

Teoria maliciosa dos conjuntos

© 2023 por Renan M. Mezabarba¹

Última Atualização: 26 de novembro de 2023.

DRAFT (RM 2023)

¹Copyright © 2023 de Renan Maneli Mezabarba. Autorizo reprodução e distribuição do texto para fins não-lucrativos desde que a autoria seja citada. Sugestões, correções, etc. podem ser enviadas para (preferencialmente) <rmmezabarba@gmail.com> ou <rmmezabarba@uesc.br>.

DRAFT (RMM 2023)

DRAFT (RMM 2023)

Look how they massacred my boy.

Don Corleone (*The Godfather*, 1972).

DRAFT (RMM 2023)

Sumário

Prefácio	7
A Zermelo-Fraenkel-Choice	9
A.1 Noções elementares	11
A.2 O problema do infinito	23
A.3 Produtos arbitrários e o Axioma da Escolha	26
A.4 Dois axiomas técnicos	29
A.5 Classes (próprias)	30
Exercícios adicionais	32
B Relações	35
B.1 Equivalências e partições	35
B.1.1 Partições e representantes	37
B.1.2 Funções quocientes e projeções	39
B.1.3 Opcional: inteiros e racionais	41
B.2 Um mínimo sobre ordens	47
B.3 Adiável: recursão	52
Exercícios adicionais	56
C Cardinalidade	59
C.1 A ideia de número cardinal	59
C.2 Cardinais enquanto <i>façon de parler</i>	63
C.3 Opcional: cardinalidades clássicas	66
C.4 Ordinais	69
C.5 Adiável: algumas recursões longas	74
Exercícios adicionais	81
D Escolhas intangíveis	83
D.1 Partições e representantes	83
D.2 Boa ordenação de Zermelo	86
D.3 Recursões de bolso: o Lema de Zorn	88
Exercícios adicionais	96
E Cardinais	99
E.1 A escadaria dos alephs	99
E.2 Dizia eu que a aritmética	103
E.2.1 Operações cardinais (caso geral)	104
E.2.2 Aritmética transfinita	107
E.2.3 Opcional: aritmética ordinal	111
E.3 Cofinalidade e pombos	113
E.4 Cardinais singulares, regulares e além	118
Exercícios adicionais	120
F Opcional: metamatemática	125
F.1 Linguagens e estruturas	126
F.2 Subestruturas, núcleos e quocientes	129
F.3 Interpretação e satisfabilidade	132
F.3.1 Pausa dramática: estruturas livres	135
F.3.2 De volta ao itinerário: fórmulas e modelos	139
F.4 Verdade ou consequência	146
F.5 Dobrando a meta	154
Exercícios adicionais	162
Lista de símbolos e siglas	168
Referências Bibliográficas	170
Índice Remissivo	171

DRAFT (RMM 2023)

Prefácio

O presente material surgiu, originalmente, como o *préâmbulo* do meu primeiro *projeto literário*, *Fundamentos de Topologia Geral* [26]. Seu objetivo era apenas fornecer os (muitos) pré-requisitos teóricos para um bom aproveitamento dos capítulos topológicos. Porém, a extensão da redação levou à sua inevitável cisão: o preâmbulo original se tornou o presente texto, enquanto [26] passou a conter uma versão muito resumida do *mesmo*. Tal separação teve um preço: reestruturar a apresentação dos conteúdos, de modo a amalgamá-los em torno de um fio condutor, algo desnecessário em sua concepção original de preâmbulo, mas imprescindível num volume individual. No caso, o fio condutor adotado foi a *axiomática* conhecida como Zermelo-Fraenkel-Choice (ZFC) para a Teoria dos Conjuntos¹. Nesse processo, alguns tópicos interessantes precisaram ser *omitidos*, como categorias, mas a coesão resultante justificou a dor.

Com isso dito, neste livro o leitor² encontrará um curso básico de Teoria dos Conjuntos sob a axiomática ZFC, com a discussão de algumas aplicações em Álgebra, Análise e *afins*. Exceto por essas eventuais aplicações, o único pré-requisito para o texto é ter algum traquejo com o *modus operandi* matemático e seus jargões, o que costuma ser adquirido num bom curso de Cálculo ou Álgebra Linear. Os conteúdos abordados se distribuem da seguinte forma:

- o Capítulo A cumpre a tarefa de postular todos os axiomas de ZFC, bem como utilizá-los para *justificar* as definições básicas (produtos cartesianos, funções, relações, etc.) que serão utilizadas nos capítulos seguintes;
- já o Capítulo B desenvolve o básico das relações de equivalência e ordem, além de fazer um primeiro contato com as definições recursivas;
- por sua vez, o Capítulo C discute as nuances entre as noções de cardinalidade e número cardinal, além de introduzir os importantes números ordinais, utilizados em recursões transfinitas *longas*;
- fica a cargo do Capítulo D enunciar e demonstrar as principais encarnações do Axioma da Escolha, além de apresentar algumas aplicações em Álgebra e Análise (nem todas evidentes);
- o Capítulo E, que encerra a parte “obrigatória” do texto, utiliza todos os ferramentais anteriores para obter resultados de aritmética cardinal, principalmente em contextos infinitários;
- ao longo do texto, os títulos de algumas (sub)seções e exemplos são precedidos dos termos “Adiável” (por serem menos urgentes) ou “Opcional” (por envolverem pré-requisitos não tratados no texto, mas que podem interessar aos leitores com bagagem prévia); não obstante, o Capítulo F, que aborda questões *metamatemáticas* inevitáveis a quem se propõe estudar *conjuntos com conjuntos*, é inteiramente opcional (embora não tenha tantos pré-requisitos).

Por ser fruto de [26], os agradecimentos lá feitos ainda se aplicam ao presente material. Apesar disso, algumas redundâncias são convenientes: Gisele T. Paula, pelas sugestões indiretas que levaram à *cisão*; Fabíola Loterio e Pedro Schineider, responsáveis pela mobilização que me permitiu ministrar uma disciplina baseada na versão original deste texto; Marcos M. Rodrigues e Priscilla S. F. Silva, pelas excelentes discussões; e Alícia Marques, pela ajuda na simplificação de diversos argumentos.

Renan M. Mezabarba

Ilhéus-BA, 26 de novembro de 2023.

¹Diga-se de passagem, “Teoria maliciosa dos conjuntos” – referência ao clássico *Naive set theory* (Teoria ingênuia dos conjuntos) de Paul Halmos – seria o título do *capítulo* dedicado a ZFC.

²Ao escrever coisas como “o leitor”, penso tão somente no gênero neutro do substantivo “leitor”. As *novas* alternativas *neutrais* ainda não me parecem totalmente estabelecidas na literatura acadêmica.

DRAFT (RMM 2023)

Capítulo A

Zermelo-Fraenkel-Choice

Intuitivamente, conjuntos constituem um meio matemático para tratar de ajuntamentos de objetos, o que pode satisfazer a ideia do leitor de *definição*. Porém, veja que enquanto uma sala S habitada por pessoas chamadas de P_0 , P_1 e P_2 é apenas uma sala, a *lista* $L := \{P_0, P_1, P_2\}$ é um conjunto *abstrato*, uma peça de informação que *diz* quais são os nomes das pessoas na sala. Tanto a sala quanto as pessoas são objetos concretos, enquanto a lista L e seus nomes são *abstraídos* desses objetos¹.

Assim, num primeiro momento, a noção de “conjunto” é quase trivial, por se tratar de um processo de generalização que antecede a própria ideia de contagem: antes de saber que uma manada M de mamutes tem 5 elementos, primeiro identifica-se o bando e abstrai-se dele um conjunto L' , para daí fazer outras considerações sobre *cardinalidades*. Contudo, conjuntos são bem mais flexíveis do que as situações *reais* sugerem: mesmo que a sala S e a manada M de mamutes estejam em continentes distintos, nada impediria que alguém considera-se o conjunto formado pelas listas descritas até aqui, digamos $T := \{L, L'\}$.

Tem-se então um primeiro ponto de confusão: as únicas coisas que pertencem a T são L e L' . Por exemplo, embora a pessoa P_1 pertença à lista L , ela própria não é uma lista e, portanto, não pertence a T , já que o último foi definido como a *coleção* das *listas anteriores*. Isso não fica tão evidente quando se pensa em conjuntos por meio de analogias físicas, o que é bastante razoável, já que conjuntos não são físicos.

De toda forma, a discussão acima sugere uma *relação de pertinência* entre *elementos* e *conjuntos*. Denotaremos tal relação com o símbolo “ \in ”, que será usado para indicar que um certo *objeto* x é *elemento* de um *conjunto* A (ou “ x pertence a A ”), caso em que se escreve “ $x \in A$ ”, bem como para indicar que um certo objeto x não é elemento de A (ou “ x não pertence a A ”), caso em que se escreve “ $x \notin A$ ”. E apesar das expressões “elemento” e “conjunto”, não haverá uma classe de objetos destinados a serem um ou outro, já que tal distinção é muito tênue num contexto que se propõe a estudar conjuntos quaisquer: veja que na discussão anterior, por exemplo, os *conjuntos* L e L' também se portaram como *elementos* do conjunto T . Dado que os objetos usuais que gostamos de tratar como elementos *puros* são números (naturais, inteiros, reais e complexos) e estes podem ser *implementados* como certos tipos de conjuntos, a preocupação com esse tipo de distinção “elemento vs. conjunto” seria irrelevante para o que faremos. Em resumo: no contexto da teoria dos conjuntos que se desenrola, *tudo é conjunto*.

¹E o emprego da expressão “abstrair” é bastante acertado, já que significa “separar”: as *propriedades* são *abstraídas* dos objetos que as possuem.

A fim de tratar matematicamente de objetos que não se definem explicitamente, precisa-se pelo menos descrever *como* tais entidades se comportam. Isso exige o estabelecimento de regras/axiomas que ditem o que pode e o que não pode ser feito com tais objetos. Em particular, por não haver uma definição do que são conjuntos, mesmo a noção de igualdade precisa ser estipulada por algum critério explícito: afinal, como decidir se duas *manifestações* se referem a um mesmo objeto se não sabemos definir o que significa ser o objeto em questão?

Para expressar a resposta de forma mais econômica, para conjuntos A e B :

- ✓ escreveremos “ $A \subseteq B$ ” para abreviar a afirmação “para todo x , se $x \in A$, então $x \in B$ ”, lida como “ A é **subconjunto** de B ”, ou “ A está contido em B ”;
- ✓ escreveremos “ $A \not\subseteq B$ ” para abreviar a negação de “ $A \subseteq B$ ”, i.e., para indicar que “existe $x \in A$ tal que $x \notin B$ ”;
- ✓ escreveremos “ $A \subsetneq B$ ” para abreviar “ $A \subseteq B$ e $A \neq B$ ” e, em tais situações, diremos que A é **subconjunto próprio** de B .

Axioma da Extensão. *Dois conjuntos são iguais se, e somente se, têm os mesmos elementos. Em notação mais econômica: $A = B \Leftrightarrow (A \subseteq B) \text{ e } (B \subseteq A)$.*

O axioma acima manifesta a ideia de que são os elementos de um conjunto, e apenas eles, que o caracterizam. Este axioma poderia não ser útil se, por exemplo, quiséssemos distinguir conjuntos não apenas por seus elementos, mas também pela cor da fonte em que eles são grafados: se fosse o caso, então os conjuntos $\{\textcolor{red}{L}, \textcolor{red}{L}'\}$ e $\{\textcolor{cyan}{L}, \textcolor{cyan}{L}'\}$ seriam distintos, mesmo com ambos sendo *compostos* pelos mesmos elementos. Da mesma forma, ao pensar em *conjuntos* como *pastas* de arquivos num computador, duas pastas distintas podem ter exatamente os mesmos arquivos (duplicados), revelando um *modelo* em que o Axioma da Extensão não é satisfeito. O Axioma da Extensão evita situações desses tipo.

Observação. *Axiomas* não devem ser tratados como *verdades absolutas* ou *inquestionáveis*, mas apenas como *convenções suposições* estabelecidas a fim de embasar deduções posteriores. Eles podem ser debatidos, mas *fora* do panorama discursivo regido por eles. Costuma ser útil pensar em axiomas como regras de um jogo: elas se discutem antes ou depois de uma partida, mas não durante. Inclusive, é lícito buscar por regras que respeitem alguma noção de *verdade*, o que evidentemente exige debate – que ocorre fora do jogo. △

Podemos agora nos dedicar a problemas mais emocionantes, como a *formação de conjuntos*. Intuitivamente, sempre que se tem algum tipo de *propriedade matemática*², é razoável considerar o conjunto das *coisas* que possuem tal propriedade. Em vista do Axioma da Extensão, para uma propriedade \mathcal{P} fixada, é único, *caso exista*, o conjunto de *todos* os elementos que possuem a propriedade \mathcal{P} : ora, se tanto A quanto B têm como elementos precisamente aqueles com a propriedade \mathcal{P} , então “ $x \in A \Leftrightarrow x \in B$ ”, acarretando $A = B$. Com isso em mente, para uma propriedade \mathcal{P} dada, ao escrever $\mathcal{P}(y)$ para indicar que y possui a propriedade \mathcal{P} e $\{x : \mathcal{P}(x)\}$ para denotar a coleção dos elementos com a propriedade \mathcal{P} , i.e., tal que $y \in \{x : \mathcal{P}(x)\}$ se, e somente se, $\mathcal{P}(y)$, a intuição clássica diz que deveria valer o seguinte

Princípio da Abstração. *Para toda propriedade \mathcal{P} existe o conjunto $\{x : \mathcal{P}(x)\}$.*

²Aqui, “propriedade matemática” é meramente uma *fórmula* escrita na *linguagem* da Teoria dos Conjuntos, possivelmente com *variáveis livres*. Porém, tais pormenores não serão discutidos... agora.

Note que tal princípio *permitiria* reduzir o estudo de conjuntos ao mero *cálculo proposicional*, já que conjuntos $\{x : \mathcal{P}(x)\}$ e $\{x : \mathcal{Q}(x)\}$ apenas *materializam* as propriedades \mathcal{P} e \mathcal{Q} . Por exemplo: uma implicação do tipo $\mathcal{P} \Rightarrow \mathcal{Q}$ se traduz na inclusão $\{x : \mathcal{P}(x)\} \subseteq \{x : \mathcal{Q}(x)\}$, enquanto *conjunções* do tipo $\mathcal{P} \wedge \mathcal{Q}$ correspondem à interseção entre $\{x : \mathcal{P}(x)\}$ e $\{x : \mathcal{Q}(x)\}$, etc. No entanto, tal princípio é perigoso: ao se considerarem a propriedade $\mathcal{R}(x)$ dada por “ $x \notin x$ ” e o conjunto $R := \{x : \mathcal{R}(x)\}$, resulta que $R \in R \Leftrightarrow R \notin R$. Este é o **ilustre paradoxo de Russell**, um sinal de alerta para que se evite o Princípio da Abstração.

Isto deixa um problema: se não podemos formar conjuntos a partir de propriedades quaisquer, então como podemos *formar* novos conjuntos? A lista de Axiomas conhecida como **Zermelo-Fraenkel-Choice (ZFC)** é um dos modos de resolver o problema revelado pelo último paradoxo: elimina-se o Princípio da Abstração e, em contrapartida, acrescentam-se axiomas auxiliares. Tais axiomas serão apresentados ao longo do capítulo.

A.1 Noções elementares

A solução paliativa para o *excesso* de poder garantido pelo *Princípio da Abstração* é restringir seu escopo para conjuntos já *conhecidos*.

Axioma da Separação³. *Dados um conjunto A e uma propriedade \mathcal{P} , existe o conjunto $B := \{x \in A : \mathcal{P}(x)\}$ formado por todos os elementos de A que possuem a propriedade \mathcal{P} , i.e., $x \in B \Leftrightarrow x \in A$ e $\mathcal{P}(x)$.*

Observação A.1.1. Ao escrever expressões do tipo “ $A := B$ ”, busca-se destacar que a igualdade $A = B$ é imposta por definição ou, em outras palavras, o símbolo B é *definido* como um *nome* alternativo para o conjunto A . \triangle

Exercício A.1. Mostre que a *notação* $\{x \in A : \mathcal{P}(x)\}$ está *bem definida*, no sentido de que ela designa um único conjunto⁴. Dica: Axioma da Extensão. ■

A diferença entre o *Princípio da Abstração* e o Axioma da Separação é sutil, mas importante: o primeiro postula a existência irrestrita de conjuntos, enquanto o segundo apenas *separa* subconjuntos a partir de um conjunto previamente conhecido. Uma consequência marcante de tal restrição é a inexistência de universo, no seguinte sentido.

Proposição A.1.2. *Não existe um conjunto \mathbb{V} tal que $x \in \mathbb{V}$ para todo x .*

Demonstração. Se existisse, então $T := \{x \in \mathbb{V} : x \notin x\}$ seria tal que “ $T \in T \Leftrightarrow T \notin T$ ”, o que leva a uma contradição. \square

Observação A.1.3. Quando lidamos com fórmulas e *variáveis*, entende-se que tais variáveis podem assumir *valores* dentro de um *universo*, frequentemente chamado de *domínio do discurso*. Embora quebre a cronologia interna deste texto, convém um exemplo rápido: a existência ou não de solução para a equação $x^2 - 2 = 0$ depende do universo de valores que a variável x pode assumir.

³A rigor, trata-se de um *esquema* de axiomas: um para cada *propriedade* \mathcal{P} . No entanto, tal sutileza só importa para leitores interessados em aspectos *metamatemáticos*, discutidos no último capítulo. Em particular, observação semelhante se aplica ao vindouro *axioma da substituição*.

⁴Em geral, a atribuição de símbolos especiais para designar um *objeto* específico visa abreviar suas eventuais menções/descrições, o que exige a garantia de algum tipo de *unicidade* no contexto considerado – justamente para que o processo de referenciamento seja efetivo. Por exemplo: numa sala em que todas as pessoas se chamam “Ariel”, seria ineficaz chamar qualquer uma delas pelo nome, dado que todas responderiam.

Nesse sentido, a última proposição diz apenas que o *universo dos conjuntos* não é um *valor* que pode ser assumido por uma variável: o *universo dos conjuntos* não é um conjunto. *Ele* existe, mas não como os objetos que existem *nele*: note que isso não é tão diferente de dizer que $\mathbb{Z} \notin \mathbb{Z}$ ou $\mathbb{R} \notin \mathbb{R}$. \triangle

Para funcionar, o Axioma da Separação depende de conjuntos previamente conhecidos. Assim, faz sentido supor a existência de *pelo menos um conjunto*⁵, digamos C , que efetivamente serve para *construir* um conjunto bastante peculiar.

Proposição A.1.4. *Existe um conjunto E tal que para todo x ocorre $x \notin E$.*

Demonstração. Para o conjunto C postulado anteriormente, defina $E := \{x \in C : x \neq x\}$, que existe em virtude do Axioma da Separação. Como não existe x satisfazendo $x \neq x$, segue que $x \notin E$ para todo x . \square

Definição A.1.5. O conjunto E definido acima é chamado de **conjunto vazio** e será denotado por \emptyset . \P

Proposição A.1.6. *Para todo conjunto A ocorre $\emptyset \subseteq A$.*

Demonstração. Dado x qualquer, a implicação “ $x \in \emptyset \Rightarrow x \in A$ ” é verdadeira por *vacuidade*, já que “ $x \in \emptyset$ ” é falso. Alternativamente: se a *inclusão* fosse falsa, deveria existir $x \in \emptyset$ com $x \notin A$, mas não existe $x \in \emptyset$. \square

Observação A.1.7 (Contido vs. pertence). O leitor deve tomar cuidado para não confundir *pertinência* e *continência*:

- “ $x \in y$ ” significa que “ x ” é um dos elementos de “ y ”;
- “ $x \subseteq y$ ” significa que “todo elemento de x é também elemento de y ”.

Embora $\emptyset \subseteq A$ ocorra para qualquer conjunto A , nem sempre ocorre $\emptyset \in A$. Veja que, por exemplo, $\emptyset \notin \emptyset$, já que o contrário diria que \emptyset tem um elemento. Mesmo assim, $\emptyset \subseteq \emptyset$. A raiz dessa confusão é, possivelmente, oriunda do fato de que muitas vezes se diz “ y contém x ”, como abuso de linguagem, a fim de expressar “ $x \in y$ ”. \triangle

Definição A.1.8. Se \mathcal{F} é um conjunto com $\mathcal{F} \neq \emptyset$, então existe $A \in \mathcal{F}$, o que permite definir

$$\bigcap \mathcal{F} := \bigcap_{F \in \mathcal{F}} F := \{x \in A : \forall F(F \in \mathcal{F} \Rightarrow x \in F)\},$$

a **interseção** dos membros do conjunto \mathcal{F} , cujos elementos são os x 's que pertencem a todos os membros de \mathcal{F} . \P

Observação A.1.9. A expressão “ $\forall F(F \in \mathcal{F} \Rightarrow x \in F)$ ” abrevia de maneira simbólica a asserção “para todo F , se $F \in \mathcal{F}$, então $x \in F$ ”, que com nossa *semântica usual* diz que “ x é membro de todos os elementos de \mathcal{F} ”. Note que para essa afirmação ser falsa, basta que *exista* um $F \in \mathcal{F}$ com $x \notin F$, o que se abreviaria com “ $\exists F(F \in \mathcal{F} \text{ e } x \notin F)$ ”. Assume-se tacitamente que o leitor tenha certa familiaridade com esse tipo de linguagem, pelo menos a nível intuitivo. \triangle

⁵O leitor afobado pode preferir postular *um* “Axioma da Existência”, que garanta a existência de pelo menos um conjunto. Leitores menos afoitos podem esperar pelo *Axioma do Infinito*.

Há quem prefira escrever $\mathcal{F} := \{F_i : i \in \mathcal{I}\}$ para algum conjunto \mathcal{I} , para daí definir $\bigcap_{i \in \mathcal{I}} F_i$ como acima. Isto é apenas fruto do vício por índices: não está errado, mas não é imprescindível. Dito isso, também é comum criar uma distinção linguística e chamar $\mathcal{F} := \{F_i : i \in \mathcal{I}\}$ de *família*, como se a hipótese de \mathcal{F} ser *parametrizado* por um conjunto \mathcal{I} fosse algo especial e digno de nota, mas não é: qualquer conjunto pode ser escrito dessa forma, como veremos após a introdução de funções. Neste texto, *família*, *coleção* e *conjunto* serão sinônimos.

Observação A.1.10. Precisa-se supor $\mathcal{F} \neq \emptyset$ a fim de definir $\bigcap \mathcal{F}$ pois, do contrário, para todo x ocorreria $x \in \bigcap \emptyset$, uma violação da Proposição A.1.2. Note ainda que a definição não depende do elemento $A \in \mathcal{F}$ escolhido, i.e., se $A, A' \in \mathcal{F}$, então

$$\{x \in A : \forall B (B \in \mathcal{F} \Rightarrow x \in B)\} = \{x \in A' : \forall B (B \in \mathcal{F} \Rightarrow x \in B)\},$$

igualdade garantida pelo Axioma da Extensão. Isso mostra que a notação $\bigcap \mathcal{F}$ *funciona*, pois depende apenas de \mathcal{F} . \triangle

Definição A.1.11. Dados dois conjuntos A e B , a **diferença** entre A e B é o conjunto $A \setminus B := \{x \in A : x \notin B\}$, também chamado de **complementar de B com respeito à A** , verbalmente a coleção dos elementos de A que não pertencem a B . \P

Até agora, os *poucos* axiomas postulados apenas garantem a existência de subconjuntos de conjuntos conhecidos: a princípio, não há qualquer garantia de que conjuntos não relacionados possam interagir entre si para *gerar* conjuntos novos. Os axiomas a seguir vão diminuir tal limitação.

Axioma do Par. *Dados conjuntos A e B , existe um conjunto que tem A e B como elementos. Simbolicamente: existe C tal que $A \in C$ e $B \in C$.*

Pelo Axioma da Separação, o conjunto C acima pode ser usado na definição de $D := \{x \in C : x = A \text{ ou } x = B\}$. Explicitamente, isso diz que os únicos elementos de D são A e B . O leitor certamente já conhece a notação apropriada.

Definição A.1.12. Dados a e b , denota-se por $\{a, b\}$ o conjunto cujos únicos elementos são a e b , que será chamado de **par não-ordenado**. \P

Exemplo A.1.13. Para cada x , o Axioma do Par permite considerar o conjunto $\{x, x\}$: pela definição, $z \in \{x, x\}$ se, e somente se, $z = x$ ou $z = x$. Em outros termos: o único elemento de $\{x, x\}$ é x , o que justifica a

Definição A.1.14. Denota-se por $\{x\}$ o conjunto cujo único elemento é x , chamado de **conjunto unitário**⁶. \P

Em particular, note que $\emptyset \neq \{\emptyset\}$: enquanto o primeiro é o conjunto vazio, sem elementos, o conjunto $\{\emptyset\}$ tem, precisamente, o conjunto vazio como elemento. É como comparar uma pasta vazia e uma pasta que contém uma pasta vazia. \blacktriangle

Observação A.1.15. Diz-se que o par $\{x, y\}$ é *não-ordenado* pois $\{x, y\} = \{y, x\}$: de fato,

$$z \in \{x, y\} \Leftrightarrow z = x \text{ ou } z = y \Leftrightarrow z = y \text{ ou } z = x \Leftrightarrow z \in \{y, x\},$$

já que a *disjunção* (“ou”) é *comutativa*, donde a igualdade sugerida segue do Axioma da Extensão. \triangle

⁶Geralmente xingado de *singleton* em referências anglófonas.

O fenômeno observado acima pode soar como uma limitação *linguística*, já que seria interessante *poder* distinguir a *ordem* com que os elementos são listados. Em outras palavras, poderia ser útil *ter* um *dispositivo formal* $\langle \cdot, \cdot \rangle$ para o qual uma identidade do tipo $\langle a, b \rangle = \langle c, d \rangle$ ocorra se, e somente se, $a = c$ e $b = d$. Embora, em contextos mais elementares, faça sentido postular tal notação, o *modus operandi* axiomático desta seção obriga que isso seja feito por meio de conjuntos.

Definição A.1.16. Dados x e y quaisquer, o conjunto $\langle x, y \rangle := \{\{x\}, \{x, y\}\}$ será chamado de **par ordenado**. ¶

Proposição A.1.17. Para quaisquer a, b, c e d ocorre $\langle a, b \rangle = \langle c, d \rangle$ se, e somente se, $a = c$ e $b = d$.

Exercício A.2. Demonstre a proposição acima. Dica: considere separadamente os casos em que $a = b$ e $a \neq b$. ■

Os elementos a e b num par ordenado $\langle a, b \rangle$ são chamados, respectivamente, de **primeira** e **segunda coordenadas** do par. Há outras nomenclaturas, como *ordenadas* e *abscissas*, mas eu as detesto e não vejo razão para propagá-las.

Observação A.1.18. A definição explícita do par ordenado $\langle x, y \rangle$, embora importante por razões técnicas, não costuma ser *relevante* para o dia a dia. Dito de outra forma: o leitor não deve se preocupar muito com a *construção* do par ordenado, mas sim com seu comportamento; o que realmente importa é saber que ao se escrever $\langle a, b \rangle$, manifesta-se a intenção de declarar “ a ” como primeira coordenada e “ b ” como segunda coordenada. Em particular, se $a \neq b$, então $\langle a, b \rangle \neq \langle b, a \rangle$. △

Observação A.1.19 (Notação). Quem preferir *pode* escrever (x, y) em vez de $\langle x, y \rangle$. Porém, a adoção desses *brackets* angulados aqui busca apenas evitar futuras ambiguidades com o uso dos parênteses enquanto *símbolos de pontuação*. △

Com o que já se estabeleceu acima, *somos* capazes de formar *novos* conjuntos a partir de velhos, mas por métodos ainda muito engessados. Por exemplo, dados x, y e z , podemos considerar os conjuntos $\{x, y\}$ e $\{z\}$ num primeiro momento, para daí definir $\{\{x, y\}, \{z\}\}$, quase como na definição de par ordenado. Porém, nenhuma variação de tais procedimentos resultaria em algo moralmente parecido com $\{x, y, z\}$, que deveria ser o conjunto cujos únicos elementos são x, y e z (pense a respeito). Tal *falha* se corrige com o próximo

Axioma da União. *Dada uma família \mathcal{F} , existe o conjunto*

$$\bigcup \mathcal{F} := \bigcup_{F \in \mathcal{F}} F := \{x : \exists F (F \in \mathcal{F} \text{ e } x \in F)\},$$

que será chamado de (re)união da família \mathcal{F} . Verbalmente, $\bigcup \mathcal{F}$ tem como elementos todos aqueles que pertencem a algum membro de \mathcal{F} .

Observação A.1.20. Equivalentemente, o Axioma da União pode ser postulado assim: para todo \mathcal{F} , existe Z tal que se existe $F \in \mathcal{F}$ com $x \in F$, então $x \in Z$. Daí o conjunto $\bigcup \mathcal{F}$ definido acima se obtém de Z por meio do Axioma da Separação. △

Definição A.1.21. Dados conjuntos A e B , denotam-se

$$\begin{aligned} A \cap B &:= \{x : x \in A \text{ e } x \in B\} \text{ e} \\ A \cup B &:= \{x : x \in A \text{ ou } x \in B\}, \end{aligned}$$

chamados, respectivamente, de **interseção** e **(re)união** dos conjuntos A e B . Em particular, A e B são **disjuntos** se $A \cap B = \emptyset$. ¶

Exemplo A.1.22. Os conjuntos $A \cap B$ e $A \cup B$ acima são instâncias particulares de $\bigcap \mathcal{F}$ e $\bigcup \mathcal{F}$ para uma família \mathcal{F} apropriada. De fato, para A e B fixados, pode-se considerar o conjunto $\mathcal{F} := \{A, B\}$ (pelo Axioma do Par), donde o Axioma da União permite que se tome $\bigcup \mathcal{F}$.

Exercício A.3. Para $\mathcal{F} := \{A, B\}$, mostre que $\bigcup \mathcal{F} = A \cup B$. ■

Analogamente, como $\{A, B\} \neq \emptyset$, o Axioma da Separação permite definir $\bigcap \{A, B\}$ como na Definição A.1.8, e o leitor não terá dificuldades em observar a ocorrência da igualdade $\bigcap \{A, B\} = A \cap B$. ▲

Portanto, dados a, b e c , é legítimo considerar o conjunto $\{a, b, c\}$: basta fazer $\{a, b\} \cup \{c\}$, pois tal conjunto tem como elementos todos os z 's com $z \in \{a, b\}$ ou $z \in \{c\}$, o que em outras palavras significa dizer que seus únicos elementos são a, b e c . Daí, para *formar* $\{a, b, c, d\}$, pode-se fazer $\{a, b, c\} \cup \{d\}$, e *assim sucessivamente*.

Observação A.1.23. O leitor rigoroso e preciosista pode ter se incomodado com a expressão “assim sucessivamente” enfatizada acima. Em breve, esse tipo de expressão fará sentido *dentro* da teoria que estamos construindo, por meio das noções de *indução* e *recursão*. Porém, tal leitor deve se perguntar o seguinte: a Matemática não existia antes dos métodos axiomáticos conjuntistas iniciados a partir do século XIX?

Seja como fruto da percepção humana ou como descoberta de mentes iluminadas, o raciocínio lógico-matemático antecede a *formalização* do raciocínio lógico-matemático (leia com calma). No caso do “assim sucessivamente”, por exemplo, argumentações indutivas e construções recursivas não dependem da formalização que se faz para elas, mas o contrário: a descrição formal de tais métodos dentro de uma teoria deve mimetizar os comportamentos informais *já conhecidos*.

Longe de ser uma defesa dos argumentos *circulares* ou por *abanação de mão*, isto é um alerta de que o rigor é importante, mas não a ponto de invalidar a discussão sobre o rigor. Quem discordar disso, infelizmente, terá pesadelos com Ouroboros. △

Exercício A.4. Sejam A, B e C conjuntos. Mostre as identidades, inclusões, equivalências e implicações a seguir.

- a) $A \cup B = B \cup A$ e $A \cap B = B \cap A$.
- b) $A \cup (B \cup C) = (A \cup B) \cup C$ e $A \cap (B \cap C) = (A \cap B) \cap C$.
- c) $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$ e $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$.
- d) $A \subseteq A$.
- e) $A \subseteq B$ e $B \subseteq C \Rightarrow A \subseteq C$.
- f) $A \subseteq B \Leftrightarrow A \cap B = A \Leftrightarrow A \cup B = B$.
- g) $A \subseteq B \Rightarrow C \setminus B \subseteq C \setminus A$.
- h) $A \setminus B = \emptyset \Leftrightarrow A \subseteq B$.
- i) $A \setminus A = \emptyset$, $A \setminus \emptyset = A$ e $A \setminus (A \setminus B) = A \cap B$. ■

Exercício A.5. Sejam \mathcal{A} e \mathcal{B} conjuntos.

- a) Mostre que se $A \in \mathcal{A}$, então $A \subseteq \bigcup \mathcal{A}$.
- b) Mostre que se $\mathcal{A} \neq \emptyset$, então para todo $A \in \mathcal{A}$ ocorre $\bigcap \mathcal{A} \subseteq A$.
- c) Suponha que para todo $A \in \mathcal{A}$ existe $B \in \mathcal{B}$ com $A \subseteq B$. Mostre que deve ocorrer $\bigcup \mathcal{A} \subseteq \bigcup \mathcal{B}$.
- d) Mostre que se $\mathcal{A} \neq \emptyset$ e $\mathcal{A} \subseteq \mathcal{B}$, então $\bigcap \mathcal{B} \subseteq \bigcap \mathcal{A}$. ■

Exercício A.6. Mostre que o item (d) do exercício anterior deixa de ser válido se, em vez da hipótese “ $\mathcal{A} \subseteq \mathcal{B}$ ”, assumir-se apenas que “para todo $A \in \mathcal{A}$ existe algum $B \in \mathcal{B}$ com $A \subseteq B$ ”. ■

Antes de avançar para os próximos axiomas, convém demonstrar uma *lei* que será importante em diversos contextos.

Lema A.1.24 (Leis de De Morgan). *Sejam X e Y conjuntos, com $Y \neq \emptyset$.*

- (i) $\bigcap_{B \in Y} (X \setminus B) = X \setminus \bigcup_{B \in Y} B$.
- (ii) $\bigcup_{B \in Y} (X \setminus B) = X \setminus \bigcap_{B \in Y} B$.

Demonstração. Se $x \in \bigcap_{B \in Y} (X \setminus B)$, então para todo $B \in Y$ tem-se $x \in X \setminus B$, donde segue que $x \in X$ e não existe $B \in Y$ tal que $x \in B$, i.e., $x \in X$ e $x \notin \bigcup_{B \in Y} B$, precisamente $x \in X \setminus \bigcup_{B \in Y} B$. Pela arbitrariedade do x tomado, segue que $\bigcap_{B \in Y} (X \setminus B) \subseteq X \setminus \bigcup_{B \in Y} B$. A recíproca é análoga.

Para provar o segundo item, note que se $x \in \bigcup_{B \in Y} (X \setminus B)$, então existe $B' \in Y$ com $x \in X \setminus B'$, i.e., $x \in X$ e $x \notin B'$ para algum $B' \in Y$, donde se infere que $x \notin \bigcap_{B \in Y} B$ e, por conseguinte, $x \in X \setminus \bigcap_{B \in Y} B$. Logo, $\bigcup_{B \in Y} (X \setminus Y) \subseteq X \setminus \bigcap_{B \in Y} B$. Novamente, a recíproca é análoga. □

A fim de coletar *todos* os subconjuntos de um dado conjunto, postula-se o próximo

Axioma das Partes. *Para todo X existe um conjunto Z tal que se $B \subseteq X$, então $B \in Z$.*

Explicitamente, o conjunto Z acima contém como elementos todos os subconjuntos de X . Como pode haver excessos, faz-se a seguinte

Definição A.1.25. Fixado X , denota-se por $\wp(X) := \{B : B \subseteq X\}$ a coleção de todos os subconjuntos de X , que será chamada de **conjunto das partes** de X . ¶

Observação A.1.26. Novamente, o Axioma da Separação garante a existência de $\wp(X)$: basta *separar* o subconjunto $C := \{B \in Z : B \subseteq X\}$ a fim de ignorar os possíveis excessos de Z . A igualdade $C = \wp(X)$ segue então pelo Axioma da Extensão. △

Exercício A.7. Mostre que para todo X ocorre $\wp(X) \neq \emptyset$. ■

Tem-se $\wp(\emptyset) \neq \emptyset$ como uma instância particular do exercício acima, que por sua vez é apenas um modo equivalente de escrever $\{\emptyset\} \neq \emptyset$. Porém, note que diferente dos axiomas anteriores, o conjunto $\wp(X)$ é *bem maior* do que X . Um primeiro modo de perceber isso é observar os *estágios iniciais* de conjuntos das partes: $\wp(\emptyset) = \{\emptyset\}$; $\wp(\wp(\emptyset)) = \{\emptyset, \{\emptyset\}\}$; $\wp(\wp(\wp(\emptyset))) = \{\emptyset, \{\emptyset\}, \{\{\emptyset\}\}, \{\emptyset, \{\emptyset\}\}\}$.

Exercício A.8. Convença-se das igualdades anteriores. ■

Assim que tivermos ferramental para discutir cardinalidades, veremos que $\wp(X)$ tem, *invariavelmente*, mais elementos do que X . Como aquecimento, pode ser interessante pensar sobre o próximo

Exercício A.9. Mostre que $\wp(X) \not\subseteq X$. Dica: o contrário significa que para todo $B \subseteq X$ existe $b \in X$ com $B = b$; pergunte-se então o que ocorre com $B := \{x \in X : x \notin x\}$. ■

Por ora, o Axioma das Partes permitirá tratar de produtos cartesianos e outros tipos de relações binárias *dentro* da teoria. Como diria Jack, vamos por *partes...*

Definição A.1.27. Dados X e Y , denota-se por $X \times Y := \{\langle x, y \rangle : x \in X \text{ e } y \in Y\}$ o conjunto de todos os pares ordenados da forma $\langle x, y \rangle$, com $x \in X$ e $y \in Y$, que será chamado de **produto cartesiano** entre X e Y . ¶

Exercício A.10. Mostre que o produto cartesiano *existe*. Dica: use o Axioma da Separação para *realizar* $X \times Y$ como um subconjunto apropriado de $\wp(\wp(X \cup Y))$. ■

Definição A.1.28. Dados conjuntos X e Y , uma **relação (binária)** R entre X e Y é um subconjunto de $X \times Y$.

- (i) Para um par $\langle x, y \rangle \in X \times Y$, costuma-se escrever $x R y$ para indicar que $\langle x, y \rangle$ é membro da relação R , situação em que x e y serão ditos **R -relacionados**. A ocorrência de $\langle x, y \rangle \notin R$ será indicada por $x \not R y$.
- (ii) O **domínio** da relação R é o subconjunto $\text{dom}(R) := \{x \in X : \exists y(x R y)\}$.
- (iii) A **imagem** da relação R é o subconjunto $\text{im}(R) := \{y \in Y : \exists x(x R y)\}$. ¶

Exemplo A.1.29 (Relação de igualdade). Fixado um conjunto X , pode-se considerar $\Delta_X := \{\langle x, y \rangle \in X \times X : x = y\}$, a *relação de igualdade* em X . Daí, de acordo com a definição anterior, pode-se escrever $x \Delta_X y$ para indicar que $\langle x, y \rangle \in \Delta_X$, i.e., $x = y$. ▲

Exemplo A.1.30 (Relação de inclusão). Fixado um conjunto X , faz sentido considerar a família $\wp(X)$ e, por conseguinte, $\wp(X) \times \wp(X)$. Desse último, pode-se *separar* o subconjunto $C_X := \{\langle A, B \rangle \in \wp(X) \times \wp(X) : A \subseteq B\}$. Em outras palavras, tem-se a *relação de inclusão* para os subconjuntos de X . ▲

Observação A.1.31. Embora banais, os exemplos acima chamam a atenção para um detalhe sutil: a rigor, não faz *sentido* tratar das relações de igualdade e inclusão *per se*, i.e., como relações binárias legítimas. Isto se deve ao fato de que como tais “símbolos” se aplicam a quaisquer conjuntos, o domínio de tais relações seria o *universo* \mathbb{V} , que não é um conjunto. △

Exemplo A.1.32 (Curvas e gráficos). Em posse das *estruturas algébricas elementares*, é possível utilizar expressões algébricas a fim de *relacionar variáveis*. Por exemplo, a expressão polinomial $x^2 + y^2 = 1$ induz a relação binária $S := \{\langle x, y \rangle \in \mathbb{R} \times \mathbb{R} : x^2 + y^2 = 1\}$. Quando se representa $\mathbb{R} \times \mathbb{R}$ graficamente como o *plano cartesiano usual*, o subconjunto S passa a corresponder aos pontos do plano que *distam* precisamente 1 da origem $\langle 0, 0 \rangle$. ▲

Uma relação binária R recebe xingamentos adicionais a depender de outras condições satisfeitas por ela. Dois tipos especiais de relações, as de *equivalência* e as de *ordem*, serão discutidas no próximo capítulo. Por ora, é mais urgente introduzir o tipo mais importante de relação.

Definição A.1.33. Uma **função** (ou **mapa** ou **aplicação**) é uma relação binária f tal que $y = z$ sempre que $x \in f y$ e $x \in f z$, para quaisquer $x \in \text{dom}(f)$ e $y, z \in \text{im}(f)$. ¶

Em outras palavras, para cada $x \in \text{dom}(f)$ existe um único y tal que $x \in f y$, embora para um mesmo y possam existir $x, x' \in \text{dom}(f)$ distintos com $x \in f y$ e $x' \in f y$. Assim, faz sentido escrever “ $f(x) := y$ ” a fim de indicar que $x \in f y$, i.e., para denotar o *valor* de f em $x \in X$, ou a **imagem** de x pela função f . Também é usual escrever $x \xrightarrow{f} y$ para indicar a identidade $f(x) := y$.

A intuição por trás da definição de função é a seguinte: pensa-se numa função f como um dispositivo que associa a cada $x \in \text{dom}(f)$ um único y , independentemente do *usuário* da função f . Tal intuição sugere o uso de funções para descrever *regras* que associam elementos de um conjunto X fixado a elementos de outro conjunto Y .

Definição A.1.34. Uma **função de X em Y** é uma função f com $\text{dom}(f) = X$ e $\text{im}(f) \subseteq Y$. ¶

Formalmente, tal objeto consiste de uma *terna* ordenada⁷ $\langle X, f, Y \rangle$ de *informações*:

- $f \subseteq X \times Y$ é uma função como na Definição A.1.33,
- $\text{dom}(f) = X$,
- $\text{im}(f) \subseteq Y$, e Y passa a ser xingada de **codomínio** da função f de X em Y .

Definição A.1.35. Escreve-se $f: X \rightarrow Y$ ou $X \xrightarrow{f} Y$ para denotar a função f de X em Y , i.e., a terna $\langle X, f, Y \rangle$ com as propriedades acima. ¶

Exemplo A.1.36. Ao escrever

$$\begin{aligned} F: X &\rightarrow \wp(X) \\ x &\mapsto \{x\} \end{aligned}$$

indica-se a função F definida pela *regra* $F(x) := \{x\}$: note que cada $x \in X$ é “levado” ao subconjunto $\{x\} \in \wp(X)$, de modo que todas as notações se encaixam. Apelando-se para entidades ainda não definidas no texto,

$$\begin{aligned} g: \mathbb{R} &\times \mathbb{R} \rightarrow \mathbb{R} \\ \langle x, y \rangle &\mapsto x \cdot y \end{aligned}$$

indica a função que a cada par $\langle x, y \rangle$ associa $x \cdot y$, i.e., $g(\langle x, y \rangle) := x \cdot y$. ▲

É claro que uma função f determina a função $f: \text{dom}(f) \rightarrow \text{im}(f)$. Na verdade, é legítimo escrever $f: \text{dom}(f) \rightarrow Y$ para qualquer conjunto Y com $\text{im}(f) \subseteq Y$. Por sua vez, para uma função $g: X \rightarrow Y$ pode-se associar o conjunto

$$\text{graf}(g: X \rightarrow Y) := \{\langle x, g(x) \rangle : x \in X\} \subseteq X \times Y,$$

chamado de **gráfico** da função $g: X \rightarrow Y$. Claramente, o gráfico de $g: X \rightarrow Y$ é a *própria função* g na terna $\langle X, g, Y \rangle$.

Em muitas situações é irrelevante especificar um codomínio particular para uma função f . Porém, há contextos nos quais é extremamente vantajoso ter um codomínio Y fixado. Como usaremos ambas as definições a depender do caso, o leitor deve estar ciente das diferenças sutis que existem entre as afirmações “ f é uma função” e “ f é uma função de X em Y ”. Há duas situações que exemplificam bem tais sutilezas.

⁷Pode-se definir a terna ordenada $\langle a, b, c \rangle$ como sendo o par ordenado $\langle \langle a, b \rangle, c \rangle$. *Upas* ordenadas serão discutidas mais adiante

Exemplo A.1.37 (Igualdade de funções). Se f e g são funções, então pelo Axioma da Extensão tem-se

$$f = g \Leftrightarrow \forall x \forall y (\langle x, y \rangle \in f \Leftrightarrow \langle x, y \rangle \in g).$$

Como “ $\langle x, y \rangle \in f$ ” consiste em afirmar que $x \in \text{dom}(f)$ e $f(x) = y$, segue que

$$f = g \Leftrightarrow \text{dom}(f) = \text{dom}(g) \text{ e } \forall x \in \text{dom}(f) \quad f(x) = g(x).$$

Por outro lado, se $f: X \rightarrow Y$ é uma função de X em Y e $g: X' \rightarrow Y'$ é uma função de X' em Y' , então deve-se ter

$$(f: X \rightarrow Y) = (g: X' \rightarrow Y') \Leftrightarrow \langle X, f, Y \rangle = \langle X', g, Y' \rangle \Leftrightarrow X = X', f = g \text{ e } Y = Y',$$

mostrando que em tal contexto só faz sentido falar de igualdade entre funções que, pelo menos, *compartilham* mesmos domínio e codomínio.

Exercício A.11. Mostre que duas funções $f: X \rightarrow Y$ e $g: X \rightarrow Y$ são iguais se, e somente se, $\text{graf}(f: X \rightarrow Y) = \text{graf}(g: X \rightarrow Y)$. ■

Isso dá margem para situações estranhas. Por exemplo, para um conjunto X , é fácil ver que $\text{Id}_X := \{\langle x, x \rangle : x \in X\}$ é uma função, chamada de **identidade** de X . Pode-se expressar a função Id_X como uma função de X em X , uma vez que se verifica $X = \text{dom}(\text{Id}_X) = \text{im}(\text{Id}_X)$.

Contudo, se Y é um conjunto com $X \subseteq Y$, então também faz sentido definir a **inclusão** $i: X \rightarrow Y$ dada por $i(x) := x$ para cada $x \in X$. Em outras palavras, $i: X \rightarrow Y$ é a terna $\langle X, \text{Id}_X, Y \rangle$. Agora, se ocorrer $X \subsetneq Y$, então as funções $\text{Id}_X: X \rightarrow X$ e $i: X \rightarrow Y$ são formalmente distintas, embora ambas sejam definidas pela *mesma regra*, i.e.,

$$\text{Id}_X(x) := x =: i(x)$$

para todo $x \in X$. ▲

O segundo estranhamento entre as duas noções de função se dá com as chamadas relações *inversas*.

Definição A.1.38. Dada uma relação binária R , a **relação inversa** de R , denotada por R^{-1} , é a relação $R^{-1} := \{\langle y, x \rangle : x R y\}$, ou $R^{-1} := \{\langle y, x \rangle \in \text{im}(R) \times \text{dom}(R) : x R y\}$ para os mais céticos. ¶

O importante sobre R^{-1} é que para quaisquer x e y vale

$$x R y \Leftrightarrow y R^{-1} x$$

De tal relação, segue que $\text{dom}(R) = \text{im}(R^{-1})$, $\text{im}(R) = \text{dom}(R^{-1})$ e $(R^{-1})^{-1} = R$.

Agora, se f é uma função, então é razoável perguntar: quando f^{-1} é uma função? A resposta para essa pergunta depende da definição de função que se utiliza.

Definição A.1.39. A função f é dita **injetora** se para quaisquer $x, x' \in \text{dom}(f)$, a ocorrência de $f(x) = f(x')$ acarretar $x = x'$. ¶

Lema A.1.40. Uma função f é injetora se, e somente se, f^{-1} é função.

Demonstração. A condição de injetividade para f diz que para $x, x' \in \text{dom}(f)$ e $y \in \text{im}(f)$ deve valer

$$x f y \text{ e } x' f y \Rightarrow x = x',$$

ou, equivalentemente,

$$y f^{-1} x \text{ e } y f^{-1} x' \Rightarrow x = x',$$

que por sua vez significa dizer que f^{-1} é função, pois $\text{im}(f) = \text{dom}(f^{-1})$. \square

No caso em que se tem uma função $f: X \rightarrow Y$, i.e., uma função $f \subseteq X \times Y$ com $\text{dom}(f) = X$ e $\text{im}(f) \subseteq Y$, a injetividade não basta para garantir que a inversa f^{-1} seja uma função $f^{-1}: Y \rightarrow X$. Mais precisamente, vamos dizer que a função $f: X \rightarrow Y$ é **injetora** se a função $f \subseteq X \times Y$ for injetiva.

Definição A.1.41. Dizemos que uma função $f: X \rightarrow Y$ é

- **sobrejetora** se $\text{im}(f) = Y$ e
- **bijetora** se $f: X \rightarrow Y$ for injetora e sobrejetora. ¶

Proposição A.1.42. Uma função $f: X \rightarrow Y$ é bijetora se, e somente se, f^{-1} é uma função de Y em X .

Demonstração. Primeiramente, note que as identidades e inclusões $\text{dom}(f^{-1}) = \text{im}(f) \subseteq Y$ e $\text{im}(f^{-1}) = \text{dom}(f) = X$ valem em geral. Agora, pelo Lema A.1.40, sabe-se que a função $f \subseteq X \times Y$ é injetiva se, e somente se, f^{-1} é uma função. Por outro lado, $f: X \rightarrow Y$ é sobrejetora se, e somente se, $\text{im}(f) = Y$, e isto equivale a dizer que $\text{dom}(f^{-1}) = Y$. Logo, $f: X \rightarrow Y$ é bijetora se, e somente se, $f^{-1} \subseteq Y \times X$ é uma função e $\text{dom}(f^{-1}) = Y$, i.e., f^{-1} é uma função de Y em X . \square

Corolário A.1.43. Se $f: X \rightarrow Y$ é uma função bijetora, então $f^{-1}: Y \rightarrow X$ é bijetora.

Demonstração. Já vimos que $f^{-1}: Y \rightarrow X$ é função. Como $(f^{-1})^{-1} = f$ é uma função de X em Y , o resultado segue da proposição anterior. \square

Em suma, toda função f define uma função $f: \text{dom}(f) \rightarrow \text{im}(f)$ sobrejetora e, em particular, toda função injetora $f: X \rightarrow Y$ induz uma função bijetora $f: X \rightarrow \text{im}(f)$. Em geral, quando não se especifica o codomínio de uma função f , implicitamente se usa a definição de f como um conjunto de pares ordenados satisfazendo a Definição A.1.33. Por outro lado, quando conjuntos X e Y estiverem fixados pelo contexto, uma função $f: X \rightarrow Y$ é, formalmente, a terna $\langle X, f, Y \rangle$ como na Definição A.1.34. Nesse último caso, é comum se referir à função $f: X \rightarrow Y$ simplesmente como “a função f ”.

Por último, mas não menos importante, conceitos definidos para funções *da forma* $X \rightarrow Y$ são facilmente traduzíveis para funções quaisquer, de modo que o cuidado com tais detalhes ficará a cargo do leitor implicante.

Observação A.1.44. Há outras variações típicas de vocabulário, que costumam ajudar a tornar o texto menos repetitivo: funções injetoras também serão chamadas de *injeções*, funções *injetivas*, etc., enquanto funções sobrejetoras também podem ser chamadas de *sobrejeções*, funções *sobrejetivas*, etc. Em consonância com tais neologismos, funções bijetoras também serão chamadas de *bijeções*, funções *bijetivas*, etc. \triangle

Definição A.1.45. Se $f: X \rightarrow Y$ e $g: Y \rightarrow Z$ são funções, então fica *bem definida* uma função $g \circ f: X \rightarrow Z$ dada pela identidade

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

para todo $x \in X$, que chamamos de (função) **composta** entre f e g . ¶

A definição acima faz sentido pois $f(x) \in Y$ para todo $x \in X$ e $Y = \text{dom}(g)$, de modo que g *sabe* o que *fazer* com $f(x)$.

Exercício A.12. Sejam $f: W \rightarrow X$, $g: X \rightarrow Y$ e $h: Y \rightarrow Z$ funções. Mostre que as funções $h \circ (g \circ f)$ e $(h \circ g) \circ f$ são iguais. ■

Exercício A.13. Mostre que $f: X \rightarrow Y$ é bijetora se, e somente se, existe uma função $g: Y \rightarrow X$ satisfazendo $g \circ f = \text{Id}_X$ e $f \circ g = \text{Id}_Y$. Em particular, a função g é, necessariamente, a inversa de f . Dica: use a Proposição A.1.42. ■

Definição A.1.46. Para uma função $f: X \rightarrow Y$ e um subconjunto $W \subseteq X$, a **restrição** da função f ao subconjunto W é a função $f|_W: W \rightarrow Y$ definida por $f|_W(w) := f(w)$ para todo $w \in W$. Mais ainda, para subconjuntos $A \subseteq X$ e $B \subseteq Y$ fixados, consideram-se:

- a **imagem direta** de A por f , definida como $f[A] := \{f(a) : a \in A\}$;
- a **pré-imagem** de B por f , definida como $f^{-1}[B] := \{x \in X : f(x) \in B\}$. ¶

Observação A.1.47. Explicitamente, $y \in f[A]$ se, e somente se, existe *algum* $a \in A$ com $f(a) = y$, enquanto $x \in f^{-1}[B]$ se, e somente se, $f(x) \in B$. △

Proposição A.1.48. Sejam $f: X \rightarrow Y$ uma função e considere famílias $\mathcal{U} \subseteq \wp(X)$ e $\mathcal{V} \subseteq \wp(Y)$. Então:

- (i) $f \left[\bigcup \mathcal{U} \right] = \bigcup_{U \in \mathcal{U}} f[U]$ e $f \left[\bigcap \mathcal{U} \right] \subseteq \bigcap_{U \in \mathcal{U}} f[U]$, com igualdade garantida no último caso se f for injetora;
- (ii) $f^{-1} \left[\bigcup \mathcal{V} \right] = \bigcup_{V \in \mathcal{V}} f^{-1}[V]$ e $f^{-1} \left[\bigcap \mathcal{V} \right] = \bigcap_{V \in \mathcal{V}} f^{-1}[V]$.

Exercício A.14. Demonstre a proposição anterior. ■

Exemplo A.1.49. Supondo conhecidos os *números inteiros* e a *potenciação usual*, para $f: \mathbb{Z} \rightarrow \mathbb{Z}$ dada por $f(x) := x^2$ e $C := \{0, 1, 4\}$ tem-se $f[C] = \{0, 1, 16\}$ e $f^{-1}[C] = \{0, -1, 1, -2, 2\}$. Para $g: \mathbb{Z} \rightarrow \mathbb{Z}$ dada por $g(x) := x^2 + 1$ ocorre

$$(f \circ g)(x) := f(x^2 + 1) = (x^2 + 1)^2 = x^4 + 2x^2 + 1 \text{ e } (g \circ f)(x) := g(f(x)) = (x^2)^2 + 1 = x^4 + 1.$$

O leitor certamente deve conhecer exemplos menos artificiais. De qualquer forma, os capítulos posteriores estarão fartos de casos mais desafiadores. ▲

Definição A.1.50. Para conjuntos X e Y fixados, denotaremos por

$$Y^X := \{f \subseteq X \times Y : f \text{ é função de } X \text{ em } Y\},$$

o conjunto de todas as funções da forma $X \rightarrow Y$. ¶

Exercício A.15. Convença-se de que Y^X existe. Dica: mostre que Y^X é um subconjunto de $\wp(X \times Y)$ obtido por meio do Axioma da Separação⁸. ■

Observação A.1.51 (Função vazia). Por mais estranho que possa parecer, existe uma função da forma $\emptyset \rightarrow \emptyset$. De fato, pela definição, uma função $f: \emptyset \rightarrow \emptyset$ deve ser

- ✓ um subconjunto $f \subseteq \emptyset \times \emptyset$ e
- ✓ tal que para todo $x \in \emptyset$ exista um único $y \in \emptyset$ com $\langle x, y \rangle \in \emptyset \times \emptyset$.

Não é difícil perceber que $\emptyset \times \emptyset = \emptyset$, donde segue que seu único subconjunto é \emptyset , que satisfaz as condições acima por vacuidade. Logo, $\emptyset^\emptyset = \{\emptyset\} \neq \emptyset$. Trívia: como \emptyset é uma bijeção (por vacuidade), obtém-se aí uma justificativa para a identidade $0! = 1$. △

Ao se escrever $\mathcal{A} := \{A_i : i \in \mathcal{I}\}$, para algum conjunto \mathcal{I} , indica-se implicitamente a existência de uma função sobrejetora $\varphi: \mathcal{I} \rightarrow \mathcal{A}$ que faz $i \mapsto A_i$ para todo $i \in \mathcal{I}$. É um modo prático de expressar a seguinte afirmação: para todo $A \in \mathcal{A}$, existe pelo menos um $i \in \mathcal{I}$ satisfazendo $A = A_i$. Esse tipo de notação costuma ser interessante quando se quer, por qualquer motivo, estabelecer algum tipo de *ordenação* aparente sobre os elementos de \mathcal{A} , posto que geralmente o conjunto de *índices* \mathcal{I} é um subconjunto de *números naturais*. Contudo, qualquer conjunto pode ser indexado por *algum conjunto*, afinal de contas, para $\mathcal{I} := \mathcal{A}$ e $\varphi := \text{Id}_{\mathcal{A}}$ tem-se $\mathcal{A} = \{A_i : i \in \mathcal{I}\}$.

Exemplo A.1.52. Parametrizar conjuntos a partir de uma coleção de índices é algo tão natural que, frequentemente, isso é feito sem que se perceba. Note que nas Definições A.1.27 e A.1.46, os conjuntos $X \times Y$ e $f[A]$ foram expressos por meio de famílias indexadas. De fato, uma escrita mais explícita é

$$\begin{aligned} X \times Y &:= \{z : \exists x \in X \exists y \in Y \text{ tal que } z = \langle x, y \rangle\}, \text{ e} \\ f[A] &:= \{y \in Y : \exists a \in A \text{ tal que } f(a) = y\}, \end{aligned}$$

o que evidentemente não colabora para transmitir a ideia. ▲

Observação A.1.53 (“ $\{\cdot\}$ ” vs. “ $\langle \cdot \rangle$ ”). Supondo conhecidos os *números naturais* 0, 1, 2, 3 e 4 e fazendo $\mathcal{I} := \{0, 1, 2, 3, 4\}$, pode-se tomar, para cada $i \in \mathcal{I}$, um certo conjunto A_i para daí definir $\mathcal{A} := \{A_i : i \in \mathcal{I}\}$. Ocorre que pelo significado explícito da notação, tem-se

$$\mathcal{A} = \{A_0, A_1, A_2, A_3, A_4\} = \{A_1, A_2, A_3, A_4, A_0\} = \{A_2, A_3, A_4, A_0, A_1\} = \dots$$

i.e., a indexação de \mathcal{A} , embora feita por um *conjunto ordenado*, não pressupõe que o próprio \mathcal{A} seja *ordenado* de modo *compatível* com a *ordem* de \mathcal{I} . Se a ordem for importante, é preferível escrever $\langle A_0, \dots, A_4 \rangle$ ou $\langle A_i : i \in \mathcal{I} \rangle$, o que ficará mais claro na Seção A.3. △

Exercício A.16. Sejam $\mathcal{A} := \{A_i : i \in \mathcal{I}\}$ e $\mathcal{J} := \bigcup \mathcal{J}$, para uma certa família \mathcal{J} .

a) Mostre que $\bigcup_{i \in \mathcal{I}} \mathcal{A} := \bigcup_{i \in \mathcal{I}} A_i = \bigcup_{J \in \mathcal{J}} \left(\bigcup_{j \in J} A_j \right)$.

b) Mostre que se $\mathcal{A} \neq \emptyset$, então $\bigcap_{i \in \mathcal{I}} \mathcal{A} := \bigcap_{i \in \mathcal{I}} A_i = \bigcap_{J \in \mathcal{J}} \left(\bigcap_{j \in J} A_j \right)$. ■

⁸Leitores preciosistas podem preferir, alternativamente, considerar $\{X\} \times Y^X \times \{Y\}$, já que os elementos de tal conjunto são as ternas da forma $\langle X, f, Y \rangle$.

