

Definição 14.8. Diz-se que A é *finito* sse $|A|$ é um número natural, isto é, $|A| \in \omega$. Caso contrário, diz-se que A é *infinito*.

Note-se que esta definição é apresentada no contexto de **ZFC**. Afinal, precisamos do Bom ordenamento para garantir que todo o conjunto tem uma cardinalidade. E, de fato, sem o Bom ordenamento, pode existir um conjunto que nem é finito nem infinito de Dedekind. Voltaremos a este tipo de questão no Capítulo 16. Por ora, continuamos a assumir o Bom ordenamento.

Passemos agora dos cardinais finitos aos cardinais infinitos. Eis dois pontos elementares:

Corolário 14.9. ω *é o menor cardinal infinito.*

Prova. ω é um cardinal, pois ω é infinito de Dedekind e se $\omega \approx n$ para algum $n \in \omega$ então n seria infinito de Dedekind, contrariando Proposição 9.8. Ora, ω é o menor cardinal infinito por definição. □

Corolário 14.10. *Todo o cardinal infinito é um ordinal limite.*

Prova. Seja α um ordinal sucessor infinito, logo $\alpha = \beta + 1$ para algum β. Por Proposição 14.5, β também é infinito, logo $\beta \approx \beta+1$ por Lema 13.10. Ora, $|\beta| = |\beta + 1| = |\alpha|$ por Lema 14.3, logo $\alpha \neq |\alpha|$. □

Já na Definição 5.2, assinalamos que podemos distinguir conjuntos infinitos enumerável de conjuntos infinitos não enumerável. Essa definição conduz naturalmente ao seguinte:

Proposição 14.11. *A é enumerável sse $|A| \leq \omega$, e A é não enumerável sse $\omega < |A|$.*

Prova. Por tricotomia, as duas afirmações são equivalentes, bastando provar que A é enumerável sse $|A| \leq \omega$. Para *da direita para a esquerda*: se $|A| \leq \omega$, então $A \preceq \omega$ por Lema 14.3 e Corolário 14.9. Para *da esquerda para a direita*: suponha-se A enumerável; então, por Definição 5.2, há três casos possíveis:

1. se $A = \emptyset$, então $|A| = 0 \in \omega$, por Corolário 14.6 e Lema 14.3.
2. se $n \approx A$, então $|A| = n \in \omega$, por Corolário 14.6 e Lema 14.3.
3. se $\omega \approx A$, então $|A| = \omega$, por Corolário 14.9.

Em todos os casos, $|A| \leq \omega$. □

De fato, ω tem um lugar especial. Embora existam muitos ordinais contáveis:

Corolário 14.12. *ω é o único cardinal infinito enumerável.*

Prova. Seja $\mathfrak{a}$ um cardinal infinito enumerável. Como $\mathfrak{a}$ é infinito, $\omega \leq \mathfrak{a}$. Como $\mathfrak{a}$ é um cardinal infinito enumerável, $\mathfrak{a} = |\mathfrak{a}| \leq \omega$. Logo $\mathfrak{a} = \omega$ por tricotomia. □

Claro que existem infinitos cardinais. Poderíamos perguntar: *Quantos cardinais há?* Os resultados seguintes mostram que convém reconsiderar essa pergunta.

Proposição 14.13. *Se todo o membro de X é um cardinal, então $\bigcup X$ é um cardinal.*

Prova. É fácil verificar que $\bigcup X$ é um ordinal. Seja $\alpha \in \bigcup X$ um ordinal; então $\alpha \in \mathfrak{b} \in X$ para algum cardinal $\mathfrak{b}$. Como $\mathfrak{b}$ é um cardinal, $\alpha \prec \mathfrak{b}$. Como $\mathfrak{b} \subseteq \bigcup X$, tem-se $\mathfrak{b} \preceq \bigcup X$, e portanto $\alpha \not\approx \bigcup X$. Generalizando, $\bigcup X$ é um cardinal. □

Teorema 14.14. *Não existe um cardinal máximo.*

Prova. Para qualquer cardinal $\mathfrak{a}$, o teorema de Cantor (Teorema 5.16) e Lema 14.2 implicam que $\mathfrak{a} < |\wp(\mathfrak{a})|$. □

Teorema 14.15. *O conjunto de todos os cardinais não existe.*

Prova. Por redução ao absurdo, suponha-se $C = \{\mathfrak{a} : \mathfrak{a} \text{ é um cardinal}\}$. Ora, $\bigcup C$ é um cardinal por Proposição 14.13, logo por Teorema 14.14 existe um cardinal $\mathfrak{b} > \bigcup C$. Por definição $\mathfrak{b} \in C$, logo $\mathfrak{b} \subseteq \bigcup C$, donde $\mathfrak{b} \leq \bigcup C$, uma contradição. □

Compare isto com o paradoxo de Russell e com Burali-Forti.

14.5 Apêndice: Princípio de Hume

Na Seção 14.1, descrevemos o Princípio de Cantor. Era:

$$|A| = |B| \text{ sse } A \approx B.$$

Isto é muito semelhante ao que hoje se chama *Princípio de Hume*, que diz:

$$\#x\, F(x) = \#x\, G(x) \text{ sse } F \sim G$$

onde '$F \sim G$' abrevia que há tantos F quantos G, isto é, os F podem pôr-se em bijeção com os G, isto é:

$$\begin{aligned}\exists R(&\forall v \forall y (Rvy \rightarrow (Fv \wedge Gy)) \wedge \\ &\forall v (Fv \rightarrow \exists! y\, Rvy) \wedge \\ &\forall y (Gy \rightarrow \exists! v\, Rvy))\end{aligned}$$

Mas há uma diferença de tipo entre o Princípio de Hume e o Princípio de Cantor. No enunciado do Princípio de Cantor, as variáveis "A" e "B" são termos de primeira ordem que denotam *conjuntos*. No enunciado do Princípio de Hume, "F", "G" e "R" *não* são termos de primeira ordem; estão em *posição de predicado*. (Talvez denotem *propriedades*.) Assim, o Princípio de Hume pode parafrasear-se em português: o número de Fs é o número de Gs sse os Fs estão em bijeção com os Gs. Chama-se *Princípio de Hume* porque Hume escreveu o seguinte:

> Quando dois números se combinam de tal modo que a cada unidade de um corresponde sempre uma unidade do outro, dizemos que são iguais. (Hume, 1740, Pt.III Bk.1 §1)

E o Princípio de Hume foi trazido à proeminência matemático-lógica contemporânea por Frege (1884, §63), que citou esta passagem de Hume antes de (em efeito) esboçar (o que chamamos) Princípio de Hume.

Note a semelhança estrutural entre o Princípio de Hume e a Lei Básica V. A formulamos na Seção 8.6 do seguinte modo:

$$\epsilon x\, F(x) = \epsilon x\, G(x) \text{sse } \forall x\, (F(x) \leftrightarrow G(x)).$$

E, neste ponto, alguns comentários e comparações podem ajudar.

Há duas maneiras de entender um princípio como o Princípio de Hume ou a Lei Básica V: *predicativamente* ou *impredicativamente* (recorde Seção 8.3). Na leitura impredicativa da Lei Básica V, para cada F, o objeto $\epsilon x\, F(x)$ pertence ao domínio de quantificação usado na própria formulação da Lei Básica V. Do

mesmo modo, na leitura impredicativa do Princípio de Hume, para cada F, o objeto $\#x\,F(x)$ pertence ao domínio de quantificação usado na formulação do Princípio de Hume. Em contraste, na compreensão *predicativa*, os objetos $\epsilon x\,F(x)$ e $\#x\,F(x)$ seriam entidades de um domínio *diferente*.

Agora, se a Lei Básica V se lê impredicativamente, conduz à inconsistência, via Compreensão ingênua (para pormenores, ver Seção 8.6). Tal como a Compreensão ingênua, pode tornar-se consistente se se lê *predicativamente*. Mas provavelmente não fará tudo o que desejávamos.

O Princípio de Hume, porém, *pode* ser lido de modo consistente impredicativamente. E, assim lido, é bastante poderoso.

Para ilustrar: considere-se o predicado "$x \neq x$", que obviamente nada satisfaz. O Princípio de Hume fornece agora um objeto $\#x(x \neq x)$. Podemos tratá-lo como o número 0. Agora, na compreensão *impredicativa*—mas *só* na impredicativa—esta entidade 0 pertence ao nosso domínio original de quantificação. Logo podemos aplicar sensatamente o Princípio de Hume com o predicado "$x = 0$" para obter um objeto $\#x(x = 0)$. Podemos tratá-lo como o número 1. Além disso, o Princípio de Hume implica que $0 \neq 1$, pois não pode haver bijeção dos objetos não autoidênticos para os objetos idênticos a 0 (não há dos primeiros, mas há um dos segundos). Agora, de novo impredicativamente, 1 pertence ao domínio original de quantificação. Logo podemos aplicar o Princípio de Hume com o predicado "$(x = 0 \vee x = 1)$" para obter um objeto $\#x(x = 0 \vee x = 1)$. Podemos tratá-lo como o número 2, e mostrar que $0 \neq 2$ e $1 \neq 2$, e assim por diante.

Em suma, tomado impredicativamente, o Princípio de Hume implica que há *infinitamente muitos objetos*. E isto encorajou os *logicistas neo-fregeanos* a tomar o Princípio de Hume como fundação da aritmética.

O próprio Frege, contudo, não tomou o Princípio de Hume como fundação da aritmética. Em vez disso, Frege provou o Princípio de Hume a partir de uma definição explícita: $\#x\,F(x)$ é definido como a extensão do conceito $F \sim \Phi$. Em termos modernos, poder-se-ia tentar render isto como $\#x\,F(x) = \{G : F \sim G\}$;

mas isto volta a arrastar-nos para os problemas da Compreensão ingênua.

CAPÍTULO 15

Aritmética Cardinal

No Capítulo 14, desenvolvemos uma teoria dos cardinais. O passo seguinte é esboçar uma teoria da aritmética cardinal. Este capítulo resume brevemente alguns fatos elementares e depois aponta algumas das dificuldades, que se revelam fascinantes e filosoficamente densas.

15.1 Definindo as Operações Básicas

Como não precisamos de controlar a ordem, a aritmética cardinal é bastante mais fácil de definir do que a aritmética ordinal. Definiremos adição, multiplicação e exponenciação simultaneamente.

Definição 15.1. Quando $\mathfrak{a}$ e $\mathfrak{b}$ são cardinais:

$$\mathfrak{a} \oplus \mathfrak{b} = |\mathfrak{a} \sqcup \mathfrak{b}|$$
$$\mathfrak{a} \otimes \mathfrak{b} = |\mathfrak{a} \times \mathfrak{b}|$$
$$\mathfrak{a}^{\mathfrak{b}} = \left|{}^{\mathfrak{b}}\mathfrak{a}\right|$$

onde ${}^{X}Y = \{f : f \text{ é uma função } X \to Y\}$. (É fácil mostrar que ${}^{X}Y$ existe para quaisquer conjuntos X e Y; deixamos isto como exercício.)

Pode ajudar explicar esta definição. Quanto à adição: usa-se a noção de soma disjunta, $\sqcup$, tal como definida em Definição 13.1; e é fácil ver que esta definição dá o resultado certo nos casos finitos. Quanto à multiplicação: Proposição 2.27 nos diz que, se A tem n membros e B tem m membros, então $A \times B$ tem $n \cdot m$ membros, pelo que a nossa definição generaliza simplesmente a ideia à multiplicação transfinita. A exponenciação é semelhante: estamos simplesmente a generalizar o pensamento do finito para o transfinito. De fato, sob certos aspectos, a aritmética cardinal transfinita parece muito mais a aritmética "ordinária" do que a aritmética ordinal transfinita:

Proposição 15.2. *$\oplus$ e $\otimes$ são comutativas e associativas.*

Prova. Para a comutatividade, por Lema 14.3 basta observar que $(\mathfrak{a} \sqcup \mathfrak{b}) \approx (\mathfrak{b} \sqcup \mathfrak{a})$ e $(\mathfrak{a} \times \mathfrak{b}) \approx (\mathfrak{b} \times \mathfrak{a})$. Deixamos a associatividade como exercício. □

Proposição 15.3. *A é infinito se e só se $|A| \oplus 1 = 1 \oplus |A| = |A|$.*

Prova. Como em Teorema 14.7, a partir de Lema 13.10 e Lema 14.3. □

Isto explica porque precisamos de símbolos distintos para adição/multiplicação ordinal versus cardinal: são operações genuinamente *diferentes*. Este par de resultados mostra que a exponenciação ordinal versus cardinal também são operações distintas. (Recordar que Definição 9.7 implica que $2 = \{0,1\}$):

Lema 15.4. $|\wp(A)| = 2^{|A|}$, *para qualquer* A.

Prova. Para cada subconjunto $B \subseteq A$, seja $\chi_B \in {}^A2$ dado por:

$$\chi_B(x) = \begin{cases} 1 & \text{se } x \in B \\ 0 & \text{caso contrário.} \end{cases}$$

Seja agora $f(B) = \chi_B$; isto define uma bijeção $f : \wp(A) \to {}^A2$. Logo $\wp(A) \approx {}^A2$. Daí $\wp(A) \approx {}^{|A|}2$, pelo que $|\wp(A)| = \left|{}^{|A|}2\right| = 2^{|A|}$. □

Esta prova sucinta subsume essencialmente a discussão em Seção 5.6. Ali, mostramos como "reduzir" a não enumerabilidade de $\wp(\omega)$ à não enumerabilidade do conjunto das sequências binárias infinitas, $\mathbb{B}^\omega$. Na prática, $\mathbb{B}^\omega$ é apenas ${}^\omega 2$; e a prova precedente mostrou que o raciocínio que seguimos na Seção 5.6 continua a valer usando qualquer conjunto A no lugar de ω. O resultado também fornece um fato rápido sobre exponenciação cardinal:

Corolário 15.5. $\mathfrak{a} < 2^{\mathfrak{a}}$ *para qualquer cardinal* $\mathfrak{a}$.

Prova. Pelo Teorema de Cantor (Teorema 5.16) e Lema 15.4. □

Logo $\omega < 2^\omega$. Mas nota: isto é um resultado sobre exponenciação *cardinal*. Deve contrastar-se com a exponenciação *ordinal*, pois neste último caso $\omega = 2^{(\omega)}$ (ver Seção 13.5).

Enquanto estamos no tema da exponenciação cardinal, podemos também ser um pouco mais precisos sobre a "maneira" pela qual $\mathbb{R}$ é não enumerável.

Teorema 15.6. $|\mathbb{R}| = 2^\omega$

Esboço de prova. Há muitas maneiras de provar isto. A mais direta é argumentar que $\wp(\omega) \preceq \mathbb{R}$ e $\mathbb{R} \preceq \wp(\omega)$, e depois usar Schröder-Bernstein para inferir que $\mathbb{R} \approx \wp(\omega)$, e Lema 15.4 para inferir que $|\mathbb{R}| = 2^\omega$. Deixamos como exercício (esclarecedor) definir injecções $f : \wp(\omega) \to \mathbb{R}$ e $g : \mathbb{R} \to \wp(\omega)$. □

15.2 Simplificando Adição e Multiplicação

Acontece que a adição e a multiplicação cardinais transfinitas são *extremamente* fáceis. Isto segue do fato de os cardinais serem (certos) ordinais, logo bem ordenados, e poderem ser manipulados de certa forma. Mostrar isto, porém, *não* é tão fácil. Para começar, precisamos de uma definição hábil:

Definição 15.7. Definimos uma *ordenação canónica*, $\lhd$, em pares de ordinais, estipulando que $\langle \alpha_1, \alpha_2 \rangle \lhd \langle \beta_1, \beta_2 \rangle$ se e só se se verifica uma das seguintes condições:

1. $\max(\alpha_1, \alpha_2) < \max(\beta_1, \beta_2)$; ou
2. $\max(\alpha_1, \alpha_2) = \max(\beta_1, \beta_2)$ e $\alpha_1 < \beta_1$; ou
3. $\max(\alpha_1, \alpha_2) = \max(\beta_1, \beta_2)$ e $\alpha_1 = \beta_1$ e $\alpha_2 < \beta_2$

Lema 15.8. *$\langle \alpha \times \alpha, \lhd \rangle$ é uma boa ordem, para qualquer ordinal α.*

Prova. Evidentemente $\lhd$ é conexa em $\alpha \times \alpha$. De fato, suponhamos que nem $\langle \alpha_1, \alpha_2 \rangle$ nem $\langle \beta_1, \beta_2 \rangle$ é menor do que a outra no sentido de $\lhd$. Então $\max(\alpha_1, \alpha_2) = \max(\beta_1, \beta_2)$ e $\alpha_1 = \beta_1$ e $\alpha_2 = \beta_2$, pelo que $\langle \alpha_1, \alpha_2 \rangle = \langle \beta_1, \beta_2 \rangle$.