Exercício A.17. Mostre que valem as identidades

$$(\bigcap \mathcal{A}) \cup (\bigcap \mathcal{B}) = \bigcap_{\langle A, B \rangle \in \mathcal{A} \times \mathcal{B}} (A \cup B) \text{ e } (\bigcup \mathcal{A}) \cap (\bigcup \mathcal{B}) = \bigcup_{\langle A, B \rangle \in \mathcal{A} \times \mathcal{B}} (A \cap B)$$

para quaisquer coleções não-vazias \mathcal{A} e \mathcal{B} . ■

Observação A.1.54 (“*boa definição*”). Digamos que X e Y sejam conjuntos dotados de uma relação binária $R \subseteq X \times Y$. Nas frequentes situações em que R é uma função da forma $X \rightarrow Y$, é comum encontrar frases como “a função R está bem definida” ou ainda “... isto prova a boa definição da função R ” (confira a Subseção B.1.2, por exemplo). Tais afirmações são apenas abusos de linguagem, e querem dizer que a relação binária R , *definida* de alguma forma apropriada dentro do contexto, é uma função: no caso, a “boa definição” não se refere à *existência* de R , mas sim ao elemento “ $y := R(x)$ ”, que fica *bem definido* em função de x se R for uma função – afinal de contas, se existissem y e y' distintos com $x R y$ e $x R y'$, ambos poderiam competir pelo título de “ $R(x)$ ”, tornando a coisa “mal-definida”. △

A.2 O problema do infinito

Conforme já se mencionou, a lista de axiomas apresentada até aqui faz parte do sistema usualmente chamado de Zermelo-Fraenkel-Choice, abreviado como ZFC em referência a Ernst Zermelo e Abraham Fraenkel, que propuseram as primeiras versões dos axiomas listados. Eles não foram os únicos envolvidos nesse processo de axiomatização, e tampouco houve uma terceira pessoa chamada *Choice*: o termo, neste caso, é a expressão inglesa para “*escolha*”, e faz referência ao *Axioma da Escolha*, um postulado *não-construtivo* sem o qual fazer Matemática com *conjuntos infinitos* se torna algo muito (mais) desafiador. Mas o que são conjuntos infinitos? E finitos?

Embora isto vá ser tratado *mais profundamente* nos Capítulos C e E, o leitor *já deve saber* o que significa dizer que uma certa *coleção informal* (e possivelmente concreta) de objetos é *finita*. Porém, tal conhecimento é *metateórico* externo à *teoria* e, em certa medida, empírico. O que se faz ao *definir* finitude em ZFC é estabelecer uma *descrição* por meio dos elementos do *universo do discurso* (os conjuntos) que remete àquilo que já se conhece: é quase como escrever um programa numa certa linguagem de programação a fim de realizar um determinado procedimento já conhecido pelo programador. A ideia é muito simples:

Definição A.2.1. Um conjunto X é **finito** se existe um *número natural* n em bijeção com X . Caso contrário, X é dito **infinito**. ¶

A definição acima foi, propositalmente, desonesta: afinal de contas, não se definiu (ainda) o que são números naturais neste contexto conjuntista. Para piorar a estranheza, a definição fala da bijeção entre um *conjunto* X e um *número...* mas o que são números?

Definição A.2.2. Para cada X , denotaremos por $X_+ := X \cup \{X\}$ o conjunto que será chamado de **sucessor de X** . ¶

Observação A.2.3. É comum encontrar “ X_+ ” denotado como “ $X + 1$ ” nas principais referências. Neste texto, isto só será adotado após a Observação E.2.30. △

Exemplo A.2.4 (Alguns números naturais). O conjunto *vazio* \emptyset , caracterizado como o único para o qual $x \notin \emptyset$ ocorre para todo x , é um conjunto sem elementos. De nossa experiência empírica com coleções vazias (como carteiras, contas bancárias ou listas de amigos platonistas) bem como com números naturais, faz sentido dizer que \emptyset tem “0” elementos, onde o termo “0” faz menção ao zero que conhecemos *aqui, fora de ZFC*, na *metateoria*. Façamos então seguinte: vamos atribuir o símbolo “0” ao conjunto \emptyset , i.e., $0 := \emptyset$. Com tal terminologia, e no sentido da Definição A.2.2, o que é o sucessor de 0? Ora:

$$0_+ := 0 \cup \{0\} = \emptyset \cup \{0\} = \{0\}.$$

O leitor deve se atentar para algo sutil: o sucessor de 0 (definido em ZFC) tem precisamente *um* elemento, onde o “um” se refere ao numeral um (1) que conhecemos *aqui, fora de ZFC*, na *metateoria*. Por que não definir $1 := \{0\}$? Feito isso, quem seria o sucessor de 1? Novamente: $1_+ := 1 \cup \{1\} = \{0\} \cup \{1\} = \{0, 1\}$, um conjunto que tem *dois* elementos, já que $0 \neq 1$. Por que parar?

$$\begin{aligned} 3 &:= 2_+ = \{0, 1, 2\}. \\ 4 &:= 3_+ = \{0, 1, 2, 3\}. \\ 5 &:= 4_+ = \{0, 1, 2, 3, 4\} \dots \end{aligned}$$

e, *mais geral e informalmente*, se um número natural “ n ” já estiver implementado em ZFC, então a implementação de seu sucessor se dará por meio da “regra” $n + 1 := n_+$, em que “ $n + 1$ ” denota o entendimento *metateórico* que temos sobre o que *deveria ser* $n + 1$.

Por exemplo, para implementar o número *dezenove* em ZFC, escreve-se $19 = 18_+$, o que depende da implementação do número 18, que por sua vez é $17_+ \dots$ donde chega-se a

$$19 = \{0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18\},$$

algo estranho, certamente, mas que contém, precisamente, *dezenove* elementos.

Em particular, se chamarmos cada um dos conjuntos *construídos* acima de *número natural*, a definição dada para conjuntos finitos se aplica perfeitamente e pode ser usada, por exemplo, para demonstrar que para quaisquer X e Y dados, $\{X, Y\}$ é um conjunto finito, por estar em bijeção com 1 (se ocorrer $X = Y$) ou 2 (se ocorrer $X \neq Y$). ▲

Observação A.2.5. Note que a descrição para o procedimento do exemplo anterior ocorre *fora de ZFC*: é preciso saber o que é o número dezenove, digamos, a fim de *implementá-lo* em ZFC. Porém, é difícil negar que os conjuntos construídos acima de fato se comportam como esperaríamos que números naturais se comportassem. Isto será discutido com mais calma ao longo do texto. △

Com tempo e paciência, seria possível utilizar o procedimento acima para definir *todos* os números naturais. Porém, há um problema: o tempo e a paciência necessários seriam *infinitos*, já que $n + 1$ é um número natural sempre que n é um número natural. Como, classicamente, *demonstrações* devem ser argumentos com *finitos passos*, não seria possível garantir a existência de todos os números naturais (enquanto conjuntos) com os axiomas já listados. Uma solução seria impor axiomáticamente a existência de *todos* eles, o que deixaria outro problema: como garantir que todos eles podem ser *coletados* num único conjunto, i.e., como garantir a existência do *conjunto dos números naturais*? Os dois problemas se resolvem com um único axioma:

Axioma do Infinito. *Existe um conjunto \mathcal{I} tal que $0 := \emptyset \in \mathcal{I}$ e para todo X vale que $X_+ \in \mathcal{I}$ sempre que $X \in \mathcal{I}$.*

Definição A.2.6. Diz-se que um conjunto \mathcal{J} é **indutivo** se $0 \in \mathcal{J}$ e $X_+ \in \mathcal{J}$ sempre que $X \in \mathcal{J}$. \P

Em outras palavras, o Axioma do Infinito postula a existência de um conjunto *indutivo*. O leitor já familiarizado com o *princípio da indução* deve estar com uma pulga atrás da orelha. E com razão.

Definição A.2.7. $\omega := \{x \in \mathcal{I} : x \in \mathcal{J} \text{ para todo conjunto indutivo } \mathcal{J}\}$. Diremos que um conjunto x é um **número natural** se $x \in \omega$. *Naturalmente*, ω será chamado de **conjunto dos números naturais**. \P

Lema A.2.8. ω é um conjunto indutivo.

Demonstração. Por definição de conjunto indutivo, tem-se $0 \in \mathcal{J}$ para todo conjunto indutivo \mathcal{J} e, como o próprio \mathcal{I} é indutivo, resulta que $0 \in \omega$. Analogamente, se $n \in \omega$, então $n \in \mathcal{J}$ para todo conjunto indutivo \mathcal{J} e, por conseguinte, $n_+ \in \mathcal{J}$ para todo \mathcal{J} indutivo, mostrando que $n_+ \in \omega$. \square

Teorema A.2.9 (Princípio da Indução). *Seja \mathcal{P} uma “propriedade” com $\mathcal{P}(0)$ verdadeira e tal que para todo $n \in \omega$ tenha-se $\mathcal{P}(n) \Rightarrow \mathcal{P}(n_+)$. Então $\mathcal{P}(n)$ vale para todo $n \in \omega$.*

Demonstração. Considere $N := \{n \in \omega : \mathcal{P}(n)\}$. Por construção tem-se $N \subseteq \omega$. Por outro lado, N é conjunto indutivo: $0 \in N$ e, se $n \in N$, então $n \in \omega$ com $\mathcal{P}(n)$, donde a hipótese acerca de \mathcal{P} acarreta em $\mathcal{P}(n_+)$, donde segue que $n_+ \in N$. Portanto, $\omega \subseteq N$. \square

Observação A.2.10. Mais uma vez utilizou-se a noção vaga de *propriedade*. Como propriedades não são conjuntos, não se deve esperar que uma definição para propriedade seja estabelecida *em ZFC*, mas sim fora.

Grosso modo, as *fórmulas* em ZFC são definidas recursivamente a partir de *fórmulas atômicas* da forma $x \in y$ e $u = v$:

- (i) fórmulas atômicas são fórmulas;
- (ii) se φ é uma fórmula, então $\neg\varphi$ é uma fórmula;
- (iii) se φ e ψ são fórmulas e ∇ é um dos símbolos \wedge , \vee , \rightarrow ou \leftrightarrow , então $\varphi\nabla\psi$ é uma fórmula;
- (iv) se φ for uma fórmula, então $\exists x(\varphi)$ e $\forall x(\varphi)$ são fórmulas.

Note que o algoritmo recursivo acima é descrito na *metateoria*. Porém, no próximo capítulo, veremos como implementar *recursões* em ZFC, o que pode causar vertigens em leitores iniciantes. De toda forma, com essa breve discussão, é possível expressar de modo um pouco mais preciso o que se entende por uma “propriedade de x ”: trata-se de uma fórmula φ como acima, em que a *variável* x não ocorre no escopo dos quantificadores de φ .

Por exemplo, a fórmula ψ expressa por “ $\exists y(x = y_+)$ ”, lida como “existe y tal que x é o sucessor de y ”, tem duas variáveis, x e y , mas apenas y está *ligada* ao quantificador \exists . É nesse sentido que ψ expressa uma *propriedade de x* , já que x pode ou não ser o sucessor de algum y . A fim de indicar a substituição da variável x por um conjunto de fato, digamos z , pode-se escrever $\psi(z)$:

- ✓ tem-se $\psi(1)$ verdadeira⁹ pois para $y := 0$ ocorre $y_+ = 1$;
- ✗ tem-se $\psi(0)$ falsa, pois para qualquer y ocorre $y_+ \neq 0 := \emptyset$.

Dito isso, o leitor não precisa se preocupar por enquanto: o nível de rigor acima não será usado na maior parte do texto. Porém, é importante *saber* que isso pode ser feito. Mais detalhes podem ser encontrados no livro de Kenneth Kunen, *The Foundations of Mathematics*¹⁰ [23]. \triangle

O Teorema A.2.9 garante que argumentos por *indução finita* podem ser realizados em ZFC. Note que neste contexto, a *indução finita* não é um *axioma* da teoria, mas sim um teorema. Como aquecimento, o leitor está convidado a lidar com o

Exercício A.18. Mostre que um conjunto n é tal que $n \in \omega$ se, e somente se, $n = 0$ ou existe $m \in \omega$ tal que $n = m_+$. ■

O leitor que já conhece os *Axiomas de Peano* e se sentir incomodado com a ausência (explícita) deles deve esperar a discussão apropriada, que será feita *oportunamente*.

A.3 Produtos arbitrários e o Axioma da Escolha

Fixados conjuntos X_0 e X_1 , um elemento de $X_0 \times X_1$ é um par $\langle x_0, x_1 \rangle$ com $x_0 \in X_0$ e $x_1 \in X_1$. Secretamente, isto define uma função $\dot{x}: \{0, 1\} \rightarrow X_0 \cup X_1$ tal que $\dot{x}(i) \in X_i$ para cada $i \in \{0, 1\}$: basta fazer $\dot{x}(0) := x_0$ e $\dot{x}(1) := x_1$. Ao escrever $\prod_{i \in 2} X_i$ para indicar o conjunto das funções $c: \{0, 1\} \rightarrow X_0 \cup X_1$ tais que $c(i) \in X_i$ para todo $i \in \{0, 1\}$, segue que a regra $\langle x_0, x_1 \rangle \mapsto \dot{x}$ define uma função $X_0 \times X_1 \rightarrow \prod_{i \in 2} X_i$. Agora, dada uma função $c \in \prod_{i \in \{0, 1\}} X_i$, pode-se determinar um par em $X_0 \times X_1$: basta fazer $p(a) := \langle c(0), c(1) \rangle$.

Exercício A.19. Com as notações acima, mostre que as funções

$$\begin{aligned} \varphi: X_0 \times X_1 &\rightarrow \prod_{i \in 2} X_i & \psi: \prod_{i \in 2} X_i &\rightarrow X_0 \times X_1 \\ &\quad \text{e} & & \\ \langle x_0, x_1 \rangle &\mapsto \dot{x} & c &\mapsto \langle c(0), c(1) \rangle \end{aligned}$$

são inversas uma da outra. ■

O exercício acima sugere que pares ordenados em $X_0 \times X_1$ podem ser pensados como funções que *escolhem* um elemento X_i para cada $i \in \{0, 1\}$.

Exercício A.20. Para conjuntos X_0 , X_1 e X_2 , defina $X_0 \times X_1 \times X_2 := (X_0 \times X_1) \times X_2$. Mostre que tal conjunto está em bijeção com $\prod_{i \in 3} X_i$, este definido como a família de funções da forma $c: \{0, 1, 2\} \rightarrow X_0 \cup X_1 \cup X_2$ tais que $c(i) \in X_i$ para todo $i \in \{0, 1, 2\}$. ■

Procedendo *recursivamente*, não é difícil se convencer de que os elementos do que *deveria* ser o produto cartesiano $X_0 \times \dots \times X_n$ podem ser interpretados como funções da forma $c: \{0, \dots, n\} \rightarrow \bigcup_{i \in n_+} X_i$ tais que $c(i) \in X_i$. Ora, em vez de apelar para uma receita *metateórica* de construção dos produtos cartesianos, é bem mais econômico fazer a seguinte:

⁹A atribuição de valor lógico (verdadeiro ou falso, no caso *clássico*) para uma fórmula também se faz do modo recursivo usual: ψ é verdadeira se, e somente se, $\neg \psi$ é falsa; $\psi \wedge \varphi$ só é verdadeira se ambas ψ e φ forem verdadeiras; $\psi \vee \varphi$ só é falsa se ambas forem falsas, etc.

¹⁰Note o plural.

Definição A.3.1. Para um conjunto \mathcal{I} qualquer, considere fixado um conjunto X_i para cada $i \in \mathcal{I}$. A família

$$\prod_{i \in \mathcal{I}} X_i := \left\{ f \in \left(\bigcup_{i \in \mathcal{I}} X_i \right)^{\mathcal{I}} : \text{para todo } i \in \mathcal{I} \text{ ocorre } f(i) \in X_i \right\}$$

será chamada de **produto cartesiano** generalizado (dos conjuntos X_i 's). ¶

Exercício A.21. Convença-se de que os axiomas já postulados garantem que o produto cartesiano $\prod_{i \in \mathcal{I}} X_i$ existe. Dica: o conjunto $(\bigcup_{i \in \mathcal{I}} X_i)^{\mathcal{I}}$ de todas as funções da forma $\mathcal{I} \rightarrow \bigcup_{i \in \mathcal{I}} X_i$ já existe, de modo que basta separar as *funções certas*. ■

Observação A.3.2. A rigor, o produto cartesiano não se define em termos da *família indexada* $\mathcal{X} := \{X_i : i \in \mathcal{I}\}$, mas sim *em função* da função $\varphi: \mathcal{I} \rightarrow \mathcal{X}$ dada por $\varphi(i) := X_i$ para cada $i \in \mathcal{I}$. Com efeito, se não fosse o caso, então para $X_0 := X_1 := X$, $\prod_{i \in \{0,1\}} X_i$ seria essencialmente idêntico ao conjunto X , já que $\{X_0, X_1\} = \{X\}$. A ideia é que os elementos do produto $\prod_{i \in \mathcal{I}} X_i$ não dependem apenas de *quem* são os X_i 's, mas também dos *i*'s: é a generalização do comportamento que *caracteriza* os pares ordenados. A próxima proposição deve explicitar a mensagem. △

Proposição A.3.3. Sejam $f, g \in \prod_{i \in \mathcal{I}} X_i$. Então $f = g$ se, e somente se, para todo $i \in \mathcal{I}$ ocorrer $f(i) = g(i)$.

Demonstração. Ora, f e g são funções da forma $\mathcal{I} \rightarrow \bigcup_{i \in \mathcal{I}} X_i$. Logo, elas são iguais se, e somente, coincidirem em todos os pontos do domínio (Exemplo A.1.37), como desejado. □

Definição A.3.4. Em vista da proposição anterior, faz sentido denotar um elemento $f \in \prod_{i \in \mathcal{I}} X_i$ por $\langle f(i) : i \in \mathcal{I} \rangle$ ou $\langle f_i \rangle_{i \in \mathcal{I}}$. Por analogia com o comportamento de pares, ternas e “*n-uplas*” ordenadas, tais funções serão chamadas de \mathcal{I} -uplas¹¹. ¶

Exercício A.22. Sejam $\langle f_i \rangle_{i \in \mathcal{I}}, \langle g_i \rangle_{i \in \mathcal{I}} \in \prod_{i \in \mathcal{I}} X_i$. Mostre que $\langle f_i \rangle_{i \in \mathcal{I}} = \langle g_i \rangle_{i \in \mathcal{I}}$ se, e somente se, para todo $i \in \mathcal{I}$ ocorrer $f_i = g_i$. Dica: isto é apenas a proposição anterior. ■

Exercício A.23. Mostre que se $X_i := \emptyset$ para algum $i \in \mathcal{I}$, então $\prod_{i \in \mathcal{I}} X_i = \emptyset$. ■

Navegamos águas perigosas agora...

Proposição A.3.5. Seja $n \in \omega$. Então $\prod_{i \in n_+} X_i \neq \emptyset$ se, e somente se, para todo $i \in n_+$ ocorrer $X_i \neq \emptyset$.

Demonstração. Esta será a primeira prova por indução do texto. Para $n := 0$, os elementos de $\prod_{i \in 0_+} X_i$ são funções da forma $\{0\} \rightarrow X_0$, donde o leitor pode concluir a equivalência desejada. Supondo a equivalência verdadeira para $n \in \omega$ arbitrário, deve-se provar o resultado para o caso n_+ .

(\Rightarrow) Se $X_i \neq \emptyset$ para todo $i \in (n_+)_+ := n_+ \cup \{n_+\}$, então em particular $X_i \neq \emptyset$ para todo $i \in n_+$, donde segue que existe $\langle f_i \rangle_{i \in n_+} \in \prod_{i \in n_+} X_i$. Como também ocorre $X_{n_+} \neq \emptyset$, existe $t \in X_{n_+}$, o que permite definir uma função $g: (n_+)_+ \rightarrow \bigcup_{i \in (n_+)_+} X_i$ declarando-se $g(i) := f_i$ se $i \in n_+$ e $g(n_+) := t$. Pelo modo como f e t foram tomados, resulta $g \in \prod_{i \in (n_+)_+} X_i$.

¹¹Upelas também costumam ser chamadas como **funções escolha**.

(\Leftarrow) Vale em geral, como indicado no Exercício A.23.

Logo, em virtude do *princípio da indução* (Teorema A.2.9), resulta que a afirmação deve ser válida para todo $n \in \omega$. \square

Moralmente, a Proposição A.3.5 diz que se X_0, \dots, X_n são todos não-vazios, então pode-se *escolher* um elemento em cada X_i , o que é feito precisamente por uma n_+ -upla em $\prod_{i \in n_+} X_i$. Note que a heurística da demonstração consiste tão somente em *escolher* testemunhas da não-vacuidade dos X_i 's. Isso pode ser feito de maneira honesta justamente por n_+ ser finito.

Pergunta. Se tivéssemos $X_n \neq \emptyset$ para todo $n \in \omega$, seria possível usar o mesmo argumento para provar que $\prod_{n \in \omega} X_n \neq \emptyset$?

Resposta. Não.

O leitor pode contra-argumentar, naturalmente. Uma primeira tentativa poderia ser a seguinte: para cada $n \in \omega$, a Proposição A.3.5 garante pelo menos uma função $f_n \in \prod_{i \in n_+} X_i$, e daí bastaria definir $f(n) := f_n(n)$. O problema dessa abordagem está num detalhe muito sutil, escondido pela notação empregada: ao utilizar o subíndice “ n ” em “ f_n ”, pode não parecer, mas se fez uma escolha! De fato, ao escrever apenas $\prod_{i \in n_+} X_i \neq \emptyset$, deixa de ser trivial definir o valor da função f em n , já que para isso precisamos escolher alguma função g em $\prod_{i \in n_+} X_i$. Explicitamente, ao *fixar* uma f_n para cada n , o que se faz é supor a existência de pelo menos uma ω -upla no produto

$$\prod_{n \in \omega} \left(\prod_{i \in n_+} X_i \right),$$

tornando a coisa toda circular.

O problema é muito simples, na verdade: a não-vacuidade de $\prod_{i \in n_+} X_i$ se verifica por meio de “ $n + 1$ ” escolhas arbitrárias de elementos, um em cada X_i ; se existirem infinitos X_i 's, então a não-vacuidade de $\prod_{i \in \mathcal{I}} X_i$ dependerá de infinitas escolhas arbitrárias. Então, de duas uma: ou muda-se à *lógica* subjacente, a fim de permitir fórmulas infinitas¹², ou *postula-se a existência de uma escolha*.

Axioma da Escolha. Para qualquer conjunto $\mathcal{I} \neq \emptyset$ ocorre $\prod_{i \in \mathcal{I}} X_i \neq \emptyset$ se, e somente se, para todo $i \in \mathcal{I}$ ocorrer $X_i \neq \emptyset$.

Este é o polêmico Axioma da Escolha, que postula a *existência* de uplas em produtos cartesianos generalizados ou, em outras palavras, a *existência* de escolhas arbitrárias, mesmo que infinitas. É justamente o seu caráter *não-construtivo* que incomoda diversos matemáticos: não há qualquer menção sobre como tais escolhas são feitas, elas apenas existem. Em contextos voltados para questões computacionais, isso não costuma ajudar. Porém, como veremos ao longo do texto, diversos resultados *razoáveis* dependem em maior ou menor grau de consequências do Axioma da Escolha.

Exemplo A.3.6 (Opcional: convergência na reta). Leitores versados em Cálculo ou Análise devem se lembrar de que $r \in \mathbb{R}$ é **ponto aderente** de um subconjunto $A \subseteq \mathbb{R}$ se existe uma **sequência** $\langle a_n \rangle_{n \in \omega}$ em A com $a_n \rightarrow r$, que por sua vez significa que para todo $\varepsilon > 0$ existe $N \in \omega$ com $|r - a_n| < \varepsilon$ sempre que $n \geq N$. Claramente, em tal situação, r satisfaz a seguinte condição: para todo $\varepsilon > 0$, existe $a \in A$ satisfazendo $|a - r| < \varepsilon$.

¹²O leitor interessado deve procurar por *Lógicas Infinitárias*.

A recíproca da implicação observada acima ilustra uma típica situação que depende de infinitas escolhas: para cada $n \in \omega$, o conjunto $A_n := A \cap (r - \frac{1}{2^n}, r + \frac{1}{2^n})$ é não-vazio, o que sugere *definir* uma sequência $\langle a_n \rangle_{n \in \omega}$ com $a_n \in A_n$ para cada n . Entretanto, não há critério disponível que permita escolher $a_n \in A_n$ “uniformemente”: para cada n , sabe-se apenas que existem elementos em A_n , de modo que a *escolha* de algum deles é *arbitrária*, i.e., depende do *arbítrio* de alguém – possivelmente o ente hipotético que *executa* a demonstração. Entra em cena o Axioma da Escolha: como $A_n \neq \emptyset$ para todo $n \in \omega$, tem-se $\prod_{n \in \omega} A_n \neq \emptyset$, que significa, precisamente, a existência de uma ω -upla $\langle a_n \rangle_{n \in \omega} \in \prod_{n \in \omega} A_n$, i.e., com $a_n \in A_n$ para todo $n \in \omega$, como desejado. ▲

A.4 Dois axiomas técnicos

Restam dois axiomas para finalizar a descrição dos axiomas de ZFC. Embora relativamente técnicos, seria anticlimático não apresentá-los.

Axioma da Fundação. *Para todo X não-vazio existe $x \in X$ tal que $x \cap X = \emptyset$.*

Este axioma responde a perguntas que ninguém fez ~~me recuso a alertar sobre sareasmo~~.

Exercício A.24. Mostre que para todo x deve ocorrer $x \notin x$. Dica: o conjunto $X := \{x\}$ é não-vazio. ■

Exercício A.25. Mostre que não existem x e y satisfazendo $x \in y$ e $y \in x$. Dica: suponha $x \in y$ e considere $X := \{x, y\}$. ■

O Axioma da Fundação não prova só isso: na verdade, ele *serves* para *estruturar* o *universo* dos conjuntos numa *cadeia* ascendente de famílias da forma

$$\emptyset \subseteq \wp(\emptyset) \subseteq \wp^2(\emptyset) := \wp(\wp(\emptyset)) \subseteq \dots \subseteq \wp^n(\emptyset) \subseteq \dots \subseteq \wp^\omega(\emptyset) := \bigcup_{n \in \omega} \wp^n(\emptyset) \subseteq \wp(\wp^\omega(\emptyset)) \subseteq \dots$$

como veremos no final do capítulo. Em contextos mais aprofundados da Teoria dos Conjuntos, tal hierarquia é essencial. Porém, fora desses cenários, saber ou não desse axioma costuma ser irrelevante. Apesar disso, convém chamar a atenção para o conjunto $\wp^\omega(\emptyset)$ definido acima.

Implicitamente, a ideia é definir $\wp^n(\emptyset)$ para cada $n \in \omega$, o que se faz por recursão¹³. Entretanto, há um problema escondido: para definir $\bigcup_{n \in \omega} \wp^n(\emptyset)$, precisa-se da existência do conjunto $\{\wp^n(\emptyset) : n \in \omega\}$, que por sua vez deve ser a imagem da função que a cada n faz corresponder $\wp^n(\emptyset)$. Pergunta-se: por que tal função existe?

A rigor, uma função f é um subconjunto de um produto cartesiano $D \times C$ de conjuntos D (domínio) e C (codomínio) previamente conhecidos. No caso acima, não se tem um codomínio, mas apenas um domínio (ω) e uma regra ($n \mapsto \wp^n(\emptyset)$).

Diremos que uma *propriedade* \mathcal{P} sobre variáveis x e y é uma **fórmula funcional** em x se para cada A existir um único B que torne $\mathcal{P}(A, B)$ verdadeira.

Exemplo A.4.1. A fórmula $\mathcal{P}(x, y)$ expressa por “ y é o conjunto das partes de x ” é funcional, já que para cada x existe um único y que torna a afirmação verdadeira. Também é funcional a fórmula $\mathcal{Q}(x, y)$ dada por “ y é o sucessor de x ”. Por sua vez, a fórmula $\mathcal{R}(x, y)$ definida como “ $x \cap y = \emptyset$ ” não é funcional, já que para um mesmo x abundam os y ’s satisfazendo $x \cap y = \emptyset$. ▲

¹³O que será discutido em mais detalhes no capítulo seguinte.

Nos casos em que $\mathcal{P}(x, y)$ é uma fórmula funcional em x , faz sentido escrever $\mathcal{P}(x)$ para indicar o único y que torna $\mathcal{P}(x, y)$ verdadeira.

Axioma da Substituição. *Se $\mathcal{P}(x, y)$ é uma fórmula funcional e A é um conjunto, então existe a família $\mathcal{P}[A] := \{\mathcal{P}(x) : x \in A\}$, chamada de imagem de A por \mathcal{P} . Explicitamente, $\mathcal{P}[A]$ é formado por todos os y que tornam $\mathcal{P}(x, y)$ verdadeira para $x \in A$.*

Moralmente, uma fórmula funcional é uma função $\mathcal{P} : \mathbb{V} \rightarrow \mathbb{V}$, onde \mathbb{V} é o universo de todos os conjuntos. O problema, evidentemente, é que \mathbb{V} não é um conjunto e, portanto, \mathcal{P} não pode ser interpretada como uma função em ZFC. Nesse sentido, o Axioma da Substituição diz apenas que restrição de \mathcal{P} a qualquer conjunto *bona fide* é uma função de fato: a ideia é que cada $x \in A$ pode ser *substituído* pelo único y associado a x por \mathcal{P} , o que não deve resultar em um conjunto *grande demais* já que, como veremos, a cardinalidade da imagem de uma função nunca excede a cardinalidade do domínio.

A.5 Classes (próprias)

Embora o Paradoxo de Russell ensine que coisas do tipo $\{x : \mathcal{P}(x)\}$ devam ser consideradas com cuidado, a notação por si só parece muito conveniente para ser abandonada sem protesto. Por isso, em vez de proibir, pode ser mais razoável *permitir* que a notação $X := \{x : \mathcal{P}(x)\}$ seja usada, mas com cautela, pois X pode não ser um conjunto, o que sugere a pergunta: se X não é conjunto, então o que é X ?

Do ponto de vista de ZFC, uma resposta razoável poderia ser **Error404** – ou então um *not even wrong*: os objetos sobre os quais ZFC discursa são aqueles cuja existência é demonstrável pelos axiomas de ZFC, ou seja, conjuntos. Logo, *coisas* desse tipo não existem *formalmente*, como $\mathbb{V} := \{x : x = x\}$. Ainda assim, do lado de fora de ZFC, onde vivemos e nos comunicamos por meio de nossa (meta)linguagem, somos capazes de discursar sobre \mathbb{V} , afinal de contas é justamente o que estamos fazendo.

Definição A.5.1. Vamos chamar $\{x : \mathcal{P}(x)\}$ de **classe**, situação em que $y \in \{x : \mathcal{P}(x)\}$ será entendido como uma *abreviação* da sentença “ y é um conjunto e $\mathcal{P}(y)$ é verdadeira.” Adicionalmente:

- (i) diremos que duas classes $\{x : \mathcal{P}(x)\}$ e $\{x : \mathcal{Q}(x)\}$ são iguais se, e somente se, valer $\mathcal{P} \Leftrightarrow \mathcal{Q}$;
- (ii) diremos que $\{x : \mathcal{P}(x)\}$ é uma **classe própria** se não existir um conjunto Z satisfazendo $Z = \{x : \mathcal{P}(x)\}$, o que significa, explicitamente, que não existe conjunto Z tal que para todo x tenha-se $x \in Z \Leftrightarrow \mathcal{P}(x)$;
- (iii) classes *impróprias* são os bons e velhos conjuntos. ¶

ZFC não vem de fábrica com suporte para tratar formalmente de classes próprias: oficialmente, *quantificar* sobre classes próprias seria quantificar sobre *fórmulas* (i.e., sobre *propriedades*), e a estrutura de nossa linguagem permite apenas a quantificação sobre objetos do universo do discurso, i.e., conjuntos¹⁴. Mas há muitas coisas que podem ser feitas com classes, algumas mais ilegais do que outras.

¹⁴O jargão “quantificar sobre...” diz respeito ao escopo de “valores” que as variáveis *quantificadas* (pelos quantificadores “existe” e “para todo”) numa fórmula podem assumir.

Exemplo A.5.2. Todo conjunto é uma classe (*imprópria*) pois, para um conjunto X fixado, tem-se $y \in \{x : x \in X\}$ se, e somente se, $y \in X$. Por outro lado, $\mathbb{V} := \{x : x = x\}$ é uma classe própria (Proposição A.1.2), que será chamada de **universo**. \blacktriangle

Exemplo A.5.3. Dadas duas classes S e T , é natural considerar o *produto cartesiano* das classes S e T , denotado $S \times T$: seus elementos são os pares ordenados cujas coordenadas pertencem às classes correspondentes. Mais precisamente, para classes $S := \{x : \mathcal{P}(x)\}$ e $T := \{x : \mathcal{Q}(x)\}$, denota-se $S \times T := \{\langle x, y \rangle : \mathcal{P}(x) \text{ e } \mathcal{Q}(y)\}$. Em particular, para *conjuntos* S e T , mostra-se que tal classe é imprópria, já que existe um conjunto Z que possui, precisamente, os elementos de $S \times T$. \blacktriangle

Também faz sentido tratar de *subclasses*, adaptando a ideia geral do que significa ser um subconjunto: dadas classes C e D , diz-se que C é **subclasse** de D se valer a implicação “ $x \in C \Rightarrow x \in D$ ” para todo x . Note que isso apenas abrevia situações do tipo “ $\mathcal{P}(x) \Rightarrow \mathcal{Q}(x)$ ”, em que \mathcal{P} e \mathcal{Q} são *propriedades*.

Definição A.5.4. Diremos que uma subclasse $\mathcal{F} \subseteq S \times T$ é uma **função de classe**, escrita $\mathcal{F} : S \rightarrow T$, se para cada $x \in S$ existir um único $y \in T$ com $\langle x, y \rangle \in \mathcal{F}$, caso em que se escreve $\mathcal{F}(x) := y$. \P

Exemplo A.5.5. Para todo conjunto x existe um único y tal que $y = \{x\}$. Assim, tem-se uma função de classe $\mathcal{F} : \mathbb{V} \rightarrow \mathbb{V}$ em que $\mathcal{F}(x) := \{x\}$ para todo x . \blacktriangle

Observação A.5.6. A rigor, funções de classes não são a mesma coisa que as *funções* (entre conjuntos). Formalmente, funções de classe abreviam *propriedades* que se comportam como funções, i.e., as fórmulas funcionais. \triangle

Definição A.5.7. Dada uma função de classe $\mathcal{F} : S \rightarrow T$ e uma classe $C \subseteq S$, denota-se por $\{\mathcal{F}(c) : c \in C\}$ a subclasse de T dada pela *propriedade de* y “existe $x \in C$ tal que $\mathcal{F}(x) = y$ ”, que passa a ser chamada de **imagem de C por \mathcal{F}** e denotada por $\mathcal{F}[C]$. \P

Exemplo A.5.8. Com a notação acima, $\{\{x\} : x \in \mathbb{V}\}$ é tão somente a classe de *todos* os conjuntos unitários: $y \in \{\{x\} : x \in \mathbb{V}\}$ se, e somente se, existe x tal que $y = \{x\}$. Por sua vez, $\{\wp(X) : X \in \mathbb{V}\}$ é a classe de *todos* os conjuntos das partes. \blacktriangle

Com isso em mente, o Axioma da Substituição impõe que se X é um conjunto e $\mathcal{F} : \mathbb{V} \rightarrow \mathbb{V}$ é uma função de classe, então $\mathcal{F}[X]$ é um conjunto, o que explicitamente significa que existe um conjunto Z tal que para todo z , $z \in Y$ se, e somente se, existe $x \in X$ com $\mathcal{F}(x) = z$. Nesse sentido, cabe uma observação aparentemente inócuia, mas que será bastante importante em capítulos posteriores:

Lema A.5.9 (Princípio da Classe dos Pombos). *Seja $\mathcal{F} : S \rightarrow T$ uma função de classes. Se T é um conjunto e S é classe própria, então \mathcal{F} não é injetora¹⁵.*

Demonstração. Se fosse, então $\mathcal{G} := \{\langle t, s \rangle : \mathcal{F}(s) = t\} \subseteq T \times S$ seria uma função de classe da forma $\mathcal{F}[S] \rightarrow S$: por um lado, dado $t \in \mathcal{F}[S]$, existe (por hipótese) $s \in S$ com $\mathcal{F}(s) = t$, mostrando que $\langle t, s \rangle \in \mathcal{G}$; por outro lado, a injetividade assegura que tal s é único. Logo, o Axioma da Substituição garante que $\mathcal{G}[T]$ é um conjunto, o que é absurdo: se $s \in S$, então $t := \mathcal{F}(s) \in \mathcal{F}[S]$ e daí $\mathcal{G}(t) = s$, mostrando que $S \subseteq \mathcal{G}[T]$ e, portanto, é um conjunto¹⁶. \square

¹⁵A definição de injetividade para funções de classe é uma extensão óbvia da injetividade de funções.

¹⁶Esta conclusão é apenas o Axioma-esquema da Separação refraseado em termos de classes: toda subclasse de um conjunto é um conjunto (Exercício A.40).

Exercícios adicionais

Exercício A.26. Mostre que se $\mathcal{F} \neq \emptyset$, então

$$\bigcap \mathcal{F} = \left\{ x \in \bigcup \mathcal{F} : \exists F \in \mathcal{F} \text{ tal que } x \in F \right\}.$$

Conclua que seria legítimo *definir* $\bigcap \mathcal{F}$ por meio da identidade acima. ■

Observação A.5.10. A identidade anterior tem a vantagem de dar sentido a $\bigcap \mathcal{F}$ mesmo nas situações em que $\mathcal{F} = \emptyset$. Porém, ela é incompatível com a implicação

$$\mathcal{A} \subseteq \mathcal{B} \Rightarrow \bigcap \mathcal{B} \subseteq \bigcap \mathcal{A},$$

válida sempre que $\mathcal{A} \neq \emptyset$: como $\emptyset \subseteq \mathcal{B}$ para todo \mathcal{B} , *seria natural* esperar que $\bigcap \mathcal{B} \subseteq \bigcap \emptyset$, o que não é compatível com a proposta do último exercício, que daria $\bigcap \emptyset = \emptyset$. Em certo sentido, isso mostra a artificialidade da definição¹⁷. △

Exercício A.27. Para uma função $h: X \rightarrow Y$, mostre que $\text{Id}_Y \circ h = h = h \circ \text{Id}_X$. ■

Exercício A.28. Sejam $f: X \rightarrow Y$ e $g: Y \rightarrow Z$ funções.

- a) Mostre que se g e f são injetoras, então $g \circ f$ é injetora.
- b) Mostre que se g e f são sobrejetoras, então $g \circ f$ é sobrejetora.
- c) Mostre que se g e f são bijetoras, então $g \circ f$ é bijetora.
- d) Determine a inversa de $g \circ f$. ■

Exercício A.29. Sejam $f: X \rightarrow Y$ e $g: Y \rightarrow Z$ funções.

- a) Mostre que se $g \circ f$ é injetiva, então f é injetiva.
- b) Mostre que se $g \circ f$ é sobre, então g é sobre.
- c) Mostre que se $g \circ f$ é bijetora, então f é injetora e g é sobrejetora. ■

Exercício A.30. Sejam $f: X \rightarrow Y$ e $g: Y \rightarrow X$ funções. Mostre que se $g \circ f$ e $f \circ g$ são bijeções, então f e g são bijeções. Adicionalmente, se $f \circ g = \text{Id}_Y$ ou $g \circ f = \text{Id}_X$, então $g = f^{-1}$. ■

Exercício A.31. Seja $f: X \rightarrow Y$ uma função. Mostre que a função f é injetora se, e somente se, existe uma (sobrejeção) $g: Y \rightarrow X$ tal que $g \circ f = \text{Id}_X$. ■

Exercício A.32. Seja $f: X \rightarrow Y$ uma função.

- a) Mostre que $f[A] \subseteq f[A']$ e $f^{-1}[B] \subseteq f^{-1}[B']$ sempre que $A \subseteq A' \subseteq X$ e $B \subseteq B' \subseteq Y$, respectivamente.
- b) Mostre que $f[\emptyset] = \emptyset$, $f^{-1}[\emptyset] = \emptyset$ e $f^{-1}[Y] = X$.
- c) Para $A \subseteq X$ e $B \subseteq Y$ quaisquer, mostre que valem as inclusões $A \subseteq f^{-1}[f[A]]$ e $f[f^{-1}[B]] \subseteq B$, com as igualdades garantidas se f for injetora (para o primeiro caso) ou sobrejetora (para o segundo caso). ■

Exercício A.33 (Colagem de funções). Seja \mathcal{F} uma família de funções.

¹⁷Para leitores versados em frações, seria como introduzir a notação $\frac{1}{0} := 0$ em \mathbb{Z} , apenas para dar sentido a expressões do tipo ab^{-1} para quaisquer $a, b \in \mathbb{Z}$.

- Mostre que $F := \bigcup \mathcal{F}$ é uma *relação binária* que satisfaz $\text{dom}(F) = \bigcup_{f \in \mathcal{F}} \text{dom}(f)$ e $\text{im}(F) = \bigcup_{f \in \mathcal{F}} \text{im}(f)$.
- Mostre que F é uma função se, e somente se, $f(x) = g(x)$ sempre que $x \in \text{dom}(f) \cap \text{dom}(g)$, para quaisquer $f, g \in \mathcal{F}$.
- Conclua que se $\text{dom}(f) \cap \text{dom}(g) = \emptyset$ sempre que $f, g \in \mathcal{F}$ forem distintas, então F é uma função. Se cada f for injetiva, o que exigir a fim de garantir a injetividade de F ? ■

Exercício A.34 (Indução a partir de k). Mostre que se $\mathcal{P}(x)$ for uma propriedade verificada para algum $k \in \omega$ e, para todo $m \geq k$ valer que $\mathcal{P}(m) \Rightarrow \mathcal{P}(m+1)$, então $\mathcal{P}(n)$ vale para todo $n \geq k$. ■

Exercício A.35. Mostre que não existe sequência $\langle x_n \rangle_{n \in \omega}$ com $x_{n+1} \in x_n$ para todo $n \in \omega$. Dica: se existisse, o que o Axioma da Fundação diria a respeito do conjunto $\{x_n : n \in \omega\}$? ■

Exercício A.36 (Propriedade universal de $\prod_{i \in \mathcal{I}}$). Dada uma \mathcal{I} -upla $\mathcal{X} := \langle X_i : i \in \mathcal{I} \rangle$ de conjuntos não-vazios, o produto cartesiano $\prod \mathcal{X} := \prod_{i \in \mathcal{I}} X_i$ vem de fábrica com uma \mathcal{I} -upla de funções $\langle \pi_i : i \in \mathcal{I} \rangle$, chamadas de **projeções**, onde para cada $j \in \mathcal{I}$ tem-se $\pi_j : \prod \mathcal{X} \rightarrow X_j$ dada pela regra $\pi_j(\langle x_i \rangle_{i \in \mathcal{I}}) := x_j \in X_j$ para cada \mathcal{I} -upla $\langle x_i \rangle_{i \in \mathcal{I}} \in \prod \mathcal{X}$. Mostre que se X é um conjunto munido de uma função $f_i : X \rightarrow X_i$ para cada $i \in \mathcal{I}$, então existe uma única função $F : X \rightarrow \prod \mathcal{X}$ satisfazendo $\pi_j \circ F = f_j$ para todo $j \in \mathcal{I}$. Dica: basta fazer $F(x) := \langle f_i(x) \rangle_{i \in \mathcal{I}}$. ■

Observação A.5.11 (Funções diagonais vs. funções produtos). Costuma-se chamar a função F do exercício anterior como **produto (diagonal)** das funções f_i , que será denotada por $\langle f_i \rangle_{i \in \mathcal{I}}$. Mais geralmente, se para cada $i \in \mathcal{I}$ existir uma função $g_i : X_i \rightarrow Y_i$ fixada, pode-se definir

$$\prod_{i \in \mathcal{I}} g_i : \prod_{i \in \mathcal{I}} X_i \rightarrow \prod_{i \in \mathcal{I}} Y_i \quad (\text{A.1})$$

que a cada \mathcal{I} -upla $\langle x_i \rangle_{i \in \mathcal{I}}$ de $\prod_{i \in \mathcal{I}} X_i$ associa a \mathcal{I} -upla $\langle g_i(x_i) \rangle_{i \in \mathcal{I}}$ de $\prod_{i \in \mathcal{I}} Y_i$. Naturalmente, ela se obtém do caso anterior tomando-se $X := \prod_{i \in \mathcal{I}} X_i$ e $f_i := \pi_i \circ g_i$ para cada $i \in \mathcal{I}$. Esta sim é chamada de **produto (cartesiano)** das funções g_i , também denotada como $g_0 \times \dots \times g_n$ caso se tenha $\mathcal{I} := \{0, \dots, n\}$. △

Exercício A.37. O leitor pode achar a notação anterior inadequada, por permitir que se confunda a função $\langle f_i \rangle_{i \in \mathcal{I}} : X \rightarrow \prod \mathcal{X}$ com a \mathcal{I} -upla de funções $\langle f_i \rangle_{i \in \mathcal{I}} \in \prod_{i \in \mathcal{I}} X_i^X$, onde cada $f_i \in X_i^X$ – i.e., cada $f_i : X \rightarrow X_i$ é uma função. Ocorre que isto é intencional.

- Mostre que o processo de associar uma \mathcal{I} -upla $\langle f_i \rangle_{i \in \mathcal{I}} \in \prod_{i \in \mathcal{I}} X_i^X$ ao produto diagonal $\langle f_i \rangle_{i \in \mathcal{I}} : X \rightarrow \prod_{i \in \mathcal{I}} X_i$ pode ser visto como uma função

$$\varphi : \prod_{i \in \mathcal{I}} X_i^X \rightarrow \left(\prod_{i \in \mathcal{I}} X_i \right)^X. \quad (\text{A.2})$$

- Mostre que uma função $G : X \rightarrow \prod_{i \in \mathcal{I}} X_i$ induz *naturalmente* uma \mathcal{I} -upla em $\prod_{i \in \mathcal{I}} X_i^X$, o que descreve uma função

$$\psi : \left(\prod_{i \in \mathcal{I}} X_i \right)^X \rightarrow \prod_{i \in \mathcal{I}} X_i^X. \quad (\text{A.3})$$

- Mostre que as funções φ e ψ são inversas uma da outra. ■

Exercício A.38. Suponha $X := X_i$ e $g_i := \text{Id}_X$ para todo $i \in \mathcal{I}$ e escreva

$$\Delta_{i \in \mathcal{I}} : X \rightarrow \prod_{i \in \mathcal{I}} X \quad (\text{A.4})$$

para indicar o produto diagonal das funções g_i 's. Mostre que se $f_i : X \rightarrow X_i$ é uma função para cada $i \in \mathcal{I}$, então $\langle f_i \rangle_{i \in \mathcal{I}} = \prod_{i \in \mathcal{I}} f_i \circ \Delta_{i \in \mathcal{I}}$. ■

Observação A.5.12. A função $\Delta_{i \in \mathcal{I}}$ passa a ser chamada de (função) **diagonal** de X em $\prod_{i \in \mathcal{I}} X$, mesmo nome dado à imagem da função $\Delta_{i \in \mathcal{I}}$. △

Exercício A.39. Sejam X , Y e Z conjuntos.

- a) Mostre que $X^Z = \prod_{z \in Z} X$.
- b) Dada uma função $f : X \rightarrow Y$, mostre que existe uma única função $f^Z : X^Z \rightarrow Y^Z$ satisfazendo $\pi_{z'}(f^Z(\langle x_z \rangle_{z \in Z})) = f(z')$ para cada $z' \in Z$. ■

Exercício A.40. Sejam S e T classes, com S subclasse de T . Mostre que se T é um conjunto, então S é um conjunto. ■

Exercício A.41. Use o Axioma da Substituição (entre outros que você achar importantes) para provar a existência de $A \times B$, para certos conjuntos A e B . Dica: para $b \in B$ fixado, mostre que existe um único conjunto, digamos F_b , cujos membros são da forma $z = \langle a, b \rangle$, com $a \in A$; depois, “olhe” $A \times B$ como uma reunião de “fibras”. ■

Exercício A.42. Seja $\psi(x, y)$ um fórmula em que x e y sejam livres. Diremos que ψ é **parcialmente funcional** em x se para todo x existe *no máximo* um y tal que $\psi(x, y)$, o que pode ser expresso assim: $\forall x \forall y \forall z (\psi(x, y) \wedge \psi(x, z) \Rightarrow y = z)$.

- a) Chamemos de *Axioma da Substituição Parcial* o seguinte *esquema* de axiomas: para uma fórmula $\psi(x, y)$ parcialmente funcional em x e para todo conjunto A existe um conjunto B tal que para todo y , $y \in B$ se, e somente se, existe $x \in A$ com $\psi(x, y)$. Simbolicamente: se $\psi(x, y)$ é parcialmente funcional, então

$$\forall A \exists B \forall y (y \in B \Leftrightarrow \exists x (x \in A \wedge \psi(x, y)))$$

é um axioma. Mostre que o Axioma da Substituição Parcial implica o Axioma da Substituição.

- b) Assuma o Axioma da Extensão juntamente com o “Axioma do Vazio”, este último que postula a existência do conjunto vazio. Mostre que com a adição destes axiomas, a recíproca do item anterior se verifica. Dica: considere, separadamente, o caso em que não existe $x \in A$ tal que $\psi(x, y)$, e o caso em que existe pelo menos um $x \in A$ tal que $\psi(x, y)$. ■

Exercício A.43. Mostre que o Axioma da Separação segue do (s):

- a) Axioma da Substituição + Axioma da Extensão + Axioma do Vazio;
- b) Axioma da Substituição Parcial.

Dica: a ideia é separar de um conjunto A dado todos os elementos x que satisfazem uma certa fórmula $\varphi(x)$; dito isso, cozinhe uma fórmula funcional $\psi(x, y)$ auxiliar cuja imagem, quando restrita ao conjunto A , resulte em $\{x \in A : \varphi(x)\}$. Meta-dica: prove apenas o segundo caso e use o exercício anterior para derivar o primeiro. ■

Exercício A.44. Mostre que o Axioma do Par segue do Axioma da Substituição aliado ao Axioma das Partes – por comodidade, use também o Axioma da Extensão. Dica: use o exercício anterior para cozinhar “o” conjunto vazio¹⁸, e o Axioma das Partes para cozinhar $\wp(\emptyset)$ e, depois, $B := \wp(\wp(\emptyset))$; use então B como domínio de uma fórmula funcional apropriada. ■

¹⁸O Axioma da Extensão entra apenas para garantir um sono tranquilo e o uso de artigos definidos.

Capítulo B

Relações

Embora o capítulo anterior já tenha cumprido a tarefa de apresentar a axiomática de ZFC e discutir a *implementação* de animais importantes, como as relações binárias e funções, não se deu atenção para a *taxonomia* e, muito menos, para as particularidades de cada *espécie*. Dada a impossibilidade de aprofundar a análise das noções de cardinalidade sem saber o que são *relações de equivalência*, *ordens*, *boas ordens* e *recursão*, o presente capítulo apresenta *um mínimo* sobre tais assuntos. Leitores versados em tais tópicos podem avançar para o próximo capítulo.

B.1 Equivalências e partições

Definição B.1.1. Uma relação binária \sim num conjunto X é dita uma **relação de equivalência** se \sim for

- (i) **reflexiva**, i.e., se para todo $x \in X$ ocorrer $x \sim x$,
- (ii) **simétrica**, i.e., se para quaisquer $x, y \in X$, a ocorrência de $x \sim y$ acarretar $y \sim x$, e
- (iii) **transitiva**, i.e., se para quaisquer $x, y, z \in X$, a ocorrência simultânea de $x \sim y$ e $y \sim z$ acarretar $x \sim z$.

Diremos também que x e y são \sim -**equivalentes** sempre que ocorrer $x \sim y$, com a omissão do sufixo “ \sim ” quando a relação \sim estiver clara pelo contexto. ¶

Em certo sentido, uma relação de equivalência \sim estabelece um critério por meio do qual objetos a princípio distintos podem ser vistos como iguais, ao mesmo tempo em que separa outros objetos distintos pelo mesmo critério. Dessa forma, não espanta que a *relação de igualdade* ($x \sim y$ se, e somente se, $x = y$) seja o exemplo óbvio de equivalência.

Exemplo B.1.2 (Horóscopo). Frequentemente, praticantes da (pseudociênciça chamada de) **Astrologia** fazem uso implícito das relações de equivalência. De fato, de um ponto de vista informal, signos determinam uma “relação de equivalência” no “conjunto” de todas as pessoas:

- ✓ toda pessoa tem o mesmo signo de si mesma;
- ✓ se P tem o mesmo signo de P' , então P' tem o mesmo signo de P ;
- ✓ se P tem o mesmo signo de P' e esta tem o mesmo signo de P'' , então P e P'' têm o mesmo signo.

Se a coisa parasse por aí, a *Astrologia* seria inofensiva. No entanto, é comum encontrar asserções do tipo “o comportamento X é característico do signo Y ”, o que sugere duas alternativas: ou a afirmação é falsa, ou *toda pessoa* do signo Y apresenta o comportamento X . Esse tipo de máxima ajuda a entender um dos usos mais comuns das relações de equivalência: a simplificação. Com efeito, se tais afirmações *astrológicas* fossem verdadeiras, então para entender os padrões comportamentais de *toda a humanidade* bastaria estudar os comportamentos de doze *representantes*, um de cada signo, algo bem mais simples do que estimar o comportamento individual dos oito bilhões de habitantes do planeta. Por sorte, signos estimam tão somente as datas de nascimento de seus portadores. ▲

Exemplo B.1.3 (Paridade). O leitor já familiarizado com *aritmética natural* sabe que números naturais podem ser classificados como pares ou ímpares: pares são os múltiplos de dois, ímpares são os outros. Isso pode ser usado para determinar uma relação \sim de equivalência em ω : $m, n \in \omega$ serão ditos \sim -equivalentes se tiverem a mesma *paridade*. Assim, $0, 2, 4, 6, \dots$ são \sim -equivalentes entre si, $1, 3, 5, 7, \dots$ são \sim -equivalentes entre si, enquanto 0 e 1 não são \sim -equivalentes, por exemplo. Agora, há certos comportamentos *algebricos* que não dependem dos números escolhidos, e sim de suas paridades: a soma de quaisquer dois ímpares é *par*, o produto entre quaisquer ímpares é *ímpar*, etc. Isto sugere a possibilidade de realizar operações diretamente com as *classes* dos pares e dos ímpares em vez de lidar com seus infinitos representantes. Voltaremos a isso oportunamente. ▲

Exemplo B.1.4 (Restos da divisão por n). Para generalizar o exemplo anterior, pode-se considerar a seguinte relação binária: para $n \in \omega$ fixado e $x, y \in \omega$, escreveremos $x \sim_n y$ a fim de indicar que x e y têm o mesmo *resto* na *divisão* por n . Verificar que tal relação \sim_n é reflexiva, simétrica e transitiva é um bom exercício para quem já sabe fazer divisões. Ocorre que, como antes, certos comportamentos algébricos não dependem dos representantes escolhidos: por exemplo, a soma de quaisquer dois números com resto 2 na divisão por 3 terá resto 1, enquanto o produto de quaisquer dois números com resto 1 na divisão por 3 ainda terá resto 1. ▲

Um efeito colateral inevitável das relações de equivalência é a segregação dos elementos do conjunto em *classes de equivalência*¹. Mais precisamente:

Definição B.1.5. Para uma relação de equivalência \sim sobre um conjunto X , diremos que o conjunto $\{y \in X : x \sim y\}$ é a \sim -**classe de equivalência** de x . ¶

Com *relação* aos exemplos anteriores:

- (i) a classe de equivalência de uma pessoa P com respeito aos signos astrológicos seria a coleção de todas as pessoas que têm o mesmo signo de P , consequentemente, existem apenas doze classes de equivalência (correspondentes aos signos possíveis);
- (ii) a classe de equivalência de um número $n \in \omega$ com respeito à paridade é a coleção dos naturais que têm a mesma paridade de n , logo, existem apenas duas classes, a dos pares e a dos ímpares;
- (iii) a classe de equivalência de um número $p \in \omega$ com respeito aos restos da divisão por n é a coleção dos números naturais que têm o mesmo resto na divisão, o que leva à conclusão de que existem precisamente n classes de equivalência (correspondentes aos restos possíveis na divisão por n).

¹Conotações políticas ficam a cargo do leitor.

Observação B.1.6. A notação para a classe de equivalência de x varia de acordo com o contexto. Frequentemente, escreve-se \bar{x} para denotá-la. Apesar disso, para a discussão a seguir, pode ser menos traumático escrever C_x em vez de \bar{x} . \triangle

Mais geralmente, verifica-se a seguinte

Proposição B.1.7. *Sejam X um conjunto e \sim uma relação de equivalência sobre X . Então:*

- (i) $X = \bigcup_{x \in X} C_x$, i.e., para todo $y \in X$, existe $x \in X$ com $y \in C_x$;
- (ii) para quaisquer $x, y \in X$ ocorre $C_x = C_y$ ou $C_x \cap C_y = \emptyset$;
- (iii) para quaisquer $x, y \in X$, $C_x = C_y$ se, e somente se, $x \sim y$.

Demonstração. O primeiro item decorre da reflexividade de \sim : como $x \sim x$, segue que $x \in C_x$ e, portanto, $X \subseteq \bigcup_{x \in X} C_x$, com a inclusão oposta automática em virtude da definição de cada C_x . Os dois itens seguintes ficam a cargo do leitor (confira o exercício a seguir). \square

Exercício B.1. Sejam \sim uma relação binária em X e $x, y \in X$ elementos quaisquer.

- a) Mostre que se \sim é transitiva, então “ $x \sim y \Rightarrow C_x \subseteq C_y$ ”.
- b) Mostre que se \sim é simétrica e transitiva, então “ $x \sim y \Rightarrow C_x = C_y$ ”.
- c) Mostre que se \sim é reflexiva, então “ $C_x \subseteq C_y \Rightarrow x \sim y$ ”.
- d) Conclua que valem as condições (ii) e (iii) da proposição anterior. Dica: para (ii), o que ocorre com C_z se $z \in C_x \cap C_y$? \blacksquare

B.1.1 Partições e representantes

A última proposição mostra que $X/\sim := \{C_x : x \in X\}$, chamado de **quociente** de X por \sim , é uma família de subconjuntos não-vazios de X que se enquadra como exemplo de **partição**.

Definição B.1.8. Uma família \mathcal{P} de subconjuntos não-vazios de X é uma **partição** de X se valerem as seguintes condições:

- (i) $X = \bigcup \mathcal{P}$; e
- (ii)² se $P, Q \in \mathcal{P}$ e $P \neq Q$, então $P \cap Q = \emptyset$. \P

Exercício B.2. Mostre que se \sim é uma relação de equivalência sobre X , então X/\sim é uma partição de X . \blacksquare

Como o nome sugere, uma partição de X *partitiona* o conjunto X em *partes* ou blocos *dois a dois disjuntos*, de modo que cada elemento de X está precisamente em apenas um membro de \mathcal{P} . Assim, faz sentido dizer que dois elementos de X são \mathcal{P} -equivalentes se pertencerem ao mesmo membro de \mathcal{P} . Como o leitor atento certamente suspeita, isto define uma relação de equivalência legítima.

²Costuma-se expressar a condição (ii) como “os membros de \mathcal{P} são dois a dois disjuntos”.

Proposição B.1.9. Se \mathcal{P} for uma partição de X , então a relação $\sim_{\mathcal{P}}$ definida por

$$u \sim_{\mathcal{P}} v \Leftrightarrow \exists P \in \mathcal{P} \text{ tal que } \{u, v\} \subseteq P$$

é uma relação de equivalência em X . Além disso, $\mathcal{P} = X/\sim_{\mathcal{P}}$.

Demonstração. A relação $\sim_{\mathcal{P}}$ é

- ✓ reflexiva, pois dado $x \in X$ existe $P \in \mathcal{P}$ com $x \in P$, e daí $\{x\} = \{x, x\} \subseteq P$,
- ✓ simétrica, pois se $\{x, y\} \subseteq P \in \mathcal{P}$, então $\{x, y\} = \{y, x\} \subseteq P \in \mathcal{P}$, e
- ✓ transitiva, pois se $\{x, y\} \subseteq P \in \mathcal{P}$ e $\{y, z\} \subseteq P' \in \mathcal{P}$, então $P \cap P' \neq \emptyset$, acarretando $P = P'$ e, por conseguinte, $\{x, z\} \subseteq \{x, y\} \cup \{y, z\} \subseteq P \in \mathcal{P}$.

A igualdade $\mathcal{P} = X/\sim_{\mathcal{P}}$ segue pois $P = [x]_{\sim_{\mathcal{P}}}$ para quaisquer $x \in P$ com $x \in P \in \mathcal{P}$. \square

O leitor com a impressão de que relações de equivalência e partições são apenas nomes diferentes para a mesma coisa pode passar a ter certeza com o

Exercício B.3. Para X fixado, sejam $\text{Eqv}(X) := \{R \subseteq X \times X : R \text{ é rel. de equivalência}\}$ e $\text{Part}(X) := \{\mathcal{P} \subseteq \wp(X) : \mathcal{P} \text{ é partição de } X\}$. Mostre que as regras $R \mapsto X/R$ e $\mathcal{P} \mapsto \sim_{\mathcal{P}}$ definem funções da forma $\text{Eqv}(X) \rightarrow \text{Part}(X)$ e $\text{Part}(X) \rightarrow \text{Eqv}(X)$, respectivamente, que são inversas uma da outra. ■

Exemplo B.1.10. Para a relação \sim_n do Exemplo B.1.4, as classes de equivalência correspondem precisamente a todos os possíveis restos pela divisão por n . Assim, chamando por R_i a coleção dos naturais que têm resto i na divisão por n , segue que $\omega/\sim_n = \{R_0, R_1, \dots, R_{n-1}\}$. Dito isso, observe que ao chamar por \bar{i} a classe de equivalência de i , verifica-se $\bar{i} = R_i$. Desse modo, seria lícito escrever, por exemplo, $\omega/\sim_2 = \{\bar{0}, \bar{1}\}$, ou ainda $\omega/\sim_2 = \{\bar{12}, \bar{13}\}$, já que $\bar{0} = \bar{12}$ na relação \sim_2 (0 e 12 são divisíveis por 2) e $\bar{1} = \bar{13}$ (1 e 13 têm resto 1 na divisão por 2). Como o leitor deve suspeitar, trata-se de um fenômeno mais geral. ▲

Definição B.1.11. Um subconjunto $R \subseteq X$ é uma **classe de representantes** de uma

- (i) relação (de equivalência) \sim se para todo $x \in X$ existe um único $r \in R$ tal que $x \sim r$,
- (ii) partição \mathcal{P} de X se R for classe de representantes da relação $\sim_{\mathcal{P}}$. ¶

Exercício B.4. Nas condições anteriores, mostre que $X/\sim = \{C_r : r \in R\}$, onde R é uma classe de representantes de \sim . ■

Observação B.1.12. A única relação de equivalência sobre \emptyset é \emptyset , o que poderia levar o leitor incauto a afirmar que $\{\emptyset\}$ é uma partição de \emptyset . Mas não é o caso: uma partição é, antes de qualquer outra coisa, uma família de conjuntos *não-vazios*. Assim, a única partição de \emptyset é \emptyset , por *vacuidade*. Diariamente, porém, não convém se preocupar com partições do vazio: o leitor não sofrerá grandes prejuízos em preferir restringir a definição de partição para conjuntos não-vazios. △

De um jeito ou de outro, uma relação de equivalência partitiona o seu domínio em “blocos” não-vazios dois a dois disjuntos: cada bloco contém precisamente todos os elementos do domínio que são equivalentes entre si. Em particular, para $X \neq \emptyset$, como cada classe de equivalência $C \in X/\sim$ é um conjunto não-vazio, segue (do Axioma da Escolha) que existe uma upla

$$f \in \prod_{C \in X/\sim} C.$$

Explicitamente, f é uma função que para cada classe de equivalência C escolhe $f(C) \in C$. Daí, o conjunto $R := \{f(C) : C \in X/\sim\}$ é uma classe de representantes para \sim : de fato, se $f(C) \neq f(D)$ para certas classes $C, D \in X/\sim$, então deve-se ter $C \neq D$, caso contrário ocorreria $f(C) = f(D)$; logo, $C \cap D = \emptyset$ e, consequentemente, $f(C)$ e $f(D)$ não estão relacionados entre si. Em suma, demonstrou-se o

Teorema B.1.13. *Se \sim é uma relação de equivalência sobre um conjunto, então existe uma classe de representantes para \sim .*

O teorema acima depende profundamente do Axioma da Escolha. Na verdade, ao se assumir seu enunciado como axioma juntamente com os demais axiomas de ZFC (mas sem assumir o Axioma da Escolha) é possível *derivar* o Axioma da Escolha como um *teorema*. O modo profissional de expressar isso é dizer que as duas afirmações são *equivalentes em ZF*³. Isto será discutido com mais calma no Capítulo D.

B.1.2 Funções quocientes e projeções

A situação em que se precisa definir alguma função cujo domínio é um quociente determinado por uma relação de equivalência é tão frequente que vale o esforço de dedicar uma subseção inteira para o problema.

Exemplo B.1.14 (Paridade). Assumindo-se novamente conhecidas as notações básicas de aritmética, mas desta vez com *números inteiros*, considere a função $\psi: \omega \rightarrow \{-1, 1\}$ que faz $\psi(n) := (-1)^n$. Ao recordar as propriedades usuais de *potenciação*, não é difícil perceber que $\psi(0) = \psi(2) = \psi(4) = \dots = 1$, enquanto $\psi(1) = \psi(3) = \psi(5) = \dots = -1$. Como todos os pares são *enviados* em 1 e todos os ímpares são *enviados* em -1, a *regra* $\bar{m} \mapsto \psi(m)$ define uma função da forma $\omega/\sim_2 \rightarrow \{-1, 1\}$:

- (i) a princípio, ω/\sim_2 tem apenas dois elementos, digamos P e I , que correspondem às classes dos pares e ímpares, respectivamente;
- (ii) por sua vez, a *regra* $\bar{m} \mapsto \psi(m)$ relaciona uma classe de equivalência à imagem pela ψ de algum representante, de modo que, a rigor, tem-se apenas a relação binária

$$\bar{\psi} := \{\langle \bar{m}, \psi(m) \rangle : m \in \omega\};$$

- (iii) note que se ocorrer $P \bar{\psi} i$ e $P \bar{\psi} j$, então $i = \psi(m)$ e $j = \psi(n)$ para certos $m, n \in P$, e pelo que se observou no começo, resulta $i = j = 1$;
- (iv) analogamente, se $I \bar{\psi} i$ e $I \bar{\psi} j$, então $i = j = -1$.

³ZF é a sigla para a axiomática Zermelo-Fraenkel, que consiste de ZFC *sem* o Axioma da Escolha. Note bem: isto não significa dizer que o Axioma da Escolha é falso em ZF; apenas não se assume a afirmação como axioma.

Em suma, a relação binária $\bar{\psi}$ é, na verdade, uma função cujo domínio é ω/\sim_2 e que tem a mesma imagem que a função ψ .

Por outro lado, se $\varphi: \omega \rightarrow Y$ for uma função que não *respeita* a paridade, digamos que com $\varphi(0) := y$, $\varphi(2) := y'$ e $y \neq y'$, então a relação binária $\bar{\varphi} := \{\langle \bar{m}, \varphi(m) \rangle : m \in \omega\}$ não é uma função, já que ocorre tanto $P\bar{\varphi}y$ quanto $P\bar{\varphi}y'$. Note que isto não invalida a definição da relação binária $\bar{\varphi}$, mas apenas mostra que ela não é uma função. Vamos ver isso num contexto um pouco mais geral a seguir. \blacktriangle

Definição B.1.15. Fixada uma relação de equivalência \sim sobre X , denota-se por π ou π_\sim à função $X \rightarrow X/\sim$ que associa cada $x \in X$ à sua classe de equivalência $\bar{x} := \{y \in X : x \sim y\}$. Profissionais costumam xingar π_\sim de **projeção canônica**. \P

A projeção π acima ajuda a entender o tipo de função que se quer definir no quociente X/\sim ou, mais precisamente, que tipo de *relação* tais *projeções* devem ter com as funções da forma $X \rightarrow Y$. Com o exemplo anterior em mente, para um função $\psi: X \rightarrow Y$ fixada, gostaríamos de definir uma função $\bar{\psi}: X/\sim \rightarrow Y$ tal que $\bar{\psi}(\bar{x}) = \psi(x)$ para qualquer $x \in X$ ou, equivalentemente, $\bar{\psi} \circ \pi = \psi$.

Teorema B.1.16 (Propriedade universal do quociente). *Nas condições acima, a função $\bar{\psi}$ procurada existe se, e somente se, ψ é \sim -compatível, no seguinte sentido: $\psi(x) = \psi(y)$ sempre que $x \sim y$. Além disso, $\bar{\psi}$, quando existe, é única.*

Demonstração. Note que se $g, h: X/\sim \rightarrow Y$ satisfazem $g \circ \pi = h \circ \pi$, então para qualquer $\alpha \in X/\sim$ pode-se tomar $x \in X$ com $\alpha = \bar{x}$, de modo que

$$g(\alpha) = g(\bar{x}) = (g \circ \pi)(x) = (h \circ \pi)(x) = h(\bar{x}) = h(\alpha),$$

o que garante a unicidade⁴. Agora, observe que se $\bar{\psi}$ existe, então ψ é \sim -compatível: na verdade, sempre que $\varphi: X/\sim \rightarrow Y$ é uma função, tem-se $\varphi \circ \pi \sim$ -compatível, já que $x \sim y$ acarreta $\pi(x) = \pi(y)$ e, por conseguinte, $(\varphi \circ \pi)(x) = (\varphi \circ \pi)(y)$. Reciprocamente, se ψ é \sim -compatível, então a relação binária $\bar{\psi} := \{\langle \bar{x}, \psi(x) \rangle : x \in X\}$ é uma função da forma $X/\sim \rightarrow Y$ com a propriedade desejada: se $\bar{x}\bar{\psi}\psi(x)$ e $\bar{x}'\bar{\psi}\psi(x')$ com $\bar{x} = \bar{x}'$, então $x \sim x'$ e, portanto, $\psi(x) = \psi(x')$. O restante é automático. \square

Exercício B.5. Sejam $R \subseteq X \times Y$ uma relação binária e \sim uma relação de equivalência em X . Mostre que $\bar{R} := \{\langle \bar{x}, y \rangle : x R y\}$ define uma função da forma $X/\sim \rightarrow Y$ se, e somente se, R é \sim -compatível, i.e., se $y = y'$ sempre que $x R y$ e $x' R y'$ com $x \sim x'$. \blacksquare

Moral da história: as funções da forma $X \rightarrow Y$ que podem ser usadas para definir funções da forma $X/\sim \rightarrow Y$ são precisamente aquelas que tratam como iguais os elementos de X equivalentes perante \sim .



Figura B.1: Enquanto ψ manda todos os elementos de um bloco para um mesmo *ponto*, $\bar{\psi}$ manda os próprios blocos para os respectivos *pontos*.

⁴Note que apenas a sobrejetividade de π foi importante.

Observação B.1.17 (Diagramas). Frequentemente, uma ferramenta pictórica muito simples torna a o tratamento de composições de setas funções mais inteligível:

$$\begin{array}{ccc} X & & \\ \downarrow \pi & \searrow \psi & \\ X/\sim & \dashrightarrow \bar{\psi} & Y \end{array}$$

Uma figura como a anterior é o que costuma se chamar de **diagrama comutativo**. A ideia é explicitar de forma mais visual as relações entre as funções envolvidas existentes (com traçado contínuo, no caso π e ψ) e as funções procuradas (tracejadas, no caso $\bar{\psi}$) que tornam o diagrama *comutativo*: a comutatividade, neste contexto, significa dizer que todos os caminhos são iguais, i.e., ir de X até Y por meio de ψ é a mesma coisa que ir de X até X/\sim por meio de π para daí prosseguir até Y por $\bar{\psi}$. Diagramas desse tipo ocorrerão mais vezes ao longo do texto (sugestão: reinterprete os Exercícios A.36 e A.39 por meio de diagramas). \triangle

B.1.3 Opcional: inteiros e racionais

A fim de ilustrar com um pouco mais de vigor o uso de quocientes, convém fazer um breve desvio para contextos algébricos, o que será feito nesta subseção. Dado que este não é o objetivo principal do texto, o leitor que já se sentir satisfeito com as discussões pode avançar para a Seção B.2.

Observação B.1.18 (Anacronismo). Justamente por seu caráter algébrico, esta subseção pressupõe familiaridade com noções básicas de aritmética. Leitores bourbakistas devem conferir o Capítulo E (ou enfrentar o Exercício B.27) para as devidas discussões sobre aritmética cardinal, o que inclui (sem grande ênfase) aritmética em ω . \triangle

Fixado um conjunto X , uma função $*: X \times X \rightarrow X$ é chamada de **operação binária** em X . Porém, como não trataremos explicitamente de outras *aridades* (por enquanto), não há risco em chamar $*$ simplesmente de operação. Para $x, y \in X$, costuma-se escrever $x * y$ em vez de $*(x, y)$, em alusão às notações já bem estabelecidas para *adição* e *multiplicação*.

Definição B.1.19. Seja $*: X \times X \rightarrow X$ uma operação num conjunto X .

- (i) Diremos que a operação $*$ é **associativa** se para quaisquer $x, y, z \in X$ ocorrer $(x * y) * z = x * (y * z)$, o que na prática significa que pode-se escrever $x * y * z$ em vez de $(x * y) * z$ ou $x * (y * z)$.
- (ii) Diremos que a operação $*$ é **comutativa** se para quaisquer $x, y \in X$ valer a identidade $x * y = y * x$.
- (iii) Diremos que $e \in X$ é **elemento neutro**⁵ da operação $*$ se para qualquer $x \in X$ ocorrer $e * x = x = x * e$. \P

⁵Também chamado de **zero** ou **unidade** a depender do contexto.

Exemplo B.1.20 (Potência). Seja $\star: \omega \times \omega \rightarrow \omega$ a operação dada pela regra $m \star n := m^n$, em que m^n indica a *potência* de m por n . Observe que $(2 \star 1) \star 2 := (2^1 \star 2) := 2^2 = 4$, enquanto $2 \star (1 \star 2) := 2 \star 1^2 := 2^1 = 2$. Também é fácil ver que, em geral, não vale $m \star n = n \star m$. Finalmente, embora $m \star 1 = m$ ocorra para todo $m \in \omega$, não há $n \in \omega$ satisfazendo $e \star m = m$. Portanto, a operação \star em ω *falha* em todos os quesitos da definição anterior. \blacktriangle

Exemplo B.1.21. Para $X := \omega$, a projeção $\pi_1: \omega \times \omega \rightarrow \omega$ na primeira coordenada define uma operação. Nesse contexto, pode-se escrever tanto $(0 \pi_1 1) \pi_1 2$ quanto $0 \pi_1 (1 \pi_1 2)$ e, curiosamente, ambos coincidem: $(0 \pi_1 1) \pi_1 2 = 0 = 0 \pi_1 (1 \pi_1 2)$. O leitor atento certamente notou que a igualdade acima é válida mais geralmente para quaisquer $x, y, z \in \omega$, i.e., a operação é associativa. No entanto, $0 \pi_1 1 \neq 1 \pi_1 0$. \blacktriangle

Exercício B.6. A operação acima tem elemento neutro? \blacksquare

Exemplo B.1.22. Para um conjunto Z fixado, faz-se $X := \wp(Z)$ e $\cap: X \times X \rightarrow X$ a operação dada por $\langle B, C \rangle \mapsto B \cap C$. Como no exemplo anterior, a operação \cap também satisfaçõa $(B \cap C) \cap D = B \cap (C \cap D)$, para quaisquer $B, C, D \in X$. Contudo, diferente do último exemplo, para quaisquer $B, C \in X$ se verifica $B \cap C = C \cap B$. Portanto, a operação \cap em X é associativa e comutativa. Note ainda que Z é *um* elemento neutro da operação. Não poderia haver outro. \blacktriangle

Lema B.1.23. *Uma operação admite, no máximo, um único elemento neutro.*

Demonstração. Se $e, e' \in X$ são ambos elementos neutros da operação, então $e = e * e'$ por e' ser elemento neutro, enquanto $e * e' = e'$ por e ser elemento neutro. \square

Definição B.1.24. As seguintes terminologias são bastante frequentes:

- (i) um **semigrupo** é um magma cuja operação é associativa;
- (ii) um **monoide** é um semigrupo com elemento neutro;
- (iii) um monoide é **comutativo** ou **abeliano** se sua operação for comutativa, terminologia que também costuma ser usada para o caso dos *grupos* (definidos adiante).

Quando se busca algum tipo de precisão linguística sem abandonar a praticidade, escreve-se algo como “ $\langle G, * \rangle$ é *um semigrupo*” a fim de abreviar a tediosa frase “ $*: G \times G \rightarrow G$ é uma operação que faz de G um semigrupo”. Além disso, é lícito escrever $\langle G, *, e \rangle$ para destacar, caso exista, o elemento neutro $e \in G$ da operação. Na maior parte dos casos, porém, diz-se apenas coisas como “ G é um monoide” ou “ G é um grupo abeliano”, pois o tempo é curto e o contexto, quase sempre, é claro.

Exemplo B.1.25 (Simetria). Para um conjunto Z fixado, não é difícil se convencer de que a operação de composição $\circ: Z^Z \times Z^Z \rightarrow Z^Z$, que associa cada par $\langle f, g \rangle$ de funções à sua composta $f \circ g$, torna $\langle Z^Z, \circ, \text{Id}_Z \rangle$ num monoide. Entretanto, em vez de considerar todo o conjunto Z^Z , é legítimo restringir a atenção ao subconjunto $\mathbb{S}(Z) := \{f \in Z^Z : f \text{ é bijeção}\}$, o conjunto de todas as bijeções da forma $Z \rightarrow Z$, que também costumam ser chamadas de **permutações**. Como a composição de bijeções é bijeção, resulta que a restrição da operação “original” $\circ: Z^Z \times Z^Z \rightarrow Z^Z$ ao subconjunto $\mathbb{S}(Z)$ induz uma operação \circ em $\mathbb{S}(Z)$. Daí, como a associatividade já se verifica em Z^Z , segue que $\langle \mathbb{S}(Z), \circ \rangle$ é um semigrupo. Mais ainda, como o elemento neutro de $\langle Z^Z, \circ, \text{Id}_Z \rangle$ é, *coincidentemente*, uma bijeção, conclui-se que $\langle \mathbb{S}(Z), \circ, \text{Id}_Z \rangle$ é um monoide.

No entanto, o monoide $\langle \mathbb{S}(Z), \circ, \text{Id}_Z \rangle$ apresenta um comportamento diferente de todos os monoides anteriores. De fato, dado $f \in \mathbb{S}(Z)$, existe e é única a função $g \in \mathbb{S}(Z)$ tal que $f \circ g = g \circ f = \text{Id}_Z$: a saber, $g = f^{-1}$, a inversa de f , que existe precisamente por f ser uma bijeção. \blacktriangle

Definição B.1.26. Seja $\langle X, *, e \rangle$ um magma com elemento neutro e . Diremos que $y \in X$ é um ***-inverso⁶ à direita** de $x \in X$ se valer $x * y = e$. Analogamente, y será dito um ***-inverso à esquerda** de x se ocorrer $y * x = e$. Se y for simultaneamente $*$ -inverso à direita e à esquerda de x , diremos simplesmente que y é um ***-inverso** de x . \P

Naturalmente, o *prefixo* “ $*$ ” nas definições acima será abandonado sempre que o contexto permitir. Dito isso, se $*$ for uma operação associativa dotada de um elemento neutro e , então um $*$ -inverso de x , caso exista, é único.

Lema B.1.27. Se $\langle X, *, e \rangle$ é um monoide, então cada $x \in X$ admite, no máximo, um $*$ -inverso.

Demonstração. Sejam $y, y' \in X$ $*$ -inversos de x . Então

$$y = y * e = y * (x * y') = (y * x) * y' = e * y' = y',$$

pois y' é $*$ -inverso (à esquerda) de x e y é $*$ -inverso (à direita) de x . \square

Definição B.1.28. Um **grupo** $\langle X, *, e \rangle$ é um monoide em que todo elemento x admite um $*$ -inverso. \P

Exercício B.7. Seja $\mathbb{S}(Z)$ como no Exemplo B.1.25. Convença-se de que $\langle \mathbb{S}(Z), \circ, \text{Id}_Z \rangle$ é um grupo. Quando $\mathbb{S}(Z)$ é abeliano? \blacksquare

Exercício B.8. Dado um grupo G , seja $\text{inv}: G \rightarrow G$ a função que associa cada $g \in G$ ao seu $*$ -inverso. Mostre que inv é uma bijeção. \blacksquare

A inserção de tais tópicos numa seção voltada para relações de equivalência se justifica pelo seguinte problema típico: para que um monoide X *não* seja um grupo, deve existir pelo menos um $x \in X$ que *não* admita um inverso, o que sugere a possibilidade de corrigir tal fenômeno. Embora esse genérico num primeiro momento, é de tal *arquétipo* que surgem as construções dos inteiros e dos racionais:

- (i) o monoide comutativo $\langle \omega, +, 0 \rangle$, com a operação de adição usual, não é um grupo, já que *faltam* os inversos aditivos de todo $n \in \omega \setminus \{0\}$;
- (ii) chamando $\mathbb{N} := \omega \setminus \{0\}$, tem-se $m \cdot n \in \mathbb{N}$ para quaisquer $m, n \in \mathbb{N}$, de modo que $\langle \mathbb{N}, \cdot, 1 \rangle$ é um monoide comutativo que também não é um grupo, já que *faltam* os inversos multiplicativos de todo $n \in \mathbb{N} \setminus \{1\}$.

Observação B.1.29. Ao longo do texto, \mathbb{N} denotará o conjunto dos números naturais diferentes de zero⁷. \triangle

⁶Ou oposto, simétrico, etc. Tudo depende do contexto.

⁷Muita gente não está familiarizada com o uso da letra grega “ ω ” (ômega) para representar o conjunto dos naturais, e muitas outras também não pensam em 0 como natural. Dessa forma, a adoção do símbolo “ \mathbb{N} ” para denotar *apenas* os números naturais não-nulos busca irritar pessoas em ambos os lados desse conflito.

O modo pelo qual os monoides acima serão corrigidos faz uso de uma condição satisfeita por ambos.

Definição B.1.30. Diz-se que um monoide abeliano $\langle X, *, e \rangle$ satisfaz a **lei do cancelamento** se a implicação

$$x * z = y * z \Rightarrow x = y \quad (\text{B.1})$$

for válida para quaisquer $x, y, z \in X$. ¶

Agora, fixado um monoide comutativo $\langle X, *, e \rangle$ satisfazendo (B.1), estipulemos sobre o conjunto $F := X \times X$ a seguinte relação de equivalência \sim :

$$\langle x, y \rangle \sim \langle a, b \rangle \Leftrightarrow x * b = y * a. \quad (\text{B.2})$$

Exercício B.9. Convença-se de que (B.2) é uma relação de equivalência em F . Dica: use a lei do cancelamento para a transitividade. ■

O candidato a grupo será então o conjunto $F/\sim := \left\{ \overline{\langle x, y \rangle} : \langle x, y \rangle \in F \right\}$ das classes de equivalência da relação, que por simplicidade será denotado por G , junto com a operação $+$ sobre G definida da maneira *óvia*:

$$\overline{\langle x, y \rangle} + \overline{\langle x', y' \rangle} := \overline{\langle x * x', y * y' \rangle}.$$

Como apontado na Subseção B.1.2, a definição dada para a operação acima depende da escolha de representantes nas classes de equivalência, de modo que é *preciso* mostrar que tal dependência é apenas aparente. Para isso, tomam-se $\langle a, b \rangle, \langle a', b' \rangle \in F$ tais que $\langle x, y \rangle \sim \langle a, b \rangle$ e $\langle x', y' \rangle \sim \langle a', b' \rangle$, a fim de mostrar que

$$\overline{\langle x * x', y * y' \rangle} = \overline{\langle a * a', b * b' \rangle}. \quad (\text{B.3})$$

Ora, como $x * b = y * a$ e $x' * b' = y' * a'$, resulta que

$$(x * x') * (b * b') = (x * b) * (x' * b') = (y * a) * (y' * a') = (y * y') * (a * a'),$$

acarretando a validade de (B.3).

Assim, $+$ é uma operação legítima em G . Escrevendo $0 := \overline{\langle e, e \rangle}$, é fácil se convencer de que $\langle G, +, 0 \rangle$ é um monoide comutativo. Para concluir que G é um grupo, basta verificar que todo elemento $\overline{\langle x, y \rangle} \in G$ admite um $(+)$ -inverso: observe que

$$\overline{\langle x, y \rangle} + \overline{\langle y, x \rangle} = 0 = \overline{\langle y, x \rangle} + \overline{\langle x, y \rangle},$$

mostrando que o inverso de $\overline{\langle x, y \rangle}$ é $\overline{\langle y, x \rangle}$.

Como o grupo $\langle G, +, 0 \rangle$ foi construído a partir do monoide $\langle X, *, e \rangle$, é natural esperar que se encontre algum tipo de marca deixada por X : neste caso, a marca é uma *cópia* de X dentro de G . Tal cópia é produzida pela injeção $i: X \rightarrow G$ dada por $x \mapsto \overline{\langle x, e \rangle}$: note que se $\langle x, e \rangle = \langle y, e \rangle$, então $x = x * e = e * y = y$. Mas não é só isso: para $x, y \in X$ quaisquer, tem-se

$$i(x) + i(y) = \overline{\langle x, e \rangle} + \overline{\langle y, e \rangle} = \overline{\langle x * y, e \rangle} = i(x * y),$$

mostrando que as cópias (em G) dos elementos de X se operam (em G) da mesma forma que os seus originais se operam (em X). Em particular, $i(e) = 0$ por construção.

Observação B.1.31. Quando *espelhamentos* como esse ocorrem, a injecção $i: X \rightarrow G$ costuma ser pensada como uma *inclusão*. Psicologicamente, é mais útil ignorar a definição formal dos conjuntos envolvidos e assumir como verdadeira a inclusão $X \subseteq G$: formalmente, o que se faz é substituir cada $x \in X$ pelo seu *único* representante $i(x) \in G$. Na prática apaga-se o “ i ” em “ $i(x)$ ”. \triangle

O *mumble jumble*⁸ ainda não acabou: por construção, o inverso de $x \in X$ em G é $-i(x) := \overline{\langle e, x \rangle}$, pois $i(x) + \overline{\langle e, x \rangle} = \overline{\langle x, e \rangle} + \overline{\langle e, x \rangle} = 0$, sugerindo que se escreva $-x$ em vez de $-i(x)$. Finalmente, dado um elemento $\langle x, y \rangle$ de G , ocorre

$$\overline{\langle x, y \rangle} = \overline{\langle x, e \rangle} + \overline{\langle e, y \rangle} := i(x) + (-i(y)) = x + (-y),$$

e o último costuma ser escrito simplesmente como $x - y$.

Definição B.1.32. O grupo $\langle G, +, 0 \rangle$ construído acima a partir do monoide comutativo $\langle X, *, e \rangle$ é usualmente chamado de **grupo de Grothendieck** (associado ao monoide). ¶

Em certo sentido, $\langle G, +, 0 \rangle$ é o *menor* grupo que contém $\langle X, *, e \rangle$: isso é relativamente razoável no caso da construção apresentada, em que o monoide satisfaz a lei do cancelamento, mas deixa de ser verdadeiro para monoides mais gerais⁹.

Definição B.1.33. O grupo de Grothendieck associado ao monoide $\langle \omega, +, 0 \rangle$ costuma ser denotado por \mathbb{Z} e xingado de **conjunto dos números inteiros**. Sua operação também é chamada de adição (ou soma) e é representada pelo mesmo símbolo $+$. ¶

Não costuma ser inteiramente óbvio que a definição acima realmente *modela* aquilo que estamos acostumados a xingar de \mathbb{Z} , razão pela qual convém alongar a discussão. Para evitar confusões nesse processo, vamos escrever G em vez de \mathbb{Z} e $i(n) \in G$ em vez de $n \in \mathbb{Z}$.

- (i) Pelo que se observou no caso geral, $i[\omega]$ é uma *cópia* de ω em G , de modo que para cada $n \in \omega$ existe $-i(n) \in G$ tal que $i(n) + (-i(n)) = 0$.
- (ii) Além disso, os elementos de G são da forma $i(n) \in i[\omega]$ ou da forma $-i(n) \in G$, para $n \in \omega$, com $-i(n) \notin i[\omega]$ a menos que se tenha $n = 0$.
- ✓ Dados $m, n \in \omega$, uma igualdade do tipo $i(n) = -i(m)$ acarreta $i(n+m) = i(0)$ e, como i é injetora, tem-se $n+m = 0$, donde resulta $n = m = 0$.
- ✓ Os elementos de G são da forma $\overline{\langle m, n \rangle} := i(m) - i(n)$, com $m, n \in \omega$. Note que se valer $m \geq n$, então existe um único $k \in \omega$ satisfazendo $n+k = m$ (por conta das noções assumidas de aritmética). Escrevendo $k := m-n$, segue que

$$\overline{\langle m, n \rangle} = \overline{\langle m-n, 0 \rangle},$$

precisamente por se ter $m+0 = (m-n)+n$ (essa última soma realizada em ω). Se, por outro lado, valer $m < n$, então existe um único $k \in \omega$ tal que $n = m+k$, escrito como $k := n-m$, donde se obtém $\overline{\langle m, n \rangle} = \overline{\langle 0, n-m \rangle}$. Em suma, mostrou-se que

$$\begin{aligned} m \geq n &\Rightarrow i(m) - i(n) = i(m-n) \\ m < n &\Rightarrow i(m) - i(n) = -i(n-m). \end{aligned}$$

⁸Gambiarra, em inglês.

⁹Na verdade, a mesma construção pode ser realizada a partir de qualquer monoide *comutativo*, desde que se altere a definição da relação de equivalência (B.2) sobre $X \times X$. Porém, em tais casos não há garantias de que a função $i: X \rightarrow G$ seja realmente uma injecção.

- (iii) Observe ainda que uma igualdade do tipo $-i(m) = -i(n)$ não ocorre para $m \neq n$: a função $-i: \omega \rightarrow G$ que faz $n \mapsto -i(n)$ é injecção por ser composição da injecção $i: \omega \rightarrow G$ com a bijeção¹⁰ $\text{inv}: G \rightarrow G$ que faz $\text{inv}(g) := -g$ para todo $g \in G$.

Temos em mãos, portanto, o adorado conjunto dos inteiros, munido de sua adição usual, onde a nostalgia escolar reina: valem as inclusões $\mathbb{N} \subsetneq \mathbb{Z}$ e $\omega \subsetneq \mathbb{Z}$; o complementar $\mathbb{Z} \setminus \omega$ tem como elementos todos os números *negativos*; se $m, n \in \omega$ são tais que $m \geq n$, então $m + (-n) = m - n \in \omega$ e, se $m < n$, então $m + (-n) = -(n - m)$.

Exercício B.10. Imita a discussão anterior a fim de mostrar que o grupo de Grothendieck associado ao monoide $\langle \mathbb{N}, \cdot, 1 \rangle$ se comporta como o *conjunto dos números racionais maiores do que zero*. ■

Um modo mais conciso de obter *todos* os racionais consiste em, antes de qualquer outra coisa, elevar \mathbb{Z} ao patamar de *anel*.

Definição B.1.34. Um **anel** consiste de um conjunto $A \neq \emptyset$ munido de duas operações, $+$ e \cdot , e elementos $0, 1 \in A$, onde

- (i) $\langle A, +, 0 \rangle$ é um grupo abeliano, cuja operação $+$ é chamada de *adição*,
- (ii) $\langle A, \cdot, 1 \rangle$ é um monoide, cuja operação é chamada de *multiplicação*, e
- (iii) as operações $+$ e \cdot *comutam* (ou se *distribuem*) entre si, i.e., para quaisquer $a, b, c \in A$ valem as identidades $a \cdot (b + c) = (a \cdot b) + (a \cdot c)$ e $(a + b) \cdot c = (a \cdot c) + (b \cdot c)$. ¶

É comum denotar o anel A como a *estrutura algébrica* $\langle A; +, \cdot, 0, 1 \rangle$. No entanto, quase sempre as menções são feitas *ao anel* A , com as operações subentendidas no contexto. Além disso, como de costume, o símbolo “ \cdot ” será omitido, pois a vida é curta.

Observação B.1.35. A rigor, os anéis definidos acima deveriam ser chamados de **anéis com unidade**, dado que existem contextos que não exigem a existência de elemento neutro multiplicativo. Dito isso, a maioria dos anéis tratados no texto goza de mais uma propriedade: diz-se que o anel A é **comutativo** se a multiplicação $\cdot: A \times A \rightarrow A$ for comutativa. Por esta razão, daqui em diante: *anel será sinônimo de anel comutativo com unidade, a menos de (raras) exceções, que serão destacadas no texto.* △

Exemplo B.1.36 (O *anel* dos inteiros). Embora a construção anterior tenha nos dado o grupo aditivo \mathbb{Z} , a operação de multiplicação, que a princípio apresenta a propriedade de distributividade com a adição, só está definida sobre ω . Para corrigir isso, deve-se estender a multiplicação de números naturais para os inteiros. Um modo desleixado (mas funcional) de fazer isso consiste em definir

$$(m - n)(a - b) := (ma + nb) - (mb + na), \quad (\text{B.4})$$

o que faz algum sentido pois os elementos de \mathbb{Z} são da forma $m - n$. Todavia, mais uma vez, é preciso lidar com o problema da boa definição apontado na Subseção B.1.2. Sejam então $m', n', a', b' \in \omega$ tais que $m - n = m' - n'$ e $a - b = a' - b'$, igualdades em \mathbb{Z} que equivalem às seguintes igualdades em ω :

$$m + n' = m' + n \quad (\text{B.5})$$

$$a + b' = a' + b. \quad (\text{B.6})$$

¹⁰Exercício B.8.

Da primeira igualdade, tem-se $ma + n'a = m'a + na$ e $mb + n'b = m'b + nb$, acarretando

$$\underbrace{(ma + nb) + (m'b + n'a)}_{\text{em } \omega} = \underbrace{(mb + na) + (m'a + n'b)}_{\text{em } \omega} \Rightarrow$$

$$\Rightarrow \underbrace{(ma + nb) - (mb + na)}_{\text{em } \mathbb{Z}} = \underbrace{(m'a + n'b) - (m'b + n'a)},$$

mostrando $(m - n)(a - b) = (m' - n')(a - b)$.

Analogamente, por meio da identidade (B.6), verifica-se em \mathbb{Z} a identidade

$$(m' - n')(a - b) = (m' - n')(a' - b'),$$

onde segue que a regra (B.4) dá origem a uma operação bem definida em \mathbb{Z} .

Observa-se então que para $a_1, a_2, a_3 \in \mathbb{Z}$ e $m_i, n_i \in \omega$ tais que $a_i = m_i - n_i$ em \mathbb{Z} , ocorre

$$\begin{aligned} a_1(a_2 + a_3) &= (m_1 - n_1)(m_2 - n_2 + m_3 - n_3) = (m_1 - n_1)((m_2 + m_3) - (n_2 + n_3)) = \\ &= (m_1(m_2 + m_3) + n_1(n_2 + n_3)) - (m_1(n_2 + n_3) + n_1(m_2 + m_3)) = \\ &= (m_1m_2 + n_1n_2) - (m_1n_2 + n_1m_2) + (m_1m_3 + n_1n_3) - (m_1n_3 + n_1m_3) = \\ &= a_1a_2 + a_1a_3, \end{aligned}$$

como desejado. O trabalho de verificar as demais propriedades de anel será deixado a cargo do leitor. ▲

Exercício B.11. Sobre $Q := \mathbb{Z} \times (\mathbb{Z} \setminus \{0\})$, considere a relação \sim que declara $\langle a, b \rangle \sim \langle c, d \rangle$ se, e somente se, $ad = bc$.

- a) Mostre que \sim é uma relação de equivalência.
- b) Chamando por $\frac{a}{b}$ a classe de equivalência de um par $\langle a, b \rangle$, mostre que Q/\sim se comporta, essencialmente, como o **anel dos números racionais** que conhecemos na escola, e que passará a ser denotado por \mathbb{Q} . Em particular, defina as operações de adição e multiplicação e verifique as propriedades de anel. ■

B.2 Um mínimo sobre ordens

Enquanto relações de equivalência permitem tratar como iguais coisas que são diferentes, *relações de ordem* se preocupam com a comparação entre objetos diferentes.

Definição B.2.1. Uma relação binária \preceq num conjunto X é dita uma **relação de ordem (parcial)** se \preceq for *reflexiva*, *transitiva* e **antissimétrica**, onde a última condição significa que a ocorrência simultânea de $x \preceq y$ e $y \preceq x$ acarreta $x = y$, para quaisquer $x, y \in X$. Escreve-se $\langle X, \preceq \rangle$ para indicar o conjunto X dotado da ordem \preceq , que se diz (estar) **(parcialmente) ordenado** pela ordem (parcial) \preceq . ¶

Alternativamente, diz-se que \prec é uma **relação de ordem estrita** em X se \prec for transitiva mas, em vez de reflexiva e antissimétrica, for

- (i) **irreflexiva**, i.e., se para todo $x \in X$ ocorrer $x \not\prec x$, e

(ii) **assimétrica**, i.e., se para quaisquer $x, y \in X$, a ocorrência de $x \prec y$ acarretar $y \not\prec x$.

Contudo, a diferença entre (relações de) *ordens parciais* e *estritas* é meramente virtual, no seguinte sentido:

- se $\langle \mathbb{S}, \prec \rangle$ é uma ordem estrita, então a relação \preceq definida por

$$x \preceq y \Leftrightarrow (x \neq y \text{ e } x \prec y) \text{ ou } x = y$$

é uma relação de ordem parcial em \mathbb{S} ;

- se $\langle \mathbb{P}, \sqsubseteq \rangle$ é uma ordem parcial, então a relação \sqsubset definida por

$$x \sqsubset y \Leftrightarrow x \neq y \text{ e } x \sqsubseteq y$$

é uma relação de ordem estrita em \mathbb{P} .

É claro que ao aplicar o primeiro procedimento à ordem estrita \sqsubset , retorna-se à ordem parcial original \sqsubseteq , enquanto o segundo procedimento aplicado à ordem parcial \preceq resulta na ordem estrita original \prec . Assim, tem-se o direito de chamar tanto $\langle \mathbb{S}, \prec \rangle$ quanto $\langle \mathbb{P}, \sqsubseteq \rangle$ de *ordens*. Em tais situações, ficam implicitamente definidas a ordem parcial \preceq e a ordem estrita \sqsubset induzidas por \prec e \sqsubseteq , respectivamente.

Observação B.2.2 (Leitura e Inversa). Em geral, costuma-se ler uma expressão do tipo “ $x \preceq y$ ” como “ x é menor do que ou igual a y ”, enquanto “ $x \prec y$ ” é lida como “ x é (estritamente) menor do que y ” – a menos que o contexto sugira uma terminologia própria para os símbolos.

Alternativamente, lê-se “ $x \preceq y$ ” como “ y é maior do que ou igual a x ”, o que esconde um fato que será importante: escrevendo $a \succeq b$ para indicar que $b \preceq a$, segue que \succeq também é uma relação de ordem parcial sobre o conjunto em questão: explicitamente, \succeq é apenas a relação inversa de \preceq . \triangle

No que segue, fixam-se uma ordem $\langle \mathbb{P}, \leq \rangle$, um subconjunto A de \mathbb{P} e elementos $a \in A$ e $p \in \mathbb{P}$. A cada *conceito* a ser definido na ordem $\langle \mathbb{P}, \leq \rangle$ a seguir, corresponderá um conceito *dual*, ou *co-conceito*, que consiste em reescrever o conceito original na ordem inversa $\langle \mathbb{P}, \geq \rangle$. Na prática, substituem-se as ocorrências dos símbolos $<$ e \leq por $>$ e \geq , respectivamente. O leitor provavelmente já conhece algumas das definições na próxima tabela.

Conceito	Co-conceito
$a \in A$ é elemento minimal de A se não existe $x \in A$ com $x < a$	$a \in A$ é elemento maximal de A se não existe $x \in A$ com $a < x$
a é um menor elemento (ou mínimo) de A se $a \leq x$ ocorrer para todo $x \in A$	a é um maior elemento (ou máximo) de A se $x \leq a$ ocorrer para todo $x \in A$
p é um limitante inferior de A se $p \leq x$ para todo $x \in A$	p é um limitante superior de A se $x \leq p$ para todo $x \in A$
p é um ínfimo de A se p for um maior elemento do conjunto dos limitantes inferiores de A	p é um supremo de A se p for um menor elemento do conjunto dos limitantes superiores de A
A é limitado se A é limitado inferiormente e superiormente	

Na tabela acima, escreve-se *um mínimo* (e *um máximo*) por puro preciosismo: se $a, a' \in A$ são mínimos de A , então ocorre $a \leq a'$ e $a' \leq a$, donde a antissimetria de \leq acarreta $a = a'$. Como um máximo de A em $\langle \mathbb{P}, \leq \rangle$ é um mínimo de A em $\langle \mathbb{P}, \geq \rangle$, segue que máximos (quando existem) também são únicos.

Exercício B.12. Convença-se das afirmações acima. ■

Exemplo B.2.3 (Dualidade). O argumento acima segue um arquétipo frequente em *teoria da ordem*, chamado de *princípio da dualidade*: ao se provar um determinado resultado sobre ordens válido em geral, a versão dual do resultado também deverá ser *verdadeira*, posto que $\langle \mathbb{P}, \geq \rangle$ também é uma ordem parcial. Isso ficará mais claro com o passar do tempo, *não se preocupe*. ▲

Como *ínfimos* e *supremos* são definidos como certos máximos e mínimos, respectivamente, segue que eles também são únicos. A unicidade de tais elementos permite adotar notações mais práticas para designá-los.

Definição B.2.4. Sejam $\langle \mathbb{P}, \leq \rangle$ uma ordem parcial e $A \subseteq \mathbb{P}$ um subconjunto. Adotaremos as seguintes notações:

- (i) o menor elemento de A (caso exista) é denotado $\min_{a \in A} a$, $\min_{\leq} A$ ou apenas $\min A$;
- (ii) o maior elemento de A (caso exista) é denotado por $\max_{a \in A} a$, $\max_{\leq} A$ ou apenas $\max A$;
- (iii) o ínfimo de A (caso exista) é denotado por $\inf_{a \in A} a$, $\inf_{\leq} A$ ou apenas $\inf A$;
- (iv) o supremo de A (caso exista) é denotado por $\sup_{a \in A} a$, $\sup_{\leq} A$ ou apenas $\sup A$. ¶

Pode ser útil frisar que *max* e *min* abreviam *maximum* e *minimum*, os termos usuais na literatura estrangeira para se referir ao maior e menor elemento, respectivamente. Dito isso, não faria sentido associar notações para elementos maximais ou minimais pois, em geral, estes não são únicos.

Exemplo B.2.5. Dado um conjunto X , segue que $\langle \wp(X), \subseteq \rangle$ é uma ordem parcial, já que a *relação* de inclusão é reflexiva, antissimétrica e transitiva. Agora, se $\mathcal{C} \neq \emptyset$ é uma família de subconjuntos de X , então $\bigcap \mathcal{C}$ é o ínfimo de \mathcal{C} na ordem $\langle \wp(X), \subseteq \rangle$, pois

- ✓ (é limitante inferior) $\bigcap \mathcal{C} \subseteq C$ para todo $C \in \mathcal{C}$, e
- ✓ (é o maior limitante inferior) se $A \in \wp(X)$ é tal que $A \subseteq C$ para todo $C \in \mathcal{C}$, então $A \subseteq \bigcap \mathcal{C}$.

Analogamente mostra-se que $\bigcup \mathcal{C}$ é o supremo de \mathcal{C} na ordem $\langle \wp(X), \subseteq \rangle$. Em particular, ao se considerar $\mathcal{A} := \{A \in \wp(X) : A \neq \emptyset \text{ e } A \neq X\}$ com a *ordem induzida* de $\wp(X)$, verifica-se sem grandes dificuldades que $\{x\}$ é um elemento minimal de \mathcal{A} para cada $x \in X$, que não é mínimo quando $X \neq \{x\}$. Nas mesmas situações, $X \setminus \{x\}$ é um elemento maximal de \mathcal{A} que não é máximo. ▲

Observação B.2.6. Obviamente, se $\min A$ existe, então $\min A = \inf A$. Em certo sentido, o que difere um ínfimo de A de um mínimo de A é a pertinência ao conjunto. Precisamente, a é um ínfimo de A se a for o *melhor* limitante inferior de A , no seguinte sentido:

- (i) além de valer $a \leq x$ para todo $x \in A$, vale $b \leq a$ para todo b que limita A inferiormente, i.e., a é o maior elemento do conjunto dos limitantes inferiores de A .

Em particular, a condição (i) quase acarreta o seguinte:

- (ii) além de valer $a \leq x$ para todo $x \in A$, nenhum $b > a$ limita A inferiormente, i.e., se $b > a$, então existe $x \in A$ com $x < b$.

De fato, se algum $b > a$ limitasse A inferiormente, então a não seria o maior limitante inferior de A . Isso garante apenas um $x \in A$ com $x \not\leq b$. \triangle

A implicação $(i) \Rightarrow (ii)$ anterior, bem como sua recíproca, se verificam (verdadeiramente) em *ordens totais*.

Definição B.2.7. Uma ordem $\langle X, < \rangle$ é **total** se para quaisquer $x, y \in X$ ocorrer somente um dos três casos a seguir: $x = y$, $x < y$ ou $y < x$. Evidentemente, se a ordem de X for parcial, basta dizer que para quaisquer $x, y \in X$ ocorre $x \leq y$ ou $y \leq x$. \P

Agora, a implicação $(i) \Rightarrow (ii)$ é automática. Já para a sua recíproca, se $\langle X, \leq \rangle$ é uma ordem total que satisfaz a condição (ii) , então a limita A inferiormente e nenhum $b > a$ faz o mesmo, donde segue que todo limitante inferior de A deve ser menor do que ou igual a a , posto que a ordem é total.

Exercício B.13. Seja $\langle \mathbb{P}, \leq \rangle$ uma ordem total. Mostre que a é elemento minimal de A se, e somente se, $a = \min A$. Enuncie e demonstre a versão *dual* para elementos maximais. \blacksquare

Exemplo B.2.8 (A totalidade importa). A coleção \mathcal{A} discutida no Exemplo B.2.5 apresenta elementos minimais e maximais que não são nem mínimos e nem máximos. Evidentemente, $\langle \mathcal{A}, \subseteq \rangle$ não pode ser uma ordem total. \blacktriangle

Definição B.2.9. Uma ordem parcial \leq (ou $<$) sobre um conjunto \mathbb{P} é chamada de **boa ordem** se todo subconjunto não-vazio de \mathbb{P} admite menor elemento. \P

Moralmente, um conjunto está bem ordenado quando seus elementos podem ser *enfileirados*: $p_0 := \min \mathbb{P}$ é o primeiro da fila, $p_1 := \min(\mathbb{P} \setminus \{p_0\})$ é o segundo, etc.

Exemplo B.2.10. A construção dos números naturais realizada no capítulo anterior os “distribuiu” de modo a valer

$$0 \in 1 \in 2 \in 3 \in \dots \in n \in n_+ \in \dots$$

Assim, é *natural* imaginar que a relação de pertinência em ω estabeleça algum tipo de ordem razoável. Isso de fato ocorre. \blacktriangle

Lema B.2.11. $\langle \omega, \in \rangle$ é uma ordem estrita.

Demonstração. Pelos Exercícios A.24 e A.25, segue que \in é uma relação irreflexiva e assimétrica. Por indução, não é difícil perceber que \in também é transitiva. De fato, para $m, n, o \in \omega$ com $m \in n$ e $n \in o$, mostraremos, por indução em o , que deve ocorrer $m \in o$.

- ✓ Se $o := 0$, então o resultado vale por vacuidade (lembre-se: $0 := \emptyset$).
- ✓ Supondo o resultado verdadeiro para algum $o \in \omega$, deve-se verificar a validade do resultado para o_+ : ora, se $m \in n$ e $n \in o_+ := o \cup \{o\}$, então $n \in o$ ou $n = o$; se ocorrer o primeiro caso, então a hipótese de indução garante que $m \in o$; se valer o segundo caso, então é claro que $m \in o$. \square

Na verdade, a ordem em ω induzida pela relação de pertinência é uma *boa ordem*. A fim de provar isso, convém demonstrar outra forma do Princípio de Indução. A partir daqui, $m < n$ será usado como sinônimo de $m \in n$.

Exercício B.14. Sejam $k, n \in \omega$. Mostre que $k < n_+$ se, e somente se, $k < n$ ou $k = n$. Dica: qual a definição de sucessor? ■

Teorema B.2.12 (Princípio da Indução, 2^a forma). *Seja $\mathcal{P}(x)$ uma fórmula tal que para todo $n \in \omega$ se verifique a implicação*

$$\text{se } \mathcal{P}(k) \text{ vale para todo } k < n, \text{ então vale } \mathcal{P}(n).$$

Então para todo $n \in \omega$ vale $\mathcal{P}(n)$.

Demonstração. Consideremos a fórmula $\mathcal{Q}(x)$ dada por “ $\mathcal{P}(k)$ vale para todo $k < x$ ”. Como não existe $n < 0$, segue que $\mathcal{Q}(0)$ é verdadeira por vacuidade. Supondo $\mathcal{Q}(n)$ para algum $n \in \omega$, deve-se ter $\mathcal{Q}(n_+)$: de fato, se $\mathcal{Q}(n)$ é verdadeira, então $\mathcal{Q}(k)$ é verdadeira para todo $k < n$; daí, a hipótese sobre \mathcal{P} garante que $\mathcal{P}(n)$ também é verdadeira, donde o exercício anterior assegura a validade de $\mathcal{Q}(n_+)$. Por indução, segue que $\mathcal{Q}(n)$ é válida para todo $n \in \omega$. Agora, dado $m \in \omega$, a validade de $\mathcal{Q}(m_+)$ acarreta, em particular, que $\mathcal{P}(m)$ deve ser verdadeira, mostrando o resultado. □

Corolário B.2.13. $\langle \omega, \in \rangle$ é uma boa ordem.

Demonstração. O último lema já atesta que \in é uma relação de ordem estrita. Agora, fixado $S \subseteq \omega$ com $S \neq \emptyset$, mostraremos que existe $s \in S$ com $s = \min S$. Procederemos pela contrapositiva, i.e., mostraremos que se S não tem menor elemento, então $S = \emptyset$. Ora, se não existisse menor elemento, então para $n \in \omega$ fixado, a ocorrência de $k \in \omega \setminus S$ para todo $k < n$ acarretaria $n \in \omega \setminus S$, já que o contrário implicaria em $n = \min S$. Portanto, pela Segunda Forma do Princípio da Indução, resulta que $n \in \omega \setminus S$ para todo $n \in \omega$, i.e., $S = \emptyset$, como desejado. □

Exemplo B.2.14. A boa ordem nativa de ω se estende a uma ordem total em \mathbb{Z} : declararam-se os elementos da forma $-n$, para $n \in \mathbb{N} := \omega \setminus \{0\}$, como estritamente menores do que 0, e

$$-m \leq -n \Leftrightarrow m \geq n$$

para quaisquer $m, n \in \omega$. Não é difícil se convencer de que com tal relação, $\langle \mathbb{Z}, \leq \rangle$ é uma ordem total. Contudo, tal extensão não é uma boa ordem: o próprio conjunto \mathbb{Z} não tem menor elemento! ▲

Embora seja comum pensar em argumentos indutivos no contexto restrito dos números naturais, eles dependem essencialmente da noção de boa ordenação:

Proposição B.2.15 (Indução em boas ordens). *Seja $\langle \mathbb{W}, \leq \rangle$ uma boa ordem e suponha que $\mathcal{P}(x)$ seja uma fórmula na variável x tal que para todo $w \in \mathbb{W}$ seja possível verificar*

$$(\forall v \in \mathbb{W} \quad v < w \Rightarrow \mathcal{P}(v)) \Rightarrow \mathcal{P}(w). \tag{B.7}$$

Então $\mathcal{P}(w)$ é verdadeira para todo $w \in \mathbb{W}$.

Demonstração. Por fatores psicológicos, note que se $\mathcal{P}(x)$ satisfaz (B.7) para cada elemento $w \in \mathbb{W}$, então isso é válido, em particular, para $w_0 := \min \mathbb{W}$. Agora, como não existe $v \in \mathbb{W}$ com $v < w_0$, segue por vacuidade que $\mathcal{P}(v)$ é verdadeira para todo $v < w_0$ e, consequentemente, tem-se $\mathcal{P}(w_0)$.

De modo geral, se existisse $\tilde{w} \in \mathbb{W}$ com $\neg \mathcal{P}(\tilde{w})$, então o conjunto das *testemunhas* $T := \{w \in \mathbb{W} : \neg \mathcal{P}(w)\}$ seria não-vazio e, por conseguinte, seria possível tomar $w_1 = \min T$. Por construção, teria-se

$$\forall v \in \mathbb{W} \quad v < w_1 \Rightarrow \mathcal{P}(v),$$

com $\neg \mathcal{P}(w_1)$, mostrando que $\mathcal{P}(x)$ não satisfaz (B.7). □

A proposição acima seria irrelevante se, por exemplo, as únicas boas ordenações existentes fossem subconjuntos de ω . Por sorte (ou azar?), não é o caso.

Exercício B.15. Mostre que $\omega_+ := \omega \cup \{\omega\}$ é bem ordenado pela relação \in . ■

Moralmente, o elemento $\omega \in \omega_+$ funciona como um *ponto no infinito*, i.e., um elemento *artificial* que fica *acima* de todos os números naturais¹¹. O leitor com a impressão de que esse processo poderia ser *iterado recursivamente* a fim de obter *ordens* maiores se alegrará com o próximo capítulo. Antes disso, porém, convém abordar a noção de *recursão*.

B.3 Adiável: recursão

Observação B.3.1. Embora a presente seção seja adiável para leitores interessados em tópicos superficiais, ela é imprescindível para o Capítulos D, especificamente na demonstração das equivalências entre o Axioma da Escolha e os Teoremas de Zermelo e Zorn (*a.k.a.*¹² Lema de Zorn). △

Grosseiramente, o princípio da indução descreve um efeito dominó formal: dada uma coleção de dominós enfileirados, se o primeiro dominó cai e *sabe-se* que todo dominó cai desde que o dominó anterior também caia, então todos os dominós caem. Embora tal metáfora tenha a (des) vantagem de não levar o tempo em consideração, ela tem o mérito de reforçar a ideia de que a indução depende da ordenação, e não dos *números*. Porém, para a presente discussão, interessa observar outro problema que talvez tenha passado incógnito aos olhos do leitor: como *enfileirar* os dominós? Ora, conhecemos o procedimento para colocar um dominó de pé e, mais geralmente, dada uma fileira de dominós, sabemos como colocar um novo dominó no final da fila. O processo completo de enfileiramento dos dominós é então feito ao se *iterar* o procedimento, *recursivamente*.

Assim, uma (não tão) importante distinção terminológica é a seguinte:

- *indução* se refere a uma técnica de demonstração – uma vez enfileirados os dominós, prova-se que eles caem se certas hipóteses forem satisfeitas;
- *recursão* é uma técnica de “construção” – dada uma coleção de dominós, descreve-se como enfileirá-los.

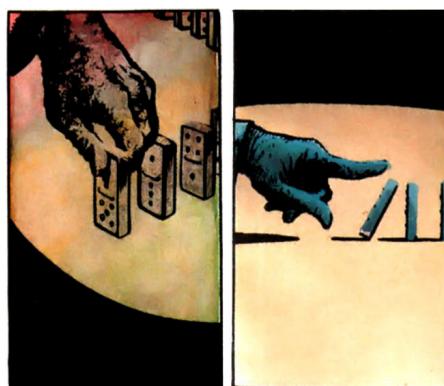


Figura B.2: Recursão vs. indução (Moore, Lloyd e Weare, 1988).

¹¹O que não é artificial?

¹²Abreviação de “*also known as*”, que por sua vez significa “também conhecido como”.

Já vimos que na presença de boas ordens somos capazes de argumentar por indução. Vejamos agora, precisamente, como *construir* por recursão. Para isso, suponha que $\langle \mathbb{W}, \leq \rangle$ seja uma boa ordem e $\mathcal{F}(x, y)$ seja uma fórmula funcional em x (lembre-se: isto significa que para cada conjunto A existe um único conjunto B com $\mathcal{F}(A, B)$ verdadeira, o que permite escrever $B = \mathcal{F}(A)$). A fórmula $\mathcal{F}(x, y)$ é o *procedimento* que será usado na definição *recursiva* de uma função f cujo domínio é \mathbb{W} . Faz-se o seguinte:

- para $w_0 := \min \mathbb{W}$, define-se $f(w_0) := \mathcal{F}(\emptyset)$;
- para $w_1 := \min (\mathbb{W} \setminus \{w_0\})$, define-se $f(w_1) := \mathcal{F}(\langle f(w_0) \rangle)$;
- para $w_2 := \min (\mathbb{W} \setminus \{w_0, w_1\})$, define-se $f(w_2) := \mathcal{F}(\langle f(w_0), f(w_1) \rangle) \dots$

Ou seja: ao menor elemento de \mathbb{W} , chamado de w_0 , associa-se $\mathcal{F}(\emptyset)$, pois não há *histórico* sobre o qual se possa aplicar o *procedimento*; no passo seguinte, toma-se o *sucessor* de w_0 em \mathbb{W} , chamado de w_1 , precisamente o menor dos elementos de \mathbb{W} maiores do que w_0 , e a ele se associa $\mathcal{F}(\langle f(w_0) \rangle)$, onde $\langle f(w_0) \rangle$ é o *histórico* da função f obtido até o passo anterior¹³; em seguida, ao sucessor de w_1 , chamado de w_2 , associa-se $\mathcal{F}(\langle f(w_0), f(w_1) \rangle)$. Mais geralmente, a função f que se busca definir deve ser dada pela regra

$$f(w) := \mathcal{F}(\langle f(v) : v < w \rangle), \quad (\text{B.8})$$

para cada $w \in \mathbb{W}$.

Como a fórmula $\mathcal{F}(x, y)$ é funcional, deveria ser verdade que a regra acima define uma única função f com domínio \mathbb{W} . Isso de fato ocorre, mas não é completamente óbvio: em (B.8), usa-se f (no lado direito) para definir f (no lado esquerdo), mas ainda nem sabemos que f existe¹⁴! Intuitivamente, do lado direito a letra f denota uma definição *parcial* de f que se usa para definir a função f no passo w . Ainda parece abstrato demais?

Exemplo B.3.2 (Fatorial). Supondo conhecida a função $\cdot : \omega \times \omega \rightarrow \omega$ que a cada par $\langle m, n \rangle \in \omega \times \omega$ associa o produto (multiplicação) $m \cdot n \in \omega$, consideremos o *problema* de definir a função $f : \omega \rightarrow \omega$ que a cada $n \in \omega$ associa o famoso *fatorial* de n , denotado por $n!$: a ideia é que $0! := 1$, $1! := 1$, $2! := 2$, $3! := 3 \cdot 2!$, $4! := 4 \cdot 3!$ e *assim por diante*.

A rigor, ainda não há garantias de que esse tipo de procedimento produza uma função $f : \omega \rightarrow \omega$, já que neste caso a definição exigiria que f fosse uma relação binária $f \subseteq \omega \times \omega$, com $\text{dom}(f) = \omega$: na melhor das hipóteses, o “assim por diante” apenas define uma relação binária $f_{n+1} \subseteq \{0, 1, \dots, n\} \times \omega$ sempre que f_n já for conhecida. Ora, como definimos f_0 , pode-se definir f_1 , o que permite definir f_2 , o que permite definir f_3 ; por *indução*, mostra-se que existe f_n para todo n , de modo que basta tomar f como a reunião das f_n ’s. Justificar melhor esse tipo de procedimento é o que será feito a seguir. ▲

Definição B.3.3. Sejam $\mathcal{F}(x, y)$ uma fórmula funcional em x e $\langle \mathbb{W}, \leq \rangle$ uma boa ordem. Diremos que uma função f é **\mathcal{F} -recursiva** em \mathbb{W} se $\text{dom}(f) = \mathbb{W}$ e a condição (B.8) se verifica para cada $w \in \mathbb{W}$. ¶

Lema B.3.4. Nas condições acima, existe no máximo uma função \mathcal{F} -recursiva em \mathbb{W} .

¹³A rigor, é a função cujo domínio é $\{0\}$ e cuja imagem é o conjunto $\{f(w_0)\}$.

¹⁴Antes de aprender recursão é preciso saber recursão.

Demonstração. Supondo que f e g sejam funções distintas e \mathcal{F} -recursivas em \mathbb{W} , seria possível tomar o menor $\tilde{w} \in \mathbb{W}$ com $f(\tilde{w}) \neq g(\tilde{w})$. Daí, teria-se $f(v) = g(v)$ para todo $v < \tilde{w}$, resultando em $\langle f(v) : v < \tilde{w} \rangle = \langle g(v) : v < \tilde{w} \rangle$ e, consequentemente,

$$f(\tilde{w}) = \mathcal{F}(\langle f(v) : v < \tilde{w} \rangle) = \mathcal{F}(\langle g(v) : v < \tilde{w} \rangle) = g(\tilde{w}),$$

uma contradição. Logo, se existir uma função \mathcal{F} -recursiva em \mathbb{W} , ela é única. \square

Teorema B.3.5 (Recursão para boas ordens). *Sejam $\langle \mathbb{W}, \leq \rangle$ uma boa ordem e $\mathcal{F}(x, y)$ uma fórmula funcional em x . Então existe uma única função \mathcal{F} -recursiva em \mathbb{W} .*

Demonstração. A unicidade segue do lema anterior. Para mostrar a existência, em vez de \mathbb{W} considera-se $Y := \mathbb{W}_\#$ como no vindouro Exercício C.10, i.e. toma-se y com $y \notin \mathbb{W}$ e estende-se a boa ordem de \mathbb{W} para $Y := \mathbb{W} \cup \{y\}$ declarando-se $w < y$ para todo $w \in \mathbb{W}$. A ideia é mostrar que para cada $v \in Y$ existe uma (única) função \mathcal{F} -recursiva em $D_v := \{u \in Y : u < v\}$, digamos f_v . Note que se isso for feito, então a função \mathcal{F} -recursiva em \mathbb{W} será, tão somente, a função \mathcal{F} -recursiva em $D_y := \{u \in Y : u < y\} = \mathbb{W}$. A estratégia é simples: provar por indução, já que Y é bem ordenado.

Para $v \in Y$, assumamos que para cada $u < v$ exista uma função f_u , \mathcal{F} -recursiva em D_u , que deve ser única em virtude do lema anterior. Com tais informações, mostraremos que existe uma função \mathcal{F} -recursiva em D_v . Há dois casos para considerar.

(i) *Existe $u \in Y$ com $v = \min\{x \in Y : u < x\}$.*

Neste caso, verifica-se a igualdade $\{w \in Y : w < v\} = \{w \in Y : w \leq u\}$, de modo que f_v se obtém com $f_v(t) := f_u(t)$ se $t < u$, e $f_v(u) := \mathcal{F}(\langle f_u(t) : t < u \rangle)$, que obviamente satisfaz as condições impostas.