Para mostrar que é boa ordem, seja $X \subseteq \alpha \times \alpha$ não vazio. Como α é um ordinal, algum δ é o menor membro de $\{\max(\gamma_1, \gamma_2) : \langle \gamma_1, \gamma_2 \rangle \in X\}$. Descartemos agora todos os pares de $\{\langle \gamma_1, \gamma_2 \rangle \in X : \max(\gamma_1, \gamma_2) = \delta\}$ exceto os de primeira coordenada mínima; entre estes, o par de segunda coordenada mínima é o elemento $\lhd$-mínimo de X. □

Segue-se uma observação mínima, mas simples:

Proposição 15.9. *Se $\alpha \approx \beta$, então $\alpha \times \alpha \approx \beta \times \beta$.*

Prova. Basta deixar que $f : \alpha \to \beta$ induza $\langle \gamma_1, \gamma_2 \rangle \mapsto \langle f(\gamma_1), f(\gamma_2) \rangle$. □

E agora aplicaremos tudo isto para provar um lema crucial:

Lema 15.10. *$\alpha \approx \alpha \times \alpha$, para qualquer ordinal infinito α*

Prova. Por redução ao absurdo, seja α o menor ordinal infinito para o qual isto falha. Proposição 5.6 mostra que $\omega \approx \omega \times \omega$, logo $\omega \in \alpha$. Além disso, α é um cardinal: suponhamos o contrário, por redução ao absurdo; então $|\alpha| \in \alpha$, pelo que $|\alpha| \approx |\alpha| \times |\alpha|$, por hipótese; e $|\alpha| \approx \alpha$ por definição; logo $\alpha \approx \alpha \times \alpha$ por Proposição 15.9.

Ora, para cada $\langle \gamma_1, \gamma_2 \rangle \in \alpha \times \alpha$, consideremos o segmento:

$$\mathrm{Seg}(\gamma_1, \gamma_2) = \{\langle \delta_1, \delta_2 \rangle \in \alpha \times \alpha : \langle \delta_1, \delta_2 \rangle \lhd \langle \gamma_1, \gamma_2 \rangle\}$$

Seja $\gamma = \max(\gamma_1, \gamma_2)$; note-se que $\langle \gamma_1, \gamma_2 \rangle \lhd \langle \gamma + 1, \gamma + 1 \rangle$. Assim, quando γ é infinito, observe-se:

$$\begin{aligned} \mathrm{Seg}(\gamma_1, \gamma_2) &\precsim ((\gamma + 1) \cdot (\gamma + 1)) \\ &\approx (\gamma \cdot \gamma), \text{ por Lema 13.10 e Proposição 15.9} \\ &\approx \gamma, \text{ pela hipótese de indução} \\ &\prec \alpha, \text{ pois } \alpha \text{ é um cardinal} \end{aligned}$$

Logo $\mathrm{ord}(\alpha \times \alpha, \lhd) \leq \alpha$, e portanto $\alpha \times \alpha \preceq \alpha$. Como obviamente $\alpha \preceq \alpha \times \alpha$, o resultado segue por Schröder-Bernstein. □

Finalmente, chegamos ao nosso resultado simplificador:

Teorema 15.11. *Se $\mathfrak{a}, \mathfrak{b}$ são cardinais infinitos, então:*

$$\mathfrak{a} \otimes \mathfrak{b} = \mathfrak{a} \oplus \mathfrak{b} = \mathit{max}(\mathfrak{a}, \mathfrak{b}).$$

Prova. Sem perda de generalidade, suponhamos $\mathfrak{a} = \max(\mathfrak{a}, \mathfrak{b})$. Então, invocando Lema 15.10, $\mathfrak{a} \otimes \mathfrak{a} = \mathfrak{a} \leq \mathfrak{a} \oplus \mathfrak{b} \leq \mathfrak{a} \oplus \mathfrak{a} \leq \mathfrak{a} \otimes \mathfrak{a}$.□

Do mesmo modo, se $\mathfrak{a}$ é infinito, uma união de cardinalidade $\mathfrak{a}$ de conjuntos de cardinalidade $\leq \mathfrak{a}$ tem cardinalidade $\leq \mathfrak{a}$:

Proposição 15.12. *Seja $\mathfrak{a}$ um cardinal infinito. Para cada ordinal $\beta \in \mathfrak{a}$, seja X_β um conjunto com $\left|X_\beta\right| \leq \mathfrak{a}$. Então $\left|\bigcup_{\beta\in\mathfrak{a}} X_\beta\right| \leq \mathfrak{a}$.*

Prova. Para cada $\beta \in \mathfrak{a}$, fixe uma injeção $f_\beta : X_\beta \to \mathfrak{a}$.[1] Defina uma injeção $g : \bigcup_{\beta\in\mathfrak{a}} X_\beta \to \mathfrak{a}\times\mathfrak{a}$ por $g(v) = \langle\beta, f_\beta(v)\rangle$, onde $v \in X_\beta$ e $v \notin X_\gamma$ para qualquer $\gamma \in \beta$. Ora, $\bigcup_{\beta\in\mathfrak{a}} X_\beta \preceq \mathfrak{a}\times\mathfrak{a} \approx \mathfrak{a}$ por Teorema 15.11. □

15.3 Algumas Simplificações com Exponenciação Cardinal

Embora definir $\lhd$ fosse um pouco trabalhoso, o resultado é um teorema útil sobre adição e multiplicação cardinais, Teorema 15.11. A exponenciação transfinita, porém, não se simplifica tão diretamente. Para explicar por que, começamos por um resultado que prolonga um padrão familiar do caso finitário (embora a sua prova esteja a um nível elevado de abstração):

Proposição 15.13. $\mathfrak{a}^{\mathfrak{b}\oplus\mathfrak{c}} = \mathfrak{a}^{\mathfrak{b}} \otimes \mathfrak{a}^{\mathfrak{c}}$ *e* $(\mathfrak{a}^{\mathfrak{b}})^{\mathfrak{c}} = \mathfrak{a}^{\mathfrak{b}\otimes\mathfrak{c}}$*, para quaisquer cardinais* $\mathfrak{a}, \mathfrak{b}, \mathfrak{c}$.

Prova. Para a primeira afirmação, considere-se uma função $f : (\mathfrak{b} \sqcup \mathfrak{c}) \to \mathfrak{a}$. Ora, "parta-se" isto, definindo $f_{\mathfrak{b}}(\beta) = f(\beta, 0)$ para cada $\beta \in \mathfrak{b}$, e $f_{\mathfrak{c}}(\gamma) = f(\gamma, 1)$ para cada $\gamma \in \mathfrak{c}$. A aplicação $f \mapsto (f_{\mathfrak{b}} \times f_{\mathfrak{c}})$ é uma bijeção ${}^{\mathfrak{b}\sqcup\mathfrak{c}}\mathfrak{a} \to ({}^{\mathfrak{b}}\mathfrak{a} \times {}^{\mathfrak{c}}\mathfrak{a})$.

Para a segunda afirmação, considere-se uma função $f : \mathfrak{c} \to ({}^{\mathfrak{b}}\mathfrak{a})$; assim, para cada $\gamma \in \mathfrak{c}$ temos alguma função $f(\gamma) : \mathfrak{b} \to \mathfrak{a}$. Ora, defina-se $f^*(\beta, \gamma) = (f(\gamma))(\beta)$ para cada $\langle\beta, \gamma\rangle \in \mathfrak{b} \times \mathfrak{c}$. A aplicação $f \mapsto f^*$ é uma bijeção ${}^{\mathfrak{c}}({}^{\mathfrak{b}}\mathfrak{a}) \to {}^{\mathfrak{b}\otimes\mathfrak{c}}\mathfrak{a}$. □

O que gostaríamos *mesmo* era de uma maneira fácil de calcular $\mathfrak{a}^{\mathfrak{b}}$ quando tratamos de cardinais infinitos. Eis um passo útil nessa direção:

[1] Como são estes "fixos"? Ver Seção 16.5.

Proposição 15.14. *Se* $2 \leq \mathfrak{a} \leq \mathfrak{b}$ *e* $\mathfrak{b}$ *é infinito, então* $\mathfrak{a}^{\mathfrak{b}} = 2^{\mathfrak{b}}$

Prova.

$$\begin{aligned} 2^{\mathfrak{b}} &\leq \mathfrak{a}^{\mathfrak{b}}, \text{ pois } 2 \leq \mathfrak{a} \\ &\leq (2^{\mathfrak{a}})^{\mathfrak{b}}, \text{ por Lema 15.4} \\ &= 2^{\mathfrak{a} \otimes \mathfrak{b}}, \text{ por Proposição 15.13} \\ &= 2^{\mathfrak{b}}, \text{ por Teorema 15.11} \end{aligned}$$ □

Não devemos esperar poder simplificar mais além disto, pois $\mathfrak{b} < 2^{\mathfrak{b}}$ por Lema 15.4. Contudo, isto não nos diz o que afirmar sobre $\mathfrak{a}^{\mathfrak{b}}$ quando $\mathfrak{b} < \mathfrak{a}$. Naturalmente, se $\mathfrak{b}$ é *finito*, sabemos o que fazer.

Proposição 15.15. *Se* $\mathfrak{a}$ *é infinito e* $n \in \omega$*, então* $\mathfrak{a}^n = \mathfrak{a}$

Prova. $\mathfrak{a}^n = \mathfrak{a} \otimes \mathfrak{a} \otimes \ldots \otimes \mathfrak{a} = \mathfrak{a}$, por Teorema 15.11. □

Adicionalmente, nalguns outros casos, podemos controlar o tamanho de $\mathfrak{a}^{\mathfrak{b}}$:

Proposição 15.16. *Se* $2 \leq \mathfrak{b} < \mathfrak{a} \leq 2^{\mathfrak{b}}$ *e* $\mathfrak{b}$ *é infinito, então* $\mathfrak{a}^{\mathfrak{b}} = 2^{\mathfrak{b}}$

Prova. $2^{\mathfrak{b}} \leq \mathfrak{a}^{\mathfrak{b}} \leq (2^{\mathfrak{b}})^{\mathfrak{b}} = 2^{\mathfrak{b} \otimes \mathfrak{b}} = 2^{\mathfrak{b}}$, raciocinando como na Proposição 15.14. □

Para além deste ponto, as coisas tornam-se bastante mais subtis.

15.4 A Hipótese do Contínuo

O resultado anterior sugere (corretamente) que a exponenciação cardinal seria bastante *fácil*, se os cardinais infinitos estivessem garantidos a "comportar-se de forma simples" com potências de 2, isto é, (por Lema 15.4) com a operação de tomar conjuntos potência. Mas não podemos supor que os cardinais infinitos *se comportem* de forma simples com os conjuntos potência.

Começando a desdobrar isto, introduzimos uma notação conveniente.

Definição 15.17. Onde $\mathfrak{a}^{\oplus}$ é o menor cardinal estritamente maior do que $\mathfrak{a}$, definimos duas sequências infinitas:

$$\begin{aligned} \aleph_0 &= \omega & \beth_0 &= \omega \\ \aleph_{\alpha+1} &= (\aleph_\alpha)^{\oplus} & \beth_{\alpha+1} &= 2^{\beth_\alpha} \\ \aleph_\alpha &= \bigcup_{\beta<\alpha} \aleph_\beta & \beth_\alpha &= \bigcup_{\beta<\alpha} \beth_\beta \quad \text{quando } \alpha \text{ é um ordinal limite.} \end{aligned}$$

A definição de $\mathfrak{a}^{\oplus}$ faz sentido, pois Teorema 14.14 nos diz que, para cada cardinal $\mathfrak{a}$, existe algum cardinal maior do que $\mathfrak{a}$, e a Indução Transfinita garante que existe um cardinal *mínimo* maior do que $\mathfrak{a}$. O resto da definição é fornecido pela recursão transfinita.

Cantor introduziu esta notação "$\aleph$"; trata-se do *alef*, a primeira letra do alfabeto hebraico e a primeira letra da palavra hebraica para "infinito". Peirce introduziu a notação "$\beth$"; trata-se do *beth*, a segunda letra do alfabeto hebraico.[2] Estas notações fornecem-nos cardinais infinitos.

Proposição 15.18. *$\aleph_\alpha$ e $\beth_\alpha$ são cardinais, para todo o ordinal α.*

Prova. Ambos os resultados valem por uma indução transfinita simples. $\aleph_0 = \beth_0 = \omega$ é um cardinal por Corolário 14.9. Supondo que $\aleph_\alpha$ e $\beth_\alpha$ são ambos cardinais, $\aleph_{\alpha+1}$ e $\beth_{\alpha+1}$ estão explicitamente definidos como cardinais. E a união de um conjunto de cardinais é um cardinal, por Proposição 14.13. □

Além disso, todo o cardinal infinito é um $\aleph$:

[2] Peirce usou esta notação numa carta a Cantor de dezembro de 1900. Infelizmente, Peirce também apresentou ali um mau argumento de que $\beth_\alpha$ não existe para $\alpha \geq \omega$.

Proposição 15.19. *Se $\mathfrak{a}$ é um cardinal infinito, então $\mathfrak{a} = \aleph_\gamma$ para um certo γ único.*

Prova. Por indução transfinita nos cardinais. Para a indução, suponhamos que se $\mathfrak{b} < \mathfrak{a}$ então $\mathfrak{b} = \aleph_{\gamma_\mathfrak{b}}$. Se $\mathfrak{a} = \mathfrak{b}^\oplus$ para algum $\mathfrak{b}$, então $\mathfrak{a} = (\aleph_{\gamma_\mathfrak{b}})^\oplus = \aleph_{\gamma_\mathfrak{b}+1}$. Se $\mathfrak{a}$ não é o sucessor de nenhum cardinal, então, por serem os cardinais ordinais, $\mathfrak{a} = \bigcup_{\mathfrak{b}<\mathfrak{a}} \mathfrak{b} = \bigcup_{\mathfrak{b}<\mathfrak{a}} \aleph_{\gamma_\mathfrak{b}}$, logo $\mathfrak{a} = \aleph_\gamma$ onde $\gamma = \bigcup_{\mathfrak{b}<\mathfrak{a}} \gamma_\mathfrak{b}$. □

Como todo o cardinal infinito é um $\aleph$, surge naturalmente a pergunta: será que todo o cardinal infinito é um $\beth$? Certamente que, se isso *se verificasse*, os cardinais infinitos "comportar-se-iam de forma simples" com a operação de tomar conjuntos potência. De fato, teríamos o seguinte:

Hipótese Generalizada do Contínuo (HGC). $\aleph_\alpha = \beth_\alpha$, para todo o α.

Além disso, se a HGC valesse, poderíamos simplificar bastante a exponenciação cardinal. Em particular, poderíamos mostrar que, quando $\mathfrak{b} < \mathfrak{a}$, o valor de $\mathfrak{a}^\mathfrak{b}$ fica encerrado entre $\mathfrak{a} \leq \mathfrak{a}^\mathfrak{b} \leq \mathfrak{a}^\oplus$. Poderíamos então dar condições precisas que determinam qual das duas possibilidades se verifica (isto é, se $\mathfrak{a} = \mathfrak{a}^\mathfrak{b}$ ou $\mathfrak{a}^\mathfrak{b} = \mathfrak{a}^\oplus$).[3]

Mas a HGC é uma *hipótese*, não um *teorema*. De fato, Gödel (1938) provou que, se **ZFC** for consistente, então também o é **ZFC** + HGC. Mas mais tarde verificou-se que podemos acrescentar igualmente ¬HGC a **ZFC**. De fato, considere-se a *instância* não trivial mais simples da HGC, nomeadamente:

Hipótese do Contínuo (HC). $\aleph_1 = \beth_1$.

Cohen (1963) provou que, se **ZFC** for consistente, então também o é **ZFC** + ¬HC. Logo a Hipótese do Contínuo é independente de **ZFC**.

[3] A condição é ditada pela *cofinalidade*.

A Hipótese do Contínuo chama-se assim porque "o contínuo" é outro nome para a reta real, $\mathbb{R}$. Teorema 15.6 nos diz que $|\mathbb{R}| = \beth_1$. Assim, a Hipótese do Contínuo afirma que não há cardinal entre a cardinalidade dos números naturais, $\aleph_0 = \beth_0$, e a cardinalidade do contínuo, $\beth_1$.

Dada a *independência* da (H)HC relativamente a **ZFC**, o que devemos dizer sobre a sua *verdade*? Há *muito* a dizer. De fato, muito trabalho recente e fecundo em teoria dos conjuntos tem sido dedicado a estas questões. Mas dois pontos rápidos valem certamente a pena enfatizar.

Primeiro: não se segue *imediatamente* destes resultados formais de independência que a HGC ou a HC sejam *indeterminadas* no valor de verdade. Afinal, talvez bastem mais axiomas, que nos pareçam naturais, para decidir a questão num sentido ou noutro. O próprio Gödel sugeriu que essa era a resposta adequada.

Segundo: a independência da HC relativamente a **ZFC** é certamente *notável*, mas não é *inacreditável* (no sentido literal). O ponto é simplesmente que, para tudo o que **ZFC** nos diz, passar de um cardinal ao seu sucessor pode envolver uma ferramenta menos direta do que simplesmente tomar conjuntos potência.

Feitas estas duas observações, quem quiser saber mais terá de recorrer aos vários filósofos e matemáticos com cavalos na corrida.[4]

15.5 Pontos Fixos de $\aleph$

No Capítulo 11, sugerimos que o Esquema de Substituição carece de justificação, pois força a hierarquia a ser bastante alta. Tendo feito alguma aritmética cardinal, podemos dar uma pequena ilustração da altura da hierarquia.

Evidentemente $0 < \aleph_0$, e $1 < \aleph_1$, e $2 < \aleph_2 \ldots$ e, de fato, a diferença de tamanho só fica *maior* a cada passo. Assim, é tentador conjecturar que $\kappa < \aleph_\kappa$ para todo o ordinal κ.

[4] Embora possa querer começar por ler Potter (2004, §15.6).

Mas esta conjectura é *falsa*, em **ZFC**. De fato, podemos provar que existem *pontos fixos de* $\aleph$, isto é, cardinais κ tais que $\kappa = \aleph_\kappa$.

Proposição 15.20. *Existe um ponto fixo de* $\aleph$.

Prova. Por recursão, defina-se:

$$\begin{aligned} \kappa_0 &= 0 \\ \kappa_{n+1} &= \aleph_{\kappa_n} \\ \kappa &= \bigcup_{n<\omega} \kappa_n \end{aligned}$$

Ora, κ é um cardinal por Proposição 14.13. Mas então:

$$\kappa = \bigcup_{n<\omega} \kappa_{n+1} = \bigcup_{n<\omega} \aleph_{\kappa_n} = \bigcup_{\alpha<\kappa} \aleph_\alpha = \aleph_\kappa$$

□

Boolos escreveu um artigo sobre precisamente o ponto fixo de $\aleph$ que acabamos de construir. Depois de notar a existência de κ, no início do artigo, disse:

> [κ é] um número *bastante grande*, aos olhos de quem não teve contato prévio com a teoria dos conjuntos, tão grande, me parece, que põe em causa a verdade de qualquer teoria que, entre as suas afirmações, inclua a de que existem pelo menos κ objetos. (Boolos, 2000, p. 257)

E concluiu o artigo perguntando:

> [será que] suspeitamos que, seja qual for o começo da história, quando chegamos tão longe as rodas giram no vazio e já não estamos a ouvir uma descrição de coisa alguma que seja o caso? (Boolos, 2000, p. 268)

Se, de fato, ultrapassamos "qualquer coisa que seja o caso", então devemos apontar o dedo da culpa diretamente ao Esquema de Substituição. Pois é este axioma que permite a nossa prova

funcionar. Nesse caso, supõe-se, Boolos teria de rever a afirmação que fizera, décadas antes, de que a Substituição não tem consequências "indesejáveis" (ver Seção 12.3).

Mas será a existência de κ assim tão má? Pode ajudar, aqui, considerar o *paradoxo de Tristram Shandy* de Russell. Tristram Shandy documenta a sua vida no diário, mas demora um ano a registrar um único dia. A cada ano que passa, Tristram fica cada vez mais atrasado: passado um ano, só registrou um dia, e viveu 364 dias por registrar; passados dois anos, só registrou dois dias, e viveu 728 dias por registrar; passados três anos, só registrou três dias, e viveu 1092 dias por registrar ...[5] Ainda assim, se Tristram é *imortal*, conseguirá registrar todos os dias, pois registrará o *n*-ésimo dia no *n*-ésimo ano da sua vida. E assim, "no fim dos tempos", Tristram terá um diário completo.

Porque é que isto difere tanto da ideia de que α é menor do que $\aleph_\alpha$—e, de fato, cada vez mais, desesperadamente menor—até κ, onde nos "apanhamos" e $\kappa = \aleph_\kappa$?

Deixando isto de lado, e supondo que aceitamos **ZFC**, fechemos com um pouco mais de diversão sobre construções de pontos fixos. Os três resultados seguintes estabelecem, intuitivamente, que existe um ponto (não trivial) em que a hierarquia é tão larga como alta:

Proposição 15.21. *Existe um ponto fixo de* $\beth$*, isto é, um* κ *tal que* $\kappa = \beth_\kappa$.

Prova. Como na Proposição 15.20, usando "$\beth$" no lugar de "$\aleph$".□

Proposição 15.22. $|V_{\omega+\alpha}| = \beth_\alpha$. *Se* $\omega \cdot \omega \leq \alpha$*, então* $|V_\alpha| = \beth_\alpha$.

Prova. A primeira afirmação vale por uma indução transfinita simples. A segunda segue-se, pois se $\omega \cdot \omega \leq \alpha$ então $\omega + \alpha = \alpha$. Para o estabelecer, usamos fatos sobre aritmética ordinal de Capítulo 13. Note-se primeiro que $\omega \cdot \omega = \omega \cdot (1 + \omega) = (\omega \cdot 1) +$

[5] Ignorando anos bissextos.

$(\omega \cdot \omega) = \omega + (\omega \cdot \omega)$. Ora, se $\omega \cdot \omega \leq \alpha$, isto é, $\alpha = (\omega \cdot \omega) + \beta$ para algum β, então $\omega + \alpha = \omega + ((\omega \cdot \omega) + \beta) = (\omega + (\omega \cdot \omega)) + \beta = (\omega \cdot \omega) + \beta = \alpha$. □

Corolário 15.23. *Existe um κ tal que $|V_\kappa| = \kappa$.*

Prova. Seja κ um ponto fixo de $\beth$, dado por Proposição 15.21. Claramente $\omega \cdot \omega < \kappa$. Logo $|V_\kappa| = \beth_\kappa = \kappa$ por Proposição 15.22.□

Há tantos estágios abaixo de V_κ quantos membros de V_κ. Intuitivamente, então, V_κ é tão largo como alto. Isto é muito à maneira de Tristram Shandy: passamos de um estágio ao seguinte tomando *conjuntos potência*, tornando a hierarquia *muito* maior a cada passo. Mas "no fim", isto é, no estágio κ, a largura da hierarquia alcança a altura.

Alguém pode perguntar: *Com que frequência a largura da hierarquia coincide com a altura?* A resposta é: *Tantas vezes quantos os ordinais.* Mas isto precisa de uma pequena explicação.

Definimos um termo τ do seguinte modo. Para qualquer A, seja:

$$\begin{aligned} \tau_0(A) &= |A| \\ \tau_{n+1}(A) &= \beth_{\tau_n(A)} \\ \tau(A) &= \bigcup_{n<\omega} \tau_n(A) \end{aligned}$$

Como na Proposição 15.21, $\tau(A)$ é um ponto fixo de $\beth$ para qualquer A, e trivialmente $|A| < \tau(A)$. Considere-se então esta definição recursiva:

$$\begin{aligned} W_0 &= 0 \\ W_{\alpha+1} &= \tau(W_\alpha) \\ W_\alpha &= \bigcup_{\beta<\alpha} W_\beta, \text{ quando } \alpha \text{ é limite} \end{aligned}$$

A construção está definida para todos os ordinais. Intuitivamente, então, W é "uma injeção" dos ordinais nos pontos fixos de $\beth$. E, exatamente como antes, V_{W_α} é tão largo como alto, para qualquer α.

Problemas

Problema 15.1. Prove em $\mathbf{Z}^-$ que ${}^X Y$ existe para quaisquer conjuntos X e Y. Trabalhando em **ZF**, calcule $\mathrm{posto}({}^X Y)$ a partir de $\mathrm{posto}(X)$ e $\mathrm{posto}(Y)$, à maneira de Lema 13.9.

Problema 15.2. Prove que $\oplus$ e $\otimes$ são associativas.

Problema 15.3. Complete a prova de Teorema 15.6, mostrando que $\wp(\omega) \preceq \mathbb{R}$ e $\mathbb{R} \preceq \wp(\omega)$.

CAPÍTULO 16

Escolha

16.1 Introdução

Em Capítulos 14 a 15, desenvolvemos uma teoria de cardinais tratando os cardinais como ordinais. Essa abordagem depende do Axioma da Boa Ordenação. Verifica-se que a Boa Ordenação é equivalente a outro princípio—o Axioma da Escolha—e houve discussão filosófica séria sobre a sua aceitabilidade. As questões deste capítulo são: Como é que o Axioma é usado, e pode ser justificado?

16.2 O Truque de Tarski–Scott

Na Definição 14.1, definimos cardinais como ordinais. Para isso, assumimos o Axioma da Boa Ordenação. Fizemo-lo apenas porque é o "padrão da indústria".

Antes de discutir questões filosóficas em torno da Boa Ordenação, importa deixar claro que *podemos* afastar-nos desse padrão e desenvolver uma teoria de cardinais *sem* assumir a Boa Ordenação. Continuamos a poder usar as definições de $A \approx B$, $A \preceq B$ e $A \prec B$, tal como apareceram no Capítulo 5. Só precisamos de uma nova noção de *cardinal*.

Uma ideia ingênua seria tentar definir a cardinalidade de A

assim:

$$\{x : A \approx x\}.$$

Talvez queira comparar isto com a definição fregeana de $\#xFx$, esboçada no final de Seção 14.5. E, pelas razões que ali adumbramos, esta definição falha. Qualquer conjunto singular é equipotente com $\{\emptyset\}$. Mas novos singulares são formados em cada estágio sucessor da hierarquia (basta considerar o singular do estágio anterior). Logo $\{x : A \approx x\}$ não existe, pois não pode ter posto.

Para contornar o problema, usamos um truque devido a Tarski e Scott:[1]

Definição 16.1 (Tarski–Scott). Para qualquer fórmula $\varphi(x)$, seja $[x : \varphi(x)]$ o conjunto de todos os x, de posto mínimo possível, tais que $\varphi(x)$ (ou $\emptyset$, se não houver φs).

Devemos verificar que esta definição é legítima. Trabalhando em **ZF**, Teorema 11.13 garante que posto(x) existe para todo o x. Agora, se houver entidades que satisfazem φ, podemos deixar α ser o posto mínimo tal que $(\exists x \subseteq V_\alpha)\varphi(x)$, isto é, $(\forall \beta \in \alpha)(\forall x \subseteq V_\beta)\neg\varphi(x)$. Podemos então definir $[x : \varphi(x)]$ por Separação como $\{x \in V_{\alpha+1} : \varphi(x)\}$.

Tendo justificado o truque de Tarski–Scott, podemos usá-lo para definir uma noção de cardinalidade:

Definição 16.2. A TS-cardinalidade de A é $\mathrm{tsc}(A) = [x : A \approx x]$.

A definição de TS-cardinal não usa a Boa Ordenação. Mas, mesmo sem esse Axioma, podemos mostrar que os *TS-cardinais* se comportam de modo semelhante aos *cardinais* tal como definidos na Definição 14.1. Por exemplo, se reformularmos Lema 14.3 e Lema 15.4 em termos de TS-cardinais, as provas funcionam em **ZF**, sem assumir a Boa Ordenação.

[1] Lembrete: todas as fórmulas podem ter parâmetros (salvo indicação em contrário).

Enquanto estamos no assunto, vale a pena notar que também podemos desenvolver uma teoria de ordinais usando o truque de Tarski–Scott. Onde $\langle A, < \rangle$ é uma boa ordenação, seja $\mathrm{tso}(A, <) = [\langle X, R\rangle : \langle A, <\rangle \cong \langle X, R\rangle]$. Para mais sobre este tratamento de cardinais e ordinais, ver Potter (2004, chs. 9–12).

16.3 Comparabilidade e Lema de Hartogs

Isto é o lado positivo. Eis o lado negativo. Sem Escolha, as coisas ficam *complicadas*. Para ver por que, eis um belo resultado devido a Hartogs (1915):

Lema 16.3 (*em* ZF). *Para qualquer conjunto A, existe um ordinal α tal que $\alpha \not\preceq A$*

Prova. Se $B \subseteq A$ e $R \subseteq B^2$, então $\langle B, R\rangle \subseteq V_{\mathrm{posto}(A)+4}$ por Lema 13.9. Logo, usando Separação, considere-se:

$$C = \{\langle B, R\rangle \in V_{\mathrm{posto}(A)+5} : B \subseteq A \text{ e } \langle B, R\rangle \text{ é uma boa ordenação}\}$$

Usando Substituição e Teorema 10.26, forme-se o conjunto:

$$\alpha = \{\mathrm{ord}(B, R) : \langle B, R\rangle \in C\}.$$

Por Corolário 10.19, α é um ordinal, pois é um conjunto transitivo de ordinais. De fato, se $\gamma \in \beta \in \alpha$, então $\beta = \mathrm{ord}(B, R)$ para algum $B \subseteq R$, donde $\gamma = \mathrm{ord}(B_b, R_b)$ para algum $b \in B$ por Lema 10.10, logo $\gamma \in \alpha$.

Por redução ao absurdo, suponhamos que existe uma injeção $f : \alpha \to A$. Então, onde:

$$B = \mathrm{Im}(f)$$
$$R = \{\langle f(\alpha), f(\beta)\rangle \in A \times A : \alpha \in \beta\}.$$

Claramente $\alpha = \mathrm{ord}(B, R)$ e $\langle B, R\rangle \in C$. Logo $\alpha \in \alpha$, o que é contradição. □

Isto implica um resultado profundo:

Teorema 16.4 (*em* ZF). *As afirmações seguintes são equivalentes:*

1. *O Axioma da Boa Ordenação*
2. *Ou $A \preceq B$ ou $B \preceq A$, para quaisquer conjuntos A e B*

Prova. (1) $\Rightarrow$ (2). Fixem-se A e B. Invocando (1), existem boas ordenações $\langle A, R\rangle$ e $\langle B, S\rangle$. Invocando Teorema 10.26, sejam $f : \alpha \to \langle A, R\rangle$ e $g : \beta \to \langle B, S\rangle$ isomorfismos. Por Proposição 10.22, ou $\alpha \subseteq \beta$ ou $\beta \subseteq \alpha$. Se $\alpha \subseteq \beta$, então $g \circ f^{-1} : A \to B$ é uma injeção, logo $A \preceq B$; analogamente, se $\beta \subseteq \alpha$ então $B \preceq A$.

(2) $\Rightarrow$ (1). Fixe-se A; por Lema 16.3 existe algum ordinal β tal que $\beta \not\preceq A$. Invocando (2), temos $A \preceq \beta$. Logo existe alguma injeção $f : A \to \beta$, e podemos usar esta injecção para bem ordenar os elementos de A, definindo uma ordem $\{\langle a, b\rangle \in A \times A : f(a) \in f(b)\}$. □

Como consequência imediata: se a Boa Ordenação falha, então alguns conjuntos são *literalmente incomparáveis* quanto ao tamanho. Assim, se a Boa Ordenação falha, a aritmética cardinal transfinita ficará complicada. Por exemplo, teremos de abandonar a ideia de que, se A e B são infinitos, então $A \sqcup B \approx A \times B \approx M$, onde M é o maior de A e B (ver Teorema 15.11). O problema é simples: se não podemos *comparar* o tamanho de A e B, então não faz sentido perguntar qual é o maior.

16.4 O Problema da Boa Ordenação

Evidentemente que muito depende de aceitarmos ou não a Boa Ordenação. Mas a discussão deste princípio tendeu a concentrar-se num equivalente, o Axioma da Escolha. Voltemos então a atenção para esse axioma (e provemos a equivalência).

Em 1883, Cantor exprimiu o seu apoio ao Axioma da Boa Ordenação, chamando-lhe "uma lei do pensamento que me parece fundamental, rica nas suas consequências e particularmente notável pela sua generalidade" (citado em Potter 2004, p. 243). Mas

Cantor acabou por convencer-se de que o "Axioma" precisava de prova. A comunidade matemática também.

O problema foi "resolvido" por Zermelo em 1904. Para explicar a solução, precisamos de algumas definições.

Definição 16.5. Uma função f é uma *função de escolha* sse $f(x) \in x$ para todo o $x \in \mathrm{dom}(f)$. Dizemos que f é uma *função de escolha para* A sse f é uma função de escolha com $\mathrm{dom}(f) = A \setminus \{\emptyset\}$.

Intuitivamente, para cada conjunto (não vazio) $x \in A$, uma função de escolha para A *escolhe* um elemento particular, $f(x)$, de x. O Axioma da Escolha é então:

Escolha (Escolha). Todo o conjunto tem uma função de escolha.

Zermelo mostrou que a Escolha implica a boa ordenação, e vice-versa:

Teorema 16.6 (em ZF). *A Boa Ordenação e a Escolha são equivalentes.*

Prova. Da esquerda para a direita. Seja A um conjunto de conjuntos. Então $\bigcup A$ existe pelo Axioma da União, e logo, pela Boa Ordenação, existe alguma relação $<$ que bem ordena $\bigcup A$. Seja agora $f(x) =$ o membro $<$-mínimo de x. Isto é uma função de escolha para A.

Da direita para a esquerda. Fixe-se A. Pela Escolha, existe uma função de escolha, f, para $\wp(A) \setminus \{\emptyset\}$. Usando Recursão Transfinita, defina-se uma função:

$$g(0) = f(A)$$
$$g(\alpha) = \begin{cases} \text{parar!} & \text{se } A = g[\alpha] \\ f(A \setminus g[\alpha]) & \text{caso contrário} \end{cases}$$

A indicação "parar!" é apenas uma abreviatura para o que seria de outro modo uma definição mais longa. Ou seja, quando $A = g[\alpha]$ pela primeira vez, seja $g(\delta) = A$ para todo o $\delta \leq \alpha$. Agora, em primeiro lugar, só podemos ter a certeza de que isto define um *termo* (ver as observações após Teorema 11.4); mas mostraremos que de fato temos uma função.

Como f é função de escolha, para cada α (quando definido) temos $g(\alpha) = f(A \setminus g[\alpha]) \in A \setminus g[\alpha]$; isto é, $g(\alpha) \notin g[\alpha]$. Logo, se $g(\alpha) = g(\beta)$ então $g(\beta) \notin g[\alpha]$, isto é, $\beta \notin \alpha$, e analogamente $\alpha \notin \beta$. Logo $\alpha = \beta$, por Tricotomia. Assim, g é injetiva.