(ii) *Não existe $u \in Y$ tal que $v = \min\{x \in Y : u < x\}$.*

Neste caso tem-se $\{w \in Y : w < v\} = \bigcup_{u < v} \{t \in Y : t < u\}$, e daí para cada elemento $t < v$ define-se $f_v(t) := f_u(t)$ para qualquer $u \in Y$ com $t < u < v$, o qual existe pela hipótese sobre v . Note que a função f_v está *bem definida*: se $u, u' < v$ são elementos distintos tais que $t < u$ e $t < u'$, então $u < u'$ ou $u' < u$; se valer $u < u'$, observe que

$$f_{u'}(s) = \mathcal{F}(\langle f_{u'}(x) : x < s \rangle)$$

para qualquer $s < u$, mostrando que a restrição $f_{u'}|_{D_u}$ é \mathcal{F} -recursiva em D_u ; como existe no máximo uma função \mathcal{F} -recursiva em D_u , segue que $f_{u'}|_{D_u} = f_u$ e, portanto, $f_{u'}(t) = f_u(t)$, mostrando que a definição de $f_v(t)$ independe da testemunha u escolhida. Enfim, a função f_v é \mathcal{F} -recursiva em D_v pois, se $w < v$, então existe $u < v$ com $w < u$ e

$$f_v(w) := f_u(w) = \mathcal{F}(\langle f_u(t) : t < w \rangle) = \mathcal{F}(\langle f_v(t) : t < w \rangle),$$

onde a última igualdade segue pois $f_u(t) = f_v(t)$ para todo $t < w$, dado que $t < u$. \square

Observação B.3.6. Secretamente, a demonstração acima depende do Axioma da Substituição. De fato, no caso (ii), mostrou-se $f_v = \bigcup_{u < v} f_u$, o que depende da existência do conjunto $\{f_u : u < v\}$: como para cada $u < v$ existe uma única função f_u , a fórmula $\mathcal{G}(x, y)$ dada por “ $y = f_x$ se $x \in \mathbb{W}$ e $x < v$ ou $y = \emptyset$ ” é funcional, donde o Axioma da Substituição garante a existência do conjunto $\mathcal{G}[D_v]$, cujos membros são todos os y 's que atestam $\mathcal{G}(x, y)$ para $x \in D_v$.

O que talvez surpreenda o leitor seja a volta: assumindo-se a validade do Teorema B.3.5 em “ZFC sem o Axioma da Substituição”, pode-se deduzir o Axioma da Substituição. Porém, como a argumentação depende do entendimento de algumas nuances do Axioma da Escolha que ainda não foram exploradas no texto, convém adiá-la. \triangle

Assim, ao nos depararmos com uma fórmula $\mathcal{F}(x, y)$ funcional em x , é lícito aplicá-la *repetidamente/recursivamente* a conjuntos bem ordenados a fim de definir funções. Um caso típico é ilustrado no próximo

Exemplo B.3.7 (Fatorial, de novo – e oráculos). Para um conjunto X qualquer, indicaremos por $\overline{\text{seq}}(X) := \bigcup_{n \in \omega} X^n$ a família de todas as funções da forma $n \rightarrow X$, com $n \in \omega$, também chamadas de **sequências finitas** de X . Diremos que uma função $\mathcal{O}: \overline{\text{seq}}(X) \rightarrow X$ é um **oráculo**.

A motivação para o nome “oráculo” se dá ao pensar nos elementos de $\overline{\text{seq}}(X)$ como todas as histórias possíveis num certo contexto: para uma sequência $s := \langle s_0, \dots, s_n \rangle \in \overline{\text{seq}}(X)$, o oráculo \mathcal{O} diz o que ocorre *em seguida* por meio de $\mathcal{O}(s) \in X$. Em particular, como \emptyset é uma sequência finita, \mathcal{O} sabe dizer *como a história deveria começar*: $\mathcal{O}(\emptyset)$, que podemos chamar de x_0 . Agora, o oráculo também sabe o que *deveria* se seguir, com $x_1 := \mathcal{O}(\langle x_0 \rangle)$, bem como o que *deveria* vir depois, com $x_2 := \mathcal{O}(\langle x_0, x_1 \rangle)$. De modo geral:

Teorema B.3.8 (ω -oráculo). *Se X é um conjunto e $\mathcal{O}: \overline{\text{seq}}(X) \rightarrow X$ é uma função, então existe uma única função \mathcal{O} -recursiva $\psi: \omega \rightarrow X$, i.e., tal que para todo $n \in \omega$ ocorre*

$$\psi(n) = \mathcal{O}(\langle \psi(m) : m < n \rangle). \quad (\text{B.9})$$

Demonstração. No Teorema B.3.5, considere $\langle \mathbb{W}, \leq \rangle := \langle \omega, \leq \rangle$ e $\mathcal{F}(x, y)$ a fórmula que declara “ $y = \mathcal{O}(x)$ se $x \in \overline{\text{seq}}(X)$, e $y = \emptyset$ caso contrário”, que é funcional em x . Dessa forma, segue que existe uma única função \mathcal{F} -recursiva em ω , justamente a função ψ procurada: com efeito, a \mathcal{F} -recursividade em ω já garante $\text{dom}(\psi) = \omega$; por sua vez, $\psi(0) = \mathcal{F}(\emptyset) = \mathcal{O}(\emptyset) \in X$, de modo que se $\psi(m) \in X$ para todo $m < n$ com $n \in \omega$ fixado, então $\langle \psi(m) : m < n \rangle$ é um habitante de $\overline{\text{seq}}(X)$, de tal forma que

$$\psi(n) = \mathcal{F}(\langle \psi(m) : m < n \rangle) = \mathcal{O}(\langle \psi(m) : m < n \rangle) \in X.$$

Logo, por indução, resulta que $\text{im}(\psi) \subseteq X$. Em particular, note que a argumentação acima também mostrou a validade da identidade (B.9). \square

Exercício B.16. Demonstre o teorema acima diretamente. Dica: imite a prova do Teorema B.3.5. \blacksquare

Corolário B.3.9 (Recursão “fatorial”). *Se X é um conjunto, $c \in X$ é um elemento fixado e $g: \omega \times X \rightarrow X$ é uma função, então existe uma única função $h: \omega \rightarrow X$ tal que $h(0) = c$ e $h(n_+) = g(n, h(n))$ para todo $n \in \omega$.*

Demonstração. A ideia é definir uma função oráculo $\mathcal{O}: \overline{\text{seq}}(X) \rightarrow X$ para a qual uma função ψ como no teorema anterior se comporte como a função h desejada. Note que ψ deve satisfazer

$$\psi(n) = \mathcal{O}(\langle \psi(j) : j < n \rangle) \quad (\text{B.10})$$

para todo $n \in \omega$. Como espera-se que $\psi(0) = c$, deve-se definir $\mathcal{O}(\emptyset) := c$. Por outro lado, como deseja-se $\psi(n_+) = g(n, \psi(n))$, a função oráculo \mathcal{O} deveria associar uma sequência finita $\langle x_j : j \leq n \rangle \in \overline{\text{seq}}(X)$ ao elemento $g(n, x_n) \in X$. Isto encerra a definição de \mathcal{O} : de fato, se $\psi: \omega \rightarrow X$ é a única função satisfazendo (B.10), então $\psi(0) = c$ e, para $n \in \omega$ qualquer, $\psi(n_+) = \mathcal{O}(\langle \psi(0), \dots, \psi(n) \rangle) := g(n, \psi(n))$, como desejado. \square

Como isso tudo ajuda a *definir* o fatorial? Para $c := 1$ e $g: \omega \times \omega \rightarrow \omega$ a função dada por $g(m, n) := (m_+) \cdot n$, o corolário anterior garante uma única função $h: \omega \rightarrow \omega$ satisfazendo as condições $h(0) = 1$ e $h(n_+) = g(n, h(n)) = (n_+) \cdot h(n)$ para todo $n \in \omega$. Dado que n_+ é aquilo que, em breve, denotaremos por $n + 1$, a função h faz, na prática:

- $h(0) = 1$,
- $h(1) = h(0 + 1) = g(0, h(0)) = (0 + 1) \cdot h(0) = 1 \cdot 1 = 1$,
- $h(2) = h(1 + 1) = g(1, h(1)) = (1 + 1) \cdot h(1) = 2 \cdot 1 = 2$,
- $h(3) = h(2 + 1) = g(2, h(2)) = (2 + 1) \cdot h(2) = 3 \cdot 2 = 6 \dots$

Agora sim, é honesto definir $h(n) := n!$, classicamente xingado como **fatorial** de n . ▲

Embora, no dia a dia típico, o Teorema B.3.8 baste para justificar a maioria das recursos que cruzam nosso caminho, existem sutilezas conjuntistas que demandam construções *parametrizadas* por conjuntos bem ordenados *arbitrariamente maiores* do que ω . Com isso em mente, pode ser preferível pensar na versão geral do Teorema B.3.5 em termos de *oráculos* (como no Teorema B.3.8). Neste caso, a fórmula funcional $\mathcal{F}(x, y)$ se torna o *super oráculo* $\mathcal{O}_{\mathcal{F}}: \mathbb{V} \rightarrow \mathbb{V}$ que a cada conjunto x no universo associa o único $y \in \mathbb{V}$ que faz de $\mathcal{F}(x, y)$ uma sentença verdadeira. Em particular, note que enquanto os oráculos do Teorema B.3.8 conhecem todas as *histórias finitas* de X (*codificadas* em $\overline{\text{seq}}(X)$), o oráculo $\mathcal{O}_{\mathcal{F}}$ certamente conhece *todas* as histórias de \mathbb{V} , inclusive aquelas que são *ordenadas* por elementos de \mathbb{W} , i.e., as uplas da forma $\langle x_v : v < w \rangle$ com $w \in \mathbb{W}$, já que o domínio de $\mathcal{O}_{\mathcal{F}}$ é o universo. É com essa versão *profética* que este capítulo se encerra.

Teorema B.3.10 (a.k.a. Teorema B.3.5). *Para uma boa ordem $\langle \mathbb{W}, \leq \rangle$ e um super oráculo $\mathcal{O}: \mathbb{V} \rightarrow \mathbb{V}$, existe uma única função \mathcal{O} -recursiva $f: \mathbb{W} \rightarrow \mathbb{V}$, i.e., tal que $f(w) = \mathcal{O}(\langle f(v) : v < w \rangle)$ para todo $w \in \mathbb{W}$.*

Exercícios adicionais

Exercício B.17. Seja $\text{Rel}(X, Y) := \{R \subseteq X \times Y : \text{dom}(R) = X\}$ a família das relações binárias definidas em X com codomínio Y .

- Mostre que $\langle \text{Rel}(X, Y), \subseteq \rangle$ é uma ordem parcial.
- Reflita: para $R, S \in \text{Rel}(X, Y)$, o que significa dizer que $R \subseteq S$?
- Mostre que existe uma bijeção entre $\text{Rel}(X, Y)$ e $\wp(Y)^X$, em que o último indica a família das funções da forma $X \rightarrow \wp(Y)$. ■

Exercício B.18. Seja $\{E_i : i \in \mathcal{I}\}$ uma família não-vazia de relações de equivalência sobre um conjunto não-vazio X . Mostre que $\bigcap_{i \in \mathcal{I}} E_i$ é uma relação de equivalência sobre X . ■

Exercício B.19. Mostre que se $S \subseteq X \times X$, então existe uma relação de equivalência \sim_S em S com as seguintes propriedades:

- se $\langle x, y \rangle \in S$, então $x \sim_S y$, e
- se \sim é uma relação de equivalência em X tal que $x \sim y$ sempre que $\langle x, y \rangle \in S$, então $\sim_S \subseteq \sim$.

Conclua que \sim_S é a menor relação de equivalência sobre X a conter S . ■

Exercício B.20. Sejam X um conjunto, $S \subseteq X \times X$ uma relação binária e considere \sim_S a menor relação de equivalência que contém S . Para $a, b \in X$, mostre que $a \sim_S b$ ocorre se, e somente se,

- $a = b$, ou
- $a \approx b$, ou
- existem $n \in \omega$ e $c_0, \dots, c_n \in X$ tais que $a \approx c_0$, $c_0 \approx c_1, \dots, c_{n-1} \approx c_n$ e $c_n \approx b$,

onde $x \approx y$ abrevia “ $x S y$ ou $y S x$ ”. ■

Exercício B.21. Sejam $\langle \mathbb{P}, \leq \rangle$ uma ordem parcial, $A \subseteq \mathbb{P}$ e $a \in \mathbb{P}$.

- Mostre que $a = \min_{\leq} A$ em $\langle \mathbb{P}, \leq \rangle$ se, e somente se, $a = \max_{\geq} A$ em $\langle \mathbb{P}, \geq \rangle$.
- Mostre que A é limitado inferiormente em $\langle \mathbb{P}, \leq \rangle$ se, e somente se, A é limitado superiormente em $\langle \mathbb{P}, \geq \rangle$.
- Mostre que $a = \inf_{\leq} A$ em $\langle \mathbb{P}, \leq \rangle$ se, e somente se, $a = \sup_{\geq} A$ em $\langle \mathbb{P}, \geq \rangle$. ■

Exercício B.22. Sejam $\langle \mathbb{S}, \leq \rangle$ e $\langle \mathbb{T}, \leq \rangle$ ordens parciais. Uma função $f: \mathbb{S} \rightarrow \mathbb{T}$ é:

- crescente** se para quaisquer $s, s' \in \mathbb{S}$ valer que $s \leq s' \Rightarrow f(s) \leq f(s')$;
- decrescente** se para quaisquer $s, s' \in \mathbb{S}$ valer que $s \leq s' \Rightarrow f(s) \geq f(s')$;
- monótona** se f for crescente ou decrescente.

Sabendo disso, suponha que a ordem de \mathbb{S} seja total.

- Mostre que se f é crescente e $f(s) < f(s')$, então $s < s'$.
- Mostre que se f é decrescente e $f(s) < f(s')$, então $s > s'$.
- Conclua que se f for monótona e bijetora, então a inversa $f^{-1}: \mathbb{T} \rightarrow \mathbb{S}$ também será monótona (na verdade, f é crescente/decrescente se, e somente se, f^{-1} também é). ■

Exercício B.23. Mostre que uma função crescente e injetora satisfaz a implicação *estrita* $s < s' \Rightarrow f(s) < f(s')$. Funções em tal condição serão chamadas de **estritamente crescentes**. A definição para funções **estritamente decrescentes** é análoga.¹⁵ ■

Exercício B.24. Seja $\langle \mathbb{P}, \leq \rangle$ uma ordem.

- Mostre que se \mathbb{P} é uma boa ordem e $f: \omega \rightarrow \mathbb{P}$ é uma função, então $\text{im}(f)$ tem menor elemento. Conclua que, neste caso, não existe função $\omega \rightarrow \mathbb{P}$ estritamente decrescente, i.e., tal que se $m, n \in \omega$ com $m < n$, então $f(m) > f(n)$.
- Suponha que \mathbb{P} seja ordem total, mas que não seja boa ordem. Mostre que existe uma função $f: \omega \rightarrow \mathbb{P}$ estritamente decrescente. Dica: recursão + Axioma da Escolha. ■

Exercício B.25. Chame por ZFC⁻ os axiomas de ZFC sem o Axioma da Fundação. Mostre que se vale o enunciado do Exercício A.35, então vale o Axioma da Fundação. ■

¹⁵A literatura também costuma xingar de *não-decrescente* as coisas que aqui foram chamadas de *crescentes*. Em tais textos, o adjetivo *crescente* se reserva para as situações de desigualdade estrita. Um comentário análogo é válido para funções *não-crescentes*.

Exercício B.26 (Recursão “paramétrica”). Para funções $a: P \rightarrow X$ e $g: \omega \times P \times X \rightarrow X$ fixadas, mostre que existe uma única função $h: \omega \times P \rightarrow X$ satisfazendo as condições

- (i) $h(0, p) = a(p)$ para todo $p \in P$, e
- (ii) $h(n_+, p) = g(n, p, h(n, p))$ para quaisquer $n \in \omega$ e $p \in P$.

Dica: use o Corolário B.3.9 para cada $p \in P$, obtendo $h_p: \omega \rightarrow X$ adequada; daí defina $H: P \rightarrow X^\omega$ da maneira óbvia e a utilize para cozinar a função h desejada. ■

Exercício B.27. Use os teoremas de recursão para mostrar que existe uma única função $+: \omega \times \omega \rightarrow \omega$ satisfazendo as seguintes condições

- (i) $+(m, 0) = m$ para todo $m \in \omega$,
- (ii) $+(m, n_+) = +(m, n)_+$ para quaisquer $m, n \in \omega$

Tal função costuma ser chamada de *adição* em ω , e escreve-se $m + n$ em vez de $+(m, n)$.

- a) Mostre que $+$ é uma operação associativa e comutativa sobre ω .
- b) A operação $+$ tem elemento neutro? Qual?
- c) Adapte as ideias acima para definir a multiplicação \cdot em ω . Verifique as propriedades de associatividade, comutatividade e existência de neutro. Em particular, mostre que $0 \cdot n = 0$ para todo $n \in \omega$.
- d) $\langle \omega, +, 0 \rangle$ ou $\langle \omega, \cdot, 1 \rangle$ são grupos? ■

Observação B.3.11. Se você enfrentou o exercício anterior, então (já) pode substituir as ocorrências de “ k_+ ” por “ $k + 1$ ” sempre que k for um número natural. Se você não enfrentou, então deverá esperar até a discussão final sobre *aritmética cardinal* (e ordinal), na Seção E.2.△

Exercício B.28. Sejam $\langle G, \cdot, 1 \rangle$ um monóide e $x \in G$. Para $n \in \omega$, defina $x^n \in G$ recursivamente fazendo $x^0 := 1$ e $x^{n+1} := x^n \cdot x$ para todo $n \in \omega$. Se x admitir inverso, para cada $n \in \omega$ com $n > 0$ defina $x^{-n} := (x^{-1})^n$.

- a) Mostre que a definição acima faz sentido.
- b) Mostre que $x^{m+n} = x^m \cdot x^n$ para quaisquer $m, n \in \omega$;
- c) Mostre que $x^{mn} = (x^m)^n$ para quaisquer $m, n \in \omega$;
- d) Mostre que se x admitir inverso, então as identidades anteriores valem para $m, n \in \mathbb{Z}$. ■

Exercício B.29. Sejam $\langle G, +, 0 \rangle$ um grupo e $x \in G$. Para $n \in \omega$, defina $nx \in G$ recursivamente, fazendo $0x := 0$ e $(n+1)x := nx + x$ para todo $n \in \omega$. Além disso, defina $(-n)x := n(-x)$ para cada $n \in \mathbb{N}$. Mostre que a definição acima faz sentido, e verifique a validade das identidades $(m+n)x = mx + nx$ e $m(nx) = (mn)x$ para quaisquer $m, n \in \mathbb{Z}$. ■

Exercício B.30 (Opcional: requer conhecimentos de Álgebra Linear e afins). Use os teoremas de recursão para definir o significado de expressões como $\sum_{i=0}^n \alpha_i v_i$ em espaços vetoriais, módulos, etc., onde $n \in \omega$ e, para cada $i \in \{0, \dots, n\}$, α_i indica um escalar do anel/corpo e v_i indica um vetor do espaço vetorial/módulo.

Capítulo C

Cardinalidade

Enquanto figuras de linguagem, conjuntos (quase?) sempre estiveram presentes na Matemática, mesmo que implicitamente. A grande novidade introduzida no século XIX foi a utilização *deles* para analisar as noções de (in)finitude e, posteriormente, abstrair e generalizar as diversas estruturas oriundas dos mais variados contextos matemáticos. O presente capítulo se propõe a fazer uma discussão um pouco mais aprofundada sobre o primeiro aspecto.

C.1 A ideia de número cardinal

Como aprendemos a lidar com *números* desde a tenra infância, há certas sutilezas sobre as noções de contagem e comparação de elementos que escapam do público leigo. A primeira delas é a de que números não são necessários nessa tarefa.



Figura C.1: É evidente que há tantos círculos quanto quadrados, mesmo sem saber quantos.

Definição C.1.1. Diremos que dois conjuntos **têm a mesma cardinalidade** (“quantidade de elementos”) se existir uma bijeção entre os dois. ¶

Proposição C.1.2. *Sejam A , B e C conjuntos.*

- (i) *Existe uma bijeção de A para A .*
- (ii) *Se existe uma bijeção de A para B , então existe uma bijeção de B para A .*
- (iii) *Se existe uma bijeção de A para B e outra bijeção de B para C , então existe uma bijeção de A para C .*

Demonstração. O primeiro item segue por Id_A ser bijeção, enquanto o terceiro item decorre do fato de que a composição de bijeções é bijeção (Exercício A.28). O segundo item é o Corolário A.1.43. □

Intuitivamente, a proposição acima diz que a relação “ $A \approx B \Leftrightarrow$ existe bijeção $A \rightarrow B$ ” define uma relação de equivalência sobre o universo \mathbb{V} de todos os conjuntos. Naturalmente, isto é apenas intuitivo, posto que \mathbb{V} não é um conjunto e relações de equivalência só podem ser definidas sobre conjuntos.

Nesse sentido, a *cardinalidade* de um conjunto X poderia ser pensada como a *classe de equivalência* de X nesta pseudo-relação \approx , i.e., a classe de todos os conjuntos em bijeção com X .

Definição provisória. A *cardinalidade* de X será denotada por $\#X$.

Observação C.1.3. A desprezível notação acima será usada apenas na discussão inicial. Em breve, ela será substituída pela notação *tradicional*¹. \triangle

Como já foi destacado, a definição acima, por enquanto, é *vazia*, posto que \approx não é uma relação de equivalência e, portanto, não há garantias de que $\#X$ seja realmente um conjunto. O leitor não deve ter qualquer esperança do contrário:

Teorema C.1.4. *Se X é um conjunto, então $\#X$ é um conjunto se, e somente se, $X = \emptyset$.*

Demonstração. Como o único conjunto em bijeção com o conjunto vazio é o próprio conjunto vazio, segue que $\#\emptyset := \{Y : Y \approx \emptyset\} = \{\emptyset\}$. Agora, se $X \neq \emptyset$ e $\#X$ fosse um conjunto, então $A := \bigcup \#X$ seria um conjunto. Ora, como $X \neq \emptyset$, existe $t \in X$. Logo, dado um y qualquer, o conjunto $Y := (X \setminus \{t\}) \cup \{y\}$ está em bijeção com X e, portanto, $Y \in \#X$ com $y \in Y$, mostrando que $y \in A$. Em outras palavras: A seria o universo, absurdo. \square

Embora não seja formalizável em ZFC, a *ideia* de interpretar cardinalidades como *classes* encapsula *toda* a intuição dos processos de contagem.

Exemplo C.1.5. Dados conjuntos disjuntos A e B , definamos $\#A + \#B := \#(A \cup B)$. Se coisas como $\#A$ e $\#B$ fizessem sentido, tal regra estaria bem definida, pelo seguinte

Exercício C.1. Sejam A, B, C e D conjuntos, com $A \cap B = C \cap D = \emptyset$. Mostre que se $A \approx C$ e $B \approx D$, então $A \cup B \approx C \cup D$. \blacksquare

Em outras palavras, a regra acima define uma *soma entre cardinalidades*, que é bastante razoável, já que, por exemplo, ocorreria $\#\{0, 1\} + \#\{\star, \bullet\} = \#\{\diamondsuit, \clubsuit, \spadesuit, \heartsuit\}$. Analogamente, num mundo sem regras, poderíamos definir o *produto* entre cardinalidades $\#A$ e $\#B$ como sendo $\#(A \times B)$. \blacktriangle

O leitor pragmático pode investigar as propriedades de comutatividade, associatividade, etc. oriundas do tratamento *metalingüístico* proposto acima. Neste texto, porém, isto será feito no contexto dos *números cardinais*, razão pela qual não convém prolongar a presente discussão. E o que seriam números cardinais?

A ideia é muito simples: como não se pode considerar a própria classe $\#X$ em ZFC, elege-se um único elemento em $\#X$, digamos $|X| \in \#X$, que será o representante de toda a classe: este será o *número cardinal* de X . Ao se fazer isso para todos os conjuntos, espera-se que conjuntos com o mesmo *número cardinal* tenham a mesma *cardinalidade*, i.e.,

$$X \approx Y \Leftrightarrow |X| = |Y|, \quad (\text{C.1})$$

ou, em outras palavras: $\mathcal{C} := \{|X| : X \text{ é conjunto}\}$ deve ser uma *classe de representantes* da pseudo-relação de equivalência \approx . Evidentemente, isso traz um problema profundo: como *escolher* tais representantes?

¹Textos sérios de Teoria dos Conjuntos costumam indicar o *número cardinal* de X por $|X|$, prática que será mantida aqui.

Não há dúvidas quanto ao que fazer no caso $X := \emptyset$, já que $\#\emptyset = \{\emptyset\}$: escolhe-se \emptyset como o número cardinal de \emptyset . Como \emptyset já foi chamado de 0, isso dá a reconfortante identidade $|\emptyset| = 0$. Curiosamente, o conjunto $0_+ := 0 \cup \{0\} := 1$ tem apenas um elemento, o que torna razoável escolhê-lo como o número cardinal de sua cardinalidade, i.e., $|1| = 1$. Tal raciocínio sugere que, pelo menos para conjuntos finitos, o problema da *escolha* dos números cardinais esteja resolvido.

Teorema C.1.6. *Se X é finito, então existe um único $n \in \omega$ em bijeção com X .*

Demonstração. A definição adotada para conjuntos finitos dá a existência do número natural n (Definição A.2.1). Para provar sua unicidade, vamos supor $m, n \in \omega$ com $X \approx m$ e $X \approx n$. Como a relação \approx é simétrica e transitiva, resulta que $m \approx n$. Desse modo, pode-se argumentar pela contrapositiva: vamos supor $m \neq n$ a fim de concluir que não existe bijeção entre m e n .

Por $\langle \omega, < \rangle$ ser uma boa ordem, não há perda de generalidade em supor $m < n$, já que um deles deve ser o menor elemento do conjunto $\{m, n\}$.

Afirmiação. *Se $X \subsetneq n$, então não existe bijeção entre n e X .*

Prova da afirmação. O argumento será feito por indução em n , sendo o caso $n := 0$ imediato. Supondo a afirmação verdadeira para $n \in \omega$, uma bijeção φ entre n_+ e $X \subsetneq n_+$ resultaria numa bijeção entre n e um subconjunto próprio de n :

- ✓ se $n \notin X$, então X já é um subconjunto de n e, portanto, a restrição de φ a n seria uma bijeção entre n e $X \setminus \{\varphi(n)\}$ (confira a Observação C.1.8);
- ✓ se $n \in X$, então existe $k \leq n$ com $\varphi(k) = n$, e daí a função $g: n \rightarrow X \setminus \{n\}$, que faz $g(i) := \varphi(i)$ para $i \neq k$ e $g(k) := \varphi(n)$ caso $k < n$, é uma bijeção;

contrariando a hipótese de indução. \square

Agora, basta ver que se $m < n$, então m é um subconjunto próprio de n , donde o resultado segue da afirmação. \square

Corolário C.1.7. *Se X é finito e $Y \subsetneq X$, então não existe bijeção entre X e Y .*

Observação C.1.8. Na demonstração do último teorema, n foi usado tanto como *elemento* do domínio de φ quanto como *subconjunto* do domínio de φ , o que é lícito em virtude da identidade $n = \{m \in \omega : m < n\}$.

De fato, pela definição da relação $<:=\in$, é claro que $\{m \in \omega : m < n\} \subseteq n$. Para a inclusão oposta, é suficiente mostrar que se $x \in n$, então $x \in \omega$, o que se faz por indução: o conjunto $\mathcal{M} := \{n \in \omega : n \subseteq \omega\}$ é tal que $0 \in \mathcal{M}$ e, se $n \in \mathcal{M}$, então pela definição de \mathcal{M} tem-se $n \subseteq \omega$, acarretando $n_+ := n \cup \{n\} \subseteq \omega$, mostrando que $n_+ \in \mathcal{M}$; portanto, $\mathcal{M} = \omega$. \triangle

Definição C.1.9. Dado um conjunto X finito, indicaremos por $|X| \in \omega$ o único número natural em bijeção com X , que será chamado de **número cardinal de X** . \P

Exercício C.2. Mostre que o conjunto ω é infinito. Dica: $n \mapsto n_+$. \blacksquare

Chega-se então a um problema de caráter prático, que pode soar trivial num primeiro momento: como *representar* a cardinalidade de conjuntos infinitos, como ω ? Ora, das tradições intuitivas e teológicas que infestam o cotidiano, o infinito é a coisa intangível e única que representa a perfeição ou qualquer outra pseudo-poiesia motivacional, antropocêntrica e mística. Com isso em mente, pode-se fazer a seguinte

Definição mística. O *número cardinal* de um conjunto infinito X será denotado por ∞ , i.e., se X for infinito, então $|X| := \infty$.

Pela discussão que embasou a *escolha* dos números naturais como representantes das cardinalidades finitas, a notação acima só faria sentido se quaisquer dois conjuntos infinitos fossem equivalentes com respeito à pseudo-relação de equivalência \approx . Em outras palavras, deveria valer o seguinte: se X e Y são infinitos, então existe uma bijeção entre X e Y . Entretanto, isto é *falso*.

Teorema C.1.10 (Cantor). *Dado um conjunto X , não existe sobrejeção $X \rightarrow \wp(X)$.*

Demonstração. Para uma função $\varphi: X \rightarrow \wp(X)$ qualquer, o conjunto

$$T := \{x \in X : x \notin \varphi(x)\}.$$

atesta a não-sobrejetividade de φ . De fato, se ocorresse $\varphi(t) = T$ para algum $t \in X$, a definição de T daria

$$t \in T \Leftrightarrow t \notin \varphi(t) = T,$$

uma contradição. Logo, não há sobrejeção $X \rightarrow \wp(X)$, como desejado. \square

Corolário C.1.11. *Se X é infinito, então $\wp(X)$ é infinito. Porém, não existe bijeção entre X e $\wp(X)$.*

Demonstração. Note que se Z é finito e $Y \subseteq Z$, então Y é finito: pode-se argumentar, sem grandes dificuldades, por indução em $|Z|$. Daí, como a correspondência $x \mapsto \{x\}$ define uma função injetora $X \rightarrow \wp(X)$, segue que se $\wp(X)$ fosse finito, então X estaria em bijeção com um subconjunto (finito) de $\wp(X)$ e, portanto, seria finito. O restante segue do Teorema de Cantor. \square

Exercício C.3. Complete a demonstração anterior. ■

Explicitamente, o Teorema de Cantor estabelece a existência de tipos distintos (e comparáveis) de infinitos: em outras palavras, *existem infinitos maiores do que outros infinitos*. Portanto, a solução trivial para o problema de ordem prática, i.e., atribuir o símbolo “ ∞ ” como representante da cardinalidade infinita, não funcionou, o que suscita refazer a pergunta: como representar as cardinalidades de conjuntos infinitos?

Existem, essencialmente, dois modos:

- (i) Escreve-se “ $|X| = |Y|$ ” como uma *abreviação* para afirmar que X e Y têm a mesma cardinalidade, i.e., “ $|X| = |Y|$ ” indica a existência de uma bijeção $X \rightarrow Y$.
- (ii) Para cada conjunto X escolhe-se um conjunto $|X|$ de alguma *maneira canônica*, assegurando-se o seguinte: para quaisquer conjuntos X e Y , verifica-se $|X| = |Y|$ se, e somente se, existe uma bijeção da forma $X \rightarrow Y$.

Para leitores adeptos da primeira abordagem, a existência dos números naturais se coloca como uma feliz exceção, pois em tais casos é possível representar explicitamente as cardinalidades finitas por meio dos números naturais. Embora seja tecnicamente limitada, essa abordagem é suficiente em diversos contextos elementares, razão pela qual a segunda alternativa será brevemente adiada.

C.2 Cardinais enquanto *façon de parler*

Definição C.2.1. Sejam X e Y conjuntos.

- (i) Escreveremos $|X| = |Y|$ a fim de indicar a existência de uma bijeção $X \rightarrow Y$, i.e., X e Y têm a mesma cardinalidade². Sua negação será indicada por $|X| \neq |Y|$.
- (ii) Escreveremos $|X| \leq |Y|$ a fim de indicar a existência de injeção $X \rightarrow Y$, que se pode ler como “a cardinalidade de X é *menor* do que a cardinalidade de Y ”.
- (iii) Escreveremos $|X| < |Y|$ para abreviar “ $|X| \leq |Y|$ e $|X| \neq |Y|$ ”, que se pode ler como “a cardinalidade de X é *estritamente menor* do que a cardinalidade de Y ”.

Acima, não se definiram os *números cardinais* de X e Y , mas apenas abreviações para afirmações referentes à existência de bijeções e injeções entre X e Y . Note ainda que no caso dos conjuntos finitos, que tiveram seus números cardinais explicitamente definidos, as abreviações acima são compatíveis com as relações existentes entre os números cardinais correspondentes.

Exercício C.4. Para X e Y conjuntos finitos, com $|X| = m$ e $|Y| = n$, mostre que $|X| \leq |Y|$ (i.e., existe injeção $X \rightarrow Y$) se, e somente se, $m \leq n$. ■

Exercício C.5. Sejam X , Y e Z conjuntos. Convença-se da validade das seguintes afirmações.

- a) $|X| = |Y| \Rightarrow |X| \leq |Y|$.
- b) $|X| \leq |Y|$ e $|Y| \leq |Z| \Rightarrow |X| \leq |Z|$.
- c) $|X| < |\wp(X)|$.
- d) $|\wp(X)| = |2^X|$, onde 2^X indica a família das funções da forma $X \rightarrow \{0, 1\}$. Dica: faça $A \in \wp(X)$ corresponder a uma função $\chi_A: X \rightarrow \{0, 1\}$ característica.
- e) $X \subseteq Y \Rightarrow |X| \leq |Y|$. ■

Os dois primeiros itens do exercício acima sugerem que a *relação de desigualdade* entre cardinalidades tem comportamento parecido com o que se esperaria de uma ordem parcial. Isto de fato ocorre, mas é menos trivial do que parece, afinal de contas, os conjuntos acima podem ser infinitos.

Definição C.2.2. Dada uma função $h: A \rightarrow A$, diremos que $a \in A$ é um **ponto fixo** da função h se ocorrer $h(a) = a$. ¶

Lema C.2.3 (Ponto Fixo de Tarski). *Se X é um conjunto e $\varphi: \wp(X) \rightarrow \wp(X)$ é uma função \subseteq -crescente, i.e., tal que $A \subseteq B \subseteq X$ acarreta $\varphi(A) \subseteq \varphi(B)$, então φ tem ponto fixo.*

Demonstração. Basta fazer $\mathcal{F} := \{S \subseteq X : S \subseteq \varphi(S)\}$ e notar que $P := \bigcup \mathcal{F}$ é um ponto fixo de φ . De fato, por φ ser \subseteq -crescente e $S \subseteq P$ para todo $S \in \mathcal{F}$, vale $\varphi(S) \subseteq \varphi(P)$. Assim, se $x \in P$, então existe $S \in \mathcal{F}$ com $x \in S \subseteq \varphi(S) \subseteq \varphi(P)$, mostrando que $P \subseteq \varphi(P)$. Logo, por φ ser \subseteq -crescente, resulta $\varphi(P) \in \mathcal{F}$, donde a construção de P permite concluir que $\varphi(P) \subseteq P$. □

²Definição C.1.1.

Teorema C.2.4 (Cantor-Bernstein³). *Se existem injeções $f: X \rightarrow Y$ e $g: Y \rightarrow X$, então existe uma bijeção $h: X \rightarrow Y$. Em outras palavras: $|X| \leq |Y|$ e $|Y| \leq |X| \Rightarrow |X| = |Y|$.*

Demonastração. Obteremos partições $\{X_0, X_1\}$ e $\{Y_0, Y_1\}$ de X e Y , respectivamente, tais que $f[X_0] = Y_0$ e $g[Y_1] = X_1$ pois, se isso for feito, então as restrições $f|_{X_0}: X_0 \rightarrow Y_0$ e $\tilde{g} := g|_{Y_1}: Y_1 \rightarrow X_1$ serão bijeções, donde bastará definir $h: X \rightarrow Y$ como $h(x) := f(x)$ se $x \in X_0$ e $h(x) := (\tilde{g})^{-1}(x)$ se $x \in X_1$.

Para obter a partição desejada, *tira-se da cartola* a função

$$\begin{aligned}\varphi: \wp(X) &\rightarrow \wp(X) \\ S &\mapsto X \setminus g[Y \setminus f[S]]\end{aligned}\tag{C.2}$$

que é \subseteq -crescente. De fato, pelo Exercício A.32,

$$\begin{aligned}S \subseteq S' \Rightarrow f[S] &\subseteq f[S'] \Rightarrow Y \setminus f[S'] \subseteq Y \setminus f[S] \Rightarrow g[Y \setminus f[S']] \subseteq g[Y \setminus f[S]] \Rightarrow \\ &\Rightarrow X \setminus g[Y \setminus f[S']] \subseteq X \setminus g[Y \setminus f[S]] \Rightarrow \varphi(S) \subseteq \varphi(S').\end{aligned}$$

Logo, pelo lema anterior, existe $X_0 \subseteq X$ tal que $X_0 = \varphi[X_0] := X \setminus g[Y \setminus f[X_0]]$, de modo que agora basta tomar $Y_0 := f[X_0]$, $X_1 := X \setminus X_0$ e $Y_1 := Y \setminus Y_0$, pois disso resulta

$$g[Y_1] = g[Y \setminus Y_0] = g[Y \setminus f[X_0]] = X \setminus X_0 = X_1,$$

como queríamos. \square

Observação C.2.5 (Cartolas não existem). A função $\varphi: \wp(X) \rightarrow \wp(X)$ usada na demonstração acima não foi *verdadeiramente* tirada da cartola. Note que qualquer subconjunto $X_0 \subseteq X$ induz tanto uma partição em X quanto uma partição em Y : basta definir $X_1 := X \setminus X_0$, $Y_0 := f[X_0]$ e $Y_1 := Y \setminus Y_0$.

Isso resolve *metade* do problema, dado que ainda precisa-se garantir a validade da condição $g[Y_1] = X_1$, essencial para a definição da bijeção h . Ao se reescrever a igualdade $g[Y_1] = X_1$ em função de X_0 , obtém-se a identidade $g[Y \setminus f[X_0]] = X \setminus X_0$, equivalente-mente exprimível como $X_0 = X \setminus g[Y \setminus f[X_0]]$. \triangle

Mesmo com tão pouco ferramental técnico, já é possível observar fenômenos estranhos referentes a *certos* conjuntos infinitos (compare com o Exercício C.29).

Teorema C.2.6. $|\omega| = |\omega \times \omega|$.

Demonstração. A função $n \mapsto \langle 0, n \rangle$ define uma injecão $\omega \rightarrow \omega \times \omega$. Por outro lado, como leitores versados em Aritmética devem saber, a correspondência $\langle m, n \rangle \mapsto 2^m 3^n$ define uma função injetora $\omega \times \omega \rightarrow \omega$. Logo, o resultado segue do Teorema de Cantor-Bernstein. \square

Observação C.2.7 (Sobrejeções e o Axioma da Escolha). Por definição, $|X| \leq |Y|$ indica a existência de uma função injetora da forma $X \rightarrow Y$. Ocorre que da experiência cotidiana com conjuntos finitos, seria natural esperar que $|X| \leq |Y|$ pudesse se traduzir por meio de funções sobrejetoras, no sentido de ser equivalente à existência de uma função $Y \rightarrow X$ sobrejetora. Isto de fato ocorre, mas não sem *custo*.

Proposição C.2.8 (Compare com o Exercício A.31). *Uma função $f: X \rightarrow Y$ é sobrejetora se, e somente se, existe uma injecão $g: Y \rightarrow X$ tal que $f \circ g = \text{Id}_Y$. Em particular, $|Y| \leq |X|$ se, e somente se, existe sobrejeção $X \rightarrow Y$.*

³Também chamado de Cantor-Bernstein-Schröder.

Demonstração. Se a função g existe como no enunciado, então a sobrejetividade de f segue do Exercício A.29 (a menos da ordem das letras). A parte delicada é a recíproca: como cozinhar uma função $g: Y \rightarrow X$ satisfazendo $f \circ g = \text{Id}_Y$?

Note que a identidade procurada se traduz em pedir $f(g(y)) = y$ para todo $y \in Y$. Na prática, isto significa dizer que $g(y) \in X$ é *algum* elemento de X que é *levado* até y por f . Certamente, *algum* $x \in X$ satisfaz $f(x) = y$, posto que f é sobrejetora por hipótese. Dessa forma, $f^{-1}[\{y\}] \neq \emptyset$ para todo $y \in Y$.

Entra em cena o Axioma da Escolha: existe $g \in \prod_{y \in Y} f^{-1}[\{y\}]$, i.e., g é uma upla da forma $\langle g(y) \rangle_{y \in Y}$ em que $g(y) \in f^{-1}[\{y\}]$ para cada $y \in Y$, o que resulta em $f(g(y)) = y$, como desejado. Em particular, a injetividade de g é consequência do Exercício A.29. \square

Exercício C.6. Mostre que se f é uma função, então $|\text{im}(f)| \leq |\text{dom}(f)|$. ■

A ideia (da prova) da última proposição é mais simples do que parece: já que se busca atribuir um *valor* $g(y) \in X$ para cada y de tal forma que $f(g(y)) = y$, basta escolher um $x_y \in X$ com tal propriedade, para cada $y \in Y$. Se Y fosse finito, seria lícito argumentar por indução em $|Y| \in \omega$. Como pode não ser o caso, apela-se para o Axioma da Escolha.

Cabe frisar que a formulação em termos de produtos cartesianos apenas buscou explorar a encarnação apresentada no primeiro capítulo. É totalmente lícito usar outras abordagens: por exemplo, ao declarar $x \sim x'$ se $f(x) = f(x')$, obtém-se uma relação de equivalência sobre X , que tem uma classe de representantes para \sim , digamos $R \subseteq X$ (Teorema B.1.13); agora, basta definir $g: Y \rightarrow X$ fazendo $g(y) := r \in R$, onde $r \in R$ é o único tal que $f(r) = y$: para $y \in Y$, existe $x \in X$ com $f(x) = y$, bem como um único $r \in R$ tal que $x \sim r$, i.e., $f(x) = f(r)$, mostrando assim que a função g está bem definida. \triangle

Corolário C.2.9. Seja $\mathcal{X} := \{X_n : n \in \omega\}$ uma família de conjuntos com $|X_n| \leq |\omega|$ para cada $n \in \omega$. Então $|\bigcup \mathcal{X}| \leq |\omega|$.

Demonstração. Para cada $n \in \omega$, o conjunto $\text{Sob}(\omega, X_n)$, das funções sobrejetoras da forma $\omega \rightarrow X_n$, é não-vazio, donde o Axioma da Escolha assegura uma upla de sobrejeções $\langle f_n \rangle_{n \in \omega} \in \prod_{n \in \omega} \text{Sob}(\omega, X_n)$. Com isso, pode-se definir $f: \omega \times \omega \rightarrow \bigcup_{n \in \omega} X_n$ a função que faz $\langle m, n \rangle \mapsto f_m(n)$. Note que se $x \in \bigcup_{n \in \omega} X_n$, então existe $m \in \omega$ com $x \in X_m$, bem como $n \in \omega$ tal que $f_m(n) = x$, i.e., $f(m, n) = x$. Logo, f é sobrejetora, acarretando $|\bigcup \mathcal{X}| \leq |\omega \times \omega| = |\omega|$. \square

É importante ressaltar que a primeira definição para conjuntos finitos e infinitos garante que um conjunto X é finito ou infinito, e *nunca* ambos, afinal de contas: ou existe uma bijeção entre X e um (único) número natural $n \in \omega$, ou não existe. Em particular, como conjuntos finitos não admitem bijeções com partes próprias (Corolário C.1.7), obtém-se o

Corolário C.2.10. Se $|\omega| \leq |X|$, então X admite bijeção com subconjunto próprio. Em particular, X é infinito.

Demonstração. Se $\psi: \omega \rightarrow X$ é injetora, então a função $g: X \rightarrow X$ dada por

$$g(x) := \begin{cases} x, & \text{se } x \notin \text{im}(\psi) \\ \psi(n_+), & \text{se } \psi(n) = x \end{cases}$$

é uma função injetora e não-sobrejetora, posto que $\psi(0) \in X \setminus \text{im}(g)$. Em outras palavras, g é uma bijeção de X sobre um subconjunto próprio de X e, portanto, não é finito. \square

O corolário acima foi, por muito tempo, um dos principais responsáveis pelo *horror infiniti* predominante nos círculos matemáticos de nossos antepassados, em virtude de sua aparente incongruência com o *princípio geométrico* de que *o todo é maior do que as partes*. Discutivelmente, Bolzano e Dedekind foram os primeiros a utilizarem tal resultado como uma caracterização para a noção de infinito, no seguinte sentido:

Teorema C.2.11 (Dedekind). *Se X é infinito, então X admite bijeção com parte própria.*

Demonstração. Em vista do corolário anterior, basta exibir uma injecção $\omega \rightarrow X$. A ideia é simples: por X ser infinito, não existe $n \in \omega$ com uma bijeção $n \rightarrow X$; logo, se $n \in \omega$ e $\varphi: n \rightarrow X$ for uma função injetora, então $X \setminus \text{im}(\varphi)$ é não-vazio, o que permite escolher, recursivamente, sequências finitas *cada vez maiores* de modo a obter uma injecção $\omega \rightarrow X$. O exercício a seguir apresenta um roteiro de formalização para o leitor interessado. \square

Exercício C.7. Sejam X um conjunto infinito, $f \in \overline{\text{seq}}(X)$ e $n \in \omega$ com $\text{dom}(f) := n$. Para tal f , defina $D_f := \{g \in X^{n+1} : f \subseteq g \text{ e } g(n) \notin \text{im}(f)\}$.

- Mostre que $D_f \neq \emptyset$. Dica: note que $\text{im}(f)$ é um subconjunto finito de X .
- Observe que $\prod_{f \in \overline{\text{seq}}(X)} D_f \neq \emptyset$.
- Seja $E \in \prod_{f \in \overline{\text{seq}}(X)} D_f$ e, para cada $f \in \overline{\text{seq}}(X)$, considere $m(f) := \max \text{dom}(E(f))$ e $\mathcal{O}: \overline{\text{seq}}(X) \rightarrow X$ a função que faz $\mathcal{O}(f) := E(f)(m(f))$. Mostre que se ψ é \mathcal{O} -recursiva (como no Teorema B.3.8), então ψ é injetora. Dica: note que se $f := \langle x_0, \dots, x_n \rangle$, então $E(f) = \langle x_0, \dots, x_n, y \rangle$ para algum $y \notin \{x_0, \dots, x_n\}$, com $m(E(f)) = n + 1$ e $\mathcal{O}(f) := y$. \blacksquare

C.3 Opcional: cardinalidades clássicas

Pode ser edificante utilizar as ferramentas anteriores para estimar as cardinalidades dos conjuntos numéricos *clássicos* (\mathbb{Z} , \mathbb{Q} e \mathbb{R}), a fim de ilustrar a *aplicabilidade*. Os dois primeiros casos são bem simples (e dependem, implicitamente, de diversos exercícios elementares propostos no final do capítulo, além dos resultados anteriores):

- como existem funções injetoras $\omega \rightarrow \mathbb{Z}$ e $\mathbb{Z} \rightarrow \mathbb{Q}$, resulta que $|\omega| \leq |\mathbb{Z}| \leq |\mathbb{Q}|$;
- por construção, \mathbb{Q} vem de fábrica com uma sobrejeção da forma $\mathbb{Z} \times (\mathbb{Z} \setminus \{0\}) \rightarrow \mathbb{Q}$, mostrando que $|\mathbb{Q}| \leq |\mathbb{Z} \times (\mathbb{Z} \setminus \{0\})|$;
- em geral, se $|A| = |A'|$ e $|B| = |B'|$, então $|A \times B| = |A' \times B'|$, de modo que por valer $|\mathbb{Z}| = |\mathbb{Z} \setminus \{0\}|$, resulta $|\mathbb{Z} \times (\mathbb{Z} \setminus \{0\})| = |\mathbb{Z} \times \mathbb{Z}|$;
- finalmente, $\mathbb{Z} = \bigcup_{n \in \omega} \{-n, n\}$, com $|\{-n, n\}| \leq |\omega|$ para todo $n \in \omega$, resultando em $|\mathbb{Z}| \leq |\omega|$ (logo, $|\mathbb{Z}| = |\omega|$) e, como acima, $|\mathbb{Z} \times \mathbb{Z}| = |\omega \times \omega| = |\omega|$.

Em suma:

$$|\omega| \leq |\mathbb{Z}| \leq |\mathbb{Q}| \leq |\mathbb{Z} \times (\mathbb{Z} \setminus \{0\})| = |\mathbb{Z} \times \mathbb{Z}| = |\omega|,$$

onde o Teorema de Cantor-Bernstein assegura as clássicas identidades $|\omega| = |\mathbb{Z}| = |\mathbb{Q}|$. Este parece ser um *momento* apropriado para recordar uma terminologia muito frequente.

Definição C.3.1. Diremos que X é **enumerável** se valer $|X| \leq \omega$ e, obviamente, *infinito enumerável* se ocorrer $|X| = \omega$. Diremos que X é **não-enumerável** se X não for enumerável (...), i.e., se não existir injeção $X \rightarrow \omega$. \blacksquare

Assim, as estimativas anteriores mostram que \mathbb{Z} e \mathbb{Q} são (infinitos) enumeráveis, enquanto o Corolário C.2.9 pode ser rephraseado como “*a reunião enumerável de conjuntos enumeráveis é enumerável*”. Por sua vez, um conjunto não-enumerável é, automaticamente, infinito, donde a prova apresentada para o Teorema de Dedekind garante uma função injetora $\omega \rightarrow X$, que não pode ser sobrejetora em virtude da definição de não-enumerabilidade. Em outras palavras, isto resolve o

Exercício C.8. Mostre que X é não-enumerável se, e somente se, $|\omega| < |X|$. \blacksquare

O exemplo mais simples de conjunto não-enumerável é $\wp(\omega)$: trata-se de uma aplicação imediata do Teorema de Cantor (ou do Corolário C.1.11, para quem estiver com preguiça de mastigar). Em particular, tal observação pode ser usada numa demonstração quase elementar de que a *reta real* é não-enumerável ou, mais precisamente:

Teorema C.3.2. Se $\langle \mathbb{K}, \leq \rangle$ é um corpo ordenado completo, então $|\mathbb{K}| = |\wp(\omega)|$.

Esboço da prova. Tanto o enunciado quanto a demonstração utilizam noções típicas de um (começo de) curso de Análise na Reta. Porém, levando-se em conta que explicitá-las poderia passar a impressão de que esta obra não apresenta uma “sistematização racional dos tópicos abordados e que os objetivos principais do autor não ficaram claros no texto”, elas serão apenas esboçadas:

- (i) dizer que $\langle \mathbb{K}, \leq \rangle$ é **corpo ordenado** significa pedir que \mathbb{K} seja um **corpo**, i.e., um anel com $0_{\mathbb{K}} \neq 1_{\mathbb{K}}$ em que todo $r \in \mathbb{K} \setminus \{0_{\mathbb{K}}\}$ tem inverso multiplicativo, munido de uma ordem total em que $a + c < b + c$ sempre que $a < b$, bem como $ab > 0_{\mathbb{K}}$ sempre que $a, b > 0_{\mathbb{K}}$, para quaisquer $a, b, c \in \mathbb{K}$;
- (ii) dizer que $\langle \mathbb{K}, \leq \rangle$ é **completo** significa exigir que todo subconjunto não-vazio e limitado superiormente tenha supremo;
- (iii) em particular, a ordenação de \mathbb{K} permite obter uma “cópia” de \mathbb{Q} em \mathbb{K} , essencialmente por valer $0_{\mathbb{K}} < 1_{\mathbb{K}}$;
- (iv) por sua vez, a completude de \mathbb{K} aliada à compatibilidade da ordem com a adição garantem que a cópia de \mathbb{N} em \mathbb{K} é ilimitada, o que por sua vez equivale a dizer que para quaisquer $x, y \in \mathbb{K}$ com $x < y$ existe $q \in \mathbb{Q}$ com $x < q_{\mathbb{K}} < y$ (onde $q_{\mathbb{K}}$ indica a “cópia” de q em \mathbb{K})⁴;
- (v) em vista da observação acima, a correspondência $r \mapsto \{q \in \mathbb{Q} : q_{\mathbb{K}} < r\}$ define uma função injetora $\mathbb{K} \rightarrow \wp(\mathbb{Q})$;
- (vi) a completude de \mathbb{K} permite definir uma injeção $\psi : 2^{\omega} \rightarrow \mathbb{K}$ por meio da correspondência $f \mapsto \sup \left\{ \sum_{n=0}^m \frac{f(n)}{10^n} : m \in \omega \right\}$;
- (vii) a bijeção desejada segue, finalmente, pelo Teorema de Cantor-Bernstein, dado que $|\wp(\omega)| = |\wp(\mathbb{Q})| = |2^{\omega}|$. \square

⁴Um corpo ordenado em tais condições é xingado de **arquimediano**.

Já que, usualmente, define-se a **reta real** (denotada por \mathbb{R}) como *um corpo ordenado completo*, o teorema acima caracteriza a cardinalidade da reta independentemente de sua construção – e o fato de quaisquer dois corpos ordenados completos serem isomorfos apenas corrobora tal definição. Evidentemente, contextos com definições distintas (mas equivalentes) para o *quê* significa ser \mathbb{R} exigem argumentos diferentes⁵.

Em particular, chamando por $\mathbb{I} := \mathbb{R} \setminus \mathbb{Q}$ a **coleção dos números irracionais**, deve-se ter \mathbb{I} não-enumerável: o contrário permitiria expressar \mathbb{R} como reunião de dois conjuntos enumeráveis, o que levaria a uma violação do teorema anterior. Observe que com tal argumento, prova-se apenas que $|\omega| < |\mathbb{I}|$ e não que $|\mathbb{I}| = |\mathbb{R}|$: grosso modo, *poderia* ocorrer $|\omega| < |\mathbb{I}| < |\mathbb{R}|$, como em $2 < 3 < 5$. Contudo, não é o caso: oportunamente, veremos que $|X| = |X \times \omega|$ sempre que X é infinito, donde segue que se Z é não-enumerável e $B \subseteq Z$ é um subconjunto com B enumerável, então $|Z| = |Z \setminus B|$.

Exercício C.9. Supondo a identidade $|X| = |X \times \omega|$ válida para qualquer X infinito, prove a afirmação anterior. Dica: note que $A := Z \setminus B$ deve ser infinito (não-enumerável!), de modo que pelo Exercício C.24, $|Z| = |A \cup B| \leq |A \times B|$; para concluir, use o Exercício C.23, juntamente com a identidade assumida (e válida para $X := A$). ■

Observação C.3.3. Não custa reforçar: dois conjuntos não-enumeráveis podem ter *cardinalidades distintas*! É o caso de $\wp(\omega)$ e $\wp(\wp(\omega))$, por exemplo. △

Como última ilustração, recordemo-nos de que para um anel A (Definição B.1.34 + Observação B.1.35), $A[x]$ indica o **anel dos polinômios na indeterminada x com coeficientes em A** , cujos habitantes são **polinômios**, i.e., *expressões polinomiais* da forma

$$a_0 + a_1x + \dots + a_nx^n, \quad (\text{C.3})$$

em que $a_0, \dots, a_n \in A$ e $n \in \omega$. Explicitamente, pode-se definir

$$A[x] := \{p \in A^\omega : |\{n \in \omega : f(n) \neq 0_A\}| < |\omega|\},$$

i.e., a coleção das *sequências* da forma $\langle p(n) \rangle_{n \in \omega}$, onde cada *coeficiente* $p(n)$ é um membro do anel A e cujo *suporte* $\{n \in \omega : p(n) \neq 0_A\}$ é finito. Sob tal encarnação, uma expressão polinomial como (C.3) corresponde à sequência $a := \langle a(m) \rangle_{m \in \omega}$ em que $a(m) := a_m$ para $m \leq n$ e $a(m) := 0_A$ para $m > n$. Em particular, o **grau** de um polinômio $p \neq \langle 0 \rangle_{n \in \omega}$ fica bem definido, em virtude da boa ordenação de ω , como o menor $n \in \omega$ para o qual ocorre $p(n) \neq 0_A$ e $p(m) = 0_A$ para todo $m > n$. Portanto, ao chamar por $A[x]_n := \{p \in A[x] : p$ tem grau $\leq n\}$, tem-se $A[x] = \{\langle 0 \rangle_{n \in \omega}\} \cup \bigcup_{n \in \omega} A[x]_n$. Consequentemente:

Proposição C.3.4. Se A é anel infinito enumerável, então $A[x]$ é infinito enumerável.

Demonstração. Primeiro, note que $|A[x]_n| = |A^{n+}|$, i.e.: polinômios de grau zero estão em bijeção com A , polinômios de grau até 1 estão em bijeção com $A \times A$, polinômios de grau até 2 estão em bijeção com $A \times A \times A$, etc. A conclusão segue, por indução, da identidade $|A \times \omega| = |A|$, que será provada (oportunamente) para qualquer A infinito⁶. □

⁵Por exemplo: único completamento do grupo topológico $(\mathbb{Q}, +, 0)$; ordem total, completa, separável e sem extremos; objeto final na categoria dos corpos arquimedianos; etc. Profissionais costumam dizer que objetos nessa situação (pertencentes a *categorias* distintas, mas equivalentes em *certo sentido*) são *criptomorfos*.

⁶Também veremos que $|A \times A| = |A|$ para qualquer A infinito. Desse modo, uma generalização desta proposição seria “ $|A[x]| = |A|$ para qualquer anel infinito A ”.

E daí? Vejamos: em contextos introdutórios de Álgebra, costuma-se verificar que se $p \in \mathbb{Q}[x]$ é um polinômio não-nulo de grau n , então p tem, no máximo n raízes; logo $\mathbb{A} := \{r \in \mathbb{R} : r \text{ é raiz de } p \in \mathbb{Q}[x] \text{ não-nulo}\}$ é um subconjunto infinito enumerável de \mathbb{R} , posto que $\mathbb{Q}[x]$ também o é e $\mathbb{A} = \bigcup_{p \in \mathbb{Q}[x] \setminus \{\emptyset\}} R_p$, em que R_p indica o conjunto das raízes de p . Conclusão: existem tantos números reais que não são raízes de polinômios com coeficientes racionais quanto pontos na reta real! Assim, embora seja tecnicamente bem mais complicado provar que certos números reais são *transcendentess*⁷, o argumento cardinal utilizado acima permite não só mostrar que eles *existem*, como existem em abundância.

Observação C.3.5 (Teologia). Evidentemente, o argumento anterior não *exibiu* um elemento em $\mathbb{R} \setminus \mathbb{A}$, mas apenas provou que as afirmações “ $\mathbb{R} \setminus \mathbb{A}$ é enumerável” e “ \mathbb{A} é enumerável” não podem ser simultaneamente verdadeiras. Dado que a segunda sentença já foi verificada e, *implicitamente*, assumimos que a primeira deve ser verdadeira ou falsa (sem uma terceira opção), resulta que “ $\mathbb{R} \setminus \mathbb{A}$ é não-enumerável” é a única opção válida, donde em particular se conclui a não-vacuidade de $\mathbb{R} \setminus \mathbb{A}$. Esse tipo de demonstração é o arquétipo de diversos argumentos que utilizam o *Princípio do Terceiro Excluído*, uma importante peça da *metateoria* subjacente. △

C.4 Ordinais

Conforme *acordado*, para números naturais $m, n \in \omega$, a notação “ $m < n$ ” é apenas uma maneira menos *ofensiva* de escrever $m \in n$, resultado da construção de n como conjunto dos números naturais anteriores (Observação C.1.8) segundo a aplicação recursiva da operação sucessor $x \mapsto x_+$. Se acordássemos escrever “ $n < \omega$ ” em vez de $n \in \omega$, também fariam sentido as expressões “ $\omega < \omega_+$ ”, “ $\omega_+ < (\omega_+)_+$ ” e *assim sucessivamente*. Em certo sentido, elevar o *conjunto* ω ao *patamar* de número abriria os portões para o tratamento de números infinitos (também chamados de *transfinitos*). Por que alguém faria isso?

Originalmente, a percepção da natureza numérica de ω se deu com Georg Cantor durante os desenvolvimentos iniciais do que passaria a ser chamado de *Topologia Geral*. Grosso modo, Cantor considerou o procedimento de *excluir pontos isolados* de subconjuntos da reta e, ao fazer isso, deparou-se com certos conjuntos peculiares de pontos com os quais esse processo poderia ser repetido *iteradamente*, uma vez para cada natural n , de modo que mesmo depois de infinitos passos, ainda era possível repetir o processo.

Nesse sentido, o fenômeno é similar ao que se descreveu na Seção A.4: mesmo após construir $\wp^n(\emptyset)$ para cada $n \in \omega$, pode-se tomar $\wp^\omega(\emptyset) := \bigcup_{n \in \omega} \wp^n(\emptyset)$ para daí prosseguir com o processo, fazendo $\wp^{\omega+} := \wp(\wp^\omega(\emptyset))$, $\wp^{\omega++} := \wp(\wp^{\omega+}(\emptyset))$ e *assim por diante*. De certa forma, o símbolo “ ω ” registra que $\wp^\omega(\emptyset)$ se obtém pelos passos estritamente anteriores, enquanto o “ ω_+ ” em $\wp^{\omega+}(\emptyset)$ indica que este deve ser o resultado de se recomeçar o processo com $\wp^\omega(\emptyset)$. A abstração desse comportamento levou Cantor a definir os *números ordinais*, que hoje em dia são descritos por meio da definição dada por von Neumann.

Definição C.4.1. Um conjunto α é chamado de **número ordinal** (para os íntimos: ordinal) se

- (i) α é um **conjunto transitivo**, i.e., se para todo x , $x \in \alpha \Rightarrow x \subseteq \alpha$, e
- (ii) $\langle \alpha, \in \rangle$ é uma boa ordem.

¶

⁷Tais como e (de Euler) e π .