Observe-se a seguir que de fato paramos!, isto é, que existe algum ordinal mínimo α tal que $A = g[\alpha]$. De fato, suponhamos o contrário; então, como g é injetiva, teríamos $\alpha \prec \wp(A) \setminus \{\emptyset\}$ para todo o ordinal α, contradizendo Lema 16.3. Daí também $\mathrm{Im}(g) = A$.

Reunindo estes fatos, g é uma bijeção de algum ordinal para A. Agora, g pode ser usada para bem ordenar A. □

Assim, a Boa Ordenação e a Escolha estão de pé ou caem juntas. Mas a questão permanece: estão de pé ou caem?

16.5 Escolha Enumerável

É fácil provar, sem usar Escolha/Boa Ordenação, que:

Lema 16.7 (em $\mathbf{Z}^-$). *Todo o conjunto finito tem uma função de escolha.*

Prova. Seja $a = \{b_1, \ldots, b_n\}$. Suponhamos, por simplicidade, que cada $b_i \neq \emptyset$. Logo existem objetos $c_1, \ldots, c_n$ tais que $c_1 \in b_1, \ldots, c_n \in b_n$. Agora, por Proposição 9.5, o conjunto $\{\langle b_1, c_1\rangle, \ldots, \langle b_n, c_n\rangle\}$ existe; e isto é uma função de escolha para a. □

Mas a situação complica-se logo que consideramos conjuntos infinitos. Por exemplo, considere-se esta extensão "mínima" do resultado acima:

Escolha Enumerável. Todo o conjunto *enumerável* tem uma função de escolha.

Isto é um caso particular da Escolha. E verifica-se que este princípio foi invocado com alguma frequência, sem uma consciência óbvia do seu uso. Eis dois bons exemplos.[2]

Exemplo 16.8. Eis um pensamento natural: para qualquer conjunto A, ou $\omega \preceq A$, ou $A \approx n$ para algum $n \in \omega$. Esta é uma maneira de exprimir a ideia intuitiva de que todo o conjunto é finito ou infinito. Cantor, e muitos outros matemáticos, fizeram esta afirmação sem a provar. Com a cautela que nos caracteriza, a provamos em Teorema 14.7. Mas nessa prova estávamos a trabalhar em **ZFC**, pois assumíamos que qualquer conjunto A pode ser bem ordenado, e logo que $|A|$ está garantidamente definido. Ou seja: assumimos explicitamente a Escolha.

De fato, Dedekind (1888) ofereceu a sua própria prova desta afirmação, do seguinte modo:

Teorema 16.9 (em $\mathbf{Z}^-$ + Escolha Enumerável). *Para qualquer A, ou $\omega \preceq A$ ou $A \approx n$ para algum $n \in \omega$.*

Prova. Suponhamos $A \not\approx n$ para todo o $n \in \omega$. Então, em particular, para cada $n < \omega$ existe um subconjunto $A_n \subseteq A$ com exatamente 2^n elementos. Usando esta sequência $A_0, A_1, A_2, \ldots$, definimos para cada n:

$$B_n = A_n \setminus \bigcup_{i<n} A_i.$$

Note-se o seguinte

$$\begin{aligned}\left|\bigcup_{i<n} A_n\right| &\leq |A_0| + |A_1| + \ldots + |A_{n-1}| \\ &= 1 + 2 + \ldots + 2^{n-1}\end{aligned}$$

[2]Devidos a Potter (2004, §9.4) e Luca Incurvati.

$$\begin{aligned} &= 2^n - 1 \\ &< 2^n = |A_n| \end{aligned}$$

Logo cada B_n tem pelo menos um membro, c_n. Além disso, os B_n são dois a dois disjuntos; assim, se $c_n = c_m$ então $n = m$. Mas cada $c_n \in A$. Logo a função $f(n) = c_n$ é uma injecção $\omega \to A$. □

Dedekind não assinalou que tinha usado a Escolha Enumerável. Mas reparou o leitor no seu uso? Olhe de novo. (De fato: *olhe de novo.*)

A prova usou a Escolha Enumerável duas vezes. A usamos uma vez, para obter a nossa sequência de conjuntos $A_0, A_1, A_2, \ldots$ Voltamos a usá-la para seleccionar os elementos c_n de cada B_n. Além disso, este uso da Escolha é irredutível. Cohen (1966, p. 138) provou que o resultado falha se não tivermos qualquer versão da Escolha. Ou seja: é consistente com **ZF** que existam conjuntos *incomparáveis* com ω.

Exemplo 16.10. Em 1878, Cantor afirmou que uma união enumerável de conjuntos enumeráveis é enumerável. Não apresentou prova, talvez por considerar a prova óbvia. Agora, com a nossa cautela, provamos uma versão mais geral deste resultado em Proposição 15.12. Mas a nossa prova assumiu explicitamente a Escolha. E mesmo a prova do resultado menos geral requer a Escolha Enumerável.

Teorema 16.11 (em $\mathbf{Z}^-$ + Escolha Enumerável). *Se A_n é enumerável para cada $n \in \omega$, então $\bigcup_{n<\omega} A_n$ é enumerável.*

Prova. Sem perda de generalidade, suponhamos que cada $A_n \neq \emptyset$. Logo, para cada $n \in \omega$ existe uma sobrejeção $f_n \colon \omega \to A_n$. Defina-se $f \colon \omega \times \omega \to \bigcup_{n<\omega} A_n$ por $f(m,n) = f_n(m)$. O resultado segue porque $\omega \times \omega$ é enumerável (Proposição 5.6) e f é uma sobrejeção. □

Notou o uso da Escolha Enumerável? Serve para escolher a nossa sequência de funções f_0, f_1, f_2, ...[3] E, de novo, o resultado falha na ausência de qualquer princípio de Escolha. Especificamente, Feferman and Levy (1963) provou que é consistente com **ZF** que uma união enumerável de conjuntos enumeráveis tenha cardinalidade $\beth_1$. Mas eis uma formulação bem mais divertida do ponto, de Russell:

> Isto ilustra-se com o milionário que comprava um par de meias sempre que comprava um par de botas, e nunca noutra ocasião, e que tinha tal paixão por comprar ambos que acabou por ter $\aleph_0$ pares de botas e $\aleph_0$ pares de meias... Nas botas podemos distinguir direita e esquerda, e portanto podemos fazer uma seleção de uma em cada par, nomeadamente, podemos escolher todas as botas direitas ou todas as esquerdas; mas com as meias não se impõe qualquer princípio de seleção, e não podemos ter a certeza, a menos que assumamos o axioma multiplicativo [isto é, em efeito, a Escolha], de que existe alguma classe constituída por uma meia de cada par. (Russell, 1919, p. 126)

Em suma, alguma forma de Escolha é necessária para provar o seguinte: se tem uma quantidade enumerável de pares de meias, então tem (apenas) uma quantidade enumerável de meias. E, de fato, sem a Escolha Enumerável (ou algo equivalente), uma união enumerável de conjuntos enumeráveis pode deixar de ser enumerável.

A moral é que a Escolha Enumerável foi usada repetidamente, sem grande consciência por parte de quem a usava. A questão filosófica é: como poderíamos *justificar* a Escolha Enumerável?

Uma tentativa de justificação intuitiva poderia apelar a uma super-tarefa. Suponhamos que fazemos a primeira escolha em 1/2

[3] Um uso semelhante da Escolha ocorreu em Proposição 15.12, quando demos a instrução "Para cada $\beta \in \mathfrak{a}$, fixe uma injeção f_β".

minuto, a segunda em $1/4$ minuto, ..., a n-ésima em $1/2^n$ minuto, ... Então, dentro de 1 minuto, teremos feito uma ω-sequência de escolhas, e definido uma função de escolha.

Mas que poderia realmente dizer-nos um tal experimento mental? Para começar, depende de tomar a ideia de "escolher" de modo bastante literal. Para além disso, parece atar a matemática a possibilidades metafísicas.

Mais importante: não nos dará justificação para a Escolha *sem reservas*, em vez da *mera* Escolha Enumerável. Porque, se precisarmos que *todo* o conjunto tenha função de escolha, então precisaremos de poder realizar uma "super-tarefa de comprimento ordinal arbitrário." Sem rodeios, essa ideia é risível.

16.6 Considerações Intrínsecas sobre a Escolha

A questão mais ampla, então, é se a Boa Ordenação, ou a Escolha, ou de fato a comparabilidade de todos os conjuntos quanto ao tamanho—não importa qual—podem ser justificadas.

Eis uma tentativa de justificação *intrínseca*. De volta a Seção 9.1, introduzimos vários princípios sobre a hierarquia. Um deles merece ser repetido:

> *Estágios-acumulam.* Para qualquer estágio S, e para quaisquer conjuntos que foram formados *antes* do estágio S: forma-se no estágio S um conjunto cujos membros são exatamente esses conjuntos. Nada mais se forma no estágio S.

De fato, muitos autores sugeriram que o Axioma da Escolha pode ser justificado via (algo como) este princípio. Daremos brevemente um resumo dessa abordagem.

Começamos por um resultado simples, que oferece *mais uma* formulação equivalente da Escolha:

Teorema 16.12 (em ZF). *A Escolha é equivalente ao princípio seguinte. Se os membros de A são disjuntos dois a dois e não vazios, então existe algum C tal que $C \cap x$ é um singular para todo o $x \in A$. (Chamamos a tal C um conjunto de escolha para A.)*

A prova deste resultado é direta, e a deixamos como exercício para o leitor.

O ponto essencial é que um conjunto de escolha para A é precisamente o contradomínio de uma função de escolha para A. Assim, para justificar a Escolha, podemos simplesmente tentar justificar a sua formulação equivalente, em termos da existência de conjuntos de escolha. E é exatamente isso que faremos agora.

Sejam os membros de A disjuntos dois a dois e não vazios. Por *Estágios-são-chave* (ver Seção 9.1), A é formado em algum estágio S. Note-se que todos os membros de $\bigcup A$ estão disponíveis antes do estágio S. Agora, por *Estágios-acumulam*, para *quaisquer* conjuntos formados antes de S, forma-se um conjunto cujos membros são exatamente esses conjuntos. Dito de outro modo: toda a *possível* coleção de conjuntos disponíveis mais cedo existirá em S. Mas é certamente *possível* seleccionar objetos que possam ser reunidos num conjunto de escolha para A; isso é apenas um certo subconjunto muito específico de $\bigcup A$. Logo: existe tal conjunto de escolha, como requerido.

Bem, esta é uma tentativa *muito* rápida de oferecer uma justificação da Escolha por razões intrínsecas. Mas, para aprofundar esta ideia, deve ler o desenvolvimento elegante de Potter (2004, §14.8).

16.7 Paradoxo de Banach–Tarski

Também podemos tentar justificar a Escolha, como Boolos tentou justificar a Substituição, apelando a considerações *extrínsecas* (ver Seção 12.3). Afinal, adotar a Escolha tem muitas consequências desejáveis: poder comparar qualquer par de cardinais; poder bem

ordenar qualquer conjunto; poder tratar os cardinais como um tipo particular de ordinal; etc.

Por vezes, contudo, afirma-se que a Escolha tem consequências *indesejáveis*. Na maior parte dos casos, isso deve-se a um resultado de Banach and Tarski (1924).

Teorema 16.13 (Paradoxo de Banach–Tarski (em ZFC)). *Qualquer bola pode ser decomposta num número finito de peças, que podem ser remontadas (por rotação e transporte) de modo a formar duas cópias dessa bola.*

À primeira vista, isto surpreende. Claramente as duas bolas têm *o dobro* do volume da bola original. Mas os movimentos rígidos—rotação e transporte—não alteram o volume. Parece assim que Banach–Tarski nos permite fazer aparecer matéria nova por magia.

Pior ainda.[4] Raciocínio semelhante mostra que uma ervilha pode ser cortada num número finito de peças, que depois podem ser remontadas (por rotação e transporte) de modo a formar um objeto com a forma e o tamanho do Big Ben.

Nada disto, porém, vale em **ZF** isoladamente.[5] Logo temos de decidir: rejeitar a Escolha, ou aprender a viver com o "paradoxo".

Sugerimos que devemos aprender a viver com o "paradoxo". De fato, não nos parece propriamente um paradoxo. Em particular, não vemos por que haveria de ser mais ou menos paradoxal do que qualquer um dos resultados seguintes:[6]

1. Há tantos pontos no intervalo $(0,1)$ como em $\mathbb{R}$.
   Prova: considere-se $\tan(\pi(r - 1/2))$.

[4] Ver Tomkowicz and Wagon (2016, Theorem 3.12).

[5] Embora Banach–Tarski possa ser provado com princípios estritamente mais fracos do que a Escolha; ver Tomkowicz and Wagon (2016, 303).

[6] Potter (2004, 276–7), Weston (2003, 16), Tomkowicz and Wagon (2016, 31, 308–9) fazem pontos semelhantes, com outros exemplos.

2. Há tantos pontos numa reta como num quadrado.
   Ver Seção 1.3 e Seção 5.10.

3. Existem curvas que preenchem o espaço.
   Ver Seção 1.3 e Seção 5.11.

Nenhum destes três resultados requer a Escolha. De fato, já os tratamos simplesmente como matemática surpreendente e bela. Talvez devêssemos adotar a mesma atitude perante o Paradoxo de Banach–Tarski.

Com certeza, é necessária aqui uma observação técnica; mas basta manter a cabeça fria. Os movimentos rígidos preservam volume. Consequentemente, as cinco[7] peças em que a bola é decomposta não podem todas ser *mensuráveis*. Grosso modo, não faz sentido atribuir volume a estas peças individualmente. Deve pensar nelas como dispersões infinitas de pontos impossíveis de visualizar. Agora, pode ser "estranho" conceber tais conjuntos "infinitamente dispersos". Mas a sua existência parece seguir da injunção, incorporada em *Estágios-acumulam*, de que se formem *todas* as coleções possíveis de conjuntos disponíveis mais cedo.

Se nada disto convence, eis um último argumento (extrínseco) a favor de abraçar o Paradoxo de Banach–Tarski. Ele implica de imediato a melhor piada matemática de sempre:

> *Pergunta.* O que é um anagrama de "Banach–Tarski"?
>
> *Resposta.* "Banach–Tarski Banach–Tarski".

16.8 Apêndice: Paradoxo de Vitali

Para ter uma ideia real de se a construção de Banach–Tarski é aceitável ou não, devíamos examinar a sua *prova*. Infelizmente, isso exigiria muito mais álgebra do que podemos apresentar aqui.

[7] Enunciamos o Paradoxo em termos de "número finito de peças". De fato, Robinson (1947) provou que a decomposição pode ser feita com *cinco* peças (mas não menos). Para uma prova, ver Tomkowicz and Wagon (2016, pp. 66–7).

Contudo, podemos oferecer algumas observações rápidas que talvez esclareçam a prova de Banach–Tarski,[8] concentrando-nos no resultado seguinte:

Teorema 16.14 (Paradoxo de Vitali (em ZFC)). *Qualquer círculo pode ser decomposto numa quantidade enumerável de peças, que podem ser remontadas (por rotação e transporte) de modo a formar duas cópias desse círculo.*

O Paradoxo de Vitali é muito mais fácil de provar do que o Paradoxo de Banach–Tarski. Chamamos-lhe "Paradoxo de Vitali", pois segue-se da construção, em 1905, de um conjunto não mensurável. Mas os aspectos conjuntistas da prova do Paradoxo de Vitali e do Paradoxo de Banach–Tarski são muito semelhantes. A diferença essencial entre os resultados é que Banach–Tarski considera uma decomposição *finita*, enquanto o Paradoxo de Vitali considera uma decomposição *enumerável infinita*. Como Weston (2003) diz, o Paradoxo de Vitali "certamente não é tão notável como o paradoxo de Banach–Tarski, mas ilustra que paradoxos geométricos podem ocorrer mesmo em situações 'simples'."

O Paradoxo de Vitali trata de uma figura bidimensional, um círculo. Logo trabalhamos no plano, $\mathbb{R}^2$. Seja R o conjunto das rotações (no sentido dos ponteiros do relógio) de pontos em torno da origem por valores em radianos *racionais* no intervalo $[0, 2\pi)$. Eis alguns fatos algébricos sobre R (se não compreender o enunciado do resultado, a prova esclarecerá o seu significado):

Lema 16.15. *R forma um grupo abeliano sob a composição de funções.*

Prova. Escrevendo 0_R para a rotação por 0 radianos, este é um elemento neutro para R, pois $\rho \circ 0_R = 0_R \circ \rho = \rho$ para qualquer $\rho \in R$.

Todo o elemento tem inverso. Onde $\rho \in R$ roda por r radianos, $\rho^{-1} \in R$ roda por $2\pi - r$ radianos, de modo que $\rho \circ \rho^{-1} = 0_R$.

[8] Para um tratamento bem mais completo, ver Weston (2003) ou Tomkowicz and Wagon (2016).

A composição é associativa: $(\tau \circ \sigma) \circ \rho = \tau \circ (\sigma \circ \rho)$ para quaisquer $\rho, \sigma, \tau \in R$

A composição é comutativa: $\sigma \circ \rho = \rho \circ \sigma$ para quaisquer $\rho, \sigma \in R$. □

De fato, podemos partir o nosso grupo R ao meio e depois usar qualquer metade para recuperar o grupo inteiro:

Lema 16.16. *Existe uma partição de R em dois conjuntos disjuntos, R_1 e R_2, ambos dos quais são uma base para R.*

Prova. Seja R_1 o conjunto das rotações por valores racionais em radianos em $[0, \pi)$; seja $R_2 = R \setminus R_1$. Por álgebra elementar, $\{\rho \circ \rho : \rho \in R_1\} = R$. Um resultado semelhante obtém-se para R_2. □

Usaremos este fato sobre grupos para estabelecer Teorema 16.14. Seja **S** o círculo unitário, isto é, o conjunto dos pontos exatamente a 1 unidade da origem do plano, isto é, $\{\langle r, s\rangle \in \mathbb{R}^2 : \sqrt{r^2 + s^2} = 1\}$. Partiremos **S** em partes considerando a relação seguinte em **S**:

$$r \sim s \text{ sse } (\exists \rho \in R)\rho(r) = s.$$

Ou seja, os pontos de **S** estão ligados por esta relação sse se pode passar de um para o outro por uma rotação de valor racional em torno da origem. Sem surpresa:

Lema 16.17. *$\sim$ é uma relação de equivalência.*

Prova. Trivial, usando Lema 16.15. □

Invocamos agora a Escolha para obter um conjunto, C, contendo exatamente um membro de cada classe de equivalência de **S** sob $\sim$. Ou seja, consideramos uma função de escolha f no

conjunto das classes de equivalência,[9]

$$E = \{[r]_\sim : r \in \mathbf{S}\},$$

e seja $C = \mathrm{Im}(f)$. Para cada rotação $\rho \in R$, o conjunto $\rho[C]$ consiste nos pontos obtidos ao aplicar a rotação ρ a cada ponto de C. Os dois resultados seguintes mostram que estes conjuntos cobrem o círculo por completo e sem sobreposição:

Lema 16.18. $\mathbf{S} = \bigcup_{\rho \in R} \rho[C]$.

Prova. Fixe-se $s \in \mathbf{S}$; existe algum $r \in C$ tal que $r \in [s]_\sim$, isto é, $r \sim s$, isto é, $\rho(r) = s$ para algum $\rho \in R$. □

Lema 16.19. *Se $\rho_1 \neq \rho_2$ então $\rho_1[C] \cap \rho_2[C] = \emptyset$.*

Prova. Suponhamos $s \in \rho_1[C] \cap \rho_2[C]$. Logo $s = \rho_1(r_1) = \rho_2(r_2)$ para alguns $r_1, r_2 \in C$. Daí $\rho_2^{-1}(\rho_1(r_1)) = r_2$, e $\rho_2^{-1} \circ \rho_1 \in R$, logo $r_1 \sim r_2$. Logo $r_1 = r_2$, pois C selecciona exatamente um membro de cada classe de equivalência sob $\sim$. Logo $s = \rho_1(r_1) = \rho_2(r_1)$, e portanto $\rho_1 = \rho_2$. □

Aplicamos agora os fatos algébricos anteriores ao nosso círculo:

Lema 16.20. *Existe uma partição de* $\mathbf{S}$ *em dois conjuntos disjuntos, D_1 e D_2, tal que D_1 pode ser particionado numa quantidade enumerável de conjuntos que podem ser rodados de modo a formar uma cópia de* $\mathbf{S}$ *(e analogamente para D_2).*

Prova. Usando R_1 e R_2 de Lema 16.16, seja:

$$D_1 = \bigcup_{\rho \in R_1} \rho[C] \qquad\qquad D_2 = \bigcup_{\rho \in R_2} \rho[C]$$

[9] Como R é enumerável, cada membro de E é enumerável. Como $\mathbf{S}$ é não enumerável, segue-se de Lema 16.18 e Proposição 15.12 que E é não enumerável. Logo isto é um uso de Escolha *não* enumerável.

Isto é uma partição de **S**, por Lema 16.18, e D_1 e D_2 são disjuntos por Lema 16.19. Por construção, D_1 pode ser particionado numa quantidade enumerável de conjuntos, $\rho[C]$ para cada $\rho \in R_1$. E estes podem ser rodados de modo a formar uma cópia de **S**, pois $\mathbf{S} = \bigcup_{\rho\in R} \rho[C] = \bigcup_{\rho\in R_1}(\rho\circ\rho)[C]$ por Lema 16.16 e Lema 16.18. O mesmo raciocínio vale para D_2. □

Isto implica de imediato o Paradoxo de Vitali. Pois podemos gerar *duas* cópias de **S** a partir de **S**, apenas partindo-o numa quantidade enumerável de peças (os vários $\rho[C]$) e depois movendo-as rigidamente (basta rodar cada peça de D_1, e primeiro transportar e depois rodar cada peça de D_2).

Recapitulemos a estratégia de prova. Começamos com fatos algébricos sobre o grupo de rotações no plano. Usamos este grupo para partir **S** em classes de equivalência. Chegamos então a um "paradoxo", usando a Escolha para seleccionar elementos de cada classe.

Usamos exatamente a mesma estratégia para provar Banach–Tarski. A principal diferença é que os fatos algébricos usados na prova de Banach–Tarski são bastante mais complicados do que os usados no Paradoxo de Vitali. Mas esses fatos algébricos nada têm a ver com a Escolha. Resumimos-os rapidamente.

Para provar Banach–Tarski, começamos por estabelecer um análogo de Lema 16.16: qualquer *grupo livre* pode ser partido em quatro peças que, intuitivamente, podemos "mover" de modo a recuperar duas cópias do grupo inteiro.[10] Mostramos depois que podemos usar duas rotações particulares em torno da origem de $\mathbb{R}^3$ para gerar um grupo livre de rotações, F.[11] (Ainda sem Escolha.) Consideramos agora pontos na superfície da esfera como "semelhantes" sse um pode obter-se do outro por uma rotação em F. *Usamos então a Escolha* para seleccionar exatamente um ponto de cada classe de equivalência de pontos "semelhantes".

[10] O fato de podermos usar *quatro* peças deve-se a Robinson (1947). Para uma prova recente, ver Tomkowicz and Wagon (2016, Theorem 5.2). Seguimos Weston (2003, p. 3) ao descrever isto como "mover" as peças do grupo.

[11] Ver Tomkowicz and Wagon (2016, Theorem 2.1).

Aplicando a nossa divisão de F à superfície da esfera, como em Lema 16.20, partimos essa superfície em quatro peças, que podemos "mover" de modo a obter duas cópias da superfície da esfera. E isto estabelece (Hausdorff, 1914):

Teorema 16.21 (Paradoxo de Hausdorff (em ZFC)). *A superfície de qualquer esfera pode ser decomposta num número finito de peças, que podem ser remontadas (por rotação e transporte) de modo a formar duas cópias disjuntas dessa esfera.*

Ainda são precisos alguns truques algébricos para obter o Teorema completo de Banach–Tarski (que não diz só respeito à superfície da esfera, mas também ao seu interior). Francamente, porém, isto é apenas a cereja no topo do bolo algébrico. Por isso Weston escreve:

> [...] o resultado sobre grupos livres é o passo *chave* na prova do paradoxo de Banach–Tarski. Deste ponto de vista, o paradoxo de Banach–Tarski não é tanto uma afirmação sobre $\mathbb{R}^3$ quanto uma afirmação sobre a complexidade do grupo [das translações e rotações em $\mathbb{R}^3$]. (Weston, 2003, p. 16)

Ou seja: podermos ou não oferecer uma decomposição *finita* (como em Banach–Tarski) ou uma decomposição *enumerável infinita* (como no Paradoxo de Vitali) resume-se a certos fatos de teoria dos grupos sobre trabalhar em duas ou três dimensões.

Admitidamente, esta última observação estraga ligeiramente a piada no final de Seção 16.7. Por ser bidimensional, "Banach–Tarski" tem de ser dividido numa *infinidade* enumerável de peças, se se quiserem rearranjar essas peças de modo a formar "Banach–Tarski Banach–Tarski". Para consertar a piada, é preciso escrever em três dimensões. Deixamos isto como exercício para o leitor.

Um último comentário. Na Seção 16.7, mencionamos que as "peças" da esfera que se obtêm não podem ser *mensuráveis*, mas têm de ser dispersões infinitas impossíveis de visualizar. O

mesmo vale para o nosso uso da Escolha ao obter Lema 16.20. E tudo isto merece explicação.

De novo, devemos esboçar alguma base (mas isto é *apenas* um esboço; talvez queira consultar uma entrada de manual sobre *medida*). Definir uma medida para um conjunto X é atribuir um valor $\mu(E) \in \mathbb{R}$ a cada E nalguma "σ-álgebra" sobre X. Os detalhes aqui não são essenciais, exceto que a função μ deve obedecer ao princípio da aditividade enumerável: a medida de uma união enumerável de conjuntos disjuntos é a soma das medidas individuais, isto é, $\mu(\bigcup_{n<\omega} X_n) = \sum_{n<\omega} \mu(X_n)$ sempre que os X_n são disjuntos. Dizer que um conjunto é "não mensurável" é dizer que nenhuma medida pode ser atribuída de modo adequado. Agora, usando o nosso R de antes:

Corolário 16.22 (Vitali). *Seja μ uma medida tal que $\mu(\mathbf{S}) = 1$, e tal que $\mu(X) = \mu(Y)$ se X e Y são congruentes. Então $\rho[C]$ é não mensurável para todo o $\rho \in R$.*

Prova. Por redução ao absurso, suponhamos o contrário. Seja então $\mu(\sigma[C]) = r$ para algum $\sigma \in R$ e algum $r \in \mathbb{R}$. Para qualquer $\rho \in R$, $\rho[C]$ e $\sigma[C]$ são congruentes, logo $\mu(\rho[C]) = r$ para qualquer $\rho \in R$. Por Lema 16.18 e Lema 16.19, $\mathbf{S} = \bigcup_{\rho \in R} \rho[C]$ é uma união enumerável de conjuntos disjuntos dois a dois. Logo a aditividade enumerável dita que $\mu(\mathbf{S}) = 1$ é a soma das medidas de cada $\rho[C]$, isto é,

$$1 = \mu(\mathbf{S}) = \sum_{\rho \in R} \mu(\rho[C]) = \sum_{\rho \in R} r$$

Mas se $r = 0$ então $\sum_{\rho \in R} r = 0$, e se $r > 0$ então $\sum_{\rho \in R} r = \infty$. □

Problemas

Problema 16.1. Prove Teorema 16.12. Se tiver dificuldades, encontrará uma prova em (Potter, 2004, pp. 242–3).

APÊNDICE A

Biografias

A.1 Georg Cantor

Uma biografia inicial de Georg Cantor (GAY-org KAHN-tor) afirmava que ele tinha nascido e sido encontrado num navio que seguia para São Petersburgo, na Rússia, e que seus pais eram desconhecidos. Isto, porém, não é verdade; embora ele tivesse nascido em São Petersburgo em 1845.

Figura A.1: Georg Cantor

Cantor se doutorou em matemática na Universidade de Berlim em 1867. É conhecido por seu trabalho em teoria dos conjuntos e é considerado o fundador da teoria dos conjuntos como disciplina de pesquisa distinta. Foi o primeiro a provar que existem conjuntos infinitos de tamanhos diferentes. Suas teorias, e sobretudo a teoria dos infinitos, provocaram muito debate entre os matemáticos da época, e seu trabalho

foi controverso.

As convicções religiosas de Cantor e seu trabalho matemático estavam inextricavelmente ligados; ele chegou a afirmar que a teoria dos números transfinitos lhe tinha sido comunicada diretamente por Deus. Mais tarde, Cantor sofreu de doença mental. A partir de 1894, e com maior frequência nos últimos anos, Cantor foi internado. As fortes críticas a seu trabalho, incluindo um rompimento com o matemático Leopold Kronecker, levaram à depressão e à perda de interesse pela matemática. Durante episódios depressivos, Cantor se dedicava à filosofia e à literatura, e chegou a publicar uma teoria de que Francis Bacon era o autor das peças de Shakespeare.

Cantor morreu em 6 de janeiro de 1918, num sanatório em Halle.

Leitura complementar Para biografias completas de Cantor, ver Dauben (1990) e Grattan-Guinness (1971). As visões radicais de Cantor são também descritas no programa da BBC Radio 4 *A Brief History of Mathematics* (du Sautoy, 2014). Se quiser ouvir as teorias de Cantor em forma de rap, ver Rose (2012).

A.2 Kurt Gödel

Kurt Gödel (GER-dle) nasceu em 28 de abril de 1906 em Brünn, no império austro-húngaro (hoje Brno, na República Tcheca). Devido a seu espírito inquisitivo e brilhante, o jovem Kurtele era muitas vezes chamado "Der kleine Herr Warum" (o Pequeno Senhor Porquê) pela família. Destacou-se nos estudos desde a escola primária, tendo obtido menos do que a nota mais alta apenas em matemática. Gödel faltava frequentemente à escola por causa da frágil saúde e estava isento de educação física. Foi diagnosticado com febre reumática durante a infância. Ao longo da vida, acreditou que isso tinha afetado permanentemente seu coração, apesar dos pareceres médicos em contrário.

Figura A.2: Kurt Gödel

Gödel começou a estudar na Universidade de Viena em 1924 e concluiu o doutorado em 1929. Inicialmente pretendia estudar física, mas seus interesses se deslocaram rapidamente para a matemática e sobretudo para a lógica, em parte devido à influência do filósofo Rudolf Carnap. Sua dissertação, orientada por Hans Hahn, provava o teorema de completude da lógica de predicados de primeira ordem com identidade (Gödel, 1929). Apenas um ano depois, obteve seus resultados mais famosos—o primeiro e o segundo teorema da incompletude (publicados em Gödel 1931). Durante o período em Viena, Gödel participou intensamente no Círculo de Viena, um grupo de filósofos de orientação científica que incluía Carnap, cujo trabalho foi especialmente influenciado pelos resultados de Gödel.

Em 1938, Gödel se casou com Adele Nimbursky. Os pais não ficaram contentes: não só ela era seis anos mais velha e já divorciada, como trabalhava como bailarina num clube noturno. As pressões sociais não afetaram Gödel, contudo, e permaneceram felizmente casados até sua morte.

Depois de a Alemanha Nazi ter anexado a Áustria em 1938, Gödel e Adele emigraram para os Estados Unidos, onde ele assumiu um lugar no Institute for Advanced Study em Princeton, Nova Jersey. Apesar da introversão e do caráter excêntrico, o período de Gödel em Princeton foi colaborativo e frutuoso. Publicou ensaios em teoria dos conjuntos, filosofia e física. Notavelmente, estabeleceu uma amizade particularmente forte com o colega no IAS, Albert Einstein.

Nos últimos anos, a saúde mental de Gödel se deteriorou. A internação da esposa em 1977 fez com que ela não pudesse mais cozinhar as refeições dele. Tendo sofrido de problemas de saúde mental ao longo da vida, sucumbiu à paranoia. Com medo mortal de ser envenenado, Gödel se recusou a comer. Morreu de inanição em 14 de janeiro de 1978, em Princeton.

Leitura complementar Para uma biografia completa da vida de Gödel, ver John Dawson (1997). Para outros textos biográficos, bem como ensaios sobre as contribuições de Gödel para a lógica e a filosofia, ver Wang (1990), Baaz et al. (2011), Takeuti et al. (2003), e Sigmund et al. (2007).

A tese de doutorado de Gödel está disponível no alemão original (Gödel, 1929). O texto original dos teoremas da incompletude é (Gödel, 1931). Todas as publicações e textos não publicados de Gödel, bem como uma seleção de correspondência, estão disponíveis em inglês em seus *Collected Papers* Feferman et al. (1986, 1990).

Para um tratamento detalhado dos teoremas da incompletude de Gödel, ver Smith (2013). Para uma discussão informal e filosófica dos teoremas de Gödel, ver o podcast de Mark Linsenmayer (Linsenmayer, 2014).

A.3 Bertrand Russell

Bertrand Russell é celebrado como um dos fundadores da filosofia analítica moderna. Nascido em 18 de maio de 1872, Russell não era conhecido apenas pelo trabalho em filosofia e lógica, mas escreveu muitos livros de divulgação em diversas áreas. Foi também um ardente ativista político ao longo da vida.

Russell nasceu em Trellech, Monmouthshire, no País de Gales. Os pais pertenciam à nobreza britânica. Eram livrepensadores e chegaram a travar amizade com os radicais de Boston da época. Infelizmente, os pais de Russell morreram quando ele era ainda criança, e Russell foi viver com os avós. Ali recebeu

uma educação religiosa (algo que os pais tinham querido evitar a todo custo). A avó era muito rigorosa em todas as questões de moralidade. Na adolescência foi sobretudo educado em casa por professores particulares.

A influência de Russell na filosofia analítica, e sobretudo na lógica, é enorme. Estudou matemática e filosofia no Trinity College, em Cambridge, onde foi influenciado pelo matemático e filósofo Alfred North Whitehead. Em 1910, Russell e Whitehead publicaram o primeiro volume de *Principia Mathematica*, defendendo a ideia de que a matemática é redutível à lógica. Seguiu-se a publicação de centenas de livros, ensaios e panfletos políticos. Em 1950, ganhou o Prêmio Nobel de Literatura.