O leitor que nunca esbarrou com aspectos mais técnicos de Teoria dos Conjuntos certamente encara com estranheza a primeira condição, o que é natural, posto que a transitividade não costuma ter paralelos fora de contextos *conjuntistas*: “se X é um conjunto de *pontos* e $p \in X$, então nem faz sentido perguntar se $p \subseteq X$ ou não, pois p não é um conjunto, p é um ponto!” disse o geômetra. Contudo, como já foi adiantado no primeiro capítulo, a postura oficial adotada no texto é a de que *tudo é conjunto*: no caso, os *pontos* do geômetra são, oficialmente, *construídos* como conjuntos. Isto não deveria ser difícil de aceitar, dado que mesmo ω foi *obtido* a partir do “vazio”⁸. Em todo caso, ordinais são bem mais comuns do que parecem.

Exemplo C.4.2. $0 := \emptyset$ é ordinal, por vacuidade, enquanto $1 := \{0\}$, $2 := \{0, 1\}$ e $3 := \{0, 1, 2\}$ são ordinais. A proposição a seguir generaliza o fenômeno. ▲

Proposição C.4.3. Se α é ordinal, então $\alpha_+ := \alpha \cup \{\alpha\}$ é ordinal.

Demonstração. De fato, $\beta := \alpha_+$ é transitivo pois, se $\gamma \in \beta$, então $\gamma \in \alpha$ ou $\beta = \alpha$ e, de ambos os casos, decorre $\gamma \subseteq \alpha \cup \{\alpha\} := \beta$. Por outro lado, $\langle \beta, \in \rangle$ é uma boa ordem, já que, na prática, β é obtido a partir de $\langle \alpha, \in \rangle$ por meio do acréscimo de um maior elemento, como no próximo exercício. □

Exercício C.10. Seja $\langle \mathbb{P}, \leq \rangle$ uma boa ordem e $z \notin \mathbb{P}$. Para $\mathbb{P}_\# := \mathbb{P} \cup \{z\}$ e $x, y \in \mathbb{P}_\#$, defina $x \prec y$ se, e somente se, $x, y \in \mathbb{P}$ e $x < y$ ou $x \in \mathbb{P}$ e $y = z$. Mostre que $\langle \mathbb{P}_\#, \preceq \rangle$ é uma boa ordem. ■

Exercício C.11. Mostre que todo elemento de ω é um ordinal. Dica: indução. ■

Em vista do exercício acima e das considerações que virão a seguir, a boa-ordenação de ω se revela como um corolário óbvio da próxima

Proposição C.4.4. Todo conjunto de ordinais é bem-ordenado pela relação \in .

Demonstração. Com efeito, se \mathcal{O} é uma coleção de ordinais não-vazia, então existe um ordinal $\alpha \in \mathcal{O}$, o que leva a dois casos: $\alpha \cap \mathcal{O} = \emptyset$ (e daí $\alpha = \min \mathcal{O}$) ou $\alpha \cap \mathcal{O} \neq \emptyset$ (e daí o menor elemento de $\alpha \cap \mathcal{O}$ será o menor elemento de \mathcal{O}). O leitor fica a cargo de verificar os detalhes. □

Exercício C.12. Mostre que ω é um ordinal. ■

Observação C.4.5. O leitor deve meditar sobre o exercício anterior: acabamos de mostrar que o conjunto dos números naturais merece ser tratado como um número (ordinal)! Isso está longe de ser trivial ou intuitivo⁹, de modo que pode ser psicologicamente reconfortante tomar algum tempo para *café e contemplação*. △

Como ficará claro ao longo das próximas proposições e exercícios, números ordinais são conjuntos muito bem comportados perante a *hierarquia da pertinência*. Em particular, escrevendo “ $\alpha < \beta$ ” como sinônimo de “ $\alpha \in \beta$ ” para ordinais α e β , vale a identidade já observada para números naturais.

Proposição C.4.6. Se β é ordinal, então $\beta = \{\alpha : \alpha \text{ é ordinal e } \alpha < \beta\}$.

⁸Nesse sentido, confira o Teorema C.5.16.

⁹Não há relatos da definição de números ordinais nas pinturas rupestres.

Demonstração. Primeiro, note que a identidade abrevia a seguinte afirmação: se β é ordinal, então $\alpha \in \beta$ se, e somente se, α é um ordinal com $\alpha < \beta$. Dito isso, é claro que se α é um ordinal satisfazendo $\alpha < \beta$, então $\alpha \in \beta$: esta é, justamente, a definição da notação “ $\alpha < \beta$ ”. É a recíproca que carece de demonstração, já que poderia ser o caso de que algum $\alpha \in \beta$ não fosse ordinal. Contudo, tal fenômeno não ocorre. \square

Exercício C.13. Seja α um ordinal. Mostre que se $x \in \alpha$, então x é ordinal. Dica: lembre-se de que $x \subseteq \alpha$, já que α é conjunto transitivo. \blacksquare

Observação C.4.7. A identidade $\beta = \{\alpha : \alpha \text{ é ordinal e } \alpha < \beta\}$ captura a ideia de que um ordinal β *registra* todos os ordinais estritamente anteriores. \triangle

Exercício C.14. Mostre que se $\langle \mathbb{P}, < \rangle$ é uma boa ordem, então $\langle \mathbb{P}, < \rangle$ é uma ordem total. Dica: $\{x, y\} \neq \emptyset$. \blacksquare

Exercício C.15. Sejam α, β e γ ordinais.

- a) Mostre que $\alpha \not< \alpha$ sem usar o Axioma da Fundação. Dica: se $\alpha \in \alpha$, então um dos axiomas de ordem estrita com respeito à relação \in seria violado.
- b) Mostre que se $\beta \not\subseteq \alpha$, então $\alpha \in \beta$. Dica: $\beta \setminus \alpha$ é um subconjunto não-vazio de β e, por isso, deve ter um menor elemento, digamos $\gamma \in \beta$; observe então que $\gamma \subseteq \alpha$ (pela minimalidade de γ) e $\alpha \subseteq \gamma$ (o contrário daria $\delta \in \alpha$ com $\delta \notin \gamma$, e daí $\gamma \in \delta$ ou $\gamma = \delta$ pelo exercício anterior). \blacksquare

Teorema C.4.8 (Tricotomia para ordinais). *Para ordinais α e β , um e somente um dos casos a seguir ocorre: $\alpha = \beta$, $\alpha < \beta$ ou $\beta < \alpha$.*

Demonstração. Note que $\alpha \cap \beta$ é um ordinal (Exercício C.34) que satisfaz $\alpha \cap \beta \subseteq \alpha$ e $\alpha \cap \beta \subseteq \beta$. Porém, não pode ocorrer simultaneamente $\alpha \cap \beta \subsetneq \alpha$ e $\alpha \cap \beta \subsetneq \beta$: se ocorresse, o exercício anterior daria $\alpha \cap \beta \in \alpha$ e $\alpha \cap \beta \in \beta$, i.e., $\alpha \cap \beta \in \alpha \cap \beta$, absurdo. \square

Os *números cardinais*, representantes canônicos das classes de cardinalidade, serão definidos como tipos especiais de ordinais.

Definição C.4.9. Um ordinal α é dito

- (i) **sucessor** se existe $\beta < \alpha$ tal que $\alpha = \beta_+$,
- (ii) **limite** se para todo $\beta < \alpha$ ocorrer $\beta_+ < \alpha$,
- (iii) **cardinal ordinal inicial** se não existe $\beta < \alpha$ com uma bijeção $\beta \rightarrow \alpha$. \P

Assim, todos os números naturais n com $n \neq 0$ são sucessores, bem como o ordinal ω_+ . Por outro lado, tanto 0 quanto ω são ordinais limite: 0 é por vacuidade, ao passo que ω é limite pois, se $n \in \omega$, então $n_+ \in \omega$ (e $\omega \in \omega$ não pode ocorrer). Finalmente, todo $\alpha \leq \omega$ é inicial, enquanto ω_+ não é. De fato:

- (i) se $n \in \omega$ não fosse inicial, haveria $m < n$ e uma bijeção entre m e n , o que não pode ocorrer, já que m seria então um subconjunto próprio de n dotado de uma bijeção com n (reveja a demonstração do Teorema C.1.6);
- (ii) se ω não fosse inicial, então existiria uma bijeção $\omega \rightarrow n$ para algum $n \in \omega$, o que levaria a concluir que n tem uma bijeção com um subconjunto próprio;

(iii) por fim, ω_+ não é um inicial pois existe uma bijeção entre ω e ω_+ , como descrito pela já desgastada anedota do *Hotel de Hilbert*¹⁰.

Observação C.4.10. A definição de ordinal inicial não torna óbvia a resposta para a pergunta: existe algum que seja não-enumerável? Assim como ω_+ não é inicial (por estar em bijeção com ω), também não são iniciais $\omega_+, \omega_{++}, \omega_{+++}, \dots$. Fazendo $\omega + 0 := \omega$ e $\omega + (n_+) := (\omega + n)_+$ para cada $n \in \omega$, não é difícil perceber que $\omega + n$ é um ordinal sucessor enumerável, de modo que mesmo

$$\omega + \omega := \bigcup_{n \in \omega} \omega + n$$

se revela um ordinal (confira o Exercício C.35), enumerável (por ser reunião enumerável de conjuntos enumeráveis). Em particular, $\omega + \omega$ é um ordinal limite (pois não é sucessor de seus antecessores) que não é inicial (por estar em bijeção com diversos antecessores). Pode-se repetir o processo para obter $\omega + \omega + \omega$, $\omega + \omega + \omega + \omega$ e assim, sucessivamente, até que se chegue a $\omega \cdot \omega$, que por ser reunião enumerável de conjuntos enumeráveis ainda é enumerável e, portanto, não é inicial! O leitor incansável já deve ter cogitado $\omega \cdot \omega \cdot \omega$, \dots , $\omega \cdot \dots \cdot \omega$ até algo que possa ser xingado de ω^ω , mas sem sucesso: como reunião enumerável de conjuntos enumeráveis é enumerável, ainda não se excedeu a *cardinalidade* de ω ! O mesmo fenômeno ocorre com ω^{ω^ω} , $\omega^{\omega^{\omega^\omega}}$, etc.

Grosso modo, uma das formas de escapar do peggioso Corolário C.2.9 (que estabeleceu o resultado sobre reuniões enumeráveis) é considerar *todos* os ordinais enumeráveis simultaneamente, de forma *intencional* (em vez da *extensional*, que tenta descrever *cada um deles*)¹¹, o que será feito na próxima seção (confira o Exemplo C.5.5). \triangle

Em certo sentido, com tal definição, um ordinal inicial é o primeiro a exceder uma certa noção de *cardinalidade* anterior e, justamente por isso, usa-se *ele* para representar a sua própria *cardinalidade*. É por tal razão que, futuramente, fará sentido xingá-los de cardinais.

Exercício C.16. Para ordinais α e β , escreva “ $\alpha \leq \beta$ ” como abreviação para “ $\alpha < \beta$ ou $\alpha = \beta$ ”. Mostre que $\alpha \leq \beta$ se, e somente se, $\alpha \subseteq \beta$. Conclua que $\alpha = \beta$ se, e somente se, $\alpha \leq \beta$ e $\beta \leq \alpha$. \blacksquare

Proposição C.4.11. Se α é um ordinal, então existe um único ordinal inicial α_0 com uma bijeção sobre α tal que $\alpha_0 \leq \gamma$ para todo ordinal γ que também tem bijeção com α .

Demonstração. Considerando o subconjunto $A := \{\beta < \alpha : \text{existe bijeção } \beta \rightarrow \alpha\}$, basta tomar $\alpha_0 := \alpha$ se $A = \emptyset$, ou $\alpha_0 := \min A$ caso contrário. A minimalidade de α_0 garante que se trata de um ordinal inicial, donde o restante segue pelo exercício anterior. \square

Definição C.4.12. Em vista da última proposição, torna-se lícito declarar $|\alpha| := \alpha_0$ como **único cardinal** de α , precisamente o menor ordinal em bijeção com α . \P

Exercício C.17. Mostre que se α é um ordinal, então $|\alpha| \leq \alpha$, com igualdade válida se, e somente se, α é inicial. \blacksquare

¹⁰Faz-se $0 \mapsto \omega$ e $n \mapsto n_+$ para cada $n \in \omega \setminus \{0\}$, o que define uma bijeção $\omega \rightarrow \omega \cup \{\omega\}$.

¹¹Definições intencionais são aquelas em que um objeto é *definido* a partir de certas propriedades. Definições extensionais são aquelas que explicitam a *extensão* do objeto, i.e., suas componentes. No caso específico de conjuntos: definições intencionais são da forma $\{x : P(x)\}$, enquanto as extensionais consistem em listar os elementos do conjunto, como em $\{0, 1, 2, \dots, n\}$.

Exercício C.18. Mostre que se κ é ordinal inicial infinito, então $|\kappa_+| = \kappa$. ■

Observação C.4.13. Em particular, se λ é outro ordinal munido de uma bijeção $f: \lambda \rightarrow \alpha$, então vale $|\lambda| = |\alpha|$: por meio de composições, obtém-se facilmente uma bijeção entre $|\lambda|$ e α , donde a minimalidade de $|\alpha|$ acarreta $|\alpha| \leq |\lambda|$; analogamente, infere-se $|\lambda| \leq |\alpha|$, donde resulta a igualdade desejada.

Assim, se para um conjunto X qualquer fosse possível estabelecer uma bijeção com *algum* ordinal, digamos γ , então faria sentido definir o *número cardinal* de X , denotado $|X|$, como sendo $|\gamma|$, já que qualquer outro ordinal levaria ao mesmo ordinal inicial. É justamente isso o que será feito pelo *Teorema da Boa Ordenação de Zermelo* no próximo capítulo, uma das manifestações do Axioma da Escolha: mostraremos que para cada X existe *um* número ordinal γ em bijeção com X . △

A demonstração do Teorema da Boa Ordenação, mencionado acima, consiste em usar o Axioma da Escolha para *escolher*, recursivamente, *elementos* $x_0, x_1, x_2, \dots, x_\alpha, \dots$ num conjunto $X \neq \emptyset$, conforme α varia na classe $\text{ORD} := \{\alpha : \alpha \text{ é ordinal}\}$. Isto sugere, inicialmente, duas perguntas: ORD é um conjunto e, se for, é bem ordenado?

Proposição C.4.14. *Toda subclasse não-vazia $S \subseteq \text{ORD}$ tem \in -menor elemento, i.e., existe $\alpha \in S$ tal que $\alpha \leq \beta$ para todo $\beta \in S$.*

Demonstração. Se \mathcal{P} é uma *propriedade* tal que $S := \{x : \mathcal{P}(x)\} \subseteq \text{ORD}$ e existe $\alpha \in S$, então $T := \{\beta \in \alpha : \mathcal{P}(\beta)\}$ satisfaz $\min S = \min T$ (se $T \neq \emptyset$) ou $\min S = \alpha$ (se $T = \emptyset$). □

Corolário C.4.15 (Paradoxo de Burali-Forti). *ORD é classe própria.*

Demonstração. Se ORD fosse um conjunto, então $\langle \text{ORD}, \in \rangle$ seria uma boa ordem. Como todo ordinal é um conjunto de ordinais, resultaria que ORD é transitivo e, portanto, um ordinal, o que violaria o primeiro item do Exercício C.15, i.e., $\text{ORD} \in \text{ORD}$. □

A princípio, a constatação de que ORD é uma classe própria impediria a utilização *imediata* do Teorema da Recursão para, digamos, obter uma (única) função \mathcal{G} -recursiva $\mathcal{F}: \text{ORD} \rightarrow \mathbb{V}$ a partir de uma função de classe $\mathcal{G}: \mathbb{V} \rightarrow \mathbb{V}$ (confira o Teorema B.3.10). Mas isto é contornável.

Teorema C.4.16 (Recursão em ordinais). *Seja $\mathcal{G}: \mathbb{V} \rightarrow \mathbb{V}$ uma função de classe. Então existe uma única função de classe \mathcal{G} -recursiva em ORD , i.e., uma função de classe $\mathcal{F}: \text{ORD} \rightarrow \mathbb{V}$ tal que para todo ordinal α se verifica $\mathcal{F}(\alpha) = \mathcal{G}(\langle \mathcal{F}(\beta) : \beta < \alpha \rangle)$.*

Demonstração. Como cada ordinal α é um conjunto bem ordenado pela relação de pertinência, o Teorema B.3.5 garante que existe uma única função \mathcal{G} -recursiva em α , digamos t_α . Com isso em mente, considera-se $\mathcal{F} := \bigcup_{\alpha \in \text{ORD}} t_{\alpha+}$, o que corresponde, explicitamente, ao seguinte: $\langle x, y \rangle \in \mathcal{F}$ se, e somente se, $x \in \text{ORD}$ e $y = t_{x+}(x)$. Pela existência e unicidade das funções \mathcal{G} -recursivas, segue que a classe \mathcal{F} é, de fato, uma função de classe da forma $\text{ORD} \rightarrow \mathbb{V}$. Por construção, para cada ordinal α se tem

$$\mathcal{F}(\alpha) = t_{\alpha+}(\alpha) = \mathcal{G}(\langle t_{\alpha+}(\beta) : \beta < \alpha \rangle)$$

em virtude da \mathcal{G} -recursividade da função $t_{\alpha+}$. Como $\beta < \alpha$ acarreta $\beta_+ < \alpha_+$, resulta que $t_{\alpha+}|_{\beta_+} = t_{\beta_+}$, donde se infere $t_{\alpha+}(\beta) = t_{\beta_+}(\beta) = \mathcal{F}(\beta)$ e, portanto,

$$\mathcal{F}(\alpha) = \mathcal{G}(\langle \mathcal{F}(\beta) : \beta < \alpha \rangle),$$

como desejado. Os argumentos usados no Lema B.3.4 para mostrar a unicidade de funções recursivas podem ser reciclados aqui, em virtude da Proposição C.4.14. □

C.5 Adiável: algumas recursões longas

Um dos grandes problemas da recursão é que *precisa-se saber recursão para aprender recursão*¹². Dado que já se sabe recursão, convém ilustrar um pouco mais o seu uso, a fim de que se possa *aprender* recursão. Todavia, leitores ávidos para discutir o Axioma da Escolha podem adiar a leitura desta seção, e retornar nas poucas ocasiões em que algum resultado depender das “construções” apresentadas aqui.

Frequentemente, as recursões podem ser realizadas sem apelar explicitamente para fórmulas funcionais (ou *super oráculos* da forma $\mathbb{V} \rightarrow \mathbb{V}$). Em tais contextos, costuma-se ter:

- (i) um conjunto Z , cujos elementos serão os *blocos* da construção;
- (ii) um número ordinal γ grande, ao longo do qual os *passos* da indução serão registrados;
- (iii) em vez de uma fórmula funcional em z (ou um super oráculo), uma função *oráculo* \mathcal{O} que sabe todas as *ramificações* possíveis de construção a partir dos passos anteriores indexados em γ .

Neste caso, o *oráculo* é uma função $\mathcal{O}: Z^{<\gamma} \rightarrow Z$, em que $Z^{<\gamma} := \bigcup_{\beta < \gamma} Z^\beta$, ou seja: uma correspondência que associa cada upla $\langle z_\alpha : \alpha < \beta \rangle$ com $\beta < \gamma$ a um *objeto* $\mathcal{O}(\langle z_\alpha : \alpha < \beta \rangle) \in Z$. Daí, como nos casos anteriores de definições por recursão, uma função $f: \gamma \rightarrow Z$ satisfazendo

$$f(\beta) = \mathcal{O}(\langle f(\alpha) : \alpha < \beta \rangle)$$

para todo $\beta < \gamma$ será chamada de **\mathcal{O} -recursiva** em γ .

Proposição C.5.1. *Dados um conjunto Z , um ordinal γ e uma função $\mathcal{O}: Z^{<\gamma} \rightarrow Z$, existe uma única função \mathcal{O} -recursiva em γ .*

Demonação. No Teorema B.3.5, faça $\mathbb{W} := \gamma$ e tome a fórmula $\mathcal{F}(z, y)$ dada por “ $z \in Z^{<\gamma}$ e $y = \mathcal{O}(z)$, ou $y = \emptyset$ caso contrário”, que é funcional em z . \square

Exercício C.19. Adapte a prova da Proposição C.5.1 para o caso em que γ e $Z^{<\gamma}$ são substituídos, respectivamente, por uma boa ordem $\langle \mathbb{B}, \leq \rangle$ e por $Z^{<\mathbb{B}} := \bigcup_{b \in \mathbb{B}} Z^{\downarrow b}$, onde $\downarrow b := \{a \in \mathbb{B} : a < b\}$. \blacksquare

O exercício anterior parece sugerir que ordinais são substitutos canônicos de conjuntos bem ordenados, não é mesmo?

Exemplo C.5.2 (Ordinais como representantes). Um **isomorfismo de ordem** entre duas ordens $\langle \mathbb{P}, \leq \rangle$ e $\langle \mathbb{P}', \preceq \rangle$ é uma bijeção $f: \mathbb{P} \rightarrow \mathbb{P}'$ tal que

$$p < q \Leftrightarrow f(p) \prec f(q)$$

para quaisquer elementos $p, q \in \mathbb{P}$. Dizemos que duas ordens $\langle \mathbb{P}, \leq \rangle$ e $\langle \mathbb{P}', \preceq \rangle$ são **isomórfas** caso exista um isomorfismo entre elas, o que será denotado com $\langle \mathbb{P}, \leq \rangle \cong \langle \mathbb{P}', \preceq \rangle$. Obviamente toda ordem é isomorfa a si mesma por meio da função identidade. No caso de ordinais, este é o único isomorfismo possível.

¹²Também é comum encontrar tal piada enunciada de modo menos preciso, como em “antes de *saber* recursão, *precisa-se saber* recursão”.

Lema C.5.3. Sejam α e β ordinais. Se $f: \alpha \rightarrow \beta$ é um isomorfismo de ordens, então $f = \text{Id}_\alpha$ e $\alpha = \beta$.

Demonstração. Por fatores psicológicos, observe que como f é isomorfismo de ordem e $0 \leq \gamma$ para todo $\gamma \in \alpha$, resulta que $f(0) \leq f(\gamma)$ e, por f ser sobrejetora, segue que $f(0) = \min \beta$, ou seja: $f(0) = 0$. De modo análogo, $f(1) = 1$, $f(2) = 2$ e assim por diante.

A rigor, a prova é por indução (Proposição B.2.15): dado $\xi \in \alpha$, supõe-se $f(\gamma) = \gamma$ para todo $\gamma < \xi$ a fim de concluir que $f(\xi) = \xi$. Ora, se $f(\xi) \neq \xi$, então $f(\xi) > \xi$: se ocorresse $f(\xi) < \xi$, teria-se $f(f(\xi)) = f(\xi)$, o que contradiz a injetividade de f . Por outro lado, a sobrejetividade de f assegura $\gamma \in \alpha$ com $f(\gamma) = \xi$. Pelas suposições feitas, deve-se ter $\gamma > \xi$, mas $f(\gamma) = \xi < f(\xi)$, contrariando a hipótese de que f é isomorfismo. Portanto, $f(\xi) = \xi$ para todo $\xi \in \alpha$ e, por f ser sobre, conclui-se $f = \text{Id}_\alpha$ e $\alpha = \beta$. \square

Teorema C.5.4. Se $\langle \mathbb{W}, \leq \rangle$ é uma boa ordem, então existe um único ordinal α com $\langle \mathbb{W}, < \rangle \cong \langle \alpha, \in \rangle$. Em particular, o isomorfismo $\mathbb{W} \rightarrow \alpha$ é único.

Demonstração. Note que se α e β são ordinais isomorfos a \mathbb{W} , então α e β são isomorfos entre si (Exercício C.32) e, pelo lema anterior, $\alpha = \beta$. Além disso, se $f, g: \mathbb{W} \rightarrow \alpha$ são dois isomorfismos, então $f \circ g^{-1}: \alpha \rightarrow \alpha$ é um isomorfismo (Exercício B.22) e, portanto, $f \circ g^{-1} = \text{Id}_\alpha$, resultando $f = g$. Isso dá conta da unicidade.

Para provar a existência, a ideia é definir $f(w_0) := 0$, $f(w_1) := 1$ e assim por diante, onde $w_0 := \min \mathbb{W}$ e $w_1 := \min (\mathbb{W} \setminus \{w_0\})$. Infelizmente, como não se conhece *a priori* um conjunto X que contenha todos os objetos que serão utilizados na recursão, precisa-se apelar para a versão geral do Teorema B.3.5: a fórmula $\text{im}(x, y)$ dada por “ $y = \text{im}(x)$ se x é uma função, $y = \emptyset$ caso contrário” é, claramente, funcional, de modo que existe uma única função im -recursiva em \mathbb{W} , i.e., com $\text{dom}(f) = \mathbb{W}$ e $f(w) = \text{im}(\langle f(v) : v < w \rangle)$ para todo $w \in \mathbb{W}$. Com isso, basta ver que $\text{im}(f)$ é um ordinal e f é um isomorfismo.

Primeiro, $\text{im}(f)$ é um ordinal: como $\langle f(v) : v < w \rangle$ é uma função para cada $w \in \mathbb{W}$, tem-se $f(w) = \{f(v) : v < w\}$, donde é relativamente fácil verificar, por indução em \mathbb{W} , que $\text{im}(f)$ é um conjunto transitivo de ordinais e, portanto, é um ordinal. Segundo: a construção de f obriga que se $v < w$ em \mathbb{W} , então $f(v) \in f(w)$. Portanto, f é um isomorfismo de ordens (confira o Exercício B.22). \square

Exercício C.20. Complete a demonstração anterior. Dica: conjuntos de ordinais são bem ordenados e, se $y \in f(v)$ com $v < w$, então $y = f(u)$ para $u < v$. \blacksquare

Uma vez que a cada boa ordem corresponde um único ordinal isomorfo a ela, faz sentido dar um nome a este animal: denota-se por $\text{ord}(\mathbb{W})$ o único ordinal isomorfo a uma boa ordem \mathbb{W} , chamado de **tipo de ordem** de \mathbb{W} . O tipo de ordem de um conjunto bem ordenado captura a *essência* de sua ordem, tal qual o número cardinal de um conjunto X busca capturar a sua *cardinalidade*. Uma aplicação interessante dessa ideia é dada no próximo exemplo. \blacktriangle

Exemplo C.5.5 (Opcional, mas nem tanto: *número de Hartogs*). Fixado um conjunto X , o teorema anterior garante que existe um conjunto $H(X)$ tal que $\alpha \in H(X)$ se, e somente se, α é um ordinal com $|\alpha| \leq |X|$. De fato, o Axioma da Separação assegura que

$$\mathcal{W}(X) := \{R \subseteq X \times X : \langle \text{dom}(R), R \rangle \text{ é boa ordem}\}$$

é um conjunto *bona fide*.

Daí, ao considerar a função $\mathcal{O}: \mathcal{W}(X) \rightarrow \text{ORD}$ que faz $\mathcal{O}(R) := \text{ord}(\langle \text{dom}(R), R \rangle)$ (bem definida pelo teorema anterior), segue que o conjunto $H(X)$ procurado é, tão somente, a imagem da função \mathcal{O} : por um lado, se α é ordinal que satisfaz $|\alpha| \leq |X|$, então existe uma injecção $f: \alpha \rightarrow X$ que permite definir $R_f := \{\langle f(\delta), f(\gamma) \rangle : \delta \leq \gamma < \alpha\}$, que por sua vez é uma boa ordem sobre $\text{im}(f)$ satisfazendo $\text{ord}(\langle \text{im}(f), R_f \rangle) = \alpha$, i.e., $\alpha \in H(X)$; reciprocamente, se $\alpha \in H(X)$, então α é isomorfo a um subconjunto *bem ordenável* de X , donde em particular resulta $|\alpha| \leq |X|$. E daí?

Teorema C.5.6. *Para um conjunto X fixado, $H(X)$ é o menor ordinal inicial satisfazendo $H(X) \not\leq |X|$, usualmente xingado como **número de Hartogs de X** .*

Demonstração. A rebuscada discussão anterior serve apenas para assegurar que ao declarar $H(X) := \{\alpha \in \text{ORD} : |\alpha| \leq |X|\}$, define-se um conjunto – e não apenas uma classe própria. Com isso, basta checar as condições na definição de ordinal inicial:

- (i) $H(X)$ é ordinal, já que
 - ✓ é bem ordenado pela relação de pertinência (por ser um conjunto de ordinais), e
 - ✓ é transitivo, pois para $\alpha \in \beta$ com $\beta \in H(X)$, tem-se $\alpha \subseteq \beta$ e $|\beta| \leq |X|$, acarretando $|\alpha| \leq |X|$, i.e., $\alpha \in H(X)$;
- (ii) $H(X)$ é inicial, pois se existisse bijeção¹³ $\alpha \rightarrow H(X)$ para $\alpha \in H(X)$, resultaria que $|H(X)| \leq |X|$, i.e., $H(X) \in H(X)$.

Em particular, $H(X) = |H(X)|$. Por fim, se κ é ordinal inicial satisfazendo $\kappa \not\leq |X|$, então pela tricotomia de ordinais, restam apenas $H(X) = \kappa$ ou $H(X) < \kappa$, já que $\kappa < H(X)$ daria $\kappa \leq |X|$ (lembre-se: para ordinais, $\gamma < \beta \Leftrightarrow \gamma \in \beta$). Portanto, $H(X) \leq \kappa$. \square

Depois de toda essa discussão, convém se indagar: o que seria $H(\omega)$? Explicitamente, trata-se do conjunto de todos os ordinais enumeráveis, que por sua vez *representam* todas as formas possíveis de *enfileirar* subconjuntos de ω . Porém, mais importante do que isso é a segunda observação: $H(\omega)$ é o menor ordinal inicial a não *injetar* em ω . Como $\omega \in H(\omega)$ por definição e ambos são iniciais, resulta $\omega = |\omega| < |H(\omega)| = H(\omega)$. \blacktriangle

Definição C.5.7. $\omega_1 := H(\omega)$. ¶

Note que ω_1 não é o mesmo animal que $\omega_+ := \omega \cup \{\omega\}$: dado que ω_+ ainda é enumerável, tem-se $\omega_+ \in H(\omega)$. Também ocorre $\omega_{++}, \omega_{+++}, \dots < \omega_1$: de modo geral, qualquer ordinal obtido recursivamente da *operação sucessor* em enumeráveis passos resulta num ordinal que ainda não é ω_1 , já que tais ordinais (enumeráveis) apenas representam as diferentes formas de se enfileirar ω . Mais geralmente, vale a seguinte

Proposição C.5.8. *Se $\langle \alpha_n \rangle_{n \in \omega}$ é sequência de ordinais com $\alpha_n < \omega_1$ para todo $n \in \omega$, então $\sup_{n \in \omega} \alpha_n < \omega_1$.*

Demonstração. A existência de supremos se verifica para qualquer família de ordinais (confira o Exercício C.35): no caso, $\alpha := \bigcup_{n \in \omega} \alpha_n$ é um ordinal que satisfaz $\alpha_n \leq \alpha$ para todo $n \in \omega$, bem como $\alpha \leq \beta$ para todo ordinal β tal que $\alpha_n \leq \beta$ para todo $n \in \omega$. Ora, como $\alpha_n < \omega_1$ é apenas um modo de dizer que $|\alpha_n| \leq |\omega|$, resulta que α é reunião enumerável de conjuntos enumeráveis e, portanto, $|\alpha| \leq |\omega|$, ou seja, $\alpha < \omega_1$. \square

¹³Bastaria injecção.

Exemplo C.5.9 (Opcional: σ -álgebras). Uma σ -álgebra num conjunto X é uma família \mathcal{A} de subconjuntos de X tal que $\emptyset, X \in \mathcal{A}$, $X \setminus A \in \mathcal{A}$ sempre que $A \in \mathcal{A}$ e $\bigcup \mathcal{M} \in \mathcal{A}$ sempre que $\mathcal{M} \subseteq \mathcal{A}$ com $|\mathcal{M}| \leq |\omega|$. Tais animais são comuns em *Teoria da Medida*, subcampo da Análise que se ocupa das diversas *teorias de integração*. Uma situação típica nesse contexto consiste em obter uma σ -álgebra $\sigma(\mathcal{E})$ a partir de uma família \mathcal{E} de subconjuntos de X , tão pequena quanto possível (pois $\wp(X)$ é uma σ -álgebra que satisfaz $\mathcal{E} \subseteq \wp(X)$). Por um lado, isto é quase trivial:

Proposição C.5.10. Se $\mathcal{F} \neq \emptyset$ é uma família de σ -álgebras sobre X , então $\bigcap \mathcal{F}$ é uma σ -álgebra em X .

Demonstração. Basta observar que por cada σ -álgebra $\mathcal{F} \in \mathcal{F}$ satisfazer as condições exigidas, necessariamente $\bigcap \mathcal{F}$ também as satisfaz. \square

Exercício C.21. Complete a prova da proposição anterior. \blacksquare

Notemos então que a coleção $\mathcal{F} := \{\mathcal{A} : \mathcal{A} \text{ é } \sigma\text{-álgebra em } X \text{ e } \mathcal{E} \subseteq \mathcal{A}\}$ é uma família nas condições da proposição anterior, donde segue que $\sigma(\mathcal{E}) := \bigcap \mathcal{F}$ é uma σ -álgebra com a seguinte propriedade: além de ocorrer $\mathcal{E} \subseteq \sigma(\mathcal{E})$, deve-se ter $\sigma(\mathcal{E}) \subseteq \mathcal{A}$ sempre que \mathcal{A} for uma σ -álgebra com $\mathcal{E} \subseteq \mathcal{A}$. Rápido, prático e limpo. Contudo, tal argumento não ajuda a responder o seguinte: quem são os elementos de $\sigma(\mathcal{E})$?

- (0) Certamente, $\mathcal{E}_0 := \mathcal{E} \cup \{\emptyset, X\} \subseteq \sigma(\mathcal{E})$, pois isto é o mínimo que se espera.
- (1) Como $\sigma(\mathcal{E})$ deve ser uma σ -álgebra, também deveria ocorrer $\mathcal{E}_1 \subseteq \sigma(\mathcal{E})$, onde

$$\mathcal{E}_1 := \left\{ \bigcup \mathcal{M} : \mathcal{M} \subseteq \mathcal{E}_0 \text{ e } |\mathcal{M}| \leq \omega \right\} \cup \{X \setminus E : E \in \mathcal{E}_0\},$$

pois $\sigma(\mathcal{E})$ deve ser “fechada” por reuniões enumeráveis e complementos.

- (2) Embora se tenha $\mathcal{E}_0 \subseteq \mathcal{E}_1$, dificilmente ocorrerá $\sigma(\mathcal{E}) = \mathcal{E}_1$, já que $\sigma(\mathcal{E})$ também deve ser fechada por reuniões enumeráveis e complementos dos “novos membros” trazidos por \mathcal{E}_1 ; em outras palavras, deve-se ter $\mathcal{E}_2 \subseteq \sigma(\mathcal{E})$, onde

$$\mathcal{E}_2 := \left\{ \bigcup \mathcal{M} : \mathcal{M} \subseteq \mathcal{E}_1 \text{ e } |\mathcal{M}| \leq \omega \right\} \cup \{X \setminus E : E \in \mathcal{E}_1\}.$$

Num primeiro momento, parece claro que $\sigma(\mathcal{E}) = \bigcup_{n \in \omega} \mathcal{E}_n$, mas não precisa ser o caso! De fato, se $E_n \in \mathcal{E}_n$ para cada $n \in \omega$, não há razão para supor que $\bigcup_{n \in \omega} E_n \in \bigcup_{n \in \omega} \mathcal{E}_n$: explicitamente, tal pertinência significaria que existe $m \in \omega$ tal que $\bigcup_{n \in \omega} E_n \in \mathcal{E}_m$! Dessa forma, somos obrigados a dar o ω -ésimo passo.

- (ω) Faz-se $\mathcal{E}_\omega := \bigcup_{n \in \omega} \mathcal{E}_n$, que será um subconjunto de $\sigma(\mathcal{E})$.
- (ω_+) Como $\sigma(\mathcal{E})$ deve ser fechada por reuniões enumeráveis e complementos dos elementos de \mathcal{E}_ω , faz sentido considerar

$$\mathcal{E}_{\omega_+} := \left\{ \bigcup \mathcal{M} : \mathcal{M} \subseteq \mathcal{E}_\omega \text{ e } |\mathcal{M}| \leq \omega \right\} \cup \{X \setminus E : E \in \mathcal{E}_\omega\}$$

e assim, sucessivamente, para ω_{++} , ω_{+++}, \dots , ω_ω, \dots , $\alpha < \omega_1$, qualquer que seja o α . Como fazer isso direito?

Com as notações da Proposição C.5.1, sejam $Z := \wp(\wp(X))$ e $\gamma := \omega_1$; no papel de oráculo, considere a função $\mathcal{O} : Z^{<\omega_1} \rightarrow Z$ que faz

- (i) $\mathcal{O}(\emptyset) := \mathcal{E}_0 \cup \{\emptyset, X\}$,
- (ii) $\mathcal{O}(\langle \mathcal{S}_\alpha : \alpha < \beta_+ \rangle) := \{\bigcup \mathcal{M} : \mathcal{M} \subseteq \mathcal{S}_\beta \text{ e } |\mathcal{M}| \leq \omega\} \cup \{X \setminus S : S \in \mathcal{S}_\beta\}$ para $\beta < \omega_1$,
- (iii) $\mathcal{O}(\langle \mathcal{S}_\alpha : \alpha < \beta \rangle) := \bigcup_{\alpha < \beta} \mathcal{S}_\alpha$ se $0 < \beta < \omega_1$ for ordinal limite.

Logo, existe uma única função \mathcal{O} -recursiva $\mathcal{E} : \omega_1 \rightarrow Z$, o que permite fazer $\mathcal{E}_\alpha := \mathcal{E}(\alpha)$. Para o que será feito adiante, convém observar o próximo

Lema C.5.11. *Se $\alpha < \beta < \omega_1$, então $\mathcal{E}_\alpha \subseteq \mathcal{E}_\beta$.*

Demonstração. Por ω_1 ser bem ordenado, pode-se argumentar por indução, no sentido da Proposição B.2.15, i.e., mostraremos que se $\gamma < \omega_1$ for tal que para todo $\beta < \gamma$, a desigualdade $\alpha < \beta$ acarretar $\mathcal{E}_\alpha \subseteq \mathcal{E}_\beta$, então para qualquer $\alpha < \gamma$ ocorre $\mathcal{E}_\alpha \subseteq \mathcal{E}_\gamma$: ora, se $\gamma := \delta_+$, então a \mathcal{O} -recursividade assegura

$$\mathcal{E}_\gamma = \mathcal{O}(\langle \mathcal{E}_\alpha : \alpha < \delta_+ \rangle) = \left\{ \bigcup \mathcal{M} : \mathcal{M} \subseteq \mathcal{E}_\delta \text{ e } |\mathcal{M}| \leq \omega \right\} \cup \{X \setminus E : E \in \mathcal{E}_\delta\},$$

acarretando $\mathcal{E}_\delta \subseteq \mathcal{E}_\gamma$ e $\mathcal{E}_\alpha \subseteq \mathcal{E}_\gamma$ sempre que $\alpha < \gamma$ (pois, neste caso, $\alpha = \delta$ ou $\alpha < \delta$, e daí $\mathcal{E}_\alpha \subseteq \mathcal{E}_\delta$ por hipótese); se, porém, γ for ordinal limite, então

$$\mathcal{E}_\gamma = \mathcal{O}(\langle \mathcal{E}(\alpha) : \alpha < \gamma \rangle) := \bigcup_{\alpha < \gamma} \mathcal{E}_\alpha,$$

onde segue a inclusão desejada. □

Finalmente:

Proposição C.5.12. *Nas notações anteriores, $\sigma(\mathcal{E}) = \bigcup_{\alpha < \omega_1} \mathcal{E}_\alpha$.*

Demonstração. Primeiro, note que $\Sigma := \bigcup_{\alpha < \omega_1} \mathcal{E}_\alpha$ é uma σ -álgebra que contém \mathcal{E} : por construção, $\mathcal{E} \subseteq \Sigma$, bem como $\emptyset, X \in \Sigma$; agora, se $\mathcal{M} \subseteq \Sigma$ com $|\mathcal{M}| \leq \omega$, então existe uma função sobrejetora $E : \omega \rightarrow \mathcal{M}$, o que permite escrever $\mathcal{M} = \{E_n : n \in \omega\}$, com cada $E_n \in \mathcal{E}_{\alpha_n}$ para algum $\alpha_n < \omega_1$, donde o lema anterior assegura que para $\alpha := \sup_{n \in \omega} \alpha_n < \omega_1$ tenha-se $E_n \in \mathcal{E}_\alpha$ para todo $n \in \omega$ e, consequentemente, $\bigcup \mathcal{M} \in \mathcal{E}_{\alpha_+}$; por fim, se $E \in \Sigma$, então $E \in \mathcal{E}_\alpha$ para algum $\alpha < \omega_1$ e, por conseguinte, $X \setminus E \in \mathcal{E}_{\alpha_+} \subseteq \Sigma$, já que $\alpha_+ < \omega_1$. Dado que $\sigma(\mathcal{E})$ é a menor σ -álgebra sobre X com $\mathcal{E} \subseteq \sigma(\mathcal{E})$, obtém-se $\sigma(\mathcal{E}) \subseteq \Sigma$. Para a inclusão oposta, basta observar que $\mathcal{E}_\alpha \subseteq \sigma(\mathcal{E})$ para todo $\alpha < \omega_1$, o que pode ser feito por indução, como no lema anterior. □

Embora pareça desnecessariamente intrincada, esta descrição *construtiva* da σ -álgebra gerada por \mathcal{E} permite estimar com *precisão* a cardinalidade de $\sigma(\mathcal{E})$. Em particular, no Capítulo E, isto será usado para mostrar que a *σ -álgebra de Borel em \mathbb{R} tem a mesma cardinalidade de \mathbb{R}* , o que, para leitores versados em Teoria da Medida, é suficiente para assegurar a existência de subconjuntos *Lebesgue-mensuráveis* que não são *Borel-mensuráveis* (confira os Exercícios E.24 e E.25). ▲

Em resumo: ω_1 , *a.k.a.* o primeiro ordinal não-enumerável, se obteve como a coleção de todos os ordinais com cardinalidade $\leq |\omega|$, i.e., $H(\omega) := \omega_1$. Como o subíndice sugere, a ideia é fazer $\omega_2 := H(\omega_1)$, $\omega_3 := H(\omega_2)$ e assim sucessivamente, *ad nauseam*.

Teorema C.5.13 (Recursão ordinal – em casos). *Se $\mathcal{C}_0(x, y)$, $\mathcal{C}_s(x, y)$ e $\mathcal{C}_l(x, y)$ são fórmulas funcionais em x , então existe uma fórmula $\mathcal{R}(x, y)$ funcional em x tal que*

- (i) $\mathcal{R}(0) = \mathcal{C}_0(0)$,
- (ii) $\mathcal{R}(\alpha_+) = \mathcal{C}_s(\mathcal{R}(\alpha))$ para todo ordinal α e
- (iii) $\mathcal{R}(\alpha) = \mathcal{C}_l(\langle \mathcal{R}(\beta) : \beta < \alpha \rangle)$ para $\alpha \neq 0$ ordinal limite.

Demonstração. Tudo se resume a considerar uma fórmula $\mathcal{G}(x, y)$ apropriada a fim de tomar \mathcal{F} como no Teorema C.4.16, de modo semelhante à definição do oráculo \mathcal{O} no exemplo anterior. Os detalhes ficam a cargo do leitor. \square

Definição C.5.14. Para um número ordinal α , define-se ω_α recursivamente como:

- (i) $\omega_0 := \omega$;
- (ii) $\omega_{\alpha_+} := H(\omega_\alpha)$;
- (iii) $\omega_\alpha := \sup_{\xi < \alpha} \omega_\xi$ se $\alpha \neq 0$ for um ordinal limite. ¶

Observação C.5.15 (Alephs). Em ZFC, um modo *alternativo* de indicar ω_α consiste em substituir a letra grega ômega (ω) pela letra hebraica *aleph* (\aleph), i.e., $\aleph_\alpha := \omega_\alpha$ para cada ordinal α . Apesar disso, a rigor, \aleph_α e ω_α cumprem papéis diferentes:

- (i) ω_{α_+} codifica todas as boas ordenações de (subconjuntos de) ω_α ;
- (ii) por sua vez, espera-se que \aleph_{α_+} seja um representante da *primeira cardinalidade maior* do que \aleph_α , no sentido de que se $|X| = \aleph_\alpha$ e $|X| < |Y|$, então $\aleph_{\alpha_+} \leq |Y|$.

Por essa razão, alephs costumam ser reservados para contextos que tratam explicitamente de cardinalidades e números cardinais. Em tempo: veremos que, em ZFC, todo conjunto infinito está em bijeção com um único aleph. Posto de outra forma: todo ordinal inicial infinito é da forma \aleph_α para algum α . No entanto, isto é assunto para o distante Capítulo E. \triangle

Encerraremos este capítulo de forma um tanto quanto poética, com a apresentação de um resultado que costuma surpreender quando visto pela primeira vez: quando interpretado da forma errada, ele nos diz a *origem* de todas as coisas (em ZFC). Para tanto, define-se

- (i) $\mathbb{V}_0 := \emptyset$,
- (ii) $\mathbb{V}_{\alpha+1} := \wp(\mathbb{V}_\alpha)$ se $\alpha \in \text{ORD}$ e
- (iii) $\mathbb{V}_\beta := \bigcup_{\alpha < \beta} \mathbb{V}_\alpha$ se $\beta \in \text{ORD}$ é um ordinal limite com $\beta > 0$.

Teorema C.5.16 (Big-Bang). *Dado um conjunto X , existe $\alpha \in \text{ORD}$ tal que $X \in \mathbb{V}_\alpha$. Em outras palavras, $\mathbb{V} = \bigcup_{\alpha \in \text{ORD}} \mathbb{V}_\alpha$.*

Para demonstrar que *tudo vem do vazio*, precisa-se estender o Axioma da Fundação para o contexto de classes.

Lema C.5.17. *Se C é uma classe não-vazia, então existe $x \in C$ tal que $x \cap C = \emptyset$.*

Demonstração. Como $C \neq \emptyset$, existe um conjunto $S \in C$. Se a classe $S \cap C$ for vazia, então acabou. Se não, então façamos $X := T \cap C$, onde $T := \bigcup_{n \in \omega} T_n$, com $T_0 := S$ e $T_{n+1} := \bigcup T_n$ para todo $n > 0$. Como $X = \bigcup_{n \in \omega} T_n \cap C$ e $T_0 \cap C \neq \emptyset$, segue que X é um conjunto não-vazio – é conjunto por ser subclasse do conjunto T (Exercício A.40). Logo, o Axioma da Fundação dá $x \in X$ tal que $x \cap X = \emptyset$, e este satisfaz a condição desejada: se existisse $y \in x \cap C$, teria-se $y \in T \cap C = X$, mostrando que $x \cap X \neq \emptyset$. \square

Demonstração do Teorema C.5.16. Seja $C := \{X : \forall \alpha \in \text{ORD} \ X \notin \mathbb{V}_\alpha\}$ e note, primeiramente, que $\emptyset \notin C$. Se a classe C não fosse vazia, o lema anterior permitiria tomar $X \in C$ com $X \cap C = \emptyset$. Logo, para cada $z \in X$ existe $\alpha_z \in \text{ORD}$ com $z \in \mathbb{V}_{\alpha_z}$. Como $\mathbb{V}_\alpha \subseteq \mathbb{V}_\beta$ se $\alpha \leq \beta$, segue que para $\gamma := \sup_{z \in X} \alpha_z$ vale $z \in \mathbb{V}_\gamma$ para todo $z \in X$ e, consequentemente, $X \subseteq \mathbb{V}_\gamma$, o que contraria o modo como X foi tomado, pois da última inclusão resulta $X \in \mathbb{V}_{\gamma+1}$. \square

Exercício C.22. Complete os detalhes da demonstração anterior. Dica: para mostrar que $\mathbb{V}_\alpha \subseteq \mathbb{V}_\beta$ sempre que $\alpha \leq \beta$, argumente por indução, no sentido da próxima proposição. ■

Proposição C.5.18 (Indução sobre ordinais). *Seja $\mathcal{P}(x)$ uma fórmula na variável x e suponha que para todo ordinal α se verifique a condição: $\forall \beta (\beta < \alpha \Rightarrow \mathcal{P}(\beta)) \Rightarrow \mathcal{P}(\alpha)$. Então $\mathcal{P}(\alpha)$ é verdadeira para todo ordinal α .*

Demonstração. De fato, se $S := \{\gamma \in \text{ORD} : \neg \mathcal{P}(\gamma)\} \neq \emptyset$, então é possível tomar $\alpha := \min S$ (Proposição C.4.14), donde segue que $\mathcal{P}(\beta)$ ocorre para todo $\beta < \alpha$, embora $\mathcal{P}(\alpha)$ não ocorra, contrariando a suposição sobre a fórmula \mathcal{P} . \square

Observação C.5.19 (*Ex nihilo nihil fit*). Recursões *muito longas*, i.e., indexadas pela classe ORD de *todos* os ordinais, costumam causar desconforto em leitores iniciantes – e com muita razão, já que elas dependem do flerte com as perigosas classes próprias. Parte da culpa é da literatura (e de quem propaga seus vícios, como eu), que insiste em utilizar o verbo “construir” nos contextos de recursão: de nossa experiência imobiliária, *construções* são procedimentos a partir dos quais coisas que não existiam passam a existir.

No caso do que se faz em Teoria dos Conjuntos e, mais geralmente, em Matemática Abstrata, seria mais correto utilizar o verbo “descrever”, já que as argumentações matemáticas ocorrem num vácuo abstrato desprovido de tempo: por mais que uma função $f: \mathbb{R} \rightarrow \mathbb{R}^3$ possa ser utilizada para *modelar* parte do trajeto de uma *partícula* no *espaço*, por exemplo, a função f não é a partícula nem o seu trajeto, mas apenas uma representação.

Nesse sentido, chamar o último teorema de Big-Bang foi uma armadilha para o leitor: em vez de *criar* o universo \mathbb{V} de todos os conjuntos a partir de \emptyset , o que se faz, na verdade, é descrevê-lo! Correndo o risco de atiçar a curiosidade do leitor para os perigos da *Metamatemática*: antes de iniciar toda a discussão sobre axiomas neste livro (ou em qualquer outro), fixamos tacitamente uma *coleção* (ingênua) \mathbb{V} de entidades não definidas que decidimos xingar de *conjuntos*, munida de uma relação binária \in apelidada como *pertinência*, e daí analisamos as consequências de se assumir que tais entidades satisfazem certas afirmações (*a.k.a.* ZFC).

Ao definir ORD com as cláusulas da Definição C.4.1, por exemplo, o que fizemos foi *separar* uma subcoleção (ingênua) da coleção \mathbb{V} , ao passo que a Proposição C.4.6 descreveu as entidades na subcoleção em termos de entidades *anteriores*: não é preciso *definir* todos os ordinais menores do que α a fim de ter α ; α existe, e é descritível como a coleção dos ordinais menores. Da mesma forma, o Teorema C.5.16 apenas mostra que se \mathbb{V} é a coleção (ingênua) que escolhemos para desenvolver ZFC, então ela se descreve como a reunião das entidades \mathbb{V}_α , conforme α percorre a subcoleção (ingênua) ORD: explicitamente, cada entidade X em \mathbb{V} *pertence* a alguma entidade da forma \mathbb{V}_α .

Em última análise, a impressão de circularidade é sintoma inevitável de se utilizar a *linguagem* de uma *teoria* para discutir a própria *teoria*. Isto não acontece quando se usam conjuntos para estudar, por exemplo, Análise na Reta ou Álgebra Comutativa, mas é inescapável quando o *objeto* de estudo são *conjuntos*. Para uma discussão mais aprofundada, o leitor *pode* conferir o Capítulo F. Mas deveria? \triangle

Exercícios adicionais

Exercício C.23. Para conjuntos A, A', B e B' com $|A| \leq |A'|$ e $|B| \leq |B'|$, mostre que vale $|A \times A'| \leq |B \times B'|$. Conclua que se $|A| = |A'|$ e $|B| = |B'|$, então $|A \times A'| = |B \times B'|$. Dica: para injecções $f: A \rightarrow A'$ e $g: B \rightarrow B'$, investigue a (injetividade da) função $f \times g: A \times A' \rightarrow B \times B'$ que faz $\langle a, b \rangle \mapsto \langle f(a), g(b) \rangle$; para a igualdade, utilize o Teorema de Cantor-Bernstein. ■

Exercício C.24. Sejam A e B conjuntos disjuntos, cada um com pelo menos dois elementos distintos. Mostre que $|A \cup B| \leq |A \times B|$. Dica: fixe $a', a'' \in A$ e $b', b'' \in B$, todos dois a dois distintos entre si, e defina uma injecção esperta da forma $A \cup B \rightarrow A \times B$. ■

Exercício C.25. Sejam A e B conjuntos com $|A| \leq |B|$. Mostre que $|\wp(A)| \leq |\wp(B)|$. Conclua que se $|A| = |B|$, então $|\wp(A)| = |\wp(B)|$. Dica: para a segunda parte, use o Teorema de Cantor-Bernstein; para a primeira parte, observe que se $f: A \rightarrow B$ é injetora, então a correspondência $C \mapsto f[C]$ define uma injecção entre $\wp(A)$ e $\wp(B)$; alternativamente, use o exercício seguinte em conjunção com o Exercício A.31. ■

Exercício C.26 (Funtores escondidos). Para uma função da forma $f: A \rightarrow B$, denote por $\wp(f): \wp(A) \rightarrow \wp(B)$ a função que faz $C \mapsto f[C]$ para cada $C \subseteq A$.

- a) Mostre que $\wp(g \circ f) = \wp(g) \circ \wp(f)$ para quaisquer funções $f: A \rightarrow B$ e $g: B \rightarrow C$.
- b) Mostre que $\wp(\text{Id}_A) = \text{Id}_{\wp(A)}$ para qualquer conjunto A . ■

Exercício C.27. Sejam X um conjunto e $k \in \omega$. Mostre que se $|X| = k_+$, então $|X \setminus \{x\}| = k$ para todo $x \in X$. Dica: indução em k ; note que para $k := 0$, tem-se $X = \{x\}$ para algum x . ■

Exercício C.28. Mostre que se X é finito e $x \notin X$, então $|X \cup \{x\}| = |X|_+$. Dica: lembre-se de que se X é finito, então existe um único $n \in \omega$ com uma bijeção $\psi: n \rightarrow X$; dado que $n_+ := n \cup \{n\}$, como definir uma bijeção $n_+ \rightarrow X \cup \{x\}$? ■

Exercício C.29. Demonstre as seguintes afirmações.

- a) Para todo $n \in \omega$ ocorre $|n| = n$.
- b) Se $X \subseteq Y$ e Y é finito, então X é finito e $|X| \leq |Y|$. Dica: indução em $|Y|$ e, possivelmente, o exercício anterior.
- c) Se X e Y são finitos, então $X \cup Y$ é finito. Dica: indução em $|X \setminus Y|$.
- d) Se \mathcal{F} é finito e todo $F \in \mathcal{F}$ é finito, então $\bigcup \mathcal{F}$ é finito. Dica: indução em $|\mathcal{F}|$.
- e) Se X e Y são finitos, então $X \times Y$ é finito. Dica: $X \times Y = \bigcup_{y \in Y} X \times \{y\}$.
- f) Se X é finito e $f: X \rightarrow Y$ é uma função, então $\text{im}(f)$ é finito e $|\text{im}(f)| \leq |X|$. Dica: indução em $|X|$.
- g) Se $X \times Y$ é finito, então X e Y são finitos.
- h) Se X é finito, então $\wp(X)$ é finito. ■

Exercício C.30. Mostre que se X é um conjunto infinito, então $|X| = |X \setminus \{x\}|$ para qualquer $x \in X$. ■

Exercício C.31. Seja $\psi: X \rightarrow Y$ uma função bijetora. Mostre que se \mathcal{P} é uma partição de X , então $\psi(\mathcal{P}) := \{\psi[P] : P \in \mathcal{P}\}$ é uma partição de Y . Em particular, tem-se $|\mathcal{P}| = |\psi(\mathcal{P})|$. ■

Exercício C.32. Sejam \mathbb{P}, \mathbb{P}' e \mathbb{P}'' ordens parciais, munidas de funções crescentes $f: \mathbb{P} \rightarrow \mathbb{P}'$ e $g: \mathbb{P}' \rightarrow \mathbb{P}''$. Mostre que $g \circ f$ é crescente. Conclua que a composta de isomorfismos de ordem é um isomorfismo de ordens. ■

Exercício C.33. Seja $\langle \mathbb{W}, \leq \rangle$ uma boa ordem. Um subconjunto $S \subseteq \mathbb{W}$ é chamado de **segmento inicial** de \mathbb{W} se $S \neq \emptyset$ e $\{x \in \mathbb{W} : x < s\} \subseteq S$ para todo $s \in S$.

- Seja S um subconjunto inicial de \mathbb{W} . Mostre que existe um único $s \in \mathbb{W}$ que verifica a identidade $S = \{x \in \mathbb{W} : x < s\}$. Dica: suponha que não e argumente por indução ou faça primeiro o próximo item.
- Sejam \mathbb{W}_0 e \mathbb{W}_1 boas ordens. Então um e somente um dos casos a seguir ocorre: $\mathbb{W}_0 \cong \mathbb{W}_1$, $\mathbb{W}_0 \cong S_1$ para algum segmento inicial $S_1 \subseteq \mathbb{W}_1$, ou $\mathbb{W}_1 \cong S_0$ para algum segmento inicial $S_0 \subseteq \mathbb{W}_0$. Dica: use ordinais. ■

Exercício C.34. Seja $\mathcal{O} \neq \emptyset$ uma família de ordinais. Mostre que $\bigcap \mathcal{O}$ é ordinal. Mais ainda: mostre que $\bigcap \mathcal{O} = \min \mathcal{O}$, i.e., $\bigcap \mathcal{O} \in \mathcal{O}$ e $\bigcap \mathcal{O} \leq \alpha$ para todo $\alpha \in \mathcal{O}$. ■

Exercício C.35. Seja X uma família não-vazia de ordinais. Mostre que $\bigcup X$ é um número ordinal com a seguinte propriedade: $\beta \leq \bigcup X$ para todo $\beta \in X$ e, se α é um ordinal tal que $\beta \leq \alpha$ para todo $\beta \in X$, então $\bigcup X \leq \alpha$. ■

Observação C.5.20. Em vista do exercício anterior, faz sentido escrever $\sup \mathcal{X}$ para indicar o ordinal obtido como reunião de \mathcal{X} . △

Exercício C.36. Sejam \mathcal{A} e \mathcal{B} conjuntos de ordinais tais que para cada $\alpha \in \mathcal{A}$ exista $\beta \in \mathcal{B}$ com $\alpha \leq \beta$. Mostre que em tais condições ocorre $\sup \mathcal{A} \leq \sup \mathcal{B}$. ■

Exercício C.37. Sejam α, β e γ ordinais.

- Mostre que se $\alpha < \beta$ e $\beta < \gamma$, então $\alpha < \gamma$.
- Mostre que se $\alpha < \beta$, então $\beta \not< \alpha$. Dica: suponha que não. ■

Exercício C.38. Dado um ordinal α , existe outro ordinal β com $\alpha < \beta < \alpha + 1$? ■

Exercício C.39. Pense rápido: dizer que um conjunto X é transitivo equivale a dizer que a relação $\in_X := \{\langle x, y \rangle \in X \times X : x \in y\}$ é transitiva? Dica: todo conjunto de ordinais é ordinal? ■

Exercício C.40. Sejam \mathbb{W}_0 e \mathbb{W}_1 ordens isomorfas.

- Mostre que \mathbb{W}_0 é boa ordem se, e somente se, \mathbb{W}_1 é boa ordem.
- Mostre que se X está em bijeção com \mathbb{W}_0 , então X admite uma ordem isomorfa a \mathbb{W}_0 . Dica: encare a definição da relação R_f no Exemplo C.5.5 até que ela te encare de volta. ■

Exercício C.41 (Cantor-Bernstein (Teorema C.2.4) revisitado). Sejam $f: X \rightarrow Y$ e $g: Y \rightarrow X$ funções.

- Mostre como usar recursão para definir uma sequência $\langle X_n \rangle_{n \in \omega}$ de modo que se tenha $X_0 := \emptyset$ e $X_{n+1} := X \setminus g[Y \setminus f[X_n]]$ para cada $n \in \omega$.
- Defina $\tilde{X} := \bigcup_{n \in \omega} X_n$ e mostre que, se g é injetiva, então $\tilde{X} = X \setminus g[Y \setminus f[\tilde{X}]]$.
- Prove o Teorema C.2.4 (Cantor-Bernstein) sem apelar para o ponto fixo de Tarski. ■

Capítulo D

Escolhas intangíveis

Certamente o axioma mais debatido de ZFC, o Axioma da Escolha apresenta diversas formulações equivalentes nas mais variadas áreas da Matemática que utilizam conjuntos como linguagem subjacente. Discutir suas encarnações mais famosas, bem como algumas aplicações importantes, é o principal objetivo deste capítulo.

D.1 Partições e representantes

Originalmente, a formulação do Axioma da Escolha feita por Zermelo no começo do século XX não tratava de produtos cartesianos ou coisas do tipo, mas sim de famílias de conjuntos não-vazios e dois a dois disjuntos. Explicitamente:

Axioma da Escolha (formulação original)¹. Se $\mathcal{A} \neq \emptyset$, $\emptyset \notin \mathcal{A}$ e para todo $A, B \in \mathcal{A}$ com $A \neq B$ valer $A \cap B = \emptyset$, então existe um conjunto C tal que $|C \cap A| = 1$ para cada $A \in \mathcal{A}$.