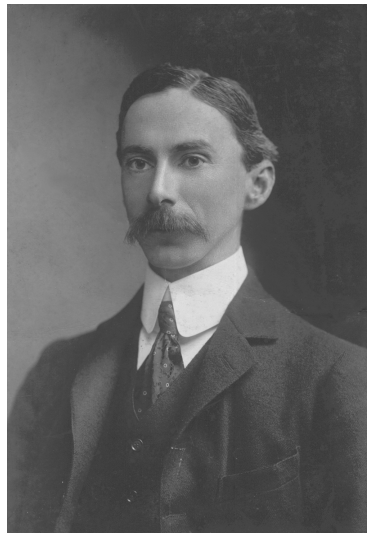


Figura A.3: Bertrand Russell

Russell se encontrava profundamente envolvido na política e no ativismo social. Durante a Primeira Guerra Mundial foi preso e condenado a seis meses de prisão por atividades pacifistas e protesto. Na prisão pôde escrever e ler, e afirma ter achado a experiência "bastante agradável." Permaneceu pacifista ao longo da vida, e voltou a ser detido por participar num comício pelo desarmamento nuclear em 1961. Sobreviveu também a um acidente de avião em 1948, no qual os únicos sobreviventes eram os que estavam sentados na área de fumantes. Por isso, Russell declarou que devia a vida ao tabagismo. Russell se casou quatro vezes, mas tinha fama de manter relações extraconjugais. Morreu em 2 de fevereiro de 1970, aos 97 anos, em Penrhyndeudraeth, no País de Gales.

Leitura complementar Russell escreveu uma autobiografia em três partes, que abrange sua vida de 1872 a 1967 (Russell, 1967, 1968, 1969). O Bertrand Russell Research Centre na McMaster University alberga os arquivos de Bertrand Russell. Ver o site em Duncan (2015), para informação sobre os volumes de suas obras reunidas (incluindo índices pesquisáveis) e projetos de arquivo. O artigo de Russell *On Denoting* (Russell, 1905) é um clássico da filosofia analítica do século XX.

A entrada da Stanford Encyclopedia of Philosophy sobre Russell (Irvine, 2015) inclui excertos de áudio de Russell a falar sobre Desejo e teoria política. Estão disponíveis online muitas entrevistas em vídeo com Russell. Para o ouvir falar de tabagismo e do envolvimento num acidente de avião, ver, por exemplo, Russell (n.d.). Algumas obras de Russell, incluindo a *Introduction to Mathematical Philosophy*, estão disponíveis como audiolivros gratuitos em LibriVox (n.d.).

A.4 Alfred Tarski

Alfred Tarski nasceu em 14 de janeiro de 1901 em Varsóvia, na Polônia (então parte do Império Russo). Descrito como "napoleônico", Tarski era ruidoso, falador e intenso. Sua energia se refletia muitas vezes nas aulas—chegou a incendiar um cesto de lixo ao jogar fora um cigarro durante uma aula, e foi proibido de lecionar novamente naquele edifício.

Tarski tinha sede de conhecimento desde cedo. Mais tarde diria aos estudantes que estudara lógica porque era a única disciplina em que tivera um B, mas os registros do colégio mostram que tinha notas máximas em tudo—até em lógica. Estudou na Universidade de Varsóvia de 1918 a 1924. Tarski pretendia inicialmente estudar biologia, mas se interessou por matemática, filosofia e lógica, pois a universidade era o centro da Escola de Lógica e Filosofia de Varsóvia. Tarski se doutorou em 1924 sob orientação de Stanisław Leśniewski.

Antes de emigrar para os Estados Unidos em 1939, Tarski

completou parte de seu trabalho mais importante enquanto professor de ensino secundário em Varsóvia. O trabalho sobre consequência lógica e verdade lógica foi escrito nesse período. Em 1939, Tarski se encontrava nos Estados Unidos numa turnê de palestras. Durante a visita, a Alemanha invadiu a Polônia, e, por causa da ascendência judaica, Tarski não pôde retornar. A esposa e os filhos permaneceram na Polônia até o fim da guerra, mas depois também puderam emigrar para os Estados Unidos. Tarski lecionou em Harvard, no College of the City of New York, no Institute for Advanced Study em Princeton, e finalmente na Universidade da Califórnia, em Berkeley. Ali fundou o programa multidisciplinar em Lógica e Metodologia da Ciência. Tarski morreu em 26 de outubro de 1983, aos 82 anos.

Leitura complementar Para mais sobre a vida de Tarski, ver a biografia *Alfred Tarski: Life and Logic* (Feferman and Feferman, 2004). As obras fundadoras de Tarski sobre consequência lógica e verdade estão disponíveis em inglês em (Corcoran, 1983). Todas as obras originais de Tarski foram reunidas numa série de quatro volumes, (Tarski, 1981).

Figura A.4: Alfred Tarski

A.5 Ernst Zermelo

Ernst Zermelo nasceu em 27 de julho de 1871 em Berlim, na Alemanha. Tinha cinco irmãs, embora a família sofresse de frágil saúde e apenas três sobrevivessem à idade adulta. Os pais também faleceram quando ele era jovem, deixando-o órfão, com os irmãos, aos dezessete anos. Zermelo tinha um profundo interesse pelas artes, e sobre-

tudo pela poesia. Era conhecido por ser perspicaz, espirituoso e crítico. Entre suas conquistas matemáticas mais celebradas estão a introdução do axioma da escolha (em 1904) e sua axiomatização da teoria dos conjuntos (em 1908).

Os interesses de Zermelo na universidade eram variados. Frequentou cursos de física, matemática e filosofia. Sob orientação de Hermann Schwarz, Zermelo concluiu a dissertação *Investigations in the Calculus of Variations* em 1894 na Universidade de Berlim. Em 1897, decidiu prosseguir estudos na Universidade de Göttingen, onde foi fortemente influenciado pelo trabalho fundacional de David Hilbert. Em 1899 tornou-se elegível para uma cátedra, mas só obteve uma onze anos depois—possivelmente devido a seu comportamento estranho e à "pressa nervosa."

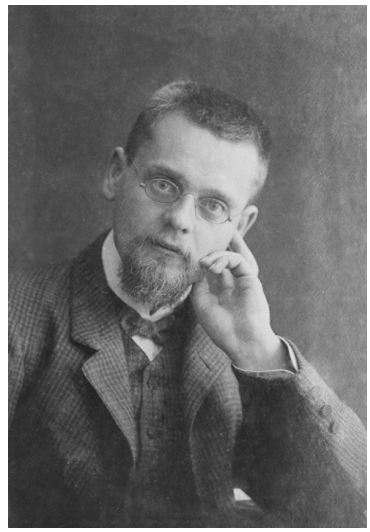


Figura A.5: Ernst Zermelo

Zermelo obteve finalmente uma cátedra remunerada na Universidade de Zurique em 1910, mas foi forçado a se aposentar em 1916 por causa da tuberculose. Depois da recuperação, foi nomeado professor honorário na Universidade de Freiburg em 1921. Nesse período dedicou-se à matemática fundacional. Irritou-se com os trabalhos de Thoralf Skolem e Kurt Gödel, e criticou publicamente suas abordagens em seus artigos. Foi demitido do cargo em Freiburg em 1935, devido a sua impopularidade e à oposição à ascensão de Hitler ao poder na Alemanha.

Os últimos anos da vida de Zermelo foram marcados pelo isolamento. Depois da demissão em 1935, abandonou a matemática. Mudou-se para o campo, onde viveu modestamente. Casou-se em 1944 e ficou completamente dependente da esposa à medida que ia ficando cego. Zermelo perdeu totalmente a visão por volta de 1951. Faleceu em Günterstal, na Alemanha, em 21 de maio de 1953.

Leitura complementar Para uma biografia completa de Zermelo, ver Ebbinghaus (2015). Os artigos fundadores de 1904 e 1908 de Zermelo podem ser lidos no alemão original (Zermelo, 1904, 1908b). As obras reunidas de Zermelo, incluindo os escritos sobre física, estão disponíveis em tradução inglesa em Ebbinghaus et al. (2010); Ebbinghaus and Kanamori (2013).

Créditos das Imagens

Georg Cantor, p. 240: Retrato de Georg Cantor por Otto Zeth, cortesia do Arquivo Universitário, Martin-Luther-Universität Halle–Wittenberg. UAHW Rep. 40-VI, Nr. 3 Bild 102.

Kurt Gödel, p. 242: Retrato de Kurt Gödel, c. 1925, fotógrafo desconhecido. Do Shelby White and Leon Levy Archives Center, Institute for Advanced Study, Princeton, NJ, EUA, em depósito na Biblioteca da Universidade de Princeton, Divisão de Manuscritos, Departamento de Livros Raros e Coleções Especiais, Kurt Gödel Papers, (C0282), Box 14b, #110000. O *Open Logic Project* obteve permissão do Arquivo do Instituto para usar esta imagem em materiais derivados do OLP não comerciais. É necessária autorização do Arquivo para qualquer outro uso.

Bertrand Russell, p. 244: Retrato de Bertrand Russell, c. 1907, cortesia da Divisão William Ready de Arquivos e Coleções de Investigação, Biblioteca da McMaster University. Arquivos Bertrand Russell, Box 2, f. 4.

Alfred Tarski, p. 246: Fotografia de passaporte de Alfred Tarski, 1939. Recortada e restaurada a partir de digitalização do passaporte de Tarski por Joel Fuller. Original cortesia da Biblioteca Bancroft, University of California, Berkeley. Alfred Tarski Papers, Banc MSS 84/49. O *Open Logic Project* obteve permissão

para usar esta imagem em materiais derivados do OLP não comerciais. É necessária autorização da Biblioteca Bancroft para qualquer outro uso.

Ernst Zermelo, p. 247: Retrato de Ernst Zermelo, c. 1922, cortesia da Abteilung für Handschriften und Seltene Drucke, Niedersächsische Staats- und Universitätsbibliothek Göttingen, Cod. Ms. D. Hilbert 754, Bl. 6 Nr. 25.

Referências Bibliográficas

Baaz, Matthias, Christos H. Papadimitriou, Hilary W. Putnam, Dana S. Scott, and Charles L. Harper Jr. 2011. *Kurt Gödel and the Foundations of Mathematics: Horizons of Truth.* Cambridge: Cambridge University Press.

Banach, Stefan and Alfred Tarski. 1924. Sur la décomposition des ensembles de points en parties respectivement congruentes. *Fundamenta Mathematicae* 6: 244–77.

Benacerraf, Paul. 1965. What numbers could not be. *The Philosophical Review* 74(1): 47–73.

Berkeley, George. 1734. *The Analyst; or, a Discourse Adressed to an Infidel Mathematician.*

Boolos, George. 1971. The iterative conception of set. *The Journal of Philosophy* 68(8): 215–31.

Boolos, George. 1989. Iteration again. *Philosophical Topics* 17(2): 5–21.

Boolos, George. 2000. Must we believe in set theory? In *Between Logic and Intuition: Essays in Honor of Charles Parsons*, eds. Gila

Sher and Richard Tieszen, 257–68. Cambridge: Cambridge University Press.

Burali-Forti, Cesare. 1897. Una questione sui numeri transfiniti. *Rendiconti del Circolo Matematico di Palermo* 11: 154–64.

Button, Tim. 2021. Level theory, part 1: Axiomatizing the bare idea of a cumulative hierarchy of sets. *The Bulletin of Symbolic Logic* 27(4): 436–460.

Cantor, Georg. 1878. Ein Beitrag zur Mannigfaltigkeitslehre. *Journal für die reine und angewandte Mathematik* 84: 242–58.

Cantor, Georg. 1883. *Grundlagen einer allgemeinen Mannigfaltigkeitslehre. Ein mathematisch-philosophischer Versuch in der Lehre des Unendlichen.* Leipzig: Teubner.

Cantor, Georg. 1892. Über eine elementare Frage der Mannigfaltigkeitslehre. *Jahresbericht der deutschen Mathematiker-Vereinigung* 1: 75–8.

Cohen, Paul J. 1963. The independence of the continuum hypothesis. *Proceedings of the National Academy of Sciences of the United States of America* 24: 556–557.

Cohen, Paul J. 1966. *Set Theory and the Continuum Hypothesis.* Reading, MA: Benjamin.

Conway, John. 2006. The power of mathematics. In *Power*, eds. Alan Blackwell and David MacKay, Darwin College Lectures. Cambridge: Cambridge University Press. URL `http://www.cs.toronto.edu/~mackay/conway.pdf`.

Corcoran, John. 1983. *Logic, Semantics, Metamathematics.* Indianapolis: Hackett, 2nd ed.

Dauben, Joseph. 1990. *Georg Cantor: His Mathematics and Philosophy of the Infinite.* Princeton: Princeton University Press.

Dedekind, Richard. 1888. *Was sind und was sollen die Zahlen?* Braunschweig: Vieweg.

du Sautoy, Marcus. 2014. A brief history of mathematics: Georg Cantor. URL `http://www.bbc.co.uk/programmes/b00ss1j0`. Audio Recording.

Duncan, Arlene. 2015. The Bertrand Russell Research Centre. URL `http://russell.mcmaster.ca/`.

Ebbinghaus, Heinz-Dieter. 2015. *Ernst Zermelo: An Approach to his Life and Work*. Berlin: Springer-Verlag.

Ebbinghaus, Heinz-Dieter, Craig G. Fraser, and Akihiro Kanamori. 2010. *Ernst Zermelo. Collected Works*, vol. 1. Berlin: Springer-Verlag.

Ebbinghaus, Heinz-Dieter and Akihiro Kanamori. 2013. *Ernst Zermelo: Collected Works*, vol. 2. Berlin: Springer-Verlag.

Feferman, Anita and Solomon Feferman. 2004. *Alfred Tarski: Life and Logic*. Cambridge: Cambridge University Press.

Feferman, Solomon, John W. Dawson Jr., Stephen C. Kleene, Gregory H. Moore, Robert M. Solovay, and Jean van Heijenoort. 1986. *Kurt Gödel: Collected Works. Vol. 1: Publications 1929–1936*. Oxford: Oxford University Press.

Feferman, Solomon, John W. Dawson Jr., Stephen C. Kleene, Gregory H. Moore, Robert M. Solovay, and Jean van Heijenoort. 1990. *Kurt Gödel: Collected Works. Vol. 2: Publications 1938–1974*. Oxford: Oxford University Press.

Feferman, Solomon and Azriel Levy. 1963. Independence results in set theory by Cohen's method II. *Notices of the American Mathematical Society* 10: 593.

Fraenkel, Abraham. 1922. Über den Begriff 'definit' und die Unabhängigkeit des Auswahlaxioms. *Sitzungsberichte der Preussischen Akadademie der Wissenschaften, Physikalisch-mathematische Klasse* 253–257.

Frege, Gottlob. 1884. *Die Grundlagen der Arithmetik: Eine logisch mathematische Untersuchung über den Begriff der Zahl.* Breslau: Wilhelm Koebner. Translation in Frege (1953).

Frege, Gottlob. 1953. *Foundations of Arithmetic*, ed. J. L. Austin. Oxford: Basil Blackwell & Mott, 2nd ed.

Giaquinto, Marcus. 2007. *Visual Thinking in Mathematics.* Oxford: Oxford University Press.

Gödel, Kurt. 1929. Über die Vollständigkeit des Logikkalküls [On the completeness of the calculus of logic]. Dissertation, Universität Wien. Reprinted and translated in Feferman et al. (1986), pp. 60–101.

Gödel, Kurt. 1931. über formal unentscheidbare Sätze der *Principia Mathematica* und verwandter Systeme I [On formally undecidable propositions of *Principia Mathematica* and related systems I]. *Monatshefte für Mathematik und Physik* 38: 173–198. Reprinted and translated in Feferman et al. (1986), pp. 144–195.

Gödel, Kurt. 1938. The consistency of the axiom of choice and the generalized continuum hypothesis. *Proceedings of the National Academy of Sciences of the United States of America* 50: 1143–8.

Gouvêa, Fernando Q. 2011. Was Cantor surprised? *American Mathematical Monthly* 118(3): 198–209.

Grattan-Guinness, Ivor. 1971. Towards a biography of Georg Cantor. *Annals of Science* 27(4): 345–391.

Hartogs, Friedrich. 1915. Über das Problem der Wohlordnung. *Mathematische Annalen* 76: 438–43.

Hausdorff, Felix. 1914. Bemerkung über den Inhalt von Punktmengen. *Mathematische Annalen* 75: 428–34.

Heck, Richard Kimberly. 2012. *Reading Frege's Grundgesetze.* Oxford: Oxford University Press.

Heijenoort, Jean van. 1967. *From Frege to Gödel: A Source Book in Mathematical Logic, 1879–1931.* Cambridge, MA: Harvard University Press.

Hilbert, David. 1891. Über die stetige Abbildung einer Linie auf ein Flächenstück. *Mathematische Annalen* 38(3): 459–460.

Hilbert, David. 2013. *David Hilbert's Lectures on the Foundations of Arithmetic and Logic 1917–1933*, eds. William Bragg Ewald and Wilfried Sieg. Heidelberg: Springer.

Hume, David. 1740. *A Treatise of Human Nature.* London.

Incurvati, Luca. 2020. *Conceptions of Set and the Foundations of Mathematics.* Cambridge: Cambridge University Press.