Em outras palavras, o conjunto C *escolhe* um elemento em cada $A \in \mathcal{A}$ por meio da interseção, já que $C \cap A$ tem, por hipótese, precisamente um elemento.

Exercício D.1. Reflita sobre a formulação acima. Dica: se precisar, faça um desenho para perceber que C , *realmente*, *escolhe* um elemento em cada $A \in \mathcal{A}$. ■

Como uma família \mathcal{A} nas hipóteses da formulação anterior constitui uma partição para o conjunto $X := \bigcup \mathcal{A}$, não é difícil ver que um conjunto como C é, tão somente, uma classe de representantes para a relação de equivalência em X determinada pela partição \mathcal{A} . Dessa forma, o Teorema B.1.13 provou, secretamente, que a formulação do Axioma da Escolha em termos de produtos cartesianos *implica* a formulação original em termos de famílias disjuntas. Mais geralmente:

Teorema D.1.1. As seguintes afirmações são equivalentes em ZF.

(AC₀) Se $\mathcal{A} \neq \emptyset$ é tal que $A \neq \emptyset$ para todo $A \in \mathcal{A}$ e $A \cap B = \emptyset$ para quaisquer $A, B \in \mathcal{A}$ distintos, então existe C com $|C \cap A| = 1$ para cada $A \in \mathcal{A}$.

(AC₁) Se $\emptyset \neq \mathcal{A} = \{A_i : i \in \mathcal{I}\}$ e $A_i \neq \emptyset$ para todo $i \in \mathcal{I}$, então $\prod_{i \in \mathcal{I}} A_i \neq \emptyset$.

(AC₂) Toda relação de equivalência/partição admite uma classe de representantes.

(AC₃) Se $\mathcal{A} \neq \emptyset$ e $\emptyset \notin \mathcal{A}$, então existe uma função $f: \mathcal{A} \rightarrow \bigcup \mathcal{A}$ tal que $f(A) \in A$ para cada $A \in \mathcal{A}$.

¹Ou quase.

Observação D.1.2. Antes de proceder com a demonstração, convém refletir sobre as diferentes instâncias apresentadas: note que em cada uma delas, escolhas arbitrárias são feitas por algum objeto cuja existência é postulada. Em (AC_0) , C faz as escolhas; em (AC_1) , uma upla $\langle a_i \rangle_{i \in I} \in \prod_{i \in I} A_i$ escolhe $a_i \in A_i$ para cada $i \in I$; já em (AC_2) , uma classe de representantes escolhe, precisamente, um representante em cada classe; finalmente, em (AC_3) , a função f escolhe $f(A)$ em A para cada $A \in \mathcal{A}$. \triangle

Demonstração. Se vale (AC_0) e \mathcal{A} é uma família como em (AC_1) , então para cada $i \in I$ pode-se definir $B_i := \{i\} \times A_i$ para daí considerar a família $\mathcal{B} := \{B_i : i \in I\}$, cuja existência segue do Axioma-Esquema da Separação ou do Axioma-Esquema da Substituição:

- (i) com a primeira abordagem, \mathcal{B} se *realiza* como subconjunto de $\wp(I \times \bigcup \mathcal{A})$;
- (ii) na segunda abordagem, \mathcal{B} é imagem de uma fórmula funcional adequada pelo conjunto I .

Agora, para $i, j \in I$ com $i \neq j$, tem-se $B_i \cap B_j = \emptyset$ e, como cada $B_i \neq \emptyset$, segue que a família \mathcal{B} satisfaz precisamente as hipóteses de (AC_0) . Logo, existe um conjunto C com $|C \cap B_i| = 1$ para cada $i \in I$. Define-se então $f := \{c \in C : \exists i \in I (c \in B_i)\}$, que também satisfaz $|f \cap B_i| = 1$ para todo $i \in I$. Mostremos que $f \in \prod_{i \in I} A_i$:

- ✓ por construção, se $c \in f$ então existe $i \in I$ com $c \in B_i = \{i\} \times A_i$, donde segue que c é um par ordenado e, mais ainda, $\text{dom}(f) \subseteq I$;
- ✓ dado $i \in I$, existe um único $c \in B_i \cap f$, o qual deve ser da forma $c = \langle i, a \rangle$ para algum $a \in A_i$, o que dá $I \subseteq \text{dom}(f)$ e, pelo item anterior, $\text{dom}(f) = I$;
- ✓ f é uma função pois se $\langle i, y \rangle, \langle i, w \rangle \in f$, então $\langle i, y \rangle, \langle i, w \rangle \in B_i$ e, da condição $|B_i \cap f| = 1$, resulta $y = w$;
- ✓ finalmente, f é função escolha, pois se $i \in I$, então $\langle i, f(i) \rangle \in B_i := \{i\} \times A_i$, donde segue $f(i) \in A_i$.

A implicação $(AC_1) \Rightarrow (AC_2)$ já foi feita: é a demonstração do Teorema B.1.13. Para $(AC_2) \Rightarrow (AC_3)$, fixa-se $\mathcal{A} \neq \emptyset$ uma família de conjuntos não-vazios e para cada $A \in \mathcal{A}$ considera-se $A' := \{A\} \times A$. Sobre $\mathcal{B} := \bigcup_{A \in \mathcal{A}} A'$, que existe, por exemplo, em virtude dos Axiomas da Separação e da União, define-se a relação \approx por

$$P \approx Q \Leftrightarrow \exists A \in \mathcal{A} \text{ tal que } P, Q \in A'.$$

Note que isso faz sentido pois se $P \in \mathcal{B}$, então existe $A \in \mathcal{A}$ com $P \in \{A\} \times A$. Como \approx é, claramente, uma relação de equivalência, (AC_2) assegura uma classe de representantes $\mathcal{R} \subseteq \mathcal{B}$ de \approx . Logo, \mathcal{R} é tal que para todo $P \in \mathcal{B}$ existe um único $R \in \mathcal{R}$ com $P \approx R$. O leitor deve se convencer de que $\bigcup \mathcal{R}$ é uma função escolha da forma $\mathcal{A} \rightarrow \bigcup \mathcal{A}$.

Finalmente, a implicação $(AC_3) \Rightarrow (AC_0)$ é quase automática: se \mathcal{A} tem as condições impostas em (AC_0) e $f: \mathcal{A} \rightarrow \bigcup \mathcal{A}$ é tal que $f(A) \in A$ para cada $A \in \mathcal{A}$, então $C := \text{im}(f)$ satisfaz a tese de (AC_0) . \square

Exercício D.2. Complete os detalhes da demonstração acima. ■

Exemplo D.1.3 (Opcional: conjuntos de Vitali). No que segue, $[0, +\infty]$ indica apenas o intervalo real $[0, +\infty)$ acrescido de um máximo artificial, chamado de $+\infty$ (de modo similar ao que se fez no Exercício C.10). Para efeitos de ilustração, considere uma função $m: \wp(\mathbb{R}) \rightarrow [0, +\infty]$ com as seguintes propriedades:

- (i) $m([0, 1]) = 1$;
- (ii) $m(A \cup B) = m(A) + m(B)$ sempre que A e B forem disjuntos;
- (iii) chamando $A + x := \{a + x : a \in A\}$ para $A \subseteq \mathbb{R}$ e $x \in \mathbb{R}$, $m(A + x) = m(A)$.
- (iv) para uma sequência $\langle A_n \rangle_{n \in \omega}$ de subconjuntos de \mathbb{R} dois a dois disjuntos, vale $m(\bigcup_{n \in \omega} A_n) = \sum_{n \in \omega} m(A_n)$.

Grosso modo, uma função com tais propriedades seria útil para definir uma *teoria de integração* para *certas* funções da forma $\mathbb{R} \rightarrow \mathbb{R}$, cujo valor da integral seria invariante por translações (o que já ocorre, por exemplo, com a integral de Riemann para *certas* funções da forma $[a, b] \rightarrow \mathbb{R}$). Embora as exigências acima soem razoáveis, não existe função $m: \wp(\mathbb{R}) \rightarrow [0, +\infty]$ satisfazendo todas as condições anteriores simultaneamente.

Teorema D.1.4. *Se uma função $m: \wp(\mathbb{R}) \rightarrow [0, +\infty]$ satisfaz as condições (i), (ii) e (iii) acima, então m não satisfaz (iv).*

Demonstração. Primeiramente, note que a segunda condição implica a monotonicidade de m , ou seja: se $A \subseteq B$, então $m(A) \leq m(B)$, pois $m(B) = m(B \setminus A) + m(A)$. Isto será útil, adiante. Agora, sobre $[0, 1]$, consideremos a relação binária \sim definida por $x \sim y$ se, e somente se, $x - y \in \mathbb{Q}$. Como \sim é uma relação de equivalência, existe $\mathcal{R} \subsetneq [0, 1]$ um conjunto de representantes de \sim , i.e., tal que para cada $x \in [0, 1]$ existe um único $r_x \in \mathcal{R}$ com $x \sim r_x$.

Façamos então $\mathcal{R}_q := \mathcal{R} + q$ para cada $q \in \mathbb{Q} \cap [-1, 1]$. Mostraremos que deve ocorrer a inclusão $[0, 1] \subseteq \bigcup_{q \in \mathbb{Q} \cap [-1, 1]} \mathcal{R}_q$: dado $x \in [0, 1]$, existe $r \in \mathcal{R}$ com $x - r := q \in \mathbb{Q}$ e, como ambos $x, r \in [0, 1]$, resulta que $q \in [-1, 1]$, com $x = r + q \in \mathcal{R}_q$. Daí, por valer $\bigcup_{q \in \mathbb{Q} \cap [-1, 1]} \mathcal{R}_q \subseteq [-1, 2]$, obtém-se

$$\begin{aligned} 0 < 1 := m([0, 1]) &\leq m\left(\bigcup_{q \in \mathbb{Q} \cap [-1, 1]} \mathcal{R}_q\right) \leq m([-1, 2]) = m([-1, 0] \cup [0, 1] \cup (1, 2]) \leq \\ &\leq m([-1, 0]) + m([0, 1]) + m([1, 2]) = 3. \end{aligned}$$

Por outro lado, deve-se ter $\mathcal{R}_p \cap \mathcal{R}_q = \emptyset$ para $p \neq q$: se $x \in \mathcal{R}_p \cap \mathcal{R}_q$, então existem $r, r' \in \mathcal{R}$ com $x = r + p$ e $x = r' + q$, donde segue que $x \sim r$ e $x \sim r'$ e, consequentemente, $r = r'$ e $p = q$. Logo,

$$m\left(\bigcup_{q \in \mathbb{Q} \cap [-1, 1]} \mathcal{R}_q\right) \neq \sum_{q \in \mathbb{Q} \cap [-1, 1]} m(\mathcal{R}_q) = \sum_{q \in \mathbb{Q} \cap [-1, 1]} m(\mathcal{R}),$$

pois $\sum_{q \in \mathbb{Q} \cap [-1, 1]} m(\mathcal{R}) = 0$ se $m(\mathcal{R}) = 0$, e $\sum_{q \in \mathbb{Q} \cap [-1, 1]} m(\mathcal{R}) = +\infty$ se $m(\mathcal{R}) \neq 0$. □

A moral da história é a seguinte: uma função m capaz de medir subconjuntos de \mathbb{R} de forma razoável não é capaz de medir todos os subconjuntos de \mathbb{R} , i.e., deve ocorrer² $\text{dom}(m) \neq \wp(\mathbb{R})$. As σ -álgebras, introduzidas no Exemplo C.5.9, costituem o tipo de domínio ideal das *boas medidas*. ▲

²A menos que se *assuma* o contrário, i.e., pode-se postular a existência de uma *medida* com tais propriedades. O preço, porém, é alto: abandonar o Axioma da Escolha.

D.2 Boa ordenação de Zermelo

Apesar do que se observou na seção anterior, há situações em que é *possível* dar uma regra explícita de escolha. O caso canônico, em certo sentido, depende de uma boa ordem. Recordemo-nos: dizer que $\langle \mathbb{P}, \leq \rangle$ é uma boa ordem consiste em afirmar que $\langle \mathbb{P}, \leq \rangle$ é uma ordem parcial em que $\min A$ existe para todo subconjunto $A \subseteq \mathbb{P}$ com $A \neq \emptyset$. Note que por valer $\min A \in A$, isso dá um método automático para que se escolha um elemento de um subconjunto não-vazio de \mathbb{P} : tome o menor! Portanto, o leitor não deve se espantar com o

Teorema D.2.1. *As seguintes afirmações são equivalentes em ZF.*

(AC₃) *Se $\mathcal{A} \neq \emptyset$ e $\emptyset \notin \mathcal{A}$, então existe uma função $f: \mathcal{A} \rightarrow \bigcup \mathcal{A}$ tal que $f(A) \in A$ para cada $A \in \mathcal{A}$.*

(AC₄) *(Teorema da Boa Ordem de Zermelo) todo conjunto admite uma boa ordem.*

Demonstração. (AC₄) implica (AC₃) de modo muito natural: toma-se uma boa ordem \leq sobre $\bigcup \mathcal{A}$ e define-se $f: \mathcal{A} \rightarrow \bigcup \mathcal{A}$ com $f(A) := \min A$.

Para a recíproca, fixado um conjunto X , basta exibir um ordinal em bijeção com X (confira o Exercício C.40). Como o caso $X := \emptyset$ é automático, supõe-se $X \neq \emptyset$, e isso permite tomar uma função escolha $f: \mathcal{A} \rightarrow \bigcup \mathcal{A}$, onde $\mathcal{A} := \wp(X) \setminus \{\emptyset\}$. Agora, para $Y \notin X$, seja $\mathcal{G}: \mathbb{V} \rightarrow \mathbb{V}$ a função de classe que faz $\mathcal{G}(x) := f(X \setminus \text{im}(x))$ se x for uma função satisfazendo $X \setminus \text{im}(x) \neq \emptyset$, ou $\mathcal{G}(x) := Y$ caso contrário. Pelo Teorema C.4.16 (da Recursão em ordinais), existe uma única função de classe \mathcal{G} -recursiva $\mathcal{F}: \text{ORD} \rightarrow \mathbb{V}$, i.e., satisfazendo $\mathcal{F}(\alpha) = \mathcal{G}(\langle \mathcal{F}(\beta) : \beta < \alpha \rangle)$ para todo $\alpha \in \text{ORD}$.

Note que, na prática

- (i) $\mathcal{F}(\alpha) = f(X \setminus \{\mathcal{F}(\beta) : \beta < \alpha\})$ se $X \setminus \{\mathcal{F}(\beta) : \beta < \alpha\} \neq \emptyset$, pois $\langle \mathcal{F}(\beta) : \beta < \alpha \rangle$ é uma função, e
- (ii) $\mathcal{F}(\alpha) = Y$ caso contrário.

Agora, o *Katzensprung*³ é observar o seguinte: *se \mathcal{F} satisfaz as duas condições acima, então $\mathcal{F}(\alpha) = Y$ para algum número ordinal α .* De fato, se ocorresse o contrário, então a condição (i) faria de \mathcal{F} uma função de classes injetiva da forma $\text{ORD} \rightarrow X$, o que não pode ocorrer em virtude do insuspeito Lema A.5.9.

Enquanto o *Katzensprung* garante a existência de *algum* ordinal γ com $\mathcal{F}(\gamma) = Y$, a Proposição C.4.14 permite assumir que γ é o menor com tal propriedade. Logo, para todo $\alpha < \gamma$, $\mathcal{F}(\alpha)$ é dada pela condição (i) acima e, portanto, a restrição de \mathcal{F} ao ordinal γ induz uma injeção $\gamma \rightarrow X$. Ocorre que tal restrição também é sobrejetora: se existisse $x \in X$ com $\mathcal{F}(\alpha) \neq x$ para todo $\alpha < \gamma$, teria-se $X \setminus \{\mathcal{F}(\alpha) : \alpha < \gamma\} \neq \emptyset$ e isso daria $\mathcal{F}(\gamma) \neq Y$, contrariando o modo como γ foi tomado. \square

Exercício D.3. Complete os detalhes na demonstração anterior. \blacksquare

Pode ser pedagogicamente edificante *implementar* o argumento acima num cenário trivial, apenas para perceber o que *realmente* está sendo feito.

³Pulo do gato, em alemão. Ou pelo menos foi o que me disse Guilherme Schultz na ocasião em que ministrei Análise II para ele.

Consideremos, por exemplo, $X := \{a, b, c\}$, com a, b e c dois a dois distintos. Neste caso, tem-se $\mathcal{A} := \wp(X) \setminus \{\emptyset\} = \{X, \{a\}, \{b\}, \{c\}, \{a, b\}, \{a, c\}, \{b, c\}\}$, e precisa-se de uma função $f: \mathcal{A} \rightarrow X$ satisfazendo $f(A) \in A$ para todo A , digamos

A	\mapsto	$f(A) \in A$
X		c
$\{a\}$		a
$\{b\}$		b
$\{c\}$		c
$\{a, b\}$		b
$\{a, c\}$		a
$\{b, c\}$		c

Agora, com $Y \notin X$, digamos $Y := \star$, a função $\mathcal{F}: \text{ORD} \rightarrow \mathbb{V}$ da demonstração faz:

- (0) $\mathcal{F}(0) = \mathcal{G}(\emptyset) = f(X \setminus \emptyset) = f(X) := c$, pois \emptyset é uma função que satisfaz $X \setminus \text{im } (\emptyset) \neq \emptyset$;
- (1) $\mathcal{F}(1) = \mathcal{G}(\langle \mathcal{F}(0) \rangle) = \mathcal{G}(\langle c \rangle) = f(X \setminus \{c\}) = f(\{a, b\}) := b$, pois $\langle c \rangle$ é uma função que satisfaz $X \setminus \text{im } (\langle c \rangle) = X \setminus \{c\} = \{a, b\} \neq \emptyset$;
- (2) $\mathcal{F}(2) = \mathcal{G}(\langle \mathcal{F}(0), \mathcal{F}(1) \rangle) = \mathcal{G}(\langle c, b \rangle) = f(X \setminus \{c, b\}) = f(\{a\}) = a$, pois $\langle c, b \rangle$ é uma função que satisfaz $X \setminus \text{im } (\langle c, b \rangle) = X \setminus \{c, b\} = \{a\} \neq \emptyset$;
- (3) $\mathcal{F}(3) = \mathcal{G}(\langle \mathcal{F}(0), \mathcal{F}(1), \mathcal{F}(2) \rangle) = \mathcal{G}(\langle c, b, a \rangle) = \star$, pois $\langle c, b, a \rangle$ é uma função que satisfaz $X \setminus \text{im } (\langle c, b, a \rangle) = \emptyset$;
- (4) $\mathcal{F}(4) = \mathcal{G}(\langle \mathcal{F}(0), \mathcal{F}(1), \mathcal{F}(2), \mathcal{F}(3) \rangle) = \mathcal{G}(\langle c, b, a, \star \rangle) = \star$, pois $\langle c, b, a, \star \rangle$ é uma função que satisfaz...

Portanto, a ordenação obtida para X , neste caso, elege $x_0 := c$, $x_1 := b$ e $x_2 := a$. Em particular, note que 3 foi o primeiro ordinal γ para o qual ocorre $\mathcal{F}(\gamma) = \star$, o que está de acordo com a ordenação sugerida, já que $3 := \{0, 1, 2\}$.

Exemplo D.2.2 (Opcional: bases em espaços vetoriais). O leitor já familiarizado com *Álgebra Linear* deve se lembrar de que uma *base* para um *espaço vetorial* V sobre um *corpo* K é um subconjunto de vetores de V , *linearmente independente* e que *gera* V . Embora seja muito comum demonstrar a existência de bases por meio do *Lema de Zorn*, pode-se apelar diretamente para a *boa ordenação*⁴.

Teorema D.2.3. *Todo espaço vetorial tem base.*

Demonstração. Fixada uma boa ordem \preceq para um espaço vetorial V , com $0_V = \min V$, mostraremos que $B := \{v \in V : v \notin \text{sp}(\{u \in V : u \prec v\})\}$ é uma base para V , em que $\text{sp}(S)$ indica o subespaço gerado por S . É fácil ver que se $B = \emptyset$, então $V = \{0_V\}$, e daí B é de fato uma base para V . Agora, supondo $B \neq \emptyset$ e uma combinação linear não-trivial $\sum_{i \leq n} a_i b_i = 0_V$, com $b_i \in B$ para cada $i \leq n$ e $b_0 \prec b_1 \prec \dots \prec b_n$, teria-se $b_n \notin B$ justamente por a_n ser invertível. Portanto, B é l.i.. Para ver que B gera V , note que, ao se assumir o contrário, o menor $u \in V \setminus \text{sp}(B)$ seria um membro de $B \subseteq \text{sp}(B)$. \square

Já a *recíproca* do teorema acima, i.e., *se todo espaço vetorial tem base, então vale o Axioma da Escolha*, é algo bem mais delicado de se provar em ZF. O leitor interessado pode conferir o original [5] ou uma versão mais amigável como a apresentada por Herrlich [18] – ou, ainda, o Exercício D.15. \blacktriangle

⁴Truque que aprendi com Leandro Aurichi.

D.3 Recursões de bolso: o Lema de Zorn

Uma das encarnações mais difundidas do Axioma da Escolha é expressa pelo *Lema de Zorn*, que essencialmente funciona como uma “recursão de bolso” – graças a ela, muitos matemáticos até hoje dormem tranquilos sem nunca terem ouvido a expressão “recursão transfinita”. Porém, tudo tem um preço.

Definição D.3.1. Para uma ordem parcial $\langle \mathbb{P}, \leq \rangle$ e um subconjunto $C \subseteq \mathbb{P}$, diremos que $C \subseteq \mathbb{P}$ é uma **cadeia**⁵ se quaisquer dois elementos de C são comparáveis segundo a relação \leq , i.e., se para quaisquer $x, y \in C$ valer $x \leq y$ ou $y \leq x$. ¶

Agora, suponha que $\langle \mathbb{P}, \leq \rangle$ seja uma ordem parcial, com $\mathbb{P} \neq \emptyset$, em que toda cadeia tenha limitante superior. Se visualizarmos \mathbb{P} como uma *árvore* cujas ramificações são determinadas pelas relações de ordem entre seus elementos, podemos imaginar cadeias como ramos *lineares* ou *caminhos*, de modo que a sentença acima se torna algo como *todo ramo da árvore admite uma extensão*.

Ora, se \mathbb{P} for uma *árvore* com tal propriedade, podemos recursivamente determinar um caminho entre os seus nós de modo a sempre permanecermos num mesmo ramo. Dessa forma, eventualmente atingiremos uma *folha*: um nó do ramo que não admite extensão, ou mais precisamente, um elemento maximal. De fato:

- escolhe-se $p_0 \in \mathbb{P}$, de modo que $\{p_0\}$ é uma cadeia de \mathbb{P} ;
 - se, por ventura, p_0 for maximal, então não há como estender a cadeia, i.e., p_0 é uma folha;
 - se não, escolhe-se $p_1 \in \mathbb{P}$ com $p_0 < p_1$ e observa-se que $\{p_0, p_1\}$ é uma cadeia;
 - * se p_1 for maximal, então atingiu-se uma folha;
 - * se não...



Figura D.1: Uma “árvore” (ordem parcial) com algumas cadeias destacadas em cinza.

Como na demonstração de que (AC_3) implica (AC_4) , a construção desse caminho se esgota em algum estágio λ , resultando numa cadeia $\{p_\alpha : \alpha < \lambda\}$ em \mathbb{P} , sem extensões. Pela hipótese sobre \mathbb{P} , existe $p \in \mathbb{P}$ tal que $p_\alpha \leq p$ para todo $\alpha < \lambda$, i.e., a cadeia $\{p_\alpha : \alpha < \lambda\}$ tem um limitante superior. Se p não fosse maximal, então a cadeia $\{p_\alpha : \alpha < \lambda\}$ poderia ser estendida, contrariando a constatação de esgotamento anterior. A menos de detalhes técnicos, isto prova a parte não-trivial do

⁵Há quem argumente que o *correto* seria traduzir como “corrente” (oriundo de *chain*). Embora eu concorde, falta-me força para quebrar mais uma... *tradição*.

Teorema D.3.2. As seguintes afirmações são equivalentes em ZF.

(AC₃) (*Axioma da Escolha*) Se $\mathcal{A} \neq \emptyset$ e $\emptyset \notin \mathcal{A}$, então existe uma função $f: \mathcal{A} \rightarrow \bigcup \mathcal{A}$ tal que $f(A) \in A$ para cada $A \in \mathcal{A}$.

(AC₅) (*Lema de Zorn*⁶) Se $\mathbb{P} \neq \emptyset$ é uma ordem parcial em que toda cadeia tem limitante superior, então \mathbb{P} tem um elemento maximal.

Demonstração. Para mostrar $(\text{AC}_3) \Rightarrow (\text{AC}_5)$, precisa-se apenas formalizar a recursão descrita na discussão anterior. Para isso, consideremos $\mathcal{A} := \wp(\mathbb{P}) \setminus \{\emptyset\}$ e fixemos uma função escolha $E: \mathcal{A} \rightarrow \bigcup \mathcal{A}$. Por fim, tomemos $Y \notin \mathbb{P}$.

Ora, sempre que $x: \xi \rightarrow \mathbb{P}$ é uma função *estritamente crescente*⁷ entre algum ordinal ξ e \mathbb{P} , sua imagem $\text{im}(x)$ é uma cadeia em \mathbb{P} , donde a hipótese acerca de \mathbb{P} assegura $\mathcal{L}(x) := \{p \in \mathbb{P} : p \text{ é limitante superior de } \text{im}(x)\} \neq \emptyset$. Logo, pode-se definir uma função de classes auxiliar $\mathcal{L}: \mathbb{V} \rightarrow \mathbb{V}$ que associa x a $\mathcal{L}(x)$ se x for uma função estritamente crescente da forma $\xi \rightarrow \mathbb{P}$ para algum ordinal ξ , enquanto faz $\mathcal{L}(x) := Y$ nos demais casos. Daí, para aplicar o Teorema da Recursão, utiliza-se $\mathcal{G}: \mathbb{V} \rightarrow \mathbb{V}$ que faz $\mathcal{G}(x) := E(\mathcal{L}(x) \setminus \text{im}(x))$ sempre que $\mathcal{L}(x) \neq Y$ e $\mathcal{L}(x) \setminus \text{im}(x) \neq \emptyset$ (i.e., \mathcal{G} escolhe, por meio de E , um limitante superior de $\text{im}(x)$ *fora* de $\text{im}(x)$), enquanto faz $\mathcal{G}(x) := Y$ nos demais casos.

Notemos que pela definição de \mathcal{G} , a ocorrência de $\mathcal{F}(\gamma) \neq Y$ acarreta duas coisas:

- (i) $x_\gamma := \langle \mathcal{F}(\alpha) : \alpha < \gamma \rangle$ é estritamente crescente, pois o contrário daria $\mathcal{L}(x_\gamma) = Y$, $\mathcal{G}(x_\gamma) = Y$ e, consequentemente, $\mathcal{F}(\gamma) = \mathcal{G}(x_\gamma) = Y$;
- (ii) $\mathcal{F}(\alpha) < \mathcal{F}(\gamma)$ para todo $\alpha < \gamma$, posto que $\mathcal{F}(\gamma) = \mathcal{G}(x_\gamma) = E(\mathcal{L}(x_\gamma) \setminus \text{im}(x_\gamma))$, que por sua vez é um limitante superior de $\text{im}(x_\gamma)$ não pertencente a $\text{im}(x_\gamma)$, e $\mathcal{F}(\alpha) \in \text{im}(x_\gamma)$ por definição.

Assim, se ocorresse $\mathcal{F}(\gamma) \neq Y$ para todo ordinal γ , $\mathcal{F}: \text{ORD} \rightarrow \mathbb{P}$ seria uma função injetiva de classes, violando o Lema A.5.9. Logo, existe um ordinal λ caracterizado como o menor tal que $\mathcal{F}(\lambda) = Y$ (Proposição C.4.14). Dado que $\mathcal{F}(\alpha) \neq Y$ para todo $\alpha < \gamma$, a argumentação anterior permite concluir que $\mathcal{F}(\alpha) < \mathcal{F}(\beta)$ sempre que $\alpha < \beta < \lambda$, donde se infere que $\{\mathcal{F}(\alpha) : \alpha < \lambda\}$ é uma cadeia em \mathbb{P} , que deve ter um limitante superior p . Finalmente, note que se p não fosse maximal e existisse $q > p$, teria-se $q \notin \{\mathcal{F}(\alpha) : \alpha < \gamma\}$ e, consequentemente, $\mathcal{F}(\gamma) \neq Y$. Portanto, p é maximal.

Reciprocamente, assumindo o *Lema de Zorn*, i.e., o enunciado de (AC₅), mostremos que vale (AC₃). Dado um conjunto $\mathcal{A} \neq \emptyset$ com $\emptyset \notin \mathcal{A}$, busca-se obter uma função escolha da forma $\mathcal{A} \rightarrow \bigcup \mathcal{A}$: a ideia aqui consiste em considerar a árvore de todas as possíveis *escolhas parciais* ordenadas pela inclusão, de modo que ao percorrer um ramo *indefinidamente*, chegaremos a uma *escolha completa*.

Mais precisamente, seja

$$\mathbb{P} := \left\{ f \subseteq \mathcal{A} \times \bigcup \mathcal{A} : f \text{ é função, } \text{dom}(f) \neq \emptyset \text{ e } f(A) \in A \text{ para todo } A \in \mathcal{A} \right\},$$

i.e., o conjunto das funções escolha definidas parcialmente em \mathcal{A} , que satisfaz $\mathbb{P} \neq \emptyset$. Note que a relação \subseteq de inclusão faz de $\langle \mathbb{P}, \subseteq \rangle$ uma ordem parcial. Além disso, toda cadeia $C \subseteq \mathbb{P}$ tem limitante superior: neste caso, basta observar que $h := \bigcup C \in \mathbb{P}$, pois daí $f \subseteq h$ para cada $f \in C$; tal pertinência ocorre pois

⁶Embora o primeiro a prová-lo tenha sido K. Kuratowski.

⁷É o que o nome sugere: dados $\gamma, \gamma' < \xi$, $\gamma < \gamma' \Rightarrow x(\gamma) < x(\gamma')$.

- ✓ como $f \subseteq \mathcal{A} \times \bigcup \mathcal{A}$ para toda $f \in C$, tem-se $h := \bigcup C \subseteq \mathcal{A} \times \bigcup \mathcal{A}$,
- ✓ se $\langle A, a \rangle, \langle A, b \rangle \in h$, então existem $f, g \in C$ com $f(A) = a$ e $g(A) = b$ (i.e., $\langle A, a \rangle \in f$ e $\langle A, b \rangle \in g$), donde a condição de cadeia satisfeita por C dá $f \subseteq g$ ou $g \subseteq f$, acarretando $a = b$ e, assim, mostrando que h é função, e
- ✓ se $A \in \text{dom}(h)$, então existe $f \in C$ com $A \in \text{dom}(f)$ e $h(A) = f(A)$, com $f(A) \in A$, mostrando que h satisfaz as condições para pertencer a \mathbb{P} .

Logo, pelo Lema de Zorn, existe $F \in \mathbb{P}$ maximal com respeito à ordem \subseteq . Se ocorresse $\text{dom}(F) \neq \mathcal{A}$, então existiriam $A \in \mathcal{A} \setminus \text{dom}(F)$ e $a \in A$, e daí $F \cup \{\langle A, a \rangle\} \in \mathbb{P}$ contradiria a maximalidade de F . Portanto, $\text{dom}(F) = \mathcal{A}$ e, por conseguinte, $F: \mathcal{A} \rightarrow \bigcup \mathcal{A}$ é uma função escolha para \mathcal{A} . \square

Como no Teorema da Boa Ordenação, convém replicar o *procedimento* sugerido na demonstração de que $(AC_5) \Rightarrow (AC_3)$ num contexto mais *palpável*. Para isso, consideremos a ordem parcial correspondente à seguinte árvore⁸:

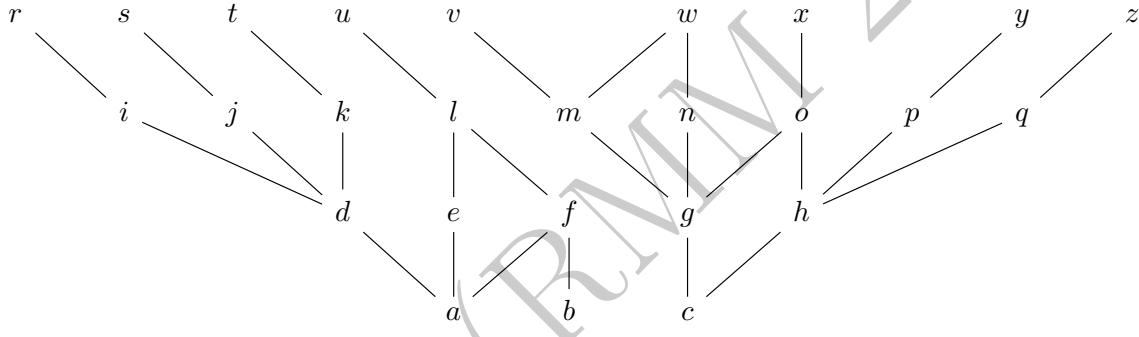


Figura D.2: Os traços indicam a ordenação dos elementos: por exemplo, $a < l$, $c < v$ e $g < o$, mas $a \not< b$, $k \not< m$ e $w \not< x$. Note que r, s, t, \dots, z são os elementos maximais.

Com alguma paciência, verifica-se que a ordem \mathbb{P} acima é tal que toda cadeia tem limitante superior, o que permite replicar o procedimento da demonstração. Para isso, precisamos de uma função escolha $E: \mathcal{A} \rightarrow \mathbb{P}$, onde $\mathcal{A} := \wp(\mathbb{P}) \setminus \{\emptyset\}$. Embora \mathcal{A} seja muito grande⁹, é bastante razoável *supor* que existe uma função E satisfazendo $E(A) \in A$ para todo $A \in \mathcal{A}$. Agora, com $Y \notin X$, digamos $Y := \star$, a função $\mathcal{F}: \text{ORD} \rightarrow \mathbb{V}$ definida na demonstração faz:

- (0) $\mathcal{F}(0) = \mathcal{G}(\emptyset) = E(\mathbb{P})$, pois \emptyset pode ser visto, por vacuidade, como uma função estritamente crescente da forma $0 \rightarrow \mathbb{P}$ e, em tal situação, \mathcal{G} usa E para escolher um limitante superior fora da imagem da função, o que significa escolher um elemento em \mathbb{P} ; digamos que $E(\mathbb{P}) := g$;
- (1) $\mathcal{F}(1) = \mathcal{G}(\langle \mathcal{F}(0) \rangle) = \mathcal{G}(\langle g \rangle) = E(\{m, n, o, v, w, x\})$, pois $\langle g \rangle$ é uma função estritamente crescente da forma $1 \rightarrow \mathbb{P}$ e, em tal situação, \mathcal{G} usa E para escolher um limitante superior fora da imagem da função, o que significa escolher um elemento do conjunto $A := \{m, n, o, v, w, x\}$; digamos que $E(A) := m$;

⁸A rigor, não se trata de uma árvore (no sentido estrito da *Teoria dos Grafos*); tampouco poderia ser chamada de *floresta*. Mas é suficientemente parecida para ser xingada assim, dado que o texto não trata de tais tópicos.

⁹Tem, precisamente, $2^{26} - 1$ elementos.

- (2) $\mathcal{F}(2) = \mathcal{G}(\langle \mathcal{F}(0), \mathcal{F}(1) \rangle) = \mathcal{G}(\langle g, m \rangle) = E(\{v, w\})$, pois $\langle g, m \rangle$ é uma função estritamente crescente da forma $2 \rightarrow \mathbb{P}$ e, em tal situação, \mathcal{G} usa E para escolher um limitante superior fora da imagem da função, o que significa escolher um elemento do conjunto $B := \{v, w\}$; digamos que $E(B) := w$;
- (3) $\mathcal{F}(3) = \mathcal{G}(\langle \mathcal{F}(0), \mathcal{F}(1), \mathcal{F}(2) \rangle) = \mathcal{G}(\langle g, m, w \rangle) = \star$, pois embora $\langle g, m, w \rangle$ seja uma função estritamente crescente da forma $3 \rightarrow \mathbb{P}$, não há limitantes superiores fora de sua imagem, situação em que \mathcal{G} é programada para devolver o valor \star ;
- (4) $\mathcal{F}(4) = \mathcal{G}(\langle \mathcal{F}(0), \mathcal{F}(1), \mathcal{F}(2), \mathcal{F}(3) \rangle) = \mathcal{G}(\langle g, m, w, \star \rangle) = \star$, pois $\langle g, m, w, \star \rangle$ não pode ser interpretada como função estritamente crescente da forma $4 \rightarrow \mathbb{P}$ (afinal, $\star \notin \mathbb{P}$!), situação em que \mathcal{G} é programada para devolver o valor \star ;
- (5) $\mathcal{F}(5) = \star$ pois...

Portanto, a cadeia $\{g, m, w\}$ obtida com a função escolha E tem w como limitante superior, que necessariamente é maximal, já que o contrário daria $\mathcal{F}(3) \neq \star$.

Observação D.3.3. A exigência “ $\mathbb{P} \neq \emptyset$ ” é redundante¹⁰. De fato, se \mathbb{P} é uma ordem em que toda cadeia tem limitante superior, então $\mathbb{P} \neq \emptyset$, pois \emptyset é uma cadeia em \mathbb{P} e, portanto, deve existir pelo menos um elemento em \mathbb{P} para limitar \emptyset superiormente. Na prática, exigir $\mathbb{P} \neq \emptyset$ desde o começo é apenas um jeito de já garantir que a cadeia vazia tenha limitante superior sem pensar muito no assunto. \triangle

Exemplo D.3.4 (Opcional: bases em espaços vetoriais (de novo)). Embora o Exemplo D.2.2 tenha apresentado uma forma bastante prática de demonstrar a existência de bases em espaços vetoriais por meio da boa ordenação, a prova que usa o Lema de Zorn tem seus méritos.

Teorema D.3.5 (Steinitz). *Sejam K um corpo e V um K -espaço vetorial. Se $G \subseteq V$ é gerador de V e $L \subseteq V$ é l.i., então $D \cup L$ é base de V para algum subconjunto $D \subseteq G$.*

Demonstração. Basta observar que $\mathbb{P} := \{C \subseteq G : C \cup L \text{ é l.i.}\}$ é uma ordem parcial (com respeito à inclusão) que satisfaz as hipóteses do Lema de Zorn. Primeiro, $\mathbb{P} \neq \emptyset$ pois $\emptyset \in \mathbb{P}$. Mais geralmente, se $\mathcal{C} \subseteq \mathbb{P}$ é uma cadeia, então $\bigcup \mathcal{C} \in \mathbb{P}$ é um limitante superior de \mathcal{C} : perceba que se $u_0, \dots, u_n \in \bigcup \mathcal{C} \cup L$ e $a_0, \dots, a_n \in K \setminus \{0\}$ são tais que $\sum_{i \leq n} a_i u_i = 0_V$, então existe $C \in \mathcal{C}$ tal que $C \cup L$ não é l.i., contrariando $C \in \mathbb{P}$. Logo, pelo Lema de Zorn, existe $D \in \mathbb{P}$ maximal. Finalmente, tudo se resume a mostrar que $D \cup L$ é uma base: se não fosse, existiria $g' \in G$ com $g' \notin \text{sp}(D \cup L)$, o que levaria a concluir que $D \cup \{g'\} \cup L$ é l.i. e, por conseguinte, $D \cup \{g'\} \in \mathbb{P}$, contrariando a maximalidade de D . \square

Em particular, o teorema acima garante que todo subconjunto gerador contém uma base (basta fazer $L := \{g\}$ para algum $g \in G \setminus \{0\}$) e também que todo subconjunto l.i. está contido num subconjunto l.i. maximal (faça $G := V$). \blacktriangle

Intuitivamente, o argumento acima consiste em escolher $g_0 \in G$ tal que $\{g_0\} \cup L$ é l.i., e então $g_1 \in G \setminus \{g_0\}$ tal que $\{g_0, g_1\} \cup L$ é l.i., e então... O que a formulação do Lema de Zorn faz é empurrar os procedimentos recursivos para debaixo do tapete, abstraídos em “toda cadeia tem limitante superior”, enquanto transforma o Lema A.5.9 (a condição de término) no elemento maximal. O próximo exemplo pode ser ainda mais ilustrativo.

¹⁰Como bem lembrado tanto pela Wikipedia quanto pelo Caio Oliveira.

Exemplo D.3.6 (Opcional: Hahn-Banach). Uma das aplicações favoritas dos analistas (funcionais) para o Lema de Zorn se dá no contexto da Análise Funcional. Em sua versão mais elementar (não a mais geral), os ingredientes são:

- (i) um \mathbb{R} -espaço vetorial X dotado de uma **norma** $\|\cdot\|: X \rightarrow \mathbb{R}$, que por sua vez é uma função que satisfaz as condições “ $\|x\| \geq 0$ ”, “ $\|x\| = 0 \Leftrightarrow x = 0_X$ ”, “ $\|\lambda x\| = |\lambda| \|x\|$ ” e “ $\|x + y\| \leq \|x\| + \|y\|$ ” para quaisquer $x, y \in X$ e $\lambda \in \mathbb{R}$;
- (ii) um **funcional linear** $\varphi: M \rightarrow \mathbb{R}$ definido num subespaço vetorial M de X , i.e., tal que $\varphi(x + \lambda y) = \varphi(x) + \lambda\varphi(y)$ para quaisquer $x, y \in M$ e $\lambda \in \mathbb{R}$, satisfazendo adicionalmente a *condição de dominância* “ $|\varphi(y)| \leq \|y\|$ para todo $y \in M$ ”.

Secretamente, a condição de dominância garante a *continuidade* do funcional, o que por sua vez assegura que a função é compatível com as noções de convergência do espaço. Em todo caso, com tais ingredientes, o que se busca é um funcional linear $\Phi: X \rightarrow \mathbb{R}$ satisfazendo $\Phi(y) = \varphi(y)$ para todo $y \in M$, mas sem alterar a dominância pela norma, i.e., com $|\Phi(x)| \leq \|x\|$ para todo $x \in X$.

Lema D.3.7. *Sejam X um \mathbb{R} -espaço vetorial, $M \subsetneq X$ um subespaço vetorial próprio e $\|\cdot\|: X \rightarrow \mathbb{R}$ uma norma. Se $z_0 \in X \setminus M$ e $\varphi: M \rightarrow \mathbb{R}$ é um funcional satisfazendo $|\varphi(y)| \leq \|y\|$ para todo $y \in M$, então existe um funcional linear $\Phi_{z_0}: M_{z_0} \rightarrow \mathbb{R}$ tal que $\Phi_{z_0}|_M = \varphi$ e $|\Phi_{z_0}(x)| \leq \|x\|$ para todo $x \in M_{z_0} := M \oplus \text{sp}(z_0)$.*

No enunciado acima, $M \oplus \text{sp}(z_0)$ indica o subespaço gerado pelos vetores da forma $m + \lambda z_0$, com $m \in M$ e $\lambda \in \mathbb{R}$ (e o uso do símbolo “ \oplus ” apenas alerta que a *soma* é *direta*, no sentido de que $M \cap \text{sp}(z_0) = \{0_X\}$, o que se deve ao modo como z_0 foi tomado). Uma vez que sua demonstração usa apenas noções elementares sobre espaços normados, ela será omitida. O importante a notar é o seguinte: se, por ventura, o espaço X fosse $M \oplus \text{sp}(z_0)$, então o problema inicial estaria resolvido; se não, então existiria $z_1 \notin M \oplus \text{sp}(z_0)$, o que permitiria reaplicar o lema e obter um novo funcional definido sobre $M \oplus \text{sp}(z_0) \oplus \text{sp}(z_1)$; se, por ventura, o espaço X fosse $M \oplus \text{sp}(z_0) \oplus \text{sp}(z_1)$, então o problema inicial estaria resolvido; se não...

Teorema D.3.8 (Hahn-Banach). *Sejam X um \mathbb{R} -espaço vetorial, $\|\cdot\|: X \rightarrow \mathbb{R}$ uma norma e $M \subsetneq X$ um subespaço vetorial próprio. Se $\varphi: M \rightarrow \mathbb{R}$ é um funcional linear que satisfaz $|\varphi(y)| \leq \|y\|$ para todo $y \in M$, então existe um funcional linear $\Phi: X \rightarrow \mathbb{R}$ com $\Phi|_M = \varphi$ e $|\Phi(x)| \leq \|x\|$ para todo $x \in X$.*

Demonstração. Para cada par $\langle N, \psi \rangle$, onde $N \subseteq X$ é um subespaço vetorial e $\psi: N \rightarrow \mathbb{R}$ é um funcional linear, diremos que $\langle N, \psi \rangle$ é maroto se $M \subseteq N$, $\psi|_M = \varphi$ e $|\psi(x)| \leq \|x\|$ para todo $x \in N$. Daí, consideremos a família $\mathbb{P} := \{\langle N, \psi \rangle : \langle N, \psi \rangle \text{ é maroto}\}$, parcialmente ordenada pela relação \preceq onde $\langle N_0, \psi_0 \rangle \preceq \langle N_1, \psi_1 \rangle$ se, e somente se, $N_0 \subseteq N_1$ e $\psi_1|_{N_0} = \psi_0$.

Note que $\langle M, \varphi \rangle \in \mathbb{P}$, mostrando que $\mathbb{P} \neq \emptyset$. Agora, se $\mathcal{C} \subseteq \mathbb{P}$ for uma cadeia, então $\bigcup \mathcal{C} \in \mathbb{P}$ limita superiormente todos os pares $\langle N, \psi \rangle \in \mathcal{C}$. Logo, o Lema de Zorn garante um par $\langle P, \Phi \rangle \in \mathbb{P}$ maximal com respeito à relação \preceq . Resta apenas mostrar que $P = X$: ora, se não fosse este o caso, então o lema anterior daria $\langle P_c, \Phi_c \rangle \in \mathbb{P}$ para algum $c \in X \setminus P$, com $\langle P, \Phi \rangle \prec \langle P_c, \Phi_c \rangle$, contrariando a maximalidade do par maroto $\langle P, \Phi \rangle$. \square

Antes de encerrar a discussão sobre Hahn-Banach, cabe uma provocação: qual a definição explícita do funcional Φ obtido na demonstração? Ora, por ter sido obtido como *um* elemento maximal cuja existência se garantiu sem uma descrição, não há porque esperar que Φ tenha descrição¹¹. Apesar disso, o Teorema de Hahn-Banach não é *tão forte* quanto, por exemplo, o Teorema D.2.3, no sentido de ser equivalente ao Axioma da Escolha em ZF – mesmo assim, certas máculas creditadas ao Axioma da Escolha são, na verdade, causadas por Hahn-Banach! O leitor interessado pode procurar pelas menções ao *Paradoxo de Banach-Tarski* em [31]. \blacktriangle

Exemplo D.3.9 (Opcional: ideais e filtros). De volta à Álgebra, outro uso típico do Lema de Zorn se dá no contexto dos **ideais**¹²: lembre-se de que um *ideal* M de um anel A não-trivial¹³ é dito **maximal** se M é um ideal próprio e maximal na família dos ideais próprios de A . Em outras palavras¹⁴: se $I \subsetneq A$ é ideal com $M \subseteq I$, então $M = I$.

Teorema D.3.10 (Krull). *Sejam $A \neq 0$ um anel e $J \subsetneq A$ um ideal. Então J está contido num ideal maximal de A .*

Demonstração. A família $\mathbb{P} := \{I \subseteq A : I \text{ é ideal de } A \text{ com } J \subseteq I\} \neq \emptyset$ é parcialmente ordenada pela inclusão. Como qualquer ideal C pertencente a uma cadeia $\mathcal{C} \subseteq \mathbb{P}$ está contido no ideal $\bigcup C \in \mathbb{P}$, segue pelo Lema de Zorn que existe $M \in \mathbb{P}$ maximal com respeito à inclusão. \square

Em particular, em anéis não-triviais, todo ideal próprio está contido num *primo*, já que todo ideal maximal é primo. Recordemo-nos de que um ideal $P \subsetneq A$ é **primo** se $xy \notin P$ sempre que $x, y \in A \setminus P$. Daí, se $M \subsetneq A$ é maximal e $x, y \in A \setminus M$, então a maximalidade assegura $A = \text{sp}(M \cup \{x\}) = \text{sp}(M \cup \{y\})$, acarretando a existência de $m, m' \in M$ e $a, a' \in A$ tais que $1_A = m + ax = m' + a'y$ e, consequentemente, $xy \notin M$ (já que o contrário daria $1_A = mm' + ma'y + axm' + aa'xy \in M$, que não pode ocorrer). Não obstante os diversos desdobramentos de tais considerações em contextos puramente algébricos, será importante considerar o seguinte caso particular.

Para um conjunto fixado X , a coleção $\wp(X)$ formada pelos subconjuntos de X vem de fábrica com uma estrutura natural de anel (comutativo e com unidade): a multiplicação é dada pela interseção, a adição é a diferença simétrica Δ (dada por $A \Delta B := (A \setminus B) \cup (B \setminus A)$ para cada $A, B \subseteq X$), cujos elementos neutros são, respectivamente, X e \emptyset .

Exercício D.4. Prove as afirmações acima. Dica: só aceite, as contas são chatas. \blacksquare

Em tal contexto, o que significa dizer que $\mathcal{I} \subseteq \wp(X)$ é um ideal? Vejamos:

- (i) o 0 aditivo deve pertencer a \mathcal{I} , no caso, $\emptyset \in \mathcal{I}$;
- (ii) além disso, sempre que $A, B \in \mathcal{I}$, deve-se ter $A + B \in \mathcal{I}$, o que corresponde a $(A \setminus B) \cup (B \setminus A) \in \mathcal{I}$;
- (iii) por fim, se $A \in \wp(X)$ e $I \in \mathcal{I}$, então $A \cap I \in \mathcal{I}$.

¹¹Na presença de subespaços *densos enumeráveis*, é possível explicitar extensões como Φ sem apelar para o Lema de Zorn.

¹²Subconjuntos não-vazios de um anel, fechados por combinações lineares com coeficientes do anel.

¹³Isto é, em que $0_A \neq 1_A$, o que costuma se abreviar com “ $A \neq 0$ ”.

¹⁴Também é comum encontrar a seguinte definição: “se $I \subseteq A$ é ideal com $M \subseteq I$, então $M = I$ ou $I = A$ ”. Note que as duas formulações dizem que não há ideal $I \neq A$ satisfazendo $M \subsetneq I$: a segunda apenas expressa isso de modo indireto.

Em particular, por $A \cap B \in \mathcal{I}$ sempre que $A, B \in \mathcal{I}$ e $C \cup D \in \mathcal{I}$ sempre que $C, D \in \mathcal{I}$ com $C \cap D = \emptyset$, resulta que $I \cup J \in \mathcal{I}$ sempre que $I, J \in \mathcal{I}$, já que $I \cup J = (I \Delta J) \cup (I \cap J)$. Além disso, $A \in \mathcal{I}$ sempre que $B \in \mathcal{I}$ com $A \subseteq B$, posto que $A = A \cap B$.

Exercício D.5. Mostre que se $\mathcal{I} \neq \emptyset$ é uma família de subconjuntos de X tal que $A \cup B \in \mathcal{I}$ e $C \in \mathcal{I}$ sempre que $A, B, D \in \mathcal{I}$ e $C \subseteq D$, então \mathcal{I} é um ideal de $\wp(X)$. ■

Observação D.3.11. O critério obtido acima costuma ser apresentado como a definição do que significa ser um **ideal sobre um conjunto** fixado X . △

Secretamente, os ideais de $\wp(X)$ estão intimamente ligados a um tipo de estrutura que ainda será bastante importante no texto.

Definição D.3.12. Uma família $\mathcal{F} \neq \emptyset$ de subconjuntos de X é um **filtro** em X se $F \cap G \in \mathcal{F}$ e $K \in \mathcal{F}$ sempre que $F, G, J \in \mathcal{F}$ e $J \subseteq K$. ¶

Exercício D.6. Sejam \mathcal{I} e \mathcal{F} famílias não-vazias de subconjuntos de X .

- Mostre que \mathcal{I} é um ideal de $\wp(X)$ se, e somente se, $\mathcal{I}^C := \{X \setminus I : I \in \mathcal{I}\}$ é um filtro em X .
- Mostre que \mathcal{F} é um filtro em X se, e somente se, $\mathcal{F}^C := \{X \setminus F : F \in \mathcal{F}\}$ é um ideal de $\wp(X)$. ■

Agora, se \mathcal{F} é um *filtro próprio*, i.e., com $\wp(X) \neq \mathcal{F}$, então $\mathcal{I} := \mathcal{F}^C$ é um ideal próprio de $\wp(X)$ que, pelo teorema anterior, está contido num ideal primo $\mathcal{P} \subsetneq \wp(X)$, que por sua vez induz um filtro próprio $\mathfrak{u} := \mathcal{P}^C$ satisfazendo $\mathcal{F} \subseteq \mathfrak{u}$. Filtros como \mathfrak{u} são mais importantes do que parecem.

Proposição D.3.13. Para um filtro próprio \mathcal{G} em X , são equivalentes:

- \mathcal{G} é filtro maximal na família dos filtros próprios;
- para todo $A \subseteq X$, $A \in \mathcal{G}$ ou $X \setminus A \in \mathcal{G}$;
- para quaisquer $A, B \subseteq X$, $A \cup B \in \mathcal{G}$ acarreta $A \in \mathcal{G}$ ou $B \in \mathcal{G}$
- o ideal \mathcal{G}^C é primo.

Demonstração. Para (i) \Rightarrow (ii), note que se $A \notin \mathcal{G}$, então $(X \setminus A) \cap G \neq \emptyset$ para todo $G \in \mathcal{G}$ (se não, $G \subseteq A$ para algum $G \in \mathcal{G}$, acarretando $A \in \mathcal{G}$), e daí não é difícil perceber que

$$\mathcal{H} := \{H \subseteq X : \exists G \in \mathcal{G} \text{ tal que } (X \setminus A) \cap G \subseteq H\}$$

é um filtro próprio satisfazendo $X \setminus A \in \mathcal{H}$ e $\mathcal{G} \subseteq \mathcal{H}$, donde a maximalidade de \mathcal{G} acarreta $X \setminus A \in \mathcal{G}$. Para (ii) \Rightarrow (iii), de $A \notin \mathcal{G}$ e $B \notin \mathcal{G}$ se infere $X \setminus A, X \setminus B \in \mathcal{G}$ e, consequentemente, $(X \setminus A) \cap (X \setminus B) = X \setminus (A \cup B) \in \mathcal{G}$, donde segue que $A \cup B \notin \mathcal{G}$ (pois \mathcal{G} é próprio!). Para (iii) \Rightarrow (i), se \mathcal{L} é filtro em X satisfazendo $\mathcal{G} \subsetneq \mathcal{L}$, então algum $L \in \mathcal{L} \setminus \mathcal{G}$, o que obriga a ocorrência de $X \setminus L \in \mathcal{G} \subseteq \mathcal{L}$ e, consequentemente, $\emptyset = L \cap (X \setminus L) \in \mathcal{L}$, o que obriga que se tenha $\mathcal{L} = \wp(X)$. Finalmente, (iii) \Leftrightarrow (iv) segue automático pelas leis de De Morgan aliadas aos exercícios anteriores. □

Definição D.3.14. Um filtro próprio \mathcal{G} em X é xingado de **ultrafiltro** se qualquer uma das condições acima for satisfeita. ¶

Em particular, ao combinar a caracterização anterior com o Teorema de Krull e o fato de que todo ideal maximal é primo, obtém-se o

Teorema D.3.15 (Lema do ultrafiltro). *Se \mathcal{F} é um filtro próprio em X , então existe um ultrafiltro \mathfrak{u} em X com $\mathcal{F} \subseteq \mathfrak{u}$.*

Evidentemente, o roteiro utilizado para alcançar o teorema anterior não prezou por qualquer noção de economia: na verdade, provar que todo filtro próprio está contido num filtro maximal é uma das aplicações clássicas mais simples do Lema de Zorn (exercício?). Com isso dito, o objetivo subjacente dos meandros tomados acima foi outro: ilustrar as ramificações imprevisíveis dos *princípios de escolha*.

No caso, uma consequência do Lema de Zorn em Álgebra, o Teorema de Krull¹⁵, foi utilizada para obter um resultado aparentemente alheio acerca de filtros. Por sua vez, filtros, oriundos da Topologia Geral como uma das generalizações naturais das sequências, são ferramentas versáteis que vão muito além do estudo das noções de convergência: o Lema do Ultrafiltro, por exemplo, pode ser usado para demonstrar que o produto arbitrário de espaços *compactos* de *Hausdorff* é compacto (*Teorema de Tychonoff*), que por sua vez é peça fundamental da demonstração do *Teorema de Banach-Alaoglu* (da Análise Funcional)¹⁶.

Embora o propósito do texto permita abordar mais detalhes sobre os resultados mencionados acima, fazê-lo sem pré-requisitos seria desonesto, ao passo que incluir os pré-requisitos de forma honesta tornaria este material muito próximo de [26]. Não obstante, ultrafiltros serão utilizados no último capítulo, na demonstração *semântica* do *Teorema da Compacidade* – um resultado *metamatemático* com diversas aplicações *matemáticas* que ilustram os benefícios de *encarar o abismo* dos *fundamentos* (até que ele te encare de volta). ▲

Este capítulo apresentou as encarnações mais conhecidas do Axioma da Escolha. A aparente diferença entre seus enunciados é sintetizada na *infeliz piada*¹⁷ cunhada por Jerry Bona (1977):

O Axioma da Escolha é obviamente verdadeiro; o Teorema da Boa Ordem é obviamente falso; e quem pode dizer alguma coisa sobre o Lema de Zorn?¹⁸

No entanto, as formulações equivalentes do Axioma da Escolha são *abundantes*. Diga-se de passagem, há pelo menos três livros dedicados a apresentar equivalências deste axioma:

- ✓ o sugestivo *Equivalents of the Axiom of Choice, II* (1985), de Herman Rubin e Jean Rubin,
- ✓ o organizadíssimo *Consequences of the Axiom of Choice* (1991), de Paul Howard e Jean Rubin, e
- ✓ o relativamente recente *Axiom of Choice* (2006), de Horsch Herrlich [18].

¹⁵Trívia: em ZF, ambos são equivalentes. O leitor interessado pode conferir [2].

¹⁶Na verdade, em ZF, todos esses resultados são equivalentes entre si, e nenhum deles acarreta o Axioma da Escolha.

¹⁷Ela costuma perder a graça depois que se entende que o Lema de Zorn é apenas um teorema sobre jardinagem.

¹⁸“The Axiom of Choice is obviously true; the Well Ordering Principle is obviously false; and who can tell about Zorn’s Lemma?” [31].

Em nossas considerações posteriores, será recorrente nos depararmos com outras manifestações do Axioma da Escolha (daqui em diante, abreviado como **AC**, em alusão a *Axiom of Choice*), algumas evidentes, outras nem tanto. De modo geral, usaremos AC sem menção explícita daqui em diante, salvo raras exceções.

Exercícios adicionais

Exercício D.7 (AC₆). Em ZF, assumindo a Proposição C.2.8, prove que vale AC. ■

Exercício D.8. Na demonstração de que $(AC_3) \Rightarrow (AC_5)$ no Teorema D.3.2, mostre que o ordinal λ é sucessor. ■

Exercício D.9. Demonstre os Teoremas D.2.1 e D.3.2 sem apelar para classes próprias. Dica: pode ser útil conferir a definição do número de Hartogs, no Exemplo C.5.5. ■

Exercício D.10 (AC₇). Enuncie e demonstre a versão dual do Lema de Zorn. ■

Exercício D.11. Se $\mathbb{P} \neq \emptyset$ é uma ordem parcial em que toda cadeia *enumerável* tem limitante superior, então \mathbb{P} tem elemento maximal? Dica: se estiver sem ideias, confira brevemente a Proposição E.1.11. ■

Exercício D.12 (Recursão como axioma¹⁹). Para este exercício, pode ser útil conferir a Observação B.3.6 e a demonstração do Teorema D.2.1. Seja ZC o sistema de axiomas obtido de ZFC com a exclusão do Axioma da Substituição, que chamaremos de S, e seja R o enunciado do Teorema B.3.5 (da Recursão). Mostre que em ZC, $R \Leftrightarrow S$. Dica: (para \Rightarrow) fixada uma fórmula $\mathcal{F}(x, y)$ funcional em x e um conjunto A , observe que o Axioma da Escolha (presente em ZC) juntamente com R permite provar que A admite uma boa ordem; defina então $\mathcal{G}: \mathbb{V} \rightarrow \mathbb{V}$ fazendo $\mathcal{G}(s) := y$ sempre que s for uma função com domínio da forma $\{a \in A : a < b\}$ para algum $b \in A$, com y o único a satisfazer $\mathcal{F}(b, y)$; note que a imagem da única função \mathcal{G} -recursiva em A dá, precisamente, o conjunto procurado para atestar a validade de S ■

Observação D.3.16. Moral da história: é lícito assumir os enunciados dos Teoremas de Recursão como axiomas de ZFC. Mas a que custo? △

Exercício D.13 (Opcional: requer traquejo com Álgebra). Sejam D um *domínio de ideais principais* (d.i.p.) e M um D -módulo. Mostre que se $B \subseteq M$ é uma base para M e $N \subseteq M$ é um submódulo, então N tem uma base C com $|C| \leq |B|$. Sugestão: siga o roteiro a seguir.