Irvine, Andrew David. 2015. Sound clips of Bertrand Russell speaking. URL `http://plato.stanford.edu/entries/russell/russell-soundclips.html`.

John Dawson, Jr. 1997. *Logical Dilemmas: The Life and Work of Kurt Gödel.* Boca Raton: CRC Press.

Katz, Karin Usadi and Mikhail G. Katz. 2012. Stevin numbers and reality. *Foundations of Science* 17(2): 109–23.

Kunen, Kenneth. 1980. *Set Theory: An Introduction to Independence Proofs.* New York: North Holland.

Lévy, Azriel. 1960. Axiom schemata of strong infinity in axiomatic set theory. *Pacific Journal of Mathematics* 10(1): 223–38.

LibriVox. n.d. Bertrand Russell. URL https://librivox.org/author/1508?primary_key=1508&search_category=author&search_page=1&search_form=get_results. Collection of public domain audiobooks.

Linnebo, Øystein. 2010. Predicative and impredicative definitions. *Internet Encyclopedia of Philosophy* URL http://www.iep.utm.edu/predicat/.

Linsenmayer, Mark. 2014. The partially examined life: Gödel on math. URL http://www.partiallyexaminedlife.com/2014/06/16/ep95-godel/. Podcast audio.

Maddy, Penelope. 1988a. Believing the axioms I. *The Journal of Symbolic Logic* 53(2): 481–511.

Maddy, Penelope. 1988b. Believing the axioms II. *The Journal of Symbolic Logic* 53(3): 736–64.

Montague, Richard. 1961. Semantic closure and non-finite axiomatizability I. In *Infinitistic Methods: Proceedings of the Symposium on Foundations of Mathematics (Warsaw 1959)*, 45–69. New York: Pergamon.

Montague, Richard. 1965. Set theory and higher-order logic. In *Formal systems and recursive functions*, eds. John Crossley and Michael Dummett, 131–48. Amsterdam: North-Holland. Proceedings of the Eight Logic Colloquium, July 1963.

O'Connor, John J. and Edmund F. Robertson. 2005. The real numbers: Stevin to Hilbert URL http://www-history.mcs.st-and.ac.uk/HistTopics/Real_numbers_2.html.

Peano, Giuseppe. 1890. Sur une courbe, qui remplit toute une aire plane. *Mathematische Annalen* 36(1): 157–60.

Potter, Michael. 2004. *Set Theory and its Philosophy*. Oxford: Oxford University Press.

Ramsey, Frank Plumpton. 1925. The foundations of mathematics. *Proceedings of the London Mathematical Society* 25: 338–384.

Robinson, Raphael. 1947. On the decomposition of spheres. *Fundamenta Mathematicae* 34(1): 246–60.

Rose, Daniel. 2012. A song about Georg Cantor. URL `https://www.youtube.com/watch?v=QUP5Z4Fb5k4`. Audio Recording.

Rose, Nicholas J. 2010. Hilbert-type space-filling curves. URL `https://web.archive.org/web/20151010184939/http://www4.ncsu.edu/~njrose/pdfFiles/HilbertCurve.pdf`.

Russell, Bertrand. 1905. On denoting. *Mind* 14: 479–493.

Russell, Bertrand. 1919. *Introduction to Mathematical Philosophy*. London: Allen & Unwin.

Russell, Bertrand. 1967. *The Autobiography of Bertrand Russell*, vol. 1. London: Allen and Unwin.

Russell, Bertrand. 1968. *The Autobiography of Bertrand Russell*, vol. 2. London: Allen and Unwin.

Russell, Bertrand. 1969. *The Autobiography of Bertrand Russell*, vol. 3. London: Allen and Unwin.

Russell, Bertrand. n.d. Bertrand Russell on smoking. URL `https://www.youtube.com/watch?v=80oLTiVW_lc`. Video Interview.

Scott, Dana. 1974. Axiomatizing set theory. In *Axiomatic Set Theory II*, ed. Thomas Jech, 207–14. American Mathematical Society. Proceedings of the Symposium in Pure Mathematics of the American Mathematical Society, July–August 1967.

Shoenfield, Joseph R. 1977. Axioms of set theory. In *Handbook of Mathematical Logic*, ed. Jon Barwise, 321–44. London: North-Holland.

Sigmund, Karl, John Dawson, Kurt Mühlberger, Hans Magnus Enzensberger, and Juliette Kennedy. 2007. Kurt Gödel: Das Album–The Album. *The Mathematical Intelligencer* 29(3): 73–76.

Skolem, Thoralf. 1922. Einige Bemerkungen zur axiomatischen Begründung der Mengenlehre. In *Wissenschaftliche Vorträge gehalten auf dem fünften Kongress der skandanivschen Mathematiker in Helsingfors vom 4. bis zum 7. Juli 1922*, 137–52. Akademiska Bokhandeln.

Smith, Peter. 2013. *An Introduction to Gödel's Theorems.* Cambridge: Cambridge University Press.

Takeuti, Gaisi, Nicholas Passell, and Mariko Yasugi. 2003. *Memoirs of a Proof Theorist: Gödel and Other Logicians.* Singapore: World Scientific.

Tarski, Alfred. 1981. *The Collected Works of Alfred Tarski*, vol. I–IV. Basel: Birkhäuser.

Tomkowicz, Grzegorz and Stan Wagon. 2016. *The Banach-Tarski Paradox.* Cambridge: Cambridge University Press.

Vitali, Giuseppe. 1905. *Sul problema della misura dei gruppi di punti di una retta.* Bologna: Gamberini e Parmeggiani.

von Neumann, John. 1925. Eine Axiomatisierung der Mengenlehre. *Journal für die reine und angewandte Mathematik* 154: 219–40.

Wang, Hao. 1990. *Reflections on Kurt Gödel.* Cambridge: MIT Press.

Weston, Tom. 2003. The Banach-Tarski paradox URL `http://people.math.umass.edu/~weston/oldpapers/banach.pdf`.

Whitehead, Alfred North and Bertrand Russell. 1910. *Principia Mathematica*, vol. 1. Cambridge: Cambridge University Press.

Zermelo, Ernst. 1904. Beweis, daß jede Menge wohlgeordnet werden kann. *Mathematische Annalen* 59: 514–516. English translation in (Ebbinghaus et al., 2010, pp. 115–119).

Zermelo, Ernst. 1908a. Untersuchungen über die Grundlagen der Mengenlehre I. *Mathematische Annalen* 65: 261–81.

Zermelo, Ernst. 1908b. Untersuchungen über die Grundlagen der Mengenlehre I. *Mathematische Annalen* 65(2): 261–281. English translation in (Ebbinghaus et al., 2010, pp. 189-229).

Siglas e abreviaturas

Abreviaturas editoriais e siglas de teorias listadas abaixo. Símbolos formais e convenções de escrita estão no capítulo *Notação*; vocabulário matemático em prosa, no *Glossário*.

Cf. *Conferir* (latim *confer*): indica ao leitor que deve comparar ou consultar outra passagem.

p.ex.; **P.ex.** *Por exemplo* (equivalente do inglês *e.g.*); *P.ex.* reserva-se ao início de frase.

ca. *Circa*: por volta de (datas, sobretudo em créditos de imagens).

sse *Se e somente se* (equivalente do inglês *iff* / *if and only if*).

Z Teoria de Zermelo *com* o axioma de Fundação.

Z$^-$ Teoria de Zermelo *sem* o axioma de Fundação.

ZF, **ZF**$^-$ Teoria de Zermelo–Fraenkel; a variante **ZF**$^-$ dispensa Fundação.

ZFC Zermelo–Fraenkel com o Axioma da Escolha.

LT Teoria dos Níveis (*Level Theory*).

Zr Extensão de **LT** com axiomas de Infinito e Conjunto potência, na qual todo o conjunto tem posto.

T Letra ou nome curto para uma *teoria* (conjunto de sentenças) de primeira ordem com símbolo primitivo $\in$.

HC Hipótese do Contínuo.

HGC Hipótese Generalizada do Contínuo.

(H)HC Referência conjunta à HC ou à HGC quando o raciocínio é paralelo.

TS *Tarski–Scott*: linha de definição de cardinalidade que evita o teorema do bom ordenamento (capítulo correspondente).

Notação

Símbolos e abreviaturas usados com frequência neste volume (incluindo capítulos de pré-requisitos sobre conjuntos, relações e funções). Quando o mesmo símbolo tiver mais de um uso, o contexto (ou a definição na secção indicada) desambigua. Para vocabulário e noções em prosa, ver o *Glossário*.

Lógica e teorias

$\vdash$ Relação de dedutibilidade sintática ("a teoria prova…").

$\rightarrow$, $\leftrightarrow$, $\neg$, $\wedge$, $\vee$ Implicação, bicondicional, negação, conjunção, disjunção.

$\exists$, $\forall$ Quantificadores existencial e universal.

T Letra ou nome curto para uma teoria (conjunto de sentenças) de primeira ordem com símbolo primitivo $\in$.

Conjuntos, relações e ordens

$\emptyset$ Conjunto vazio.

$\in$, $\subseteq$, $\subsetneq$ Pertence a; subconjunto; subconjunto próprio.

$\cup$, $\cap$, $\setminus$ União, intersecção, diferença.

$\wp(X)$, $\wp(X)$ Conjunto potência de X.

$\langle a, b\rangle$, $\langle a_1, \ldots, a_n\rangle$ Par ordenado; n-upla ordenada.

$\{x : \varphi(x)\}$ Conjunto dos x que verificam $\varphi(x)$ (quando definível no contexto).

$[a]_R$, $A/_R$ Classe de equivalência de a; conjunto das classes em A.

Id_A Relação identidade sobre A: $\{\langle x, x\rangle : x \in A\}$.

R^n, $R \mid S$ Potências e composição de relações (definidas no capítulo de relações).

R^+ Fecho transitivo da relação R (união das potências R^n com $n \geq 1$). *Atenção:* em algumas passagens sobre ordens, R^+ denota $R \cup \mathrm{Id}_A$ (reflexivação de uma ordem estrita); o contexto distingue.

R^* Fecho reflexivo transitivo: $R^+ \cup \mathrm{Id}_A$ (definição padrão no capítulo de relações).

Funções

$\mathrm{dom}(f)$ Domínio de f.

$\mathrm{Im}(f)$ Imagem de f (neste volume a imagem aparece sob esta forma).

$f \restriction A$ Restrição de f ao subconjunto A.

$g \circ f$ Composição (aplicar primeiro f, depois g).

${}^X Y$ Conjunto de todas as funções de X para Y.

$f[A]$ Imagem do conjunto A por f: $\{f(x) : x \in A\}$.

$\mathrm{comp}(\sigma)$ Comprimento de uma sequência finita σ.

Números habituais

$\mathbb{N}$, $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$ Naturais, inteiros, racionais, reais.

Ordinais e aritmética ordinal

ω Primeiro ordinal limite (identificado com os naturais de von Neumann).

α^+ Sucessor ordinal de α.

$+, \cdot$ **entre ordinais** Adição e multiplicação *ordinais* (símbolos iguais aos da aritmética usual; o contexto distingue).

$\alpha^{(\beta)}$ Exponenciação ordinal, quando aparece explicitamente.

$\mathrm{ord}(A, R)$ Tipo de ordem da estrutura bem ordenada $\langle A, R\rangle$.

$\lessdot$ Ordem lexicográfica restrita (usada em produtos de ordinais ordenados).

$\sqcup$ Soma ordenada disjunta de ordinais (soma como ordem em cópias disjuntas).

$\mathrm{lsub}(X)$ Para X conjunto de ordinais, o menor ordinal estritamente maior que todo o membro de X (no PDF aparece como lsub(X), convenção herdada do projeto original em inglês).

Cardinalidade e equipotência

$|X|$ Cardinalidade (tamanho) do conjunto X.

$A \approx B$, $A \preceq B$, $A \prec B$ Comparação de tamanhos: equipotência ($\approx$); não maior ($\preceq$); estritamente menor ($\prec$), no sentido dos capítulos sobre cardinalidade.

$A \not\approx B$, $\alpha \not\leq A$ Negação de equipotência; e formulações do tipo "não existe injeção de um ordinal em A" usadas, por exemplo, no lema de Hartogs.

$\prec$, $\preceq$, $\approx$ Ordem e equivalência entre cardinais (no sentido introduzido nos capítulos de cardinais).

$\aleph_\alpha$, $\beth_\alpha$ Sucessões transfinitas de cardinais infinitos (*alef* e *beth*).

$\mathfrak{a}$, $\mathfrak{b}$ Letras góticas para designar cardinais.

$\oplus$, $\otimes$, **potências cardinais** Soma, produto e exponenciação entre cardinais.

$\kappa^{\oplus}$ Sucessor cardinal, quando notado explicitamente.

Hierarquia cumulativa, posto e fecho

V_α Estágio de nível α na hierarquia cumulativa de conjuntos.

posto(x) Posto do conjunto x: o menor ordinal α com $x \subseteq V_\alpha$ (em inglês costuma dizer-se *rank*).

ft(x) Fecho transitivo do *conjunto* x. A abreviatura *ft* segue as iniciais de *fecho transitivo*; em inglês usa-se muitas vezes *trcl* (*tr*ansitive *cl*osure).

clo$_f$(o) Menor conjunto que contém o e é fechado sob a função f (estruturas de Dedekind, naturais de von Neumann, etc.).

Objetos tipo Frege e cardinalidade Tarski–Scott

$\#x\,\varphi(x)$ Notação abreviada para certos objetos definidos a partir de fórmulas (número de Hume, extensões de Frege, conforme o contexto).

tsc(A) Cardinalidade no sentido *Tarski–Scott* de um conjunto A.

Geometria e medida (paradoxos da escolha)

S Esfera (contexto de decomposição paradoxal).

R Grupo de rotações usado nas construções de Vitali e afins.

Glossário

Este glossário reúne **termos** (palavras e expressões usadas no discurso matemático do livro). Os **símbolos** e as **listas formais** de abreviaturas teóricas estão no capítulo *Notação.* As entradas agrupam-se por tema, na ordem de leitura; dentro de cada tema, vão em ordem alfabética em português.

Prelúdio e motivação histórica

Continuidade Noção de "sem saltos" para funções sobre o contínuo, afinada no século XIX com argumentos do tipo épsilon-delta e no rigor de Cauchy e Weierstrass; motiva a necessidade de fundamentos.

Curva (Jordan) Neste livro, mapa contínuo de um intervalo para o plano; contexto para curvas preenchedoras e exemplos geométricos "patológicos".

Curva preenchedora de espaço Mapa contínuo que sobrecarrega uma dimensão (por exemplo curvas de Peano ou de Hilbert), desafiando a intuição de dimensão e pressionando a noção rigorosa de função e de conjunto.

Derivada Taxa de variação instantânea, desenvolvida primeiro com ideias de infinitesimal e depois no rigor moderno.

Diferenciabilidade Propriedade de admitir derivada; contrasta com continuidade e com exemplos limite discutidos no prelúdio.

Função No prelúdio, objeto central do cálculo; mais adiante, tratada conjuntisticamente como relação funcional ou conjunto de pares ordenados.

Infinitesimal Quantidade "infinitamente pequena" usada heuristicamente antes do rigor moderno; parte da história que leva à formalização.

Limite Conceito que articula aproximação e continuidade; peça-chave na crise de rigor do século XIX.

Mitologia da matemática Expressão do texto para narrativas simplificadoras ou exageradas sobre descoberta e progresso; contrapõe-se ao estudo histórico sério.

Patologia / patológico Exemplos que violam intuições ingénuas (funções, curvas, cardinalidades), muitas vezes motores de novo rigor.

Rigor matemático Padrões de definição e prova que tornam os resultados auditáveis; o prelúdio explica parte da pressão histórica em direção à teoria dos conjuntos.

Teoria dos conjuntos Quadro no qual, no fim do livro, se modelam números, funções e análise; o prelúdio apresenta parte das razões históricas dessa adopção.

Conjuntos (teoria ingênua)

Coleção extensional Ideia de que um conjunto fica determinado pelos seus membros, não pela maneira de o descrever; base da extensionalidade.

Compreensão (princípio ingénuo) Princípio de formar o conjunto de todos os objetos que satisfazem uma condição; sob a ingenuidade total, conduz ao paradoxo de Russell.

Conjunto Coleção considerada como objeto matemático; na parte ingénua, tratada informalmente antes dos axiomas restritivos.

Conjunto vazio Conjunto sem elementos; ponto de partida de muitas construções.

Conjunto potência Conjunto de todos os subconjuntos de um conjunto dado; aparece no axioma homónimo na teoria formal.

Diferença de conjuntos Conjunto dos elementos da primeira parcela que não pertencem à segunda.

Elemento / membro Objeto que "pertence" a um conjunto; o texto distingue léxico conforme o original em inglês (*element* / *member*).

Extensionalidade (ingénua) Princípio de que conjuntos com os mesmos membros são iguais; na teoria formal torna-se o axioma de extensionalidade.

Intersecção Conjunto dos objetos que pertencem simultaneamente a todos os conjuntos de uma família; inclui intersecção de dois conjuntos e intersecção de uma família arbitrária.

Par ordenado Objeto que codifica ordem; base para relações e funções como conjuntos.

Paradoxo de Russell Contradição obtida ao considerar "o conjunto de todos os conjuntos que não são membros de si próprios"; motiva a teoria axiomática.