- Fixe uma boa ordem para B . Para $b \in B$ fixado, sejam $M'_b := \text{sp}(a : a < b)$ e $M_b := M'_b \oplus \text{sp}(b)$. Mostre que se $n \in N \cap M_b$, então existem únicos $d \in D$ e $m'_b \in M'_b$ tais que $n = m'_b + db$.
- Use o item anterior para conjurar uma *projeção* $p_b: N \cap M_b \rightarrow D$ que torne a sequência a seguir *exata*²⁰: $0 \rightarrow N \cap M'_b \rightarrow N \cap M_b \rightarrow \text{im}(p_b) \rightarrow 0$.
- Use o fato de D ser d.i.p. para obter $d_b \in D$ tal que $\text{im}(p_b) = \text{sp}(d_b)$.

¹⁹Baseado na exposição de Joel David Hamkins sobre o trabalho de Benjamin Rin [28], na postagem “*Transfinite recursion as a fundamental principle in set theory*” disponível em seu blog: <http://jdh.hamkins.org/transfinite-recursion-as-a-fundamental-principle-in-set-theory/>.

²⁰A segunda seta no diagrama deve ser injetiva (tome a inclusão), a terceira deve ser sobrejetiva (a projeção p_b) e $\ker p_b = N \cap M'_b$.

- d) Finalmente, considere $B' := \{b \in B : d_b \neq 0\}$ e, para cada $b \in B'$, seja $n_b \in N \cap M_b$ um elemento satisfazendo $p_b(n_b) = d_b$. Mostre que $C := \{n_b : b \in B'\}$ é uma base para N . Dica: para mostrar que C é l.i., use a boa ordem de B para ordenar os vetores tomados na verificação de independência linear, juntamente com a identidade $N \cap M'_b = \ker p_b$; para mostrar que C gera N , suponha que não seja o caso e tome o menor $b \in B$ para o qual exista $m \in N \cap M_b$ que não seja combinação linear de C , a fim de obter uma contradição. ■

Exercício D.14 (AC₈). Considere a seguinte afirmação, chamada de **Axioma das Escolhas Múltiplas**: se $\mathcal{I} \neq \emptyset$ e $\langle X_i : i \in \mathcal{I} \rangle$ é tal que $X_i \neq \emptyset$ para todo $i \in \mathcal{I}$, então existe uma \mathcal{I} -upla $\langle F_i : i \in \mathcal{I} \rangle$ com $F_i \neq \emptyset$, $F_i \subseteq X_i$ e $|F_i| < |\omega|$ para cada $i \in \mathcal{I}$.

- a) Mostre que, em ZFC, vale o Axioma das Escolhas Múltiplas.
- b) Mostre que, em ZF, o Axioma das Escolhas Múltiplas é equivalente à seguinte asserção: para todo conjunto $\mathcal{A} \neq \emptyset$ com $\emptyset \notin \mathcal{A}$, existe uma função $g: \mathcal{A} \rightarrow \bigcup_{A \in \mathcal{A}} \wp(A)$ tal que $g(A) \in \wp(A)$ com $0 < |g(A)| < |\omega|$ para cada $A \in \mathcal{A}$.
- c) Para cada x , defina $\text{rank}(x) := \min\{\alpha : \alpha \text{ é ordinal e } x \in \mathbb{V}_\alpha\}$, onde os \mathbb{V}_α são tomados como no Teorema C.5.16. Convença-se de que tal definição faz sentido.
- d) Para um ordinal limite $\alpha \neq 0$ fixado, suponha \mathbb{V}_β bem ordenado para cada $\beta < \alpha$. Mostre que \mathbb{V}_α é bem ordenável. Dica: para $x, y \in \mathbb{V}_\alpha$ defina $x < y$ se, e somente se, $\text{rank}(x) < \text{rank}(y)$ ou $\delta = \text{rank}(x) = \text{rank}(y)$ e $x <_\delta y$, onde $<_\delta$ é a boa ordem em \mathbb{V}_δ existente por hipótese.
- e) Seja $\langle \mathbb{W}, \leq \rangle$ uma boa ordem. Para $A, B \in \wp(\mathbb{W})$, declare $A \prec B$ se, e somente se, existir $a \in A \setminus B$ com $a < b$ para todo $b \in B \setminus A$. Mostre que $\langle \wp(\mathbb{W}), \preceq \rangle$ é uma ordem total.
- f) Assumindo apenas ZF e o Axioma das Escolhas Múltiplas, prove que \mathbb{V}_α é bem-ordenável para todo ordinal α . Dica: para o caso sucessor $\alpha := \beta_+$, seja g uma função como no item (b), com $\mathcal{A} := \wp(\mathbb{V}_\alpha) \setminus \{\emptyset\}$; agora, defina $f: \mathcal{A} \rightarrow \bigcup \mathcal{A}$ fazendo, para cada $X \in \mathcal{A}$, $f(X) := \min_{\preceq} g(X)$, onde \preceq indica a ordem total sobre $\wp(\mathbb{V}_\beta)$ do item anterior, o que faz sentido pois $g(X)$ é um subconjunto finito de $\wp(\mathbb{V}_\beta)$ posto que $g(X) \subseteq X \subseteq \mathbb{V}_\alpha := \wp(\mathbb{V}_\beta)$; use a demonstração do Teorema D.2.1 para encerrar.
- g) Conclua que $(\text{ZF} + \text{Axioma das Escolhas Múltiplas}) \Rightarrow \text{Axioma da Escolha}$. Dica: use o Axioma da Fundação. ■

Exercício D.15 (AC₉). Seja $\langle X_\lambda : \lambda \in \Lambda \rangle$ uma upla de conjuntos não-vazios e dois a dois disjuntos. Mostraremos que “se todo espaço vetorial sobre algum corpo tem base”, então vale o Axioma das Escolhas Múltiplas (Exercício D.14).

1. Considere k o seu corpo favorito²¹. Defina $X := \bigcup_{\lambda \in \Lambda} X_\lambda$ e considere $k(X)$ o corpo de frações de $k[X]$. Para um monômio $p = a \cdot x_1^{n_1} \cdots x_m^{n_m}$, defina o λ -grau de p como sendo $d_\lambda(p) = \sum_{x_j \in X_\lambda} n_j$, i.e., a soma dos expoentes cujas variáveis pertencem ao λ -ésimo conjunto X_λ . Reflita sobre tudo isso.
2. Agora, um elemento $P \in k(X)$ é uma fração da forma $\frac{p_1+\dots+p_n}{q_1+\dots+q_m}$ para certos monômios p_i e q_j , o que nos permite definir o seguinte: para $d \in \mathbb{Z}$, diremos que P é λ -homogêneo de grau d se todos os monômios q_j tiverem o mesmo λ -grau, digamos d_1 , e todos os polinômios p_i tiverem o mesmo λ -grau $d_1 + d$. Mostre que $K := \{P \in k(X) : P \text{ é } \lambda \text{ homogêneo de grau zero para todo } i \in \mathcal{I}\}$ é um subcorpo de $k(X)$.

²¹Algum que não dependa do Axioma da Escolha, como corpos finitos, \mathbb{R} ou \mathbb{C} .

3. Em vista do último item, $k(X)$ é um K -espaço vetorial e, portanto, tem uma base, digamos $B \subseteq k(X)$. Mostre que cada $x \in X$ se expressa de forma única como uma combinação K -linear $x = \sum_{b \in B(x)} a_b(x) \cdot b$, onde $B(x)$ é um subconjunto finito de B e cada $a_b(x) \in K \setminus \{0\}$.
4. Mostre que se $x, y \in X_\lambda$, então $B(x) = B(y)$ e $\frac{a_b(x)}{x} = \frac{a_b(y)}{y}$. Dica: $y = \frac{y}{x}x$.
5. Pelo item anterior, faz sentido chamar $B_\lambda := B(x)$ para qualquer $x \in X_\lambda$, bem como $a_b(\lambda) := \frac{a_b(x)}{x}$ para qualquer $b \in B(x)$ com $x \in X_\lambda$. Com isso, note que cada $a_b(\lambda)$ é λ -homogêneo de grau -1 e, portanto, o conjunto F_λ que contém os (finitos) x 's em X_λ que ocorrem no denominador de $a_b(\lambda)$ em sua forma reduzida deve ser não-vazio.
6. Conclua que vale o Axioma das Escolhas Múltiplas. ■

DRAFT (RMM 2023)

Capítulo E

Cardinais

No Capítulo C, a noção de cardinalidade foi explorada como uma “relação” entre conjuntos: X e Y têm a mesma cardinalidade (abrev. $X \approx Y$) se existe uma bijeção $X \rightarrow Y$. Uma vez que \approx se comporta como uma relação de equivalência (na classe própria \mathbb{V}), tal abordagem permitiu impor um critério para o que poderia ser uma *boa definição de número cardinal*: representantes das “classes de equivalência” da relação \approx . Explicitamente, isto significa que ao associar um *objeto* $|X|$ para cada X , dois critérios (*a.k.a. princípio de Hume*) devem ser satisfeitos:

- (i) para todo conjunto X , $X \approx |X|$;
- (ii) para quaisquer conjuntos X e Y , $X \approx Y$ se, e somente se, $|X| = |Y|$.

Evidentemente, as condições acima não definem *o quê* é o número cardinal de X , mas estabelecem como os possíveis candidatos a tal título devem se comportar. Nesse sentido, num primeiro momento, o Teorema C.1.6 demonstrou que números naturais servem como os números cardinais dos conjuntos finitos: dado que um conjunto finito X está em bijeção com um único natural n_X , as duas exigências acima se cumprem ao estipularmos $|X| := n_X$ para qualquer X finito. Mais geralmente, a combinação da Proposição C.4.11 (todo ordinal está em bijeção com um único ordinal inicial) com o Teorema C.5.4 (todo conjunto bem ordenado é isomorfo a um único ordinal) permitem estender o Teorema C.1.6 para os conjuntos *bem ordenáveis*: se X admite uma boa ordem, então existe um ordinal em bijeção com X , que por sua vez está em bijeção com um único ordinal inicial, digamos α_X , de modo que também se cumprem as exigências (i) e (ii) acima ao fazermos $|X| := \alpha_X$.

Por essa razão, a *atribuição honesta* do número cardinal de um conjunto arbitrário X precisa esperar o Teorema da Boa Ordenação de Zermelo, a fim de que se garanta que X admite uma boa ordenação. Tratar desses números cardinais, agora entidades de *carne e osso*, é o objetivo deste capítulo.

E.1 A escadaria dos alephs

Definição E.1.1. Dado um conjunto X , o **número cardinal** de X , denotado por $|X|$, é o número cardinal $|\alpha|$ de qualquer número ordinal α em bijeção com X . ¶

Dessa forma, a expressão “ordinal inicial” se revela, em ZFC, sinônimo de “número cardinal”, já que por um lado, todo número cardinal é ordinal inicial e, por outro lado, todo ordinal inicial é o seu próprio número cardinal. Em todo caso, por preciosismo, convém reforçar que a definição acima *faz sentido*.

Proposição E.1.2. *A definição acima atende ao princípio de Hume.*

Demonstração. Em vista de [\(AC₄\)](#), a classe $C_X := \{\alpha \in \text{ORD} : \exists f: \alpha \rightarrow X \text{ bijetora}\}$ é não-vazia, donde segue que existe um menor elemento, digamos γ . Pela minimalidade, não há $\beta < \gamma$ em bijeção com γ , mostrando que γ é inicial, enquanto a tricotomia garante que tal ordinal (inicial) é o único em bijeção com X . Em particular, cumpre-se a primeira condição. Para a segunda: se X e Y estão em bijeção, então qualquer ordinal em bijeção com um estará em bijeção com o outro, acarretando $|X| \leq |Y|$ e $|Y| \leq |X|$ e, em vista do Exercício [C.16](#), $|X| = |Y|$; reciprocamente, se γ é ordinal inicial com $|X| = \gamma$ e $|Y| = \gamma$, então é claro que X e Y estão em bijeção. \square

Em particular, como os números cardinais de conjuntos X e Y são números ordinais, exatamente um dos seguintes casos deve ocorrer: $|X| = |Y|$, $|X| < |Y|$ ou $|Y| < |X|$. Porém, precisa-se de *alguma* cautela aqui, já que tais expressões *também* foram introduzidas como abreviações para afirmações sobre a existência de injecções e bijeções (Definição [C.2.1](#)): poderia ser o caso de que as duas noções fossem incompatíveis. Não é.

Proposição E.1.3. *Sejam X e Y conjuntos quaisquer.*

1. *São equivalentes:*

- (i) $|X| \leq |Y|$ (*como desigualdade de ordinais*);
- (ii) *existe injecção* $X \rightarrow Y$,
- (iii) *existe sobrejeção* $Y \rightarrow X$.

2. *São equivalentes:*

- (i) $|X| < |Y|$ (*como desigualdade de ordinais*);
- (ii) *não existe injecção* $Y \rightarrow X$;
- (iii) *não existe sobrejeção* $X \rightarrow Y$.

Demonstração. Se $|X| \leq |Y|$, então o cardinal $|X|$ é subconjunto do cardinal $|Y|$, donde segue que existe uma função injetora (a inclusão!) $|X| \rightarrow |Y|$. Logo, o primeiro *grupo* de equivalências segue dos fatos corriqueiros já observados acerca de injecções, sobrejeções e composições¹. O segundo grupo de equivalências segue da tricotomia de ordinais aliada ao primeiro grupo de equivalências: por exemplo, se $|X| < |Y|$ e existisse uma injecção $Y \rightarrow X$, teria-se $|Y| \leq |X|$ e $|X| < |Y|$ como ordinais, o que não pode ocorrer. Os detalhes ficam a cargo do leitor. \square

Em particular, o Teorema [C.1.10](#) (de Cantor), que estabelece a inexistência de sobrejeções da forma $\wp(X) \rightarrow X$, agora pode ser usado para expressar que o número cardinal de X é estritamente menor do que o número cardinal de $\wp(X)$. Em particular, existe *pelo menos um* cardinal maior do que $|X|$.

Definição E.1.4. Dado um cardinal κ , denota-se por κ^+ o menor cardinal estritamente maior do que κ , chamado **sucessor de κ** . \P

Proposição E.1.5. *Todo cardinal tem um sucessor.*

¹Explicitamente: Exercícios [A.28](#), [A.31](#) e [C.5](#), além da Proposição [C.2.8](#).

Demonstração. Dado um cardinal κ , o Teorema de Cantor acarreta as desigualdades $\kappa = |\kappa| < |\wp(\kappa)| < |\wp(\wp(\kappa))| := \mu$, e daí $S := \{|\alpha| \in \mu : \kappa < |\alpha|\} \neq \emptyset$, precisamente por ocorrer $|\wp(\kappa)| \in S$. Logo, $\kappa^+ = \min S$, que existe em virtude da Proposição C.4.14. \square

Observação E.1.6 (Cardinalidade vs. ordem, de novo). É importante não confundir κ^+ com κ_+ : embora ambos coincidam no caso finito, a situação muda completamente se κ for infinito. De fato, se α é um ordinal com $\omega \leq \alpha$, então $\alpha_+ := \alpha \cup \{\alpha\}$ não é um cardinal (*a.k.a.* ordinal inicial): obtém-se uma bijeção² $\alpha_+ \rightarrow \alpha$, com

$$\alpha \mapsto 0, \quad n \mapsto n_+ \text{ se } n \in \omega \setminus \{0\} \quad \text{e} \quad \gamma \mapsto \gamma \text{ se } \gamma \in \alpha \setminus \omega,$$

mostrando que α_+ não é inicial. Note que a contrapositiva desse argumento diz que cardinais infinitos são necessariamente ordinais limite. \triangle

Exercício E.1. Mostre que se κ é cardinal infinito, então $\kappa_+ < \kappa^+$. Dica: reveja o Exercício C.18. \blacksquare

Quem enfrentou a Seção C.5 e, em particular, o Exemplo C.5.5, pode demonstrar a última proposição de modo mais *limpo*. Com efeito, ao se considerar $H(\kappa)$, o *número de Hartogs* do cardinal κ , o Teorema C.5.6 assegura que $H(\kappa)$ é o menor ordinal inicial que satisfaz $H(\kappa) \not\leq \kappa$, o que equivale a $\kappa < H(\kappa)$ pela tricotomia. Em outras palavras: $\kappa^+ = H(\kappa)$. A vantagem de tal abordagem é que, diferente da demonstração que se apresentou inicialmente, não se apela para AC a fim de *construir* $H(\kappa)$. Tamanho excesso de zelo traz recompensas.

Teorema E.1.7. As seguintes afirmações são equivalentes em ZF.

(AC₄) Todo conjunto admite uma boa ordem.

(AC₁₀) (Tricotomia de cardinalidade). Dados dois conjuntos X e Y , ocorre um e somente um dos seguintes casos:

- existe uma bijeção $X \rightarrow Y$;
- existe injeção $X \rightarrow Y$ e não existe injeção $Y \rightarrow X$;
- existe injeção $Y \rightarrow X$ e não existe injeção $X \rightarrow Y$.

(AC₁₁) Para qualquer conjunto X existe injeção $X \rightarrow H(X)$.

Demonstração. Sabemos que AC permite associar os cardinais $|X|$ e $|Y|$ a X e Y , respectivamente. Daí, a tricotomia dos ordinais dá $|X| = |Y|$, $|X| < |Y|$ ou $|Y| < |X|$, o que se traduz precisamente na sentença (AC₁₀). Supondo (AC₁₀), o número de Hartogs de X não admite injeção $H(X) \rightarrow X$, e assim (AC₁₀) obriga que exista injeção $X \rightarrow H(X)$. Finalmente, supondo (AC₁₁), segue que o conjunto X pode ser visto como subconjunto do ordinal $H(X)$, mostrando que X admite uma boa ordem e, portanto, a validade de (AC₄). \square

Dado que nem todo ordinal é cardinal, segue que $\text{CARD} := \{\kappa : \kappa \text{ é cardinal}\}$ é subclasse de ORD, e *própria* no sentido de ser diferente de ORD. Poderia ser o caso de CARD ser um conjunto? Não – e antes de prosseguir, convém conferir o Exercício C.35 e a Observação C.5.20.

²Devido ao Teorema de Cantor-Bernstein, a fim de mostrar que um certo ordinal β **não** é inicial (cardinal) basta obter uma injeção $\beta \rightarrow \gamma$ para algum $\gamma < \beta$.

Lema E.1.8. Se \mathcal{X} é um conjunto de cardinais, então $\sup \mathcal{X}$ é cardinal.

Demonstração. Note que a existência de uma injeção $\sup \mathcal{X} \rightarrow \alpha$ para algum $\alpha \in \kappa$ com $\kappa \in \mathcal{X}$ permite induzir (via restrição, já que $\kappa \leq \sup \mathcal{X}$) uma injeção $\kappa \rightarrow \alpha$, contrariando o fato de κ ser ordinal inicial. \square

Corolário E.1.9 (Paradoxo de Cantor). CARD é classe própria.

Demonstração. Ora, se CARD fosse um conjunto, então $\Omega := \sup \text{CARD}$ seria um cardinal, bem como Ω^+ , o que levaria a $\Omega^+ \in \text{CARD}$ e, por conseguinte, $\Omega^+ \leq \Omega$, absurdo. \square

Embora se tenha $\text{CARD} \subsetneq \text{ORD}$, essas duas classes próprias estão *intimamente* relacionadas por meio de suas noções de sucessão. Entram em cena os *alephs*³.

Definição E.1.10. Para um número ordinal α , define-se \aleph_α recursivamente como:

- (i) $\aleph_0 := \omega$;
- (ii) $\aleph_{\alpha_+} := (\aleph_\alpha)^+$ para qualquer ordinal α ; e
- (iii) $\aleph_\alpha := \sup_{\xi < \alpha} \aleph_\xi$ se $\alpha \neq 0$ for um ordinal limite. \P

No caso, a descrição dos alephs estipula $\aleph_0 := \omega$ como passo base, o *primeiro* número cardinal infinito. Como $1 = 0_+$, a definição impõe $\aleph_1 := (\aleph_0)^+$, o *primeiro* número cardinal maior do que \aleph_0 ou, com a terminologia clássica, o menor número cardinal não-enumerável. De modo análogo, $\aleph_2 := (\aleph_1)^+$ é o *primeiro* número cardinal maior do que \aleph_1 . Também faz sentido considerar $\aleph_\omega := \sup_{n < \omega} \aleph_n$, explicitamente o primeiro número cardinal maior do que *todos* os \aleph_n 's para $n < \omega$, enquanto $\aleph_{\omega_+} := (\aleph_\omega)^+$ é o *primeiro* número cardinal maior do que \aleph_ω .

Se fosse legítimo considerar ORD e CARD como conjuntos, a descrição dos \aleph 's seria, na verdade, uma função

$$\begin{aligned} \aleph: \text{ORD} &\rightarrow \text{CARD} \\ \alpha &\mapsto \aleph_\alpha \end{aligned}$$

que associa cada ordinal α ao α -ésimo aleph de modo a *converter* a operação de sucessão ordinal⁴ ($\alpha \mapsto \alpha_+$) em sucessão cardinal ($\kappa \mapsto \kappa^+$) por meio da identidade $\aleph_{\alpha_+} = \aleph_\alpha^+$. Porém, ORD e CARD não são conjuntos⁵, o que inviabiliza o uso formal desta interpretação em ZFC. Se não fosse isso, o próximo teorema poderia ser enunciado de maneira muito simples: \aleph é um *isomorfismo* de ordens.

Teorema E.1.11. Para todo cardinal infinito κ existe um único ordinal α tal que $\kappa = \aleph_\alpha$. Além disso, $\aleph_\alpha < \aleph_\beta$ se, e somente se, $\alpha < \beta$.

Note que por ser uma afirmação acerca do *comportamento* dos alephs, o teorema acima depende explicitamente da descrição recursiva dada na Definição E.1.10. Porém, em vez de ser um complicador, isto permite usar a boa ordenação dos ordinais a fim de argumentar por indução... em ORD.

³Aleph (lido “álef”) é o nome da letra hebraica “א”.

⁴Tenho tomado bastante cuidado para não sucumbir ao modo padrão de denotar o sucessor de um ordinal α , a saber: $\alpha + 1$. Faço isso apenas para evitar confusões com as notações aritméticas que serão discutidas em breve. O leitor pode ignorar isso e escrever “ $\alpha + 1$ ” em vez de “ α_+ ” em suas próprias anotações.

⁵Em particular, perguntas como “qual o número cardinal de $\{\aleph_\alpha : \alpha \in \text{ORD}\}?$ ” nem fazem sentido em ZFC, posto que apenas conjuntos tem números cardinais.

Teorema E.1.12 (Indução transfinita – em casos). *Seja $\mathcal{P}(x)$ uma fórmula na variável x tal que:*

- (i) *tenha-se $\mathcal{P}(0)$;*
- (ii) *para todo ordinal α , tenha-se $\mathcal{P}(\alpha_+)$ sempre que valer $\mathcal{P}(\alpha)$;*
- (iii) *para todo ordinal limite $\beta \neq 0$, tenha-se $\mathcal{P}(\beta)$ sempre que $\mathcal{P}(\alpha)$ valer para todo $\alpha < \beta$.*

Então $\mathcal{P}(\alpha)$ vale para todo ordinal α .

Demonstração. Se a classe de ordinais $T := \{\alpha \in \text{ORD} : \neg \mathcal{P}(\alpha)\}$ fosse não-vazia, então a Proposição C.4.14 asseguraria um ordinal $\alpha_0 := \min T$, porém: pela condição (i), $\alpha_0 \neq 0$; pela condição (ii), α_0 não pode ser sucessor; finalmente, pela condição (iii), α_0 não pode ser limite diferente de 0. Como todo ordinal se enquadra em pelo menos uma dessas condições, infere-se que $T = \emptyset$. \square

Usaremos indução em ORD a fim de demonstrar três afirmações intermediárias, acerca de ordinais α e β quaisquer.

Afirmiação I : $\alpha < \beta \Rightarrow \aleph_\alpha < \aleph_\beta$.

Afirmiação II : *se κ é cardinal infinito e $\kappa < \aleph_\alpha$, então existe $\beta < \alpha$ com $\kappa = \aleph_\beta$.*

Afirmiação III : $\alpha \leq \aleph_\alpha$.

Demonstração do Teorema E.1.11. Em posse da Afirmação III, tem-se $\kappa \leq \aleph_\kappa$, enquanto a Afirmação I dá $\aleph_\kappa < \aleph_{\kappa_+}$, o que permite fazer $\alpha := \kappa_+$ na Afirmação II e obter $\beta < \kappa_+$ com $\kappa = \aleph_\beta$. A unicidade do ordinal β segue da tricotomia de ordinais aliada à Afirmação I. O restante segue de uma adaptação simples do Exercício B.22. \square

Demonstração da Afirmação I. Por indução em β : para $\beta := 0$ a implicação é válida por vacuidade; se $\beta := \gamma_+$ e $\aleph_\delta < \aleph_\gamma$ para todo $\delta < \gamma$, então para $\alpha < \beta$ pode-se ter $\alpha < \gamma$ ou $\alpha = \gamma$, resultando em $\aleph_\alpha < \aleph_\gamma$ ou $\aleph_\alpha := \aleph_\gamma < \aleph_\gamma^+ := \aleph_{\gamma_+} := \aleph_\beta$; se $\beta \neq 0$ é ordinal limite e para todo $\gamma < \beta$ vale que $\delta < \gamma$ acarreta $\aleph_\delta < \aleph_\gamma$, então para $\alpha < \beta$ existe $\gamma < \beta$ com $\alpha < \gamma$, acarretando $\aleph_\alpha < \aleph_\gamma$ e $\aleph_\gamma \leq \sup_{\gamma < \beta} \aleph_\gamma := \aleph_\beta$. \square

Demonstração da Afirmação II. Por indução em α : para $\alpha := 0$ o resultado é válido por vacuidade; se $\alpha := \beta_+$ e, para todo cardinal infinito $\lambda < \aleph_\beta$ existe $\xi < \beta$ com $\lambda = \aleph_\xi$, então para um cardinal infinito $\kappa < \aleph_\alpha := \aleph_\beta^+ := \aleph_\alpha$ pode-se ter $\kappa < \aleph_\beta$ ou $\kappa = \aleph_\beta$; se $\alpha \neq 0$ é ordinal limite e para todo $\beta < \alpha$ verifica-se a afirmação, então para um cardinal infinito $\kappa < \aleph_\alpha := \sup_{\beta < \alpha} \aleph_\beta$ existe $\beta' < \alpha$ com $\kappa \leq \aleph_{\beta'}$, donde o resultado segue. \square

Demonstração da Afirmação III. Por indução em α : é óbvio que $0 \leq \aleph_0$; se $\alpha := \beta_+$ e $\beta \leq \aleph_\beta$, então $\alpha := \beta_+ \leq (\aleph_\beta)_+ < \aleph_\beta^+ := \aleph_\alpha$ (Exercício E.1); se $\alpha \neq 0$ é ordinal limite e $\beta \leq \aleph_\beta$ para todo $\beta < \alpha$, então $\alpha = \sup_{\beta < \alpha} \beta \leq \sup_{\beta < \alpha} \aleph_\beta := \aleph_\alpha$ (Exercício C.36). \square

E.2 Dizia eu que a aritmética

O *infinitamente distante* Exemplo C.1.5 sugeriu as bases para a chamada *aritmética cardinal*, que consiste no estudo das *propriedades operatórias* que envolvem a noção de cardinalidade. Porém, na *ocasião*, coisas como “ $|X|$ ” e “ $|Y|$ ” eram apenas abreviações *metalingüísticas* para lidar com cardinalidades. Agora que números cardinais foram introduzidos formalmente na teoria, convém revisar e estender tais procedimentos.

E.2.1 Operações cardinais (caso geral)

Definição E.2.1. Para cardinais κ e λ , definem-se:

- (i) (**adição ou soma**) $\kappa + \lambda := |(\kappa \times \{0\}) \cup (\lambda \times \{1\})|$;
- (ii) (**multiplicação ou produto**) $\kappa \cdot \lambda := |\kappa \times \lambda|$;
- (iii) (**potência**) $\kappa^\lambda := |\kappa^\lambda|$.

¶

Definição E.2.2. Dados um conjunto \mathcal{I} e cardinais κ_i para cada $i \in \mathcal{I}$, definem-se:

- (i) $\sum_{i \in \mathcal{I}} \kappa_i := \left| \bigcup_{i \in \mathcal{I}} (\kappa_i \times \{i\}) \right|$;
- (ii) $\prod_{i \in \mathcal{I}} \kappa_i := \left| \prod_{i \in \mathcal{I}} \kappa_i \right|$.

¶

É evidente que a última definição generaliza as duas primeiras cláusulas da anterior. Além disso, diferente do que se fez no Exemplo C.1.5, aqui as *operações* estão automaticamente bem definidas, pois a atribuição foi feita diretamente em termos de representantes que satisfazem o *princípio de Hume*. Ainda assim, convém destacar que *outros* representantes poderiam ser colocados no lado direito das identidades.

Proposição E.2.3. Sejam \mathcal{I} um conjunto, $\mathcal{A} := \langle A_i : i \in \mathcal{I} \rangle$ e $\mathcal{B} := \langle B_i : i \in \mathcal{I} \rangle$ uplas de conjuntos tais que $|A_i| \leq |B_i|$ para cada $i \in \mathcal{I}$.

- (i) Se cada upla é composta por termos dois a dois disjuntos, então $\left| \bigcup_{i \in \mathcal{I}} A_i \right| \leq \left| \bigcup_{i \in \mathcal{I}} B_i \right|$.
- (ii) Tem-se $\left| \prod_{i \in \mathcal{I}} A_i \right| \leq \left| \prod_{i \in \mathcal{I}} B_i \right|$.

Além disso, se $|A_i| = |B_i|$ para todo $i \in \mathcal{I}$, então ocorrem igualdade em ambos os casos.

Demonstração. No primeiro item, para cada $j \in \mathcal{I}$ existe uma injeção $f_j: A_j \rightarrow B_j$, o que permite definir $f: \bigcup_{i \in \mathcal{I}} A_i \rightarrow \bigcup_{i \in \mathcal{I}} B_i$ com $f(x) = f_j(x)$, onde $j \in \mathcal{I}$ é o único tal que $x \in A_j$. Claramente f é uma injeção. Para a segunda afirmação, para cada $i \in \mathcal{I}$ fixa-se uma função $f_i: A_i \rightarrow B_i$ a fim de definir a função⁶

$$\begin{aligned} \prod_{i \in \mathcal{I}} f_i: \prod_{i \in \mathcal{I}} A_i &\longrightarrow \prod_{i \in \mathcal{I}} B_i \\ \langle a_i \rangle_{i \in \mathcal{I}} &\longmapsto \langle f_i(a_i) \rangle_{i \in \mathcal{I}} \end{aligned}$$

Note que se cada f_i é uma injeção, então $\prod_{i \in \mathcal{I}} f_i$ é uma injeção. Em ambos os casos, a parte final segue do Teorema de Cantor-Bernstein. □

⁶Também denotada como $g_0 \times \dots \times g_n$ caso se tenha $\mathcal{I} := \{0, \dots, n\}$ (Observação A.5.11).

Observação E.2.4. Quando $\langle A_i : i \in \mathcal{I} \rangle$ é uma \mathcal{I} -upla de conjuntos não necessariamente disjuntos, garante-se apenas

$$\left| \bigcup_{i \in \mathcal{I}} A_i \right| \leq \sum_{i \in \mathcal{I}} |A_i|,$$

que segue da injeção $\bigcup_{i \in \mathcal{I}} A_i \rightarrow \bigcup_{i \in \mathcal{I}} A_i \times \{i\}$ dada por $x \mapsto \langle x, \min\{j \in \mathcal{I} : x \in A_j\} \rangle$, onde implicitamente se assume \mathcal{I} bem-ordenado. \triangle

Corolário E.2.5. Sejam \mathcal{I} um conjunto e para cada $i \in \mathcal{I}$ considere cardinais κ_i e λ_i .

(i) (Monotonicidade) Se $\kappa_i \leq \lambda_i$ para todo i , então $\sum_{i \in \mathcal{I}} \kappa_i \leq \sum_{i \in \mathcal{I}} \lambda_i$ e $\prod_{i \in \mathcal{I}} \kappa_i \leq \prod_{i \in \mathcal{I}} \lambda_i$.

(ii) (Comutatividade) Se $\psi: \mathcal{I} \rightarrow \mathcal{I}$ é uma bijeção, então

$$\sum_{i \in \mathcal{I}} \kappa_i = \sum_{\psi(i) \in \mathcal{I}} \kappa_{\psi(i)} \quad \text{e} \quad \prod_{i \in \mathcal{I}} \kappa_i = \prod_{\psi(i) \in \mathcal{I}} \kappa_{\psi(i)}.$$

(iii) (Associatividade) Se $\{J : J \in \mathcal{J}\}$ é uma partição de \mathcal{I} , então

$$\sum_{J \in \mathcal{J}} \left(\sum_{j \in J} \kappa_j \right) = \sum_{i \in \mathcal{I}} \kappa_i \quad \text{e} \quad \prod_{J \in \mathcal{J}} \left(\prod_{j \in J} \kappa_j \right) = \prod_{i \in \mathcal{I}} \kappa_i.$$

Demonstração. A monotonicidade é imediata da proposição anterior, e as demais afirmações referentes a adição são imediatas. Para o produto, a correspondência $\langle a_i \rangle_{i \in \mathcal{I}} \mapsto \langle a_{\psi(i)} \rangle_{\psi(i) \in \mathcal{I}}$ define uma bijeção entre $\prod_{i \in \mathcal{I}} \kappa_i$ e $\prod_{\psi(i) \in \mathcal{I}} \kappa_{\psi(i)}$, o que estabelece a comutatividade. Para a associatividade, note que numa \mathcal{J} -upla $\langle f_J \rangle_{J \in \mathcal{J}}$, cada f_J é uma J -upla, i.e., $f_J = \langle f_J(j) \rangle_{j \in J}$. Como \mathcal{J} é uma partição de \mathcal{I} , para cada $i \in \mathcal{I}$ existe um único $J(i) \in \mathcal{J}$ tal que $i \in J(i)$, e daí a função

$$\begin{aligned} \varphi: \prod_{J \in \mathcal{J}} \left(\prod_{j \in J} \kappa_j \right) &\longrightarrow \prod_{i \in \mathcal{I}} \kappa_i \\ \langle f_J \rangle_{J \in \mathcal{J}} &\mapsto \langle f_{J(i)}(i) \rangle_{i \in \mathcal{I}} \end{aligned}$$

é claramente uma bijeção. \square

Proposição E.2.6. Sejam A, B, C e D conjuntos com $|A| \leq |C|$ e $|B| \leq |D|$. Então vale $|A^B| \leq |C^D|$. Em particular, se $|A| = |C|$ e $|B| = |D|$, então $|A^B| = |C^D|$.

Demonstração. O resultado é trivial caso C ou D sejam vazios. Supondo ambos não-vazios, basta mostrar que toda função $B \rightarrow A$ é restrição⁷ de uma função $D \rightarrow C$, o que dá uma sobrejeção $C^D \rightarrow A^B$. \square

Corolário E.2.7. Se $\kappa_0 \leq \kappa_1$ e $\lambda_0 \leq \lambda_1$ são cardinais, então $\kappa_0^{\lambda_0} \leq \kappa_1^{\lambda_1}$.

Proposição E.2.8 (Exercício C.5, item d)). Para qualquer conjunto X vale $|\wp(X)| = 2^{|X|}$.

Corolário E.2.9 (Cantor, repaginado). Para qualquer cardinal κ ocorre $\kappa < 2^\kappa$.

⁷Pois podemos pensar em $B \subseteq D$ e $A \subseteq C$ por meio das injeções dadas pela hipótese.

Exercício E.2. A seguir, números cardinais (possivelmente finitos) serão denotados por letras gregas.

a) Demonstre as seguintes identidades.

- i) $\kappa \cdot 0 = 0, \kappa + 0 = \kappa, 1 \cdot \kappa = \kappa, \kappa^0 = 1$ e $1^\kappa = 1$.
- ii) $\kappa = \sum_{\alpha \in \kappa} 1$. Dica: $X = \bigcup_{x \in X} \{x\}$.
- iii) $\sum_{i \in \mathcal{I}} \kappa_i = 0 \Leftrightarrow \kappa_i = 0$ para todo i . Dica: $\bigcup_{i \in \mathcal{I}} X_i = \emptyset \Leftrightarrow X_i = \emptyset$ para todo i .
- iv) $\prod_{i \in \mathcal{I}} \kappa_i = 0 \Leftrightarrow$ algum $\kappa_i = 0$. Dica: quando $\prod_{i \in \mathcal{I}} X_i = \emptyset$?

b) Demonstre as seguintes propriedades distributivas.

- (i) $\kappa \cdot (\sum_{i \in \mathcal{I}} \lambda_i) = \sum_{i \in \mathcal{I}} \kappa \cdot \lambda_i$. Dica: mostre que $X \times (\bigcup_{i \in \mathcal{I}} Y_i) = \bigcup_{i \in \mathcal{I}} (X \times Y_i)$.
- (ii) $(\kappa^\lambda)^\mu = \kappa^{\lambda \cdot \mu}$. Dica: encontre uma bijeção entre $(X^Y)^Z$ e $X^{Y \times Z}$. Metadica: pense simples.
- (iii) $(\prod_{i \in \mathcal{I}} \kappa_i)^\lambda = \prod_{i \in \mathcal{I}} \kappa_i^\lambda$. Dica: encare a bijeção em (A.3) até que ela te encare de volta.
- (iv) $(\prod_{i \in \mathcal{I}} \kappa^{\lambda_i}) = \kappa^{\sum_{i \in \mathcal{I}} \lambda_i}$. Dica: se $\{X_i : i \in \mathcal{I}\}$ é uma partição de um conjunto X , então determinar uma função $X \rightarrow Y$ é determinar uma função $X_i \rightarrow Y$ para cada $i \in \mathcal{I}$.

c) Demonstre as desigualdades a seguir.

- (i) $\sup_{\alpha \in \lambda} \kappa_\alpha \leq \sum_{\alpha \in \lambda} \kappa_\alpha$. Dica: $\kappa_\beta \leq \sum_{\alpha \in \lambda} \kappa_\alpha$ para todo $\beta \in \lambda$.
- (ii) $\sum_{\alpha \in \lambda} \kappa_\alpha \leq \lambda \cdot \sup_{\alpha \in \lambda} \kappa_\alpha$. Dica: para $\kappa := \sup_{\alpha \in \lambda} \kappa_\alpha$, observe que $\kappa_\alpha \leq \kappa$ ocorre para todo $\alpha \in \lambda$, daí $\sum_{\alpha \in \lambda} \kappa_\alpha \leq \sum_{\alpha \in \lambda} \kappa = \lambda \kappa$.
- (iii) Se $\lambda > 0$, então $\kappa \leq \kappa \lambda$. Dica: procure uma injecção $\kappa \rightarrow \kappa \times \lambda$.
- (iv) Se $\lambda > 0$, então $\kappa \leq \kappa^\lambda$.
- (v) Se $\kappa > 1$, então $\lambda \leq \kappa^\lambda$. Dica: como $\kappa > 1$, tem-se $0, 1 \in \kappa$. ■

Antes de nos focarmos nas especificidades de cardinais infinitos, há uma última desigualdade que merece atenção (compare com o Exercício C.24):

Teorema E.2.10 (König). *Se $\kappa_i < \lambda_i$ para todo $i \in \mathcal{I}$, então $\sum_{i \in \mathcal{I}} \kappa_i < \prod_{i \in \mathcal{I}} \lambda_i$.*

Demonstração. Sejam $\langle A_i : i \in \mathcal{I} \rangle$ e $\langle B_i : i \in \mathcal{I} \rangle$ \mathcal{I} -uplas de conjuntos dois-a-dois disjuntos, com $|A_i| = \kappa_i$ e $|B_i| = \lambda_i$. Mostraremos que se $f: \bigcup_{i \in \mathcal{I}} A_i \rightarrow \prod_{i \in \mathcal{I}} B_i$ é uma função, então f não é sobrejetora, donde a desigualdade desejada seguirá. Se $\pi_j: \prod_{i \in \mathcal{I}} B_i \rightarrow B_j$ é a projeção em j , então $X_j := B_j \setminus \pi_j[f[A_j]] \neq \emptyset$ para cada $j \in \mathcal{I}$, pois $\kappa_j < \lambda_j$. Note que não há $x \in \bigcup_{i \in \mathcal{I}} A_i$ tal que $f(x) \in X := \prod_{i \in \mathcal{I}} X_i$, mas $\emptyset \neq X \subseteq \prod_{i \in \mathcal{I}} B_i$. □

Exercício E.3 (For fun⁸). Use o Teorema de König para mostrar que $\kappa < 2^\kappa$ para todo cardinal κ . ■

⁸Há outras aplicações, como o Exercício E.29.

E.2.2 Aritmética transfinita

Para cardinais κ e λ , os procedimentos para determinar $\kappa + \lambda$ e $\kappa \cdot \lambda$ podem ser bastante intrincados a depender da *natureza* de κ e λ . A situação típica em que ambos são finitos é, certamente, conhecida pelo leitor: trata-se do que se faz em *Teoria dos Números, Aritmética* e outros campos afins. Tendo em vista as propriedades já demonstradas na subseção anterior, convém apenas reforçar a seguinte

Proposição E.2.11. *Sejam X e Y conjuntos. Se X e Y são finitos, então $X \cup Y$, $X \times Y$ e $\wp(X)$ são finitos.*

Exercício E.4. Demonstre a proposição acima. Dica: proceda por indução em $|X|$. ■

Em vista disso, segue que $\kappa + \lambda$, $\kappa \cdot \lambda$ e κ^λ pertencem a ω sempre que κ e λ forem números cardinais finitos, *a.k.a.* números naturais. Em particular, ficam bem definidas funções da forma $\omega \times \omega \rightarrow \omega$ que fazem $\langle m, n \rangle \mapsto m + n$, $\langle m, n \rangle \mapsto m \cdot n$ e $\langle m, n \rangle \mapsto m^n$.

A situação em que pelo menos um dos cardinais é infinito, por outro lado, é drasticamente distinta.

Teorema E.2.12 (Tarski). *Se A é um conjunto infinito, então $|A| = |A \times A|$.*

Demonstração. Como A é infinito, existe um subconjunto $\mathcal{N} \subseteq A$ com $|\mathcal{N}| = \omega$: deve-se ter $|A| \geq \aleph_0$, o que se traduz na existência de função injetora $\varphi: \omega \rightarrow A$, de modo que basta tomar $\mathcal{N} := \varphi[\omega]$. Logo, o conjunto

$$\mathbb{P} := \{\langle M, g \rangle \mid \mathcal{N} \subseteq M \subseteq A \text{ e } g: M \rightarrow M \times M \text{ é bijeção}\}$$

é não-vazio, pois existe uma bijeção entre \mathcal{N} e $\mathcal{N} \times \mathcal{N}$ (Teorema C.2.6). Define-se então a relação binária \preceq em \mathbb{P} , dada por $\langle M, g \rangle \preceq \langle K, h \rangle \Leftrightarrow M \subseteq K$ e $g \subseteq h$, claramente uma ordem parcial. O próximo passo é apelar para o Lema de Zorn: note que se $\mathcal{C} := \{\langle M_i, g_i \rangle : i \in \mathcal{I}\}$ é uma cadeia em \mathbb{P} , então $\langle \bigcup_{i \in \mathcal{I}} M_i, \bigcup_{i \in \mathcal{I}} g_i \rangle \in \mathbb{P}$ é um limitante superior para \mathcal{C} ; logo, existe $\langle B, f \rangle \in \mathbb{P}$ maximal. Note que $\aleph_0 \leq |B|$.

Para encerrar, basta mostrar que para $\kappa := |B|$ e $\lambda := |A \setminus B|$, deve ocorrer $\lambda \leq \kappa$: com efeito, uma vez estabelecida a desigualdade, teremos $\kappa^2 = \kappa$ (pois $\langle B, f \rangle \in \mathbb{P}$) e

$$|A| \leq |A \times A| = (\lambda + \kappa)^2 = \lambda^2 + \lambda\kappa + \kappa\lambda + \kappa^2 \leq 4\kappa^2 = 4\kappa \leq \kappa^2 = \kappa \leq |A|.$$

Supondo, por absurdo, $\kappa < \lambda$, deve existir um subconjunto $D \subseteq A \setminus B$ com $|D| = \kappa$, o que permite definir $B' := B \cup D$, conjunto que contém B propriamente e satisfaz $|B'| = \kappa + \kappa$. Agora, observe que o subconjunto $E := (B' \times B') \setminus (B \times B)$ satisfaz $|E| = \kappa$, pois $D \times D \subseteq E$ e, por conseguinte,

$$\kappa = |D| \leq |D \times D| \leq |E| \leq (\kappa + \kappa)^2 = 4\kappa^2 = 4\kappa \leq \kappa^2 = \kappa.$$

Logo, existe uma bijeção $h: D \rightarrow E$, que pode ser “colada” com a bijeção $f: B \rightarrow B \times B$ a fim de obter uma bijeção $g: B' \rightarrow B' \times B'$ (como no Exercício A.33), mostrando assim que $\langle B, f \rangle \prec \langle B', g \rangle$, contrariando a maximalidade de $\langle B, f \rangle$ na ordem (\mathbb{P}, \preceq) . □

Observação E.2.13 (*Je le vois, mais je ne le crois pas*⁹). O teorema acima demonstra que *qualquer conjunto infinito* está em bijeção com seu quadrado. Assim, não apenas ocorre $|\omega| = |\omega \times \omega|$, como também $|\mathbb{R}| = |\mathbb{R}^2|$ e, por indução, $|\mathbb{R}| = |\mathbb{R}^n|$ para qualquer $n \in \mathbb{N}$ (vide o corolário a seguir). Resultados como esse motivaram um estudo mais cuidadoso das diversas noções de *dimensão* pertinentes a um *espaço*. △

⁹Do francês, “eu vejo, mas não acredito”. Trecho de uma das muitas cartas que Cantor enviou a Dedekind durante o quarto final do século XIX. Nesta, ele discutia sua demonstração de que \mathbb{R} e \mathbb{R}^n têm a mesma cardinalidade, para qualquer $n \in \mathbb{N}$ [10].

Corolário E.2.14. Se $\kappa \geq \aleph_0$, então $\kappa^n = \kappa$ para todo $n \in \mathbb{N}$.

Demonstração. Lembre-se de que neste texto, $\mathbb{N} := \omega \setminus \{0\}$. Agora, para $n := 1$ o resultado é obviamente verdadeiro. Supondo $\kappa^n = \kappa$ para $n > 1$, note que $\kappa^{n+1} = \kappa^n \kappa = \kappa \kappa = \kappa^2 = \kappa$, onde a última igualdade decorre precisamente do teorema anterior. \square

Corolário E.2.15. Sejam κ e λ cardinais, ambos diferentes de 0, com pelo menos um deles infinito. Então $\kappa + \lambda = \kappa \cdot \lambda = \max\{\kappa, \lambda\}$.

Demonstração. Assuma $\lambda \geq \aleph_0$ com $\kappa \leq \lambda$. Então $\max\{\kappa, \lambda\} = \lambda$ e

$$\lambda \leq \kappa + \lambda \leq \lambda + \lambda = 2 \cdot \lambda \leq \lambda \cdot \lambda = \lambda = 1 \cdot \lambda \leq \kappa \cdot \lambda \leq \lambda \cdot \lambda = \lambda,$$

onde a igualdade desejada segue. \square

Observação E.2.16. O Lema de Zorn é *indispensável* nos resultados anteriores, no seguinte sentido.

Teorema E.2.17. As seguintes afirmações são equivalentes em ZF.

(AC₃) Se $\mathcal{A} \neq \emptyset$ e $\emptyset \notin \mathcal{A}$, então existe uma função $f: \mathcal{A} \rightarrow \bigcup \mathcal{A}$ tal que $f(A) \in A$ para cada $A \in \mathcal{A}$.

(AC₁₂) $\kappa^2 = \kappa$ para todo cardinal $\kappa \geq \aleph_0$.

(AC₁₃) $\kappa + \lambda = \kappa \lambda$ para quaisquer cardinais $\kappa, \lambda \geq \aleph_0$.

O leitor interessado pode consultar [18] ou, alternativamente, demonstrar por conta própria: as implicações (AC₃) \Rightarrow (AC₁₂) e (AC₁₂) \Rightarrow (AC₁₃) já foram provadas; para (AC₁₃) \Rightarrow (AC₃), costuma-se usar o número de Hartogs, entre outros truques¹⁰. \triangle

Como $\kappa^2 = \sum_{\alpha < \kappa} \kappa$, a igualdade $\kappa^2 = \kappa$ permite particionar κ em κ subconjuntos de cardinalidade κ . Interpretações análogas e mais surpreendentes podem ser feitas para alguns dos próximos resultados.

Definição E.2.18. Dados um conjunto X , um ordinal α e um cardinal κ , denotam-se:

- (i) $X^{<\alpha} := \bigcup_{\gamma < \alpha} X^\gamma$;
- (ii) $X^{\leq \alpha} := X^{<\alpha} \cup X^\alpha$;
- (iii) $[X]^\kappa := \{Y \subseteq X : |Y| = \kappa\}$;
- (iv) $[X]^{<\kappa} := \{Y \subseteq X : |Y| < \kappa\}$;
- (v) $[X]^{\leq \kappa} := \{Y \subseteq X : |Y| \leq \kappa\}$.

¶

Proposição E.2.19. Sejam X um conjunto e κ um cardinal. Valem as desigualdades: $|[X]^\kappa| \leq |X^\kappa|$, $|[X]^{<\kappa}| \leq |X^{<\kappa}|$ e $|[X]^{\leq \kappa}| \leq |X^{\leq \kappa}|$. Além disso, se $\kappa \leq |X|$ com $|X| \geq \aleph_0$, então também valem as desigualdades opostas.

¹⁰Apesar da simplicidade da dica, não se engane: não são truques imediatos.

Demonstração. Para cada subconjunto $S \in [X]^\kappa$ existe uma injeção $f_S: \kappa \rightarrow X$ com $\text{im}(f_S) = S$, de modo que a correspondência $S \mapsto f_S$ define uma injeção $[X]^\kappa \rightarrow X^\kappa$ e, portanto, $|[X]^\kappa| \leq |X^\kappa|$. Se valer $\kappa \leq |X|$ com $|X| \geq \aleph_0$, então uma função $f \in X^\kappa$ é um subconjunto de $\kappa \times X$ com cardinalidade precisamente κ , o que define uma injeção $X^\kappa \rightarrow [\kappa \times X]^\kappa$ e, consequentemente, $|X^\kappa| \leq |[\kappa \times X]^\kappa| = |[X]^\kappa|$ (a última igualdade usa, entre outras coisas, o fato de que $|\kappa \times X| = |X|$). As demais desigualdades, em ambos os casos, são análogas. \square

Exercício E.5. Complete a demonstração anterior. ■

Corolário E.2.20. *Sejam $\kappa \geq \aleph_0$ e X um conjunto com $|X| \leq 2^\kappa$. Então $|X^{\leq \kappa}| \leq 2^\kappa$. Em particular, se X é infinito, então $|X^{< \aleph_0}| = |[X]^{< \aleph_0}| = |X|$, i.e., X tem $|X|$ sequências finitas e $|X|$ subconjuntos finitos.*

Demonstração. Para cada $\alpha \leq \kappa$ ocorre $|X^\alpha| \leq |X^\kappa| \leq (2^\kappa)^\kappa = 2^{\kappa \cdot \kappa} = 2^\kappa$. Daí

$$|X^{\leq \kappa}| = \left| \bigcup_{\alpha \leq \kappa} X^\alpha \right| \leq \sum_{\alpha \leq \kappa} |X^\alpha| \leq \sum_{\alpha \leq \kappa} 2^\kappa \leq \kappa \cdot 2^\kappa = 2^\kappa.$$

Em particular, um raciocínio análogo mostra que

$$|X^{< \aleph_0}| \leq \sum_{n \in \omega} |X^n| = \sum_{n \in \omega} |X| = \aleph_0 \cdot |X| \leq |X|^2 = |X| \leq |X^{< \aleph_0}|,$$

pois $|X| = |X^1|$ e $X^1 \subseteq X^{< \aleph_0}$. O restante é consequência da proposição anterior. \square

Exercício E.6. Complete a demonstração anterior. Dica: $x \mapsto \{x\}$. ■

Observação E.2.21 (Um *paradoxo de expressividade*). Seja \mathcal{L} a coleção dos símbolos usados na linguagem formal da teoria dos conjuntos ($\in, =, \wedge, \vee$, etc.) juntamente com as variáveis usadas para escrever as fórmulas, digamos x_n para cada $n \in \omega$. Assim, $|\mathcal{L}| = \aleph_0$. Como as fórmulas da linguagem são obtidas por meio da concatenação de finitos elementos de \mathcal{L} , segue que a gramática correspondente \mathcal{G} , i.e., o conjunto de todas as *palavras*, é um subconjunto de $\mathcal{L}^{< \aleph_0}$, donde segue que $\aleph_0 \leq |\mathcal{G}| \leq |\mathcal{L}^{< \aleph_0}| = |\mathcal{L}| = \aleph_0$. O leitor deve se perguntar: e daí?

Ora, enquanto o raciocínio acima diz que existem somente \aleph_0 fórmulas na linguagem, o Teorema de Cantor estabelece a existência de $2^{\aleph_0} > \aleph_0$ funções da forma $\omega \rightarrow \{0, 1\}$. Verbalmente: há (bem mais!) funções da forma $\omega \rightarrow \{0, 1\}$ do que sonha a nossa vã linguagem. Desse modo, a definição de função adotada, embora precisa (em algum sentido), cria *monstros* que não se imaginavam existir, num sentido essencialmente *lovecraftiano*: os monstros existem, mas são literalmente indescritíveis. \triangle

Teorema E.2.22. *Sejam $\lambda \geq \aleph_0$ um cardinal e $\langle \kappa_\alpha : \alpha < \lambda \rangle$ uma λ -upla de cardinais, com $\kappa_\alpha > 0$ para todo $\alpha \in \lambda$. Então $\sum_{\alpha < \lambda} \kappa_\alpha = \lambda \cdot \sup_{\alpha < \lambda} \kappa_\alpha$.*

Demonstração. Já se tem $\sum_{\alpha < \lambda} \kappa_\alpha \leq \lambda \cdot \sup_{\alpha < \lambda} \kappa_\alpha$. Para verificar a outra desigualdade, observe que por valer $1 \leq \kappa_\alpha$ para todo α , segue que $\lambda = \sum_{\alpha < \lambda} 1 \leq \sum_{\alpha < \lambda} \kappa_\alpha$. Como $\kappa := \sup_{\alpha < \lambda} \kappa_\alpha \leq \sum_{\alpha < \lambda} \kappa_\alpha$ e $\lambda \geq \aleph_0$, resulta $\lambda \cdot \kappa = \max\{\lambda, \kappa\} \leq \sum_{\alpha < \lambda} \kappa_\alpha$. \square

Corolário E.2.23 (Compare com o Corolário C.2.9). *Sejam $\kappa \geq \aleph_0$ um cardinal e \mathcal{X} uma família de conjuntos com $|X| \leq \kappa$ para todo $X \in \mathcal{X}$. Então $|\bigcup \mathcal{X}| \leq \kappa \cdot |\mathcal{X}|$. Em particular, se $|\mathcal{X}| \leq \kappa$, então $|\bigcup \mathcal{X}| \leq \kappa$.*

Exemplo E.2.24 (Estimativas cardinais). Como aplicação simples do último teorema, note que $\sum_{n<\omega} n = \aleph_0 \cdot \sup_{n<\omega} n = \aleph_0 \cdot \aleph_0 = \aleph_0$, e, para $m \in \omega \setminus \{0, 1\}$ fixado,

$$\sum_{n<\omega} m^n = \aleph_0 \cdot \sup_{n<\omega} m^n = \aleph_0 \cdot \aleph_0 = \aleph_0, \quad (\text{E.1})$$

onde a igualdade $\sup_{n<\omega} m^n = \aleph_0$ segue pois $k \leq m^k$ para qualquer $k \in \omega$. Em particular, por meio de (E.1), mostra-se que se $0 < |X| < \aleph_0$, então $|X^{<\aleph_0}| = \aleph_0$. Por sua vez, mesmo o produto infinito de cardinais finitos tende a ser incontrolável. Para ilustrar isso, note que para $n \geq 2$ tem-se $2^{\aleph_0} \leq (2^{\aleph_0})^n = 2^{\aleph_0 \cdot n} = 2^{\aleph_0} \leq n^{\aleph_0} \leq \aleph_0^{\aleph_0} \leq (2^{\aleph_0})^{\aleph_0} = 2^{\aleph_0 \cdot \aleph_0} = 2^{\aleph_0}$, donde em particular resulta que $\prod_{n \in \omega \setminus \{0\}} n = 2^{\aleph_0}$. \blacktriangle

Observação E.2.25 (Contínuo). O cardinal 2^{\aleph_0} , mencionado no último exemplo, é bastante frequente na Matemática Contemporânea: um dos motivos é que a *reta real* tem cardinalidade 2^{\aleph_0} (Teorema C.3.2). Em particular, a identidade $2^{\aleph_0} = (2^{\aleph_0})^{\aleph_0}$ traduz o fato de que \mathbb{R} e \mathbb{R}^ω têm a mesma cardinalidade, explicitamente: existe bijeção entre \mathbb{R} e a coleção de *todas* as sequências da forma $\omega \rightarrow \mathbb{R}$.

Definição E.2.26. Também é comum denotar 2^{\aleph_0} pela letra gótica \mathfrak{c} , situação em que passa a ser xingado de (cardinalidade do) **contínuo** (ou *continuum*, para os preciosistas). ¶

Note que, pelo Teorema de Cantor, $\mathfrak{c} > \aleph_0$ e, por conseguinte, $\mathfrak{c} \geq \aleph_0^+ := \aleph_1$, já que \aleph_0^+ é o menor cardinal maior do que \aleph_0 . A (realmente grande) questão é: $\mathfrak{c} > \aleph_1$ ou $\mathfrak{c} = \aleph_1$? A resposta curta é: depende. Mais precisamente, ZFC não é capaz de decidir qual das duas alternativas acima escolher. Isto possivelmente ficará um pouco mais claro no próximo capítulo – mas não é uma promessa. De toda forma, o leitor já tem o ferramental para entender o conteúdo da **Hipótese do Contínuo**, frequentemente abreviada como CH: é a afirmação de que a identidade $\mathfrak{c} = \aleph_1$ é verdadeira.¹¹ \triangle

Exemplo E.2.27 (Opcional: boa definição da dimensão). O corolário anterior também torna bastante simples o trabalho de estabelecer a *invariância da cardinalidade* das bases de espaços vetoriais. Mais precisamente:

Teorema E.2.28 (Invariância da cardinalidade de bases). *Sejam K um corpo e V um K -espaço vetorial. Se G e L são duas bases de M , então $|L| = |G|$.*

Demonstração. Naturalmente, por *simetria*, basta mostrar que $|L| \leq |G|$, o que *será feito* apenas para o caso em que G é infinito¹². Note que para cada $g \in G$ existe um subconjunto finito $H_g \subseteq L$ tal que $g \in \text{sp}(H_g)$. Logo, $H := \bigcup_{g \in G} H_g \subseteq L$ e $M = \text{sp}(G) = \text{sp}(H)$. Se ocorresse $H \neq L$, então existiria $l \in L \setminus H$ com $l \in \text{sp}(H) \subseteq \text{sp}(L \setminus \{l\})$, o que contraria a independência linear de L . Desse modo, ocorre $L = H$ e, consequentemente, $|L| = \left| \bigcup_{g \in G} H_g \right| \leq |G| \cdot \aleph_0 = |G|$, como queríamos. \square

¹¹E o que levou Cantor a propor a pergunta? Certamente uma das motivações foi o fato de que em suas investigações iniciais com subconjuntos de \mathbb{R} , as cardinalidades encontradas eram, invariavelmente, $\leq \aleph_0$ ou \mathfrak{c} : como ilustração, pergunta-se “qual a cardinalidade de um subconjunto fechado de \mathbb{R} ?”; após algum tempo, você perceberá que as respostas são $n \in \omega$, \aleph_0 ou \mathfrak{c} .

¹²O leitor interessado no caso finito pode provar (ou usar) o seguinte resultado, conhecido como *Lema de Steinitz* (compare com o enunciado do Teorema D.3.5): *se $X, Y, Z \subseteq V$ são subconjuntos de V tais que $X \cup Y$ gera V e $Y \cup Z$ é linearmente independente, então para cada $z \in Z$ existe $x \in X$ tal que $(X \setminus \{x\}) \cup (Y \cup \{z\})$ gera V .* Com tal resultado, provar o Teorema E.2.28 para o caso em que $|G| < \aleph_0$ se transforma num exercício quase simples de indução finita – e sem a necessidade de apelar para matrizes ou sistemas lineares!

Tal teorema garante a *boa definição* da dimensão de espaços vetoriais como a cardinalidade de *qualquer base*: sempre existe uma (por AC) e quaisquer duas sempre têm a mesma cardinalidade. Além disso, com alguma paciência (e quocientes), pode-se usar tal resultado para mostrar que a *dimensão de módulos livres* (sobre anéis comutativos) também é invariante (embora, neste caso, seja mais comum xingá-la de *rank*). \blacktriangle

Outra consequência (não tão) marcante do Teorema E.2.15 é a trivialização da subtração de cardinais infinitos. De fato, se $\lambda < \kappa$ e $\kappa \geq \aleph_0$, então existe um único cardinal μ tal que $\mu + \lambda = \kappa$: a saber, $\mu = \kappa$. Ora, se $\lambda < \kappa$ e $\aleph_0 \leq \nu < \kappa$, então $\lambda + \nu = \max\{\lambda, \nu\} < \kappa$, donde segue que $\mu = \kappa$ é o único cardinal possível. Assim, faz sentido escrever $\kappa - \lambda = \kappa$.¹³

E.2.3 Opcional: aritmética ordinal

Também é possível desenvolver aritmética de números infinitos explorando-se a boa ordem dos números ordinais. Em certo sentido, a ideia é a mesma que se utiliza para desenvolver aritmética de números naturais sintaticamente, i.e., sem apelar para noções de cardinalidade.