Parte própria Subconjunto estrito; usado na definição de infinito de Dedekind.

Produto cartesiano Conjunto de todos os pares ordenados cujas primeiras componentes vêm de um conjunto e as segundas de outro (e generalizações).

Subconjunto Conjunto cujos membros são todos membros de outro; subconjunto próprio se, além disso, falta pelo menos um membro do conjunto maior.

Tupla Generalização de par ordenado para várias componentes; usada em aritmetização e em estruturas ordenadas.

União Conjunto dos membros dos membros de uma família; inclui união de dois conjuntos e união de uma família arbitrária.

União disjunta Construção que "etiqueta" cópias disjuntas de conjuntos; usada em ordinais e cardinalidade para somas ordenadas.

Relações e ordens (teoria ingênua)

Anti-simetria Se dois elementos se relacionam mutuamente, então são iguais; típica de ordens parciais reflexivas.

Assimetria Nunca dois elementos distintos se relacionam mutuamente nos dois sentidos; típica de ordens estritas.

Classe de equivalência Conjunto dos elementos relacionados a um dado membro por uma relação de equivalência.

Conexa (relação) Para quaisquer dois elementos distintos, um relaciona-se com o outro num dos sentidos (ordem total).

Divisibilidade Exemplo de ordem parcial sobre os inteiros; ilustra relações não lineares.

Equivalência (relação) Relação reflexiva, simétrica e transitiva; particiona o domínio em classes.

Fecho reflexivo Acrescentar à relação todos os pares na diagonal (cada elemento relacionado consigo); atenção à sobrecarga do símbolo de fecho transitivo noutros capítulos.

Fecho reflexivo transitivo Menor relação reflexiva e transitiva que contém a relação dada; no capítulo de operações aparece construída a partir do fecho transitivo e da diagonal.

Fecho transitivo (relação) Menor relação transitiva que contém a relação dada, obtida como união das potências iteradas da relação.

Identidade (relação) Relação que relaciona cada elemento apenas consigo, sobre um conjunto dado.

Irreflexividade Nenhum elemento se relaciona consigo.

Ordem estrita Relação irreflexiva e transitiva (e, em geral, assimétrica); contrapõe-se à ordem reflexiva.

Ordem linear / total Ordem parcial conexa: quaisquer dois elementos são comparáveis.

Ordem parcial Relação reflexiva, anti-simétrica e transitiva.

Pré-ordem Relação reflexiva e transitiva (não exige anti-simetria).

Quociente Conjunto de classes de equivalência sob uma relação de equivalência.

Reflexividade Todo elemento relaciona-se consigo.

Relação binária Subconjunto de um produto cartesiano; base para ordens, equivalências e funções.

Relação inversa Relação obtida invertendo a ordem dos pares.

Restrição (relação) Relação obtida considerando apenas elementos de um subconjunto do domínio.

Simetria Se um elemento relaciona-se com outro, o segundo relaciona-se com o primeiro.

Transitividade Se o primeiro relaciona-se com o segundo e o segundo com o terceiro, então o primeiro relaciona-se com o terceiro.

Funções (teoria ingênua)

Aplicação Termo usado quando o original distingue *application* de *map*; designa o valor ou o acto de aplicar uma função a um argumento.

Bijeção Função simultaneamente injetiva e sobrejetiva; admite inversa bilateral.

Composição de funções Aplicar primeiro uma função e depois outra; exige compatibilidade entre imagem da primeira e domínio da segunda.

Contradomínio Conjunto onde a função toma valores declarados; pode ser estritamente maior que a imagem.

Domínio Conjunto de argumentos para os quais a função está definida.

Extensionalidade (funções) Duas funções com o mesmo domínio e os mesmos valores ponto a ponto são iguais.

Função Relação em que cada argumento do domínio tem exactamente um valor; identificada com o seu gráfico.

Gráfico (de função) Conjunto de pares argumento-valor; identifica a função com uma relação binária especial.

Imagem Conjunto dos valores que a função toma no seu domínio.

Injeção Função que nunca envia dois argumentos distintos para o mesmo valor.

Inversa Função que reverte outra; existência global equivalente à bijeção, ou uso de inversas à esquerda ou à direita.

Mapa / mapeamento Sinónimos estáveis de *map* / *mapping* no original; designam função.

Relação funcional Relação em que cada elemento do domínio tem exactamente um sucessor pela relação.

Restrição (função) Função obtida limitando o domínio a um subconjunto, mantendo os mesmos valores.

Sobrejeção Função cuja imagem coincide com o contradomínio declarado.

Valor Objeto associado a um argumento pela função.

Cardinalidade ingênua

Cardinalidade (informal) "Tamanho" de um conjunto, refinado depois por equipotência e pela teoria dos cardinais.

Confinito Conjunto equipotente a um subconjunto próprio de si; na presença dos axiomas usuais, alinha-se com o infinito de Dedekind.

Diagonalização Argumento que constrói um objecto diferindo de cada linha de uma lista (Cantor, conjunto potência dos naturais, sequências binárias).

Enumerável Equipotente ao conjunto dos naturais ou a um segmento inicial finito; no texto ingénuo, também caracterizado por injeção nos naturais.

Enumeração Lista ou bijeção que organiza todos os membros de um conjunto com índices naturais.

Equipotência Relação entre dois conjuntos quando existe bijeção de um para o outro; "mesmo tamanho" em sentido conjuntista.

Não enumerável Conjunto que não é enumerável; exemplo paradigmático no livro: o conjunto potência dos naturais.

Redução (enumerabilidade) Provar não enumerabilidade transportando a dificuldade a partir de outro conjunto já conhecido por não enumerável.

Teorema de Cantor Para qualquer conjunto, o seu conjunto potência tem cardinalidade estritamente maior; não existe sobrejeção do conjunto para o seu conjunto potência.

Teorema de Schröder–Bernstein Se cada um de dois conjuntos injeta no outro, então são equipotentes; a prova completa é adiada até à teoria dos cardinais.

Aritmetização e construção dos números

Construção de Cauchy Reais como classes de equivalência de sequências de Cauchy de racionais.

Cortes de Dedekind Partições dos racionais em dois conjuntos que modelam a completude da recta real.

Inteiros Construídos a partir dos naturais por equivalência em pares (diferenças formais).

Racionais Construídos como classes de equivalência de pares de inteiros (fracções formais).

Reais Construídos por cortes de Dedekind ou por Cauchy; completam a cadeia natural dos sistemas numéricos no livro.

Sucessor aritmético Operação que gera o "próximo" natural na estrutura deduzida dos axiomas de Peano codificados em teoria dos conjuntos.

Infinito, naturais e indução

Álgebra de Dedekind Estrutura com sucessor injectivo, elemento inicial fora da imagem do sucessor e princípio da indução; modelo para os naturais.

Fecho indutivo Menor conjunto que contém um gerador e é fechado sob uma operação; no tratamento de von Neumann coincide com o conjunto dos naturais.

Indução matemática Princípio que permite provar propriedades para todos os naturais a partir do caso inicial e do passo indutivo.

Infinito de Dedekind Conjunto equipotente a uma parte própria de si; caracterização do infinito antes do desenvolvimento completo da hierarquia ordinal.

Números naturais No von Neumann, o conjunto dos ordinais finitos; base para a aritmetização no livro.

Recursão Método para definir funções nos naturais (e generalizações) especificando valor inicial e passo iterativo.

Sucessor (von Neumann) Operação conjuntista que gera o próximo estágio a partir de um conjunto; usada na construção dos naturais e dos ordinais.

Concepção iterativa, paradoxos e filosofia

Compreensão impredicativa Definição de um objeto por referência a uma totalidade que inclui o próprio objeto; tema na crítica predicativista.

Concepção iterativa Visão de que os conjuntos são formados estágio a estágio, em contraste com a "totalidade" ingénua.

Extensionalidade (justificação) Motivação filosófica para identificar conjuntos pela coincidência de membros.

Grundgesetze Obra de Frege; sistema em que a compreensão irrestrita conduz à contradição.

Paradoxo de Russell (revisitado) Retorno ao paradoxo após a narrativa axiomática; liga ingenuidade, classes e diagonalização.

Predicatividade Restrição a definições que não quantificam sobre totais "maiores" que o objeto definido.

Urelemento Objecto sem membros que não é o conjunto vazio; discutido como opcional face à teoria pura de conjuntos.

Axiomas e fragmentos de teoria

Axioma da escolha Existência de função de escolha para famílias de conjuntos não vazios; caracteriza a teoria com escolha face à sem escolha.

Axioma da extensionalidade Conjuntos com os mesmos membros são iguais.

Axioma da fundamentação Cada conjunto não vazio tem um membro disjunto dele; evita ciclos de pertença e patologias associadas.

Axioma da regularidade Formulação equivalente à fundamentação no desenvolvimento do livro (equivalência demonstrada módulo o fragmento adequado da teoria).

Axioma da união Para qualquer conjunto, existe a união dos seus membros.

Axioma do conjunto potência Para qualquer conjunto, existe o conjunto de todos os seus subconjuntos.

Axioma do infinito Afirma a existência de um conjunto indutivo adequado para fundar os naturais de von Neumann.

Axioma dos pares Para quaisquer dois objectos, existe o par (não ordenado) formado por eles.

Esquema de separação Para cada propriedade definível, o subconjunto dos membros de um conjunto dado que a verificam é conjunto; restringe a compreensão.

Esquema de substituição A imagem de um conjunto por uma relação funcional definível é conjunto; distingue a força de Zermelo–Fraenkel da de Zermelo puro.

Teoria de Zermelo–Fraenkel Teoria de primeira ordem com predicado de pertença, axiomas de Zermelo e substituição; variantes sem escolha ou com escolha aparecem no texto.

Teoria de Zermelo Fragmento com os axiomas de Zermelo (sem substituição, conforme o enquadramento do livro).

Teoria dos Níveis Teoria mais fraca baseada na ideia de "níveis"; relacionada com Zermelo–Fraenkel e com extensões discutidas na parte de substituição.

Extensão Zr Teoria obtida acrescentando axiomas fortes à Teoria dos Níveis; no texto, todo conjunto tem posto e a teoria supera Zermelo puro.

Ordinais

Boa ordenação Ordem linear em que todo subconjunto não vazio tem mínimo; em certos contextos, equivalente a princípios de escolha.

Indução transfinita Generalização da indução para ordinais arbitrários; ferramenta central em provas sobre a hierarquia e sobre postos.

Limite (ordinal) Ordinal não zero que não é sucessor; acumula todos os ordinais menores.

Ordinal No sentido de von Neumann: conjunto transitivo bem ordenado pela pertença.

Ordinal inicial Menor ordinal de uma dada cardinalidade infinita; liga cardinais a ordinais escolhidos canonicamente.

Ordem bem ordenada Estrutura ordenada isomorfa a um ordinal, ou, equivalentemente, estritamente bem fundamentada na ordem.

Representação ordinal Todo bem ordenamento é isomorfo a um único ordinal.

Segmento inicial Subconjunto fechado para predecessores na ordem; usado em isomorfismos de ordens.

Sucessor ordinal Próximo estágio após um ordinal dado na construção de von Neumann.

Transitivo (conjunto) Se um membro pertence a um membro do conjunto, então pertence ao próprio conjunto; propriedade dos ordinais e dos estágios da hierarquia.

Hierarquia cumulativa, posto e fundação

Estágio da hierarquia Nível indexado por um ordinal na hierarquia cumulativa; conjunto de tudo o que aparece “antes” desse ordinal.

Fecho transitivo (conjunto) Menor conjunto transitivo que contém um conjunto dado; no volume, o texto adopta uma abreviatura consagrada na lista de notação.

Hierarquia cumulativa Família de estágios indexada pelos ordinais, acumulando conjuntos estágio a estágio; modelo formal da concepção iterativa.

Posto Menor ordinal α tal que o conjunto dado é subconjunto do estágio α da hierarquia; mede profundidade na cumulativa.

Recursão na hierarquia Definições e provas por indução nos ordinais ou nos estágios da hierarquia.

Substituição e limitação de tamanho

Limitação de tamanho Princípio que impede a formação de "conjuntos demasiado grandes" sem critério; contrapõe-se à compreensão irrestrita.

Reflexão (metatema) Tópicos sobre o que a linguagem de primeira ordem consegue exprimir sobre a hierarquia vista como totalidade.

Aritmética ordinal

Adição ordinal Operação que "cola" duas boas ordenações; em geral não comutativa.

Exponenciação ordinal Potenciação transfinita baseada na multiplicação ordinal; distinta da exponenciação cardinal.

Multiplicação ordinal Repetição ordenada de uma boa ordenação; liga-se a produtos lexicográficos.

Ordem lexicográfica Ordem em produtos que compara primeiro numa coordenada e depois noutra; usada na aritmética ordinal.

Cardinais e comparabilidade

Cardinal No livro, sobretudo como ordinal inicial ou como objeto que codifica o "tamanho" de um conjunto; as sucessões alef e beth dos infinitos aparecem na parte final.

Cardinal inicial Ver *Ordinal inicial*; identifica um cardinal infinito com o menor ordinal da sua cardinalidade.

Comparabilidade de cardinais Relações de não maior, estritamente menor e equipotência entre tamanhos; sem Axioma da Escolha, a comparabilidade falha em geral.

Hartogs (ordinal de) Menor ordinal que não se injeta num conjunto dado; ferramenta para comparar tamanhos sem escolha global.

Tipo de ordem Ordinal associado a um bem ordenamento, único a menos de isomorfismo.

Aritmética cardinal e hipóteses do contínuo

Aritmética cardinal Operações de soma, produto e potência entre cardinais; leis distintas da aritmética ordinal.

Cofinalidade Menor ordinal que serve de domínio para uma função com imagem ilimitada num cardinal dado; usada em limitações de cardinalidade.

Continuo Cardinalidade da recta real; no desenvolvimento clássico, identifica-se com a potência do conjunto dos naturais.

Hipótese do Contínuo Não existe cardinal estritamente entre o dos naturais e o do continuo; independente da teoria usual com escolha.

Hipótese Generalizada do Contínuo Identificação, para todo ordinal, das duas sucessões canónicas dos infinitos que o livro nota com alef e beth; implica a Hipótese do Contínuo.

Independência da hipótese do contínuo Resultado de que nem a hipótese nem a sua negação se demonstram só com os axiomas habituais; o livro resume o panorama (Gödel, Cohen e posteriores).

Ponto fixo de alef ou de beth Cardinal que coincide com o termo indexado pela própria cardinal na sucessão alef ou beth; existência por argumentos de limite.

Sucessor cardinal Menor cardinal estritamente maior que um cardinal dado.

Escolha e princípios fracos

Axioma da boa ordenação Todo conjunto pode ser bem ordenado por alguma relação; no livro figura na enumeração da teoria com escolha e é equivalente ao Axioma da Escolha.

Conjunto não mensurável Subconjunto da recta real sem medida de Lebesgue compatível com a invariância por translação; construção associada a escolhas (Vitali).

Decomposição paradoxal Construções que decompõem sólidos e recompõem por isometrias, possíveis com o Axioma da Escolha (contexto Banach–Tarski).

Escolha enumerável Princípio que afirma escolha apenas para famílias enumeráveis de conjuntos não vazios; capítulo dedicado.

Escolha (noção) Uso frequente do texto para designar o Axioma da Escolha quando se trabalha na teoria com escolha.

Função de escolha Associa a cada membro de uma família de conjuntos não vazios um elemento desse membro; a existência global para toda família é o Axioma da Escolha.

Teorema da boa ordenação Todo conjunto admite uma boa ordenação; equivalente ao Axioma da Escolha no tratamento clássico.

Cardinalidade no sentido de Tarski e Scott

Cardinalidade Tarski–Scott Noção de "tamanho" que dispensa o teorema da boa ordenação; permite falar de cardinalidade sem assumir o Axioma da Escolha.

Número de Hume Abstracção de cardinalidade por equinumerosidade, no enquadramento filosófico do texto.

Extensão de Frege Objeto abstrato associado a uma condição; contexto histórico e lógico ligado a paradoxos e a alternativas conjuntistas.

Sobre o Open Logic Project

O *Open Logic Text* é um livro-texto colaborativo de código aberto de meta-lógica formal e métodos formais, começando em um nível intermediário (ou seja, após um curso introdutório de lógica formal). Embora voltado a um público não matemático (em particular, estudantes de filosofia e ciência da computação), é rigoroso.

A cobertura de alguns tópicos atualmente incluídos pode ainda não estar completa, e muitas seções ainda exigem revisão substancial. Planejamos expandir o texto para cobrir mais tópicos no futuro. Também planejamos adicionar recursos ao texto, como um glossário, uma lista de leituras complementares, notas históricas, imagens, melhores explicações, seções que explicam a relevância dos resultados para filosofia, ciência da computação e matemática, e mais problemas e exemplos. Se você encontrar um erro ou tiver uma sugestão, informe a equipe do projeto.

O projeto opera no espírito do código aberto. O texto não só está disponível gratuitamente; fornecemos o código-fonte LaTeX sob a licença Creative Commons Atribuição, que dá a qualquer pessoa o direito de baixar, usar, modificar, reorganizar, converter e redistribuir nosso trabalho, desde que deem o devido crédito. Consulte o site do Open Logic Project em openlogicproject.org

para informações adicionais.