Definição E.2.29. Seja α um ordinal.

- Define-se recursivamente o ordinal $\alpha + \beta$ da seguinte forma:
 - (i) $\alpha + 0 := \alpha$;
 - (ii) $\alpha + (\beta_+) := (\alpha + \beta)_+$ para todo ordinal β ; e
 - (iii) $\alpha + \beta := \sup_{\xi < \beta} \alpha + \xi$ se $\beta \neq 0$ é ordinal limite.
- Define-se recursivamente o ordinal $\alpha \cdot \beta$ da seguinte forma:
 - (i) $\alpha \cdot 0 := 0$;
 - (ii) $\alpha \cdot (\beta_+) := \beta \cdot \alpha + \beta$ para todo ordinal β ;
 - (iii) $\alpha \cdot \beta := \sup_{\xi < \beta} \alpha \cdot \xi$ se $\beta \neq 0$ é ordinal limite.
- Define-se recursivamente o ordinal α^β da seguinte forma:
 - (i) $\alpha^0 := 1$;
 - (ii) $\alpha^{\beta+} := \alpha^\beta \cdot \beta$ para todo ordinal β ;
 - (iii) $\alpha^\beta := \sup_{\xi < \beta} \alpha^\xi$ se $\beta \neq 0$ é ordinal limite. \P

Observação E.2.30. Fazendo-se $\beta := 1 := 0_+$ em $\alpha + \beta$, obtém-se

$$\alpha + 1 := (\alpha + 0)_+ := \alpha_+ := \alpha \cup \{\alpha\},$$

mostrando que a sucessão ordinal coincide com o processo de “somar 1” em aritmética ordinal. Por isso, daqui em diante, “ α_+ ” será denotado por “ $\alpha + 1$ ” (como de costume). Dito isso, as cláusulas de estágios sucessores nas *operações* acima se reescrevem como

$$\begin{aligned}\alpha + (\beta + 1) &:= (\alpha + \beta) + 1 \\ \alpha \cdot (\beta + 1) &:= \alpha \cdot \beta + \beta \\ \alpha^{\beta+1} &:= \alpha^\beta \cdot \beta\end{aligned}$$

para quaisquer ordinais α e β . \triangle

¹³ “ \aleph_0 garrafas de cerveja no muro, \aleph_0 garrafas de cerveja! Se uma garrafa cair no chão, quantas ficarão no muro?” (*ad infinitum*).

Em certo sentido, as operações ordinais iteram de diferentes formas a operação de acrescentar um último elemento numa lista bem ordenada de objetos. Quando tal lista é finita, gera-se um *novo* conjunto cuja *cardinalidade* é estritamente maior do que a anterior, razão pela qual as operações ordinais e cardinais coincidem em contextos finitos.

Exercício E.7. Convença-se de que as afirmações acima não são mentirosas. ■

Porém, as coincidências acabam em ω .

Exemplo E.2.31. Enquanto soma cardinal, tem-se $\omega + 1 = 1 + \omega = \omega$, já que o resultado de $\omega + 1$ corresponde ao número cardinal da reunião $X \cup Y$, onde $|X| = \omega$, $|Y| = 1$ e $X \cap Y = \emptyset$. Por outro lado, $1 + \omega \neq \omega + 1$ enquanto soma ordinal. Com efeito,

$$1 + \omega = \sup_{n < \omega} 1 + n = \omega,$$

enquanto $\omega < \omega + 1$. Isso ocorre pois, moralmente, $1 + \omega$ consiste em *plugar* um menor elemento *antes* de uma fileira de *tipo* ω , o que na prática apenas empurra os elementos de ω a fim de abrir uma vaga para o novo menor elemento. Mais um exemplo: $2^\omega = \omega$ enquanto *exponenciação ordinal*¹⁴, mas $2^\omega \geq \aleph_1 > \omega$ enquanto exponenciação cardinal. ▲

Exemplo E.2.32. Como outra ilustração, observe que aplicações sucessivas da adição ordinal de 1 se dão da seguinte forma:

- ✓ $0, 1, 2, \dots, n, n + 1, \dots, \omega = \sup_{n < \omega} n;$
- ✓ $\omega, \omega + 1, \omega + 2 = (\omega + 1) + 1, \omega + 3, \dots, \omega + n, \dots, \omega + \omega := \sup_{n < \omega} \omega + n;$
- ✓ $\omega + \omega = \omega \cdot 2, \omega \cdot 2 + 1, \omega \cdot 2 + 2, \dots, \omega \cdot 2 + n, \dots, \omega \cdot 2 + \omega := \omega \cdot 3;$
- ✓ $\omega \cdot 3, \omega \cdot 3 + 1, \dots, \omega \cdot 4, \dots, \omega \cdot n, \dots, \omega \cdot \omega := \sup_{n < \omega} \omega \cdot n;$
- ✓ $\omega \cdot \omega := \omega^2, \omega^2 + 1, \dots, \omega^2 \cdot 2, \dots, \omega^3, \dots, \omega^n, \dots, \omega^\omega := \sup_{n < \omega} \omega^n;$
- ✓ $\omega^\omega, \omega^\omega + 1, \dots, \omega^\omega \cdot 2, \dots, \omega^\omega \cdot \omega := \omega^{\omega+1};$
- ✓ $\omega^{\omega+1}, \dots, \omega^{\omega^2}, \dots, \omega^{\omega^3}, \dots, \omega^{\omega^\omega}, \dots, \omega^{\omega^\omega}, \dots, \varepsilon_0 := \sup\{\omega, \omega^\omega, \omega^{\omega^\omega}, \dots\};$
- ✓ $\varepsilon_0, \varepsilon_0 + 1, \dots, \varepsilon_0 + \omega, \dots, \varepsilon_0^\omega, \dots, \varepsilon_0^{\varepsilon_0}, \dots, \varepsilon_0^{\varepsilon_0^{\varepsilon_0}}, \dots, \varepsilon_1, \dots, \varepsilon_{\varepsilon_0}, \dots$

Note que da segunda linha em diante, todos os ordinais são infinitos mas, em nenhum dos estágios seguintes ocorre algum ordinal não-enumerável, afinal de contas, \aleph_1 é o primeiro ordinal/cardinal não-enumerável. É comum que a percepção deste *fato* leve alguns leitores a perderem a *fé* na *existência* de \aleph_1 ... em ZF(C). Por essa razão, convém revisitar os dois modos que podem ser utilizados para garantir que listas enumeráveis como a ilustrada acima não esgotam todos os ordinais. ▲

- Com o Axioma da Escolha, mostra-se que existe um ordinal γ em bijeção com $\wp(\omega)$ que deve ser não-enumerável pelo Teorema de Cantor. Tomando-se o menor desses γ 's, obtém-se \aleph_1 .

¹⁴De modo geral, exponenciação ordinal deve ser tratada como exceção neste texto: logo, coisas como ω^2 , por exemplo, indicam $\omega \times \omega$ ou $|\omega \times \omega|$, a menos de menção contrária.

- (ii) Moralmente, o número de Hartogs de ω , $H(\omega)$, é a coleção de todos os ordinais enumeráveis: de certa forma, cada ordinal α em $H(\omega)$ registra um modo diferente de ordenar ω numa fila. Nesse sentido, a prova de que $H(\omega)$ é o menor ordinal não-enumerável pode ser interpretada da seguinte forma: apenas com a *coleta* de todos os tipos possíveis de ordens para ω é que se alcança o primeiro tipo de ordem não-enumerável.

Tais exemplos ajudam a ilustrar que a aritmética ordinal se ocupa de informações mais finas do que *mera* cardinalidade, razão pela qual ordinais não costumam surgir em contextos corriqueiros de Matemática, mas apenas em cenários mais delicados em que iterações *transfinitas* de procedimentos sejam relevantes.

E.3 Cofinalidade e pombos

Considere os cardinais \aleph_1 e \aleph_ω , e suponha que cada um desses cardinais seja uma *escadaria*¹⁵ cujos degraus são determinados por seus ordinais menores. Mais do que isso, suponha que por alguma falha de caráter ou coisa do tipo, duas pessoas queiram subir tais escadarias o mais alto possível com apenas \aleph_0 passos.

Evidentemente, pessoas dispostas a fazer esse tipo de atividade física extrema não costumam considerar medidas de segurança razoáveis, como subir um degrau a cada novo passo: por isso, vamos assumir que elas subirão aos *saltos*, i.e., elas darão \aleph_0 saltos ao longo dos degraus das escadas. Porém, dado que viver perigosamente ainda não é a mesma coisa que atentar contra a própria vida, as personagens desta anedota imbecil vão se impor uma única restrição: seus saltos devem levá-las a degraus pertencentes às respectivas escadarias, a fim de evitar quedas infinitamente altas.

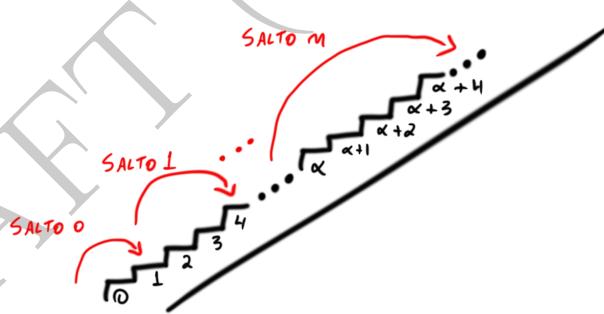


Figura E.1: Ninguém aguenta mais o Hotel de Hilbert.

A pergunta é: qual das duas pessoas chega ao final de sua respectiva escada?

- ✓ Em \aleph_ω , escala-se tão alto quanto desejado: basta estipular que o 0-ésimo salto leve até o \aleph_0 -ésimo degrau, o 1-ésimo salto leve até o \aleph_1 -ésimo degrau e, de modo geral, o n -ésimo salto deve *pousar* no \aleph_n -ésimo degrau. Posto que para qualquer ordinal $\alpha < \aleph_\omega$ existe $n < \omega$ com $\alpha < \aleph_n$, para efeitos práticos tal procedimento percorre a escadaria \aleph_ω completamente.
- ✗ Como $\aleph_1 < \aleph_\omega$, é natural imaginar que *deveria* ser possível galgar os degraus de \aleph_1 com a mesma eficiência do cenário anterior. Porém, não é o caso:

¹⁵Possivelmente construída em torno de um Hotel de Hilbert, quem sabe.

Proposição E.3.1. Sejam $\mathcal{D} \subseteq \aleph_1$ e $s: \mathcal{D} \rightarrow \aleph_1$ uma função. Se $|\mathcal{D}| \leq \aleph_0$, então $\sup_{d \in \mathcal{D}} s(d) < \aleph_1$.

Exercício E.8. Demonstre a proposição anterior. Dica: trata-se da reencarnação cardinal da Proposição C.5.8. ■

Portanto, fazendo \mathcal{D} como o conjunto de degraus escolhidos em \aleph_1 , segue que os saltos $s: \mathcal{D} \rightarrow \aleph_0$ alcançam, no máximo, o δ -ésimo degrau da escadaria \aleph_1 , com $\delta := \sup_{d \in \mathcal{D}} s(d)$.

Observação E.3.2. Na verdade, a situação é pior do que parece. Escrevendo

$$\aleph_1 = \{\gamma : \gamma \leq \delta\} \cup \{\xi : \delta < \xi < \aleph_1\}$$

com $\{\gamma : \gamma \leq \delta\} = \delta + 1$ enumerável, segue que $\{\xi : \alpha < \xi < \aleph_1\}$ é não-enumerável: se o último conjunto fosse enumerável, \aleph_1 seria enumerável, absurdo. △

A parte final do argumento acima é uma variação do chamado *Princípio da Casa dos Pombos*, em que secretamente se usam propriedades de *cofinalidade* de ordens. Classicamente, o **Princípio da Casa dos Pombos** (finito), que será abreviado como PCP, estabelece que se $n < \omega$ pombos são distribuídos em $m < \omega$ casas, com $m < n$, então em pelo menos uma das casas haverá mais de um pombo, ou, mais formalmente:

Proposição E.3.3 (PCP, finito). Se X é um conjunto finito e \mathcal{P} é uma partição de X com $|\mathcal{P}| < |X|$, então existe $P \in \mathcal{P}$ com $|P| > 1$.

Demonstração. Pela contrapositiva, se $|P| \leq 1$ para todo $P \in \mathcal{P}$, então para cada P existe $x_P \in X$ com $P = \{x_P\}$, e daí $X = \{x_P : P \in \mathcal{P}\}$, acarretando em $|X| \leq |\mathcal{P}|$. □

Exemplo E.3.4 (Baseado numa história real). Certa vez, tive a infelicidade de me deparar com uma notícia verdadeiramente triste: um acidente automobilístico havia vitimado quatro pessoas na cidade de Três Lagoas (MS). Apesar disso, não pude evitar a conclusão óbvia: uma das lagoas deve ter contido ao menos duas vítimas. ▲

Por sua vez, o argumento na Observação E.3.2 usou a seguinte versão:

Proposição E.3.5 (PCP, infinito). Sejam X um conjunto e \mathcal{P} uma partição de X . Se $|X| > \aleph_0$ e $|\mathcal{P}| \leq \aleph_0$, então existe $P_0 \in \mathcal{P}$ com $|P_0| > \aleph_0$. Em particular, se $|X| = \aleph_1$, então $|P_0| = \aleph_1$.

Dada a similaridade entre a Proposição E.3.5 e o PCP, é razoável chamar também a última de Princípio da Casa dos Pombos (infinito). Atentemo-nos para a segunda parte do enunciado: como $\aleph_0 < |P_0| \leq \aleph_1$, resta somente o caso $|P_0| = |X|$. O ponto a se destacar aqui é que a generalização não é automática para cardinalidades não-enumeráveis quaisquer: se $|X| = \aleph_\omega$, garante-se apenas $|P_0| = \aleph_n$ para algum $n \geq 1$. Como no caso das escadas, a sutileza está na *cofinalidade*.

Definição E.3.6. Seja $\langle \mathbb{P}, \leq \rangle$ uma ordem. Um subconjunto $C \subseteq \mathbb{P}$ é dito **cofinal** em \mathbb{P} (ou \leq -cofinal) se para todo $p \in \mathbb{P}$ existir $c \in C$ com $p \leq c$. ¶

Assim, por exemplo, mostrou-se que o conjunto

$$C := \{\aleph_n : n \in \omega\} \tag{E.2}$$

é cofinal em \aleph_ω , ao passo que nenhum subconjunto enumerável de \aleph_1 é cofinal em \aleph_1 . Como qualquer ordem é cofinal em si mesma, segue que

$$E_{\mathbb{P}} := \{|C| \leq |\mathbb{P}| : C \subseteq \mathbb{P} \text{ é cofinal em } \mathbb{P}\}$$

é uma família não-vazia de cardinais e, por conseguinte, admite um menor elemento.

Definição E.3.7. A **cofinalidade** de \mathbb{P} é o cardinal $\text{cof}(\mathbb{P}) := \min E_{\mathbb{P}}$, i.e., a menor cardinalidade de um subconjunto cofinal em \mathbb{P} . ¶

Exercício E.9. Mostre que $\text{cof}(\mathbb{P}) \leq |\mathbb{P}|$ se verifica para qualquer ordem \mathbb{P} . Em particular, $\text{cof}(\alpha) \leq |\alpha| \leq \alpha$ para qualquer ordinal α . ■

Como o conjunto C definido em (E.2) é enumerável, tem-se $\text{cof}(\aleph_\omega) \leq \aleph_0$. Por outro lado, como nenhum subconjunto finito pode ser cofinal em \aleph_ω , resulta $\text{cof}(\aleph_\omega) = \aleph_0$. Já para \aleph_1 , tem-se $\text{cof}(\aleph_1) \leq \aleph_1$ e $\text{cof}(\aleph_1) > \aleph_0$ (pois não existe subconjunto enumerável e cofinal em \aleph_1), e daí $\text{cof}(\aleph_1) = \aleph_1$.

Observação E.3.8. Curiosamente, há ordens infinitas com cofinalidades finitas: qualquer que seja o ordinal α , verifica-se a igualdade $\text{cof}(\alpha + 1) = 1$, já que o conjunto $\{\alpha\}$ é cofinal em $\alpha + 1$. Na verdade, vale a recíproca, i.e., se $\text{cof}(\beta) < \aleph_0$, então β é ordinal sucessor¹⁶. Por essa razão, apenas a cofinalidade de ordinais limites diferentes de 0 costuma ser interessante. △

Exercício E.10. Seja $\alpha \neq 0$ um ordinal.

- a) Mostre que α é ordinal limite se, e somente se, $\text{cof}(\alpha) \geq \aleph_0$.
- b) Se α é ordinal limite, o que é $\text{cof}(\text{cof}(\alpha))$? Dica: confira o Exercício E.27. ■

Em certo sentido, a cofinalidade dá a medida exata para generalizar a Proposição E.3.5:

Proposição E.3.9. *Sejam X um conjunto e \mathcal{P} uma partição de X . Se X é infinito e $|\mathcal{P}| < \text{cof}(|X|)$, então existe $P_0 \in \mathcal{P}$ tal que $|P_0| = |X|$.*

Demonstração. Como \mathcal{P} é uma família de subconjuntos de X dois-a-dois disjuntos tal que $X = \bigcup \mathcal{P}$, ocorre $|X| = \sum_{P \in \mathcal{P}} |P|$ e, por X ser infinito, tem-se $\text{cof}(|X|) \geq \aleph_0$. Por ω ser como Las Vegas¹⁷, se ocorrer $|\mathcal{P}| < \aleph_0$, então algum dos $P \in \mathcal{P}$ é infinito e, por conseguinte,

$$|X| = \sum_{P \in \mathcal{P}} |P| = \max_{P \in \mathcal{P}} |P|,$$

de modo que P_0 pode ser tomado como um dos finitos elementos de \mathcal{P} que realizam o máximo acima. Se, porém, ocorrer $|\mathcal{P}| \geq \aleph_0$, então

$$|X| = \sum_{P \in \mathcal{P}} |P| = |\mathcal{P}| \cdot \sup_{P \in \mathcal{P}} |P| = \max \left\{ |\mathcal{P}|, \sup_{P \in \mathcal{P}} |P| \right\}.$$

Agora, pela suposição de que $|\mathcal{P}| < \text{cof}(|X|)$ e por valer $\text{cof}(|X|) \leq |X|$, a igualdade acima acarreta $|X| = \sup_{P \in \mathcal{P}} |P|$. Por fim, observe que se valesse $|P| < |X|$ para todo $P \in \mathcal{P}$, então a última igualdade garantiria a cofinalidade do subconjunto $\{|P| : P \in \mathcal{P}\}$ em $|X|$, com $|\mathcal{P}| < \text{cof}(|X|)$, absurdo. □

Na maioria das vezes, o PCP se manifesta na forma da Proposição E.3.5: se X é não-enumerável e \mathcal{P} é uma partição enumerável de X , então existe $P_0 \in \mathcal{P}$ não-enumerável, sem a garantia de que $|P_0| = |X|$. A formulação dada pela Proposição E.3.9 (demonstrada acima) apenas visa dar um critério que garanta a igualdade $|P_0| = |X|$:

¹⁶Em particular, note que $\text{cof}(\beta) < \aleph_0$ se, e somente se, $\text{cof}(\beta) = 1$.

¹⁷Tudo o que se faz em Vegas, fica em Vegas.

- na Proposição E.3.5, pedir $|\mathcal{P}| \leq \aleph_0$ e $|X| = \aleph_1$ recai em $|\mathcal{P}| < \text{cof}(|X|)$, caso tratado pela Proposição E.3.9;
- se $|X| = \aleph_\omega$, então dizer que uma partição \mathcal{P} de X satisfaz $|\mathcal{P}| < \text{cof}(|X|)$ é afirmar que \mathcal{P} é uma partição finita – note que se $|\mathcal{P}| = \aleph_0$, então a Proposição E.3.9 não garante $P_0 \in \mathcal{P}$ com $|P_0| = \aleph_\omega$, o que é muito bom, já que tal P_0 poderia não existir.

Exemplo E.3.10 (Opcional: σ -álgebras¹⁸). Uma variação simples do PCP dá um argumento rápido para limitar inferiormente a cardinalidade de uma σ -álgebra infinita \mathcal{M} definida sobre um conjunto X .

Lema E.3.11. *Se \mathcal{N} é uma σ -álgebra infinita sobre um conjunto Y , então existe pelo menos um $A \in \mathcal{N} \setminus \{\emptyset, Y\}$ tal que $\mathcal{N}_A := \{E \cap A : E \in \mathcal{N}\}$ é subconjunto infinito de \mathcal{N} .*

Demonstração. Por \mathcal{N} ser infinito, existe $A \in \mathcal{N} \setminus \{\emptyset, Y\}$. Agora, com a notação sugerida pelo enunciado, note que \mathcal{N}_A ou $\mathcal{N}_{Y \setminus A}$ precisa ser infinito: com efeito, \mathcal{N} é subconjunto de $\mathcal{M} := \{C \cup D : \langle C, D \rangle \in \mathcal{N}_A \times \mathcal{N}_{Y \setminus A}\}$, já que para $B \in \mathcal{N}$ qualquer se verifica a identidade $B = (B \cap A) \cup (B \cap (Y \setminus A))$, e $|\mathcal{M}| \leq |\mathcal{N}_A| + |\mathcal{N}_{Y \setminus A}|$. Por fim, observe que $Y \setminus A \in \mathcal{N} \setminus \{\emptyset, Y\}$. \square

Agora, por \mathcal{N}_A ser uma σ -álgebra infinita sobre A , o argumento pode ser iterado no caso da σ -álgebra infinita \mathcal{M} sobre X , de modo a assegurar uma sequência $\langle E_n \rangle_{n \in \omega}$ com $E_n \in \mathcal{M}$ e $E_{n+1} \subsetneq E_n$ para todo $n \in \omega$. Ao fazer $U_n := E_n \setminus E_{n+1}$ para cada n , resulta que $\langle U_n \rangle_{n \in \omega}$ é uma sequência de elementos de \mathcal{M} , dois a dois disjuntos em X . Por fim, tal sequência permite definir a função $\varphi: \omega \rightarrow \mathcal{M}$ com $\varphi(N) := \bigcup_{n \in N} U_n$, injetora em virtude da disjunção entre os U_n 's. \blacktriangle

Exemplo E.3.12 (Opcional: Bolzano-Weierstrass). Um dos pilares da Análise na Reta é o *Teorema de Bolzano-Weierstrass*, que assegura a existência de subsequências convergentes (em \mathbb{R}) para sequências que sejam limitadas em \mathbb{R} . O roteiro usual da prova consiste em, primeiro, mostrar que sequências monótonas convergem¹⁹ e, depois, mostrar que sequências limitadas admitem subsequências monótonas. É na segunda etapa que o PCP pode se mostrar bastante eficaz.

Proposição E.3.13. *Toda sequência em \mathbb{R} admite subsequência monótona.*

Demonstração. Dada uma sequência real $\langle x_n \rangle_n$, pode-se considerar o conjunto $[\omega]^2$ de todos os subconjuntos de ω com precisamente dois elementos, e daí definir a função $c: [\omega]^2 \rightarrow \{\text{A, V}\}$ que faz $c(\{m, n\}) := \text{A}$ se $m < n$ e $x_m < x_n$, e $c(\{m, n\}) := \text{V}$ se $m < n$ com $x_m \geq x_n$. O leitor, com certa razão, pode se perguntar: *quê?!*

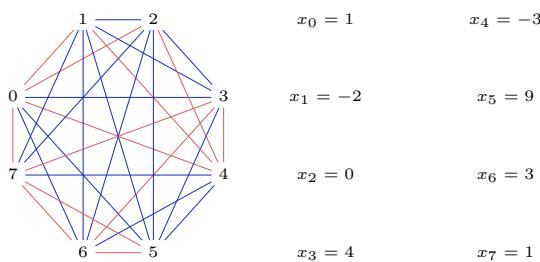


Figura E.2: Exemplo de um subconjunto finito do grafo descrito acima.

¹⁸O leitor pode recordar as definições envolvidas no Exemplo C.5.9.

¹⁹Para o supremo se forem crescentes, para o ínfimo se forem decrescentes.

Intuitivamente, a construção proposta consiste em considerar o *grafo infinito* cujos vértices são todos os números naturais e cujas arestas são todas as possíveis ligações entre eles. Nesse sentido, a função c pinta uma “aresta” $m \bullet \cdots \bullet n$ em que $m < n$: de Azul se ocorrer $x_m < x_n$; de Vermelho caso contrário²⁰.

Por que alguém faria isso? *Muito simples!* Um subconjunto infinito $M \subseteq \omega$ no qual qualquer aresta ligando seus vértices tenha a mesma cor se traduz numa subsequência monótona: (estritamente) crescente se a cor for Azul; decrescente se a cor for Vermelha. Por exemplo, se $c := A$, então $x_m < x_n$ sempre que $m, n \in M$ com $m < n$, ou seja: a subsequência $\langle x_m \rangle_{m \in M}$ é estritamente crescente. O raciocínio é análogo para $c := V$.

O passo fundamental na demonstração de que existe um subconjunto $M \subseteq \omega$ com a propriedade desejada faz uso da mesma versão do Princípio da Casa dos Pombos usado no último exemplo: *se X é infinito, $A, B \subseteq X$ são tais que $A \cap B = \emptyset$ e $A \cup B = X$, então A é infinito ou B é infinito.* Em particular, se $P \subseteq [\omega]^2$ é infinito, então P é união disjunta dos subconjuntos $\{p \in P : c(p) = A\}$ e $\{p \in P : c(p) = V\}$, donde segue que pelo menos um deles deve ser infinito.

Katzensprung. Fixados um subconjunto infinito $S \subseteq \omega$ e um elemento $s \in S$, existem um subconjunto infinito $G_{S,s} \subseteq S$ e uma cor $c_S \in \{A, V\}$ tal que $s < \min G_{S,s}$ e $c(\{s, n\}) = c_S$ para qualquer $n \in G_{S,s}$.

Demonstração. Em outras palavras, existe um subconjunto infinito de S cujas arestas que ligam seus elementos ao número s têm todas a mesma cor. Para se dar conta disso, note que o subconjunto $P := \{\{s, n\} : n > s \text{ e } n \in S\} \subseteq [\omega]^2$ é infinito e, pelo argumento do parágrafo anterior, existe uma cor $C \in \{A, V\}$ tal que $Q := \{p \in P : c(p) = C\}$ é infinito. Daí, basta tomar $G_{S,s} := (\bigcup Q) \setminus \{s\}$ e $c_S := C$. \square

Dito isso, mostraremos que existe uma sequência estritamente crescente $\langle k_n \rangle_n$ de números naturais tais que, para todo $n \in \omega$, existe $c_n \in \{A, V\}$ com $c(\{k_n, k_m\}) = c_n$ para todo $m > n$. De fato, em vista da argumentação anterior, basta proceder recursivamente:

- ✓ $S_0 := \omega$, $k_0 := \min S_0$ e $c_0 := c_{S_0}$;
- ✓ $S_1 := G_{S_0, k_0}$, $k_1 := \min S_1$ e $c_1 := c_{S_1}$;
- ✓ para $n \geq 1$, e supondo $S_0, \dots, S_n \subseteq \omega$ definidos com $S_n \subseteq S_{n-1} \subseteq \dots \subseteq S_0$, todos infinitos, com $k_i \in S_i$ para cada $i \leq n$, faz-se $S_{n+1} := G_{S_n, k_n}$, $k_{n+1} := \min S_{n+1}$ e $c_{n+1} := c_{S_{n+1}}$.

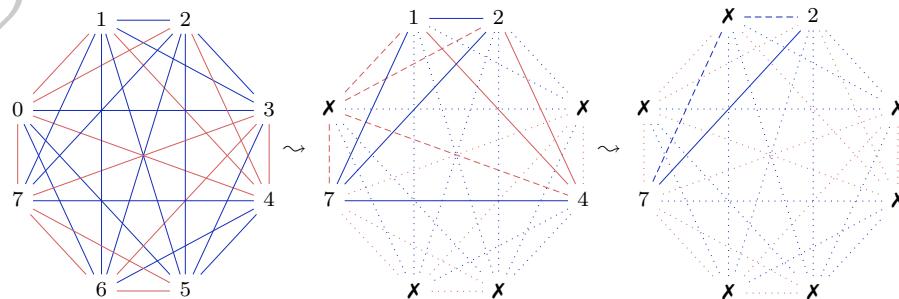


Figura E.3: Ilustração do procedimento anterior, com $c_0 = V$ e $c_1 = A$.

²⁰Evidentemente, A e V são apenas modos psicologicamente agradáveis de denotar 0 e 1, x e $\{x\}$ ou, mais geralmente, quaisquer dois conjuntos A e V com $A \neq V$.

Finalmente, a função $\varphi: \omega \rightarrow \{\text{A}, \text{V}\}$, que faz $\varphi(n) := c_n$ para cada $n \in \omega$, tem imagem finita, donde o Princípio da Casa dos Pombos garante um subconjunto infinito $T \subseteq \omega$ e $c \in \{\text{A}, \text{V}\}$ com $\varphi(t) = c$ para todo $t \in T$. Ora, isto significa, precisamente, que $c(\{k_s, k_t\}) = c$ para quaisquer $s, t \in T$ distintos. Logo, basta fazer $M := \{k_t : t \in T\}$. \square

Observação E.3.14. Essencialmente, a sequência de subconjuntos $\langle S_n \rangle_n$ seleciona, a cada passo n , um subconjunto infinito de índices $S_{n+1} \subsetneq S_n$ cujos termos correspondentes são monótonos com respeito ao primeiro termo de S_n , mas sem afetar o tipo de monotonicidade (crescente ou decrescente) que seus termos mantêm com o primeiro termo de S_{n-1} . Com base nas ilustrações da Figura E.3:

- (i) com $c_0 = \text{V}$ e $S_1 := \{1, 2, 4, 7, \dots\}$, indica-se que o conjunto de vértices S_1 é infinito e tal que todas as arestas ligando os seus vértices a 0 são vermelhas – isto é, $x_0 \geq x_s$ para todo $s \in S_1$; note que no segundo grafo restam apenas os vértices de S_1 ;
- (ii) com $k_1 = 1$ o menor elemento de S_1 , a ocorrência de $c_1 = \text{A}$ com $S_2 := \{2, 7, \dots\} \subsetneq S_1$ indica que S_2 é um subconjunto infinito de vértices (de S_1 !) cujas arestas que ligam seus vértices a 1 são todas azuis – ou seja, $x_1 < x_s$ para todo $s \in S_2$; note que no terceiro grafo restam apenas os vértices de S_2 ;
- (iii) observe que por valer $S_2 \subsetneq S_1$, ainda se tem $x_0 \geq x_s$ para todo $s \in S_2$!

Portanto, ao encontrar um subconjunto infinito de ω cujos c_n 's correspondentes coincidem, assegura-se que o tipo de monotonicidade que seus termos correspondentes mantêm um com os outros é o mesmo. \triangle

Exercício E.11. O Axioma da Escolha foi usado na demonstração acima? Dica: para dar conta desta recursão, precisa-se apenas de uma função oráculo da forma $Z^{<\omega} \rightarrow Z$, para algum Z apropriado. \blacksquare

O argumento de *coloração* acima, que aprendi com Leandro Aurichi, ilustra o *modus operandi* típico da chamada *Teoria de Ramsey*, que consiste em estudar problemas sobre cardinalidades de partições (infinitas, nos casos mais interessantes para *teoristas de conjuntos*). \blacktriangle

E.4 Cardinais singulares, regulares e além

Embora seja um tópico relativamente avançado, a discussão feita sobre cofinalidade faz deste o trecho ideal para introduzir as noções de singularidade e regularidade para ordinais.

Definição E.4.1. Um cardinal infinito κ é dito **singular** se $\text{cof}(\kappa) < \kappa$. Um cardinal infinito não-singular é chamado de **regular**. \P

Com a terminologia da seção anterior, um cardinal κ é regular se não for possível escalá-lo com menos do que κ saltos. Desse modo, \aleph_0 e \aleph_1 são exemplos de cardinais regulares, enquanto \aleph_ω é um caso clássico de cardinal singular. Mais geralmente, vale o seguinte

Teorema E.4.2. Se $\alpha \neq 0$ é ordinal, então

$$\text{cof}(\aleph_\alpha) = \begin{cases} \aleph_\alpha, & \text{se } \alpha \text{ é sucessor} \\ \text{cof}(\alpha), & \text{caso contrário} \end{cases}.$$

Em particular, todo cardinal da forma $\aleph_{\alpha+1}$ é regular.

Demonstração. Primeiro, note que se κ é um cardinal infinito, então $C \subseteq \kappa$ é cofinal se, e somente se, $\sup C = \kappa$:

- (\Rightarrow) é claro que $\sup_{\xi \in C} \xi \leq \kappa$; por outro lado, não há $\alpha < \kappa$ que limite C superiormente, posto que a cofinalidade garante um $\xi \in C$ com $\alpha < \alpha + 1 \leq \xi$;
- (\Leftarrow) se $\sup_{\xi \in C} \xi = \kappa$ e $\alpha < \kappa$, então α não pode ser limitante superior de C , donde segue que existe $\xi \in C$ com $\alpha < \xi$, mostrando que C é cofinal em κ .

Agora, se $\kappa = \lambda^+$ e $C \subseteq \kappa$ é cofinal, então $\lambda < |C|$: o contrário daria uma sobrejeção $\lambda \rightarrow C$, o que permitiria escrever $C = \{\alpha_\xi : \xi < \lambda\}$, com cada $\alpha_\xi < \kappa$; como κ é cardinal, teria-se $|\alpha_\xi| \leq \lambda < \kappa$, o que impossibilita a cofinalidade de C em κ , pois se $\gamma := \sup C = \bigcup_{\xi < \lambda} \alpha_\xi$, então

$$|\gamma| = \left| \bigcup_{\xi < \lambda} \alpha_\xi \right| \leq \sum_{\xi < \lambda} |\alpha_\xi| \leq \sum_{\xi < \lambda} \lambda = \lambda\lambda = \lambda < \kappa.$$

Logo, $|C| \geq \lambda^+ := \kappa$ e, portanto, $|C| = \kappa$. Da arbitrariedade do conjunto cofinal tomado, resulta $\text{cof}(\kappa) = \kappa$, o que prova a primeira parte da afirmação.

Agora, se α é limite, então $\alpha = \sup_{\beta < \alpha} \beta$ e $\aleph_\alpha := \sup_{\beta < \alpha} \aleph_\beta$. Daí, não é difícil perceber que um subconjunto $C \subseteq \alpha$ cofinal em α induz um subconjunto $\aleph(C) \subseteq \aleph_\alpha$ cofinal em \aleph_α com $|\aleph(C)| \leq |C|$, enquanto um conjunto $D \subseteq \aleph_\alpha$ cofinal em \aleph_α permite construir um subconjunto $A(D) \subseteq \alpha$ cofinal em α com $|A(D)| \leq |D|$, donde a igualdade desejada segue. \square

Exercício E.12. Complete a demonstração anterior. ■

Como todo número cardinal infinito κ é da forma \aleph_α para um único ordinal α , os xingamentos “sucessor” e “limite” também se aplicam a cardinais infinitos²¹: diz-se que \aleph_α é **sucessor** se o ordinal α for sucessor; caso contrário, \aleph_α é chamado de (cardinal) **limite**. Com tal terminologia, o teorema acima garante que todo cardinal sucessor é regular, mas nada se diz sobre a regularidade ou singularidade de cardinais limites.

O 0-ésimo cardinal limite, a saber \aleph_0 , é regular, posto que $\aleph_0 = \text{cof}(\aleph_0)$. Como o primeiro ordinal limite maior do que 0 é ω , segue que \aleph_ω é o 1-ésimo cardinal limite, mas este é singular pois vale $\text{cof}(\aleph_\omega) = \text{cof}(\omega) = \aleph_0$. O próximo ordinal limite é $\omega + \omega$ e, por conseguinte, $\aleph_{\omega+\omega}$ é o 2-ésimo cardinal limite, também singular já que $\text{cof}(\aleph_{\omega+\omega}) = \aleph_0$ (pense a respeito). Com um salto um pouco mais longo, percebe-se que \aleph_{ω^2} é o ω -ésimo cardinal limite, onde ω^2 é dado por exponenciação ordinal.

Um pouco mais geralmente, por valer $\text{cof}(\alpha) \leq |\alpha|$ para qualquer ordinal α (pelo Exercício E.9) e $\aleph_\beta > \aleph_0$ para todo $\beta > 0$, resulta que todo cardinal limite \aleph_α com $\alpha < \omega_1$ é tal que $\text{cof}(\aleph_\alpha) = \aleph_0$ e, portanto, é seguro dizer que todo cardinal limite menor do que \aleph_{ω_1} é singular. Ocorre que o próprio \aleph_{ω_1} é singular: pelo último teorema, tem-se $\text{cof}(\aleph_{\omega_1}) = \text{cof}(\omega_1) = \aleph_1 < \aleph_{\omega_1}$, já que $\omega_1 := \aleph_1$ é regular (pelo mesmo teorema). Parece então irresistível perguntar: existe *algum* cardinal limite $\kappa > \aleph_0$ com κ regular? Um cardinal assim certamente seria bastante *diferente* de todos os cardinais que já foram explicitados. Por exemplo:

Exercício E.13. Seja $\kappa > \aleph_0$ um cardinal.

²¹Cardinais finitos não recebem tais alcunhas.

- a) Mostre que se κ é limite e regular, então $\kappa = \aleph_\kappa$. Dica: com α limite tal que $\kappa = \aleph_\alpha$, use a regularidade de κ e a *natureza* de α para concluir que $\aleph_\alpha \leq \kappa$.
- b) Mostre que se $\kappa = \aleph_\kappa$, então κ é cardinal limite. ■

Em outras palavras, um cardinal limite regular não-enumerável seria *tão* grande quanto o seu próprio *aleph!* Curiosamente, essa última propriedade é mais comum do que parece e, por si só, não garante regularidade.

Proposição E.4.3. *Existem cardinais singulares arbitrariamente grandes satisfazendo a identidade $\kappa = \aleph_\kappa$.*

Demonstração. Dado um cardinal infinito λ , define-se $\kappa_0 := \lambda$ e, supondo κ_n definido para $n \in \omega$, estipula-se $\kappa_{n+1} := \aleph_{\kappa_n}$. Dessa forma, $\kappa := \sup_{n \in \omega} \kappa_n$ é um cardinal que satisfaz as condições impostas:

- (i) se $\beta < \kappa$, então existe $n \in \omega$ com $\beta < \kappa_n$, logo $\aleph_\beta \leq \sup_{\xi < \kappa_n} \aleph_\xi := \aleph_{\kappa_n} := \kappa_{n+1}$, o que mostra $\aleph_\kappa := \sup_{\beta < \kappa} \aleph_\beta \leq \kappa$;
- (ii) κ é singular pois, por construção, $C := \{\kappa_n : n \in \omega\}$ é um conjunto cofinal em κ ;
- (iii) λ foi tomado arbitrariamente, e o κ obtido satisfaz $\lambda < \kappa$. □

Assim, convém repetir a pergunta: existe *algum* cardinal limite $\kappa > \aleph_0$ com κ regular? A resposta pode ser surpreendente: ZFC *não pode* provar que tais cardinais existem. Em certo sentido, um cardinal com tais propriedades seria *grande demais*, grande a ponto de conseguir *modelar* todos os axiomas de ZFC, o que não costuma ser algo muito bom²². Talvez por isso, costuma-se chamar um cardinal κ que responde afirmativamente à pergunta como **fracamente inacessível** – onde o “fracamente” é colocado para distingui-lo de outro tipo de cardinal ainda mais *inacessível* (confira o Exercício F.27).

Observação E.4.4. Isto encerra o que considero ser uma introdução honesta à Teoria dos Conjuntos sob a axiomática de ZFC. O leitor que tiver interesse em se aprofundar *verdadeiramente* pode começar com as obras canônicas da área: Hrbacek & Jech [19], Jech [21] e Kunen [22, 24]; para os aspectos históricos, deliberadamente omitidos, a sugestão é o impressionante *Labyrinth of thought*, de Ferreirós [10]. Cabe destacar que existem outros sistemas axiomáticos considerados na literatura, como NBG e MK: o leitor interessado em descobrir o significado das siglas pode começar sua investigação no supracitado texto de Kunen [22]. △

Exercícios adicionais

Exercício E.14. Mostre que $\kappa + \kappa = 2\kappa$ para qualquer cardinal κ . ■

Exercício E.15. Mostre que existe uma partição $\mathcal{P} := \{P_n : n \in \omega\}$ de ω com cada P_n infinito. Enuncie e demonstre uma generalização adequada para X infinito qualquer. Dica: considere a função $\pi_0: X \times X \rightarrow X$ que faz $\pi(x, y) := x$ para cada par $\langle x, y \rangle \in X \times X$. ■

Exercício E.16. Pense rápido: se $f: \omega \rightarrow \omega$ é uma função sobrejetora, então existe $n \in \omega$ tal que $f^{-1}[\{n\}]$ é finito? ■

²²O leitor interessado em tais sutilezas *metamatemáticas* deve conferir o próximo capítulo.

Exercício E.17. Seja $\mathbb{R}[x]$ a família dos polinômios na indeterminada x e coeficientes em \mathbb{R} . Qual a cardinalidade de $\mathbb{R}[x]$? ■

Exercício E.18 (Requer noções básicas de topologia em \mathbb{R}). Suponha conhecido o seguinte fato: se $f, g: \mathbb{R} \rightarrow \mathbb{R}$ são contínuas e $f(q) = g(q)$ para todo $q \in \mathbb{Q}$, então $f = g$.

- Chamando por $\mathcal{C}(\mathbb{R}, \mathbb{R})$ o conjunto de todas as funções contínuas da forma $\mathbb{R} \rightarrow \mathbb{R}$, mostre que $|\mathcal{C}(\mathbb{R}, \mathbb{R})| = \mathfrak{c}$.
- For fun: utilizando argumentação cardinal, mostre que existem funções descontínuas da forma $\mathbb{R} \rightarrow \mathbb{R}$. Dica: $\mathfrak{c} < 2^{\mathfrak{c}}$. ■

Exercício E.19. Seja $f: X \rightarrow Y$ uma função.

- Mostre que $\mathcal{P} := \{f^{-1}[\{y\}] : y \in \text{im}(f)\}$ é uma partição de X .
- Suponha Y enumerável e X não-enumerável. Mostre que existe $y \in Y$ tal que o subconjunto $X' := \{x \in X : f(x) = y\}$ é não-enumerável. ■

Exercício E.20. Mostre que se $2 \leq \kappa \leq \lambda$ e $\lambda \geq \aleph_0$, então $\kappa^\lambda = 2^\lambda$. Dica: $\kappa < 2^\kappa$. ■

Exercício E.21. Pague a dívida contraída na Seção C.3: mostre que se X é infinito, então $|X| = |X \times \omega|$. ■

Exercício E.22. Mostre que se X é não-enumerável, então $|X \times \omega_1| = |X|$. ■

Exercício E.23 (For fun²³). Juntos temos \aleph_1 reais. Você tem \aleph_0 reais a mais do que eu. Quantos reais eu tenho? ■

Exercício E.24 (Opcional: boreelianos). Sejam \mathcal{E} uma família infinita de subconjuntos de X e $\sigma(\mathcal{E})$ a σ -álgebra gerada por \mathcal{E} . Considere também $\mathcal{E}_\alpha \subseteq \sigma(\mathcal{E})$ para todo $\alpha < \omega_1$, como na descrição feita ao longo do Exemplo C.5.9.

- Mostre que $|\sigma(\mathcal{E})| \leq \aleph_1 \cdot \sup_{\alpha < \omega_1} |\mathcal{E}_\alpha|$.
- Mostre que se $|\mathcal{E}| = \aleph_0$, então $|\mathcal{E}_\alpha| \leq \mathfrak{c}$ para cada $\alpha < \omega_1$. Conclua que $|\sigma(\mathcal{E})| \leq \mathfrak{c}$. Dica: para a primeira parte, argumete por indução em $\alpha < \omega_1$.
- Sejam $X := \mathbb{R}$ e \mathcal{E} a família dos intervalos abertos com extremos racionais. Mostre que $|\sigma(\mathcal{E})| = \mathfrak{c}$. Dica: para a desigualdade que falta, confira o Exemplo E.3.10. ■

Observação E.4.5. A σ -álgebra definida no último item do exercício anterior costuma ser chamada de **σ -álgebra de Borel** em \mathbb{R} . Por simplicidade, ela será denotada por $\mathcal{B}_{\mathbb{R}}$. △

Exercício E.25 (Opcional: Lebesgue mensuráveis). Uma **medida**²⁴ $m: \mathcal{A} \rightarrow [0, +\infty]$ definida numa σ -álgebra \mathcal{A} sobre um conjunto X é dita **completa** se para quaisquer $A, B \subseteq X$ com $A \subseteq B$ e $B \in \mathcal{A}$, a ocorrência de $m(B) = 0$ garantir que $A \in \mathcal{A}$.

- Mostre que se m é completa e existe $B \in \mathcal{A}$ com $m(B) = 0$, então $2^{|B|} \leq |\mathcal{A}|$.
- Assuma a existência de uma σ -álgebra \mathcal{L} sobre \mathbb{R} com as seguintes propriedades: $\mathcal{B}_{\mathbb{R}} \subseteq \mathcal{L}$; existe, uma medida completa $m: \mathcal{L} \rightarrow [0, +\infty]$ e um subconjunto $C \in \mathcal{B}_{\mathbb{R}}$ com $|C| := \mathfrak{c}$ e $m(C) = 0$. Nestas condições, mostre que existe $L \subseteq C$ tal que $L \in \mathcal{L} \setminus \mathcal{B}_{\mathbb{R}}$. Dica: $\wp(C) \subseteq \mathcal{L}$, mas $\wp(C) \not\subseteq \mathcal{B}_{\mathbb{R}}$ (por quê?!). ■

²³Créditos a Tiago Fonseca, a.k.a. @tejotaefe.

²⁴Confira a definição de medida aditivamente finita no Exemplo F.4.9, mas troque o seu codomínio por $[0, +\infty]$ e suponha que a condição de aditividade se verifique para quaisquer coleções enumeráveis de membros dois a dois disjuntos da σ -álgebra.

Observação E.4.6. A medida de Lebesgue, juntamente com sua σ -álgebra, constitui o exemplo canônico do que se observou no último item do exercício anterior. Em particular, o conjunto C costuma ser interpretado pelo *conjunto de Cantor*. \triangle

Exercício E.26. Para um conjunto X , exiba uma bijecção $X \times \{0, 1\} \rightarrow X \times \{0, 1\}$ sem pontos fixos. Conclua que se X é infinito, então existe uma bijecção $X \rightarrow X$ sem pontos fixos. ■

Exercício E.27. Sejam \mathbb{P} uma ordem parcial e $C, D \subseteq \mathbb{P}$ subconjuntos, com $C \subseteq D$. Mostre que se C é cofinal em D e D é cofinal em \mathbb{P} , então C é cofinal em \mathbb{P} . ■

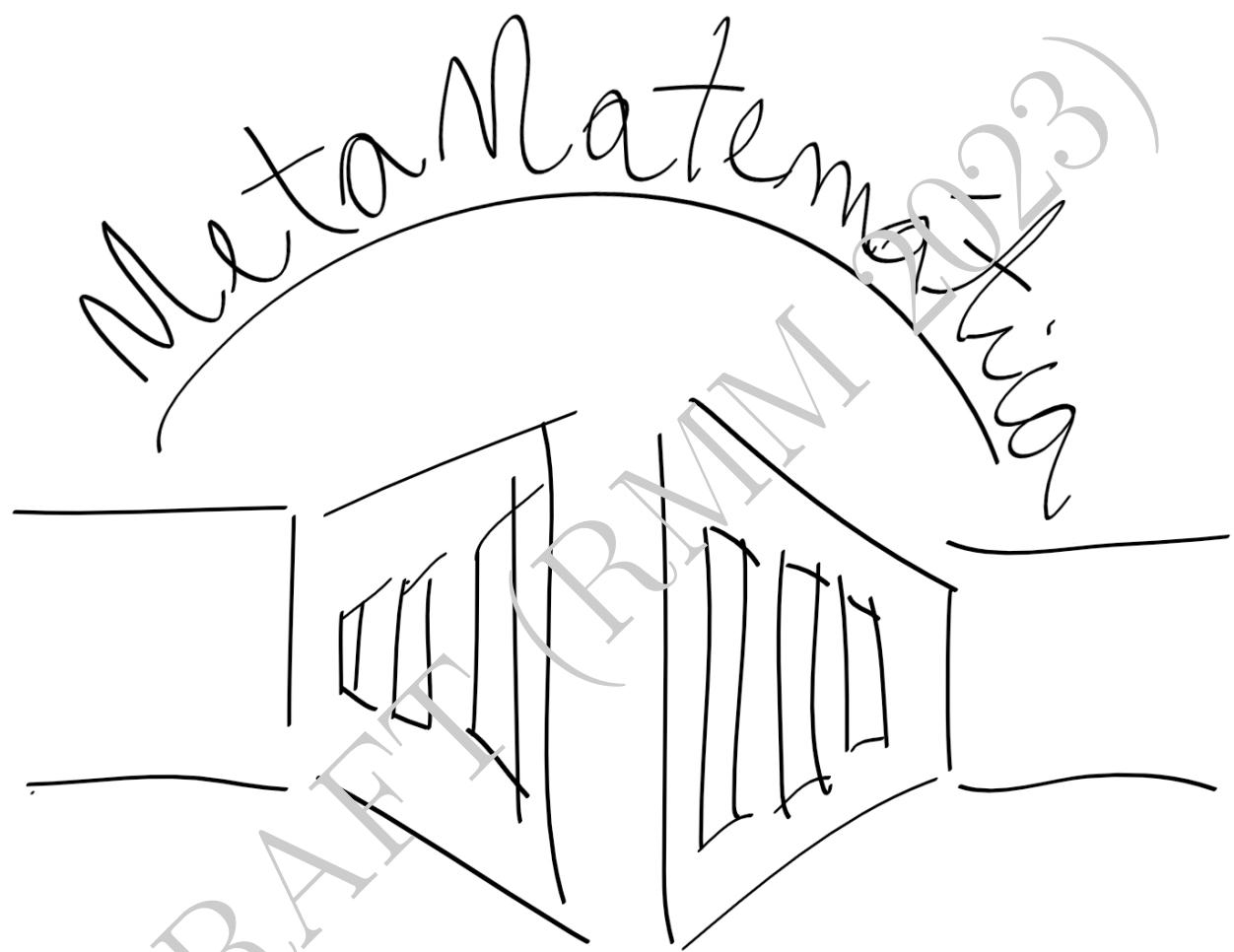
Exercício E.28. Determine quais das afirmações a seguir são *verdadeiras* (e quais são *falsas*).

- Se X é infinito e \mathcal{P} é uma partição de X com $|\mathcal{P}| < |X|$, então existe $P \in \mathcal{P}$ com $|P| = |X|$.
- Se X é infinito, então existe uma partição \mathcal{P} de X tal que $|P| = 2$ para todo $P \in \mathcal{P}$.
- Para cardinais $\kappa, \lambda > 0$, tem-se $\kappa^{\aleph_0} < \lambda^{\aleph_0}$ sempre que $\kappa < \lambda$. ■

Exercício E.29. Sejam κ e λ cardinais.

- Mostre que se $\kappa \geq \aleph_0$, então $\kappa < \kappa^{\text{cof}(\kappa)}$. Dica: encontre um subconjunto cofinal em κ de cardinalidade $\text{cof}(\kappa)$ e dê um jeito de usar o Teorema de König (Exercício E.2.10).
- Mostre que se $\kappa \geq \aleph_0$ e $\lambda \geq 2$, então $\kappa < \text{cof}(\lambda^\kappa)$. Dica: note que λ^κ deve ser infinito, de modo que o item anterior acarreta $\lambda^\kappa < (\lambda^\kappa)^{\text{cof}(\lambda^\kappa)}$; o que ocorreria se valesse $\text{cof}(\lambda^\kappa) \leq \kappa$?
- Conclua que $\text{cof}(2^{\aleph_0}) > \aleph_0$. Em particular, $2^{\aleph_0} \neq \aleph_\omega$. ■

Atenção !!!



Você está prestes a cruzar
os portões. Não há volta!

DRAFT (RMM 2023)

Capítulo F

Optional: metamatemática

Para fixar as ideias, convém recordar o leitor sobre uma modalidade de *jogo interpretativo* muito popular entre jovens com pelo menos 40 anos de idade: o RPG, abreviação para *role playing game*, que significa algo como *jogo de interpretar papéis*, também chamado de *jogo narrativo* ou *de representação* (de papéis!). Nesse tipo de jogo, um dos jogadores (**mestre**) propõe um cenário com *regras próprias* e, possivelmente, algum tipo de *conflito*, enquanto os outros jogadores interpretam *personagens* inseridas em tal cenário, que buscam solucionar o conflito nos limites das possibilidades impostas pelas regras.

Numa partida de RPG existem dois *níveis de atuação*, que podemos chamar de “jogo” e “meta-jogo”: a atuação *em jogo* ocorre *dentro* do cenário, por meio da interação entre as personagens interpretadas pelos jogadores (e.g., as personagens decidiram investir contra o monstro pois é isso que personagens com seus respectivos históricos fariam); a expressão *meta-jogo*, por sua vez, pode se aplicar a atuações ou situações *em jogo*, mas que são motivadas por interações entre os jogadores (e.g., as personagens decidiram investir contra o monstro pois já são 4 horas da manhã e a partida *precisa* acabar). Outro exemplo: pode ser que o cenário não possua *aranhas* por conta de eventos ocorridos *em jogo* numa partida anterior; mas também pode ser que o cenário não possua *aranhas* pois um dos jogadores sofre de aracnofobia, com a ausência de aranhas estipulada como regra por questões alheias ao cenário (meta-jogo).

A ideia com as noções de *Matemática* e *metamatemática* é semelhante, embora mais volátil: enquanto a primeira se ocupa de analisar *noções matemáticas* (*números, formas, padrões*, etc.), a segunda analisa a primeira como se esta fosse uma noção matemática. Para ilustrar, considere atentamente o Teorema A.2.9 e a Proposição A.3.5: enquanto o segundo apresenta um *fato* sobre a não-vacuidade de certos produtos finitos, o primeiro estabelece como garantir que certos tipos de afirmação sejam fatos! Em outras palavras: o Teorema A.2.9 discursa sobre outros teoremas, i.e., do ponto de vista de quem usa a Proposição A.3.5, o Teorema A.2.9 é um *metateorema*.

Uma vez que a maioria dos praticantes de Matemática Abstrata Clássica utiliza as noções de conjuntos para descrever seus objetos de interesse, pode-se julgar que os capítulos anteriores já tenham discutido metamatemática. Porém, a Matemática é volátil: para quem tem conjuntos como *objetos de estudo*, fez-se apenas um pouco mais de Matemática! Dado que não há definição última do *quê* é Matemática, pode-se dizer que as duas afirmações estão corretas. Mesmo assim, a segunda dá a impressão de circularidade, já que conjuntos são utilizados para descrever as *regras do jogo* de quem estuda conjuntos. O propósito deste capítulo é, justamente, desatar alguns desses nós.

F.1 Linguagens e estruturas

Definição F.1.1. Uma **linguagem** \mathcal{L} é um par $\mathcal{L} := \langle \mathcal{S}, \Omega \rangle$, onde \mathcal{S} é um conjunto de *símbolos* e $\Omega: \mathcal{S} \rightarrow \omega$ é uma função que a cada símbolo $s \in \mathcal{S}$ associa um número natural $\Omega(s) \in \omega$ chamado de **aridade** de s . A função Ω costuma ser chamada de **assinatura** ou **tipo** da linguagem \mathcal{L} . ¶

Mais precisamente, os símbolos em \mathcal{S} são particionados em dois subconjuntos disjuntos, $\mathcal{O}_{\mathcal{L}}$ e $\mathcal{R}_{\mathcal{L}}$: o primeiro composto pelos chamados **símbolos operacionais**, enquanto o segundo é composto pelos **símbolos relacionais** – em particular, linguagens sem símbolos relacionais costumam ser chamadas de **algébricas** e são *responsabilidade* da Álgebra Universal. Distinguir essas duas classes de símbolos é importante pois, como os nomes sugerem, símbolos operacionais e relacionais são usados para designar, respectivamente, operações e relações finitárias, que desempenham papéis diferentes na construção da *semântica* de uma linguagem.

Definição F.1.2. Fixados um conjunto A e um número natural $n \in \omega$, uma função $p: A^n \rightarrow A$ é chamada de **operação n -ária**. Analogamente, uma **relação n -ária** sobre A consiste num subconjunto $R \subseteq A^n$. Em ambos os casos, o número n é chamado de **aridade**, tanto da operação p quanto da relação R . ¶

Evidentemente, relações n -árias generalizam as relações *binárias* introduzidas na Definição A.1.28. O mesmo princípio gramatical empregado nesse caso justifica chamar de **binárias** as operações 2-árias, i.e., funções da forma $A \times A \rightarrow A$, bem como **unárias** as funções da forma $A \rightarrow A$. Os casos de operações e relações 0-árias são curiosos:

- ✓ como $A^0 = \{\emptyset\}$, segue que uma operação 0-ária $A^0 \rightarrow A$ faz $\emptyset \mapsto a$ para algum $a \in A$, o que na prática consiste na escolha de uma **constante** em A , nome usual para tais operações;
- ✓ como $A^0 = \{\emptyset\}$, segue que existem apenas duas relações 0-árias possíveis em A , a saber \emptyset e $\{\emptyset\}$, que correspondem precisamente aos naturais 0 e 1.

A imagem de um elemento $\langle a_1, \dots, a_n \rangle \in A^n$ em A será denotada por $p(a_1, \dots, a_n)$, mas o leitor preciosista tem o direito de escrever $p(\langle a_1, \dots, a_n \rangle)$, exceto nos casos em que $n \leq 2$: para $n := 2$, faz-se $a_1 p a_2$ em vez de $p(a_1, a_2)$; para $n := 1$ se escreve $p a$ ou algo equivalente a depender do contexto (vide o próximo exemplo); para $n := 0$, é inofensivo usar o mesmo *símbolo* p da operação $p: \{\emptyset\} \rightarrow A$ para denotar a *constante* $p(\emptyset)$. A discussão de exemplos será mais frutífera após a introdução da noção de *estruturas*.

Definição F.1.3. Sejam $\mathcal{L} := \langle \mathcal{S}, \Omega \rangle$ uma linguagem e M um conjunto. Dizemos que um par $\langle M, \mathcal{I} \rangle$ é uma **\mathcal{L} -estrutura** se \mathcal{I} for uma função que a cada $s \in \mathcal{S}$ com $\Omega(s) := n$ associa:

- (i) uma operação n -ária $s_M: M^n \rightarrow M$ se s for um símbolo operacional;
- (ii) uma relação n -ária $s_M \subseteq M^n$ se s for um símbolo relacional. ¶

Nas mesmas condições da definição acima, também é comum dizer que \mathcal{I} é *uma \mathcal{L} -estrutura* sobre o *universo* M , a fim de indicar que um mesmo conjunto M pode ter mais de uma \mathcal{L} -estrutura. Usa-se a letra “ \mathcal{I} ” em vez da letra “ \mathcal{E} ” pois a ideia é que uma estrutura permite *interpretar* os símbolos da linguagem. Apesar disso, quase sempre é seguro omitir menções explícitas à função \mathcal{I} e tratar apenas das operações e relações induzidas, como feito na próxima

Definição F.1.4. Sejam $\mathcal{L} := \langle \mathcal{S}, \Omega \rangle$ uma linguagem e M, N \mathcal{L} -estruturas. Uma função $f: M \rightarrow N$ é chamada de **\mathcal{L} -morfismo** (ou **morfismo de \mathcal{L} -estruturas**) se para todo $s \in \mathcal{S}$ com $\Omega(s) := n$ ocorrer

- (i) $f \circ s_M = s_N \circ f^n$ se s for um símbolo operacional, e
- (ii) $f^n[s_M] \subseteq s_N$ se s for um símbolo relacional,

onde $f^n: M^n \rightarrow N^n$ é dada por $f^n(m_1, \dots, m_n) := \langle f(m_1), \dots, f(m_n) \rangle$. ¶

Observação F.1.5. Embora a notação não ajude a perceber, no caso em que $n := 0$, deve-se ter $f(s_M) = s_N$ no caso operacional/constante, pois $f^0: M^0 \rightarrow N^0$ automaticamente satisfaz $f^0(\emptyset) = \emptyset$, no sentido do Exercício A.39 e, a rigor, s_N é a função que faz $\emptyset \mapsto s_N$. Em particular, a condição sobre relações 0-árias fica automaticamente satisfeita. △

Exemplo F.1.6. Linguagens e suas respectivas estruturas não impõem *restrições* de *conteúdo*, mas apenas de *forma*. Por exemplo, a linguagem algébrica $\mathcal{L}_{\text{mag}} := \langle \{\ast\}, \{\ast\}, 2 \rangle$ designa apenas um símbolo *operacional*¹ de aridade 2, de modo que *qualquer* conjunto X dotado de uma operação binária $\ast_X: X \times X \rightarrow X$ é uma \mathcal{L}_{mag} -estrutura legítima. Não interessa, por enquanto, se alguma dessas estruturas satisfaz condições adicionais que permitem xingá-la por nomes conhecidos (e.g., *monóide*, *grupo*, etc.), pois tais informações dependem do *conteúdo* particular da estrutura, enquanto ser uma estrutura depende apenas da *forma*.

Apesar disso, é certamente ao revisitar tais cenários particulares que o leitor encontrará os casos fundamentais que ajudarão a digerir a definição de morfismo: se $+_M$ e $+_N$ indicam *somas* em M e N , por exemplo, então (i) traduz a condição $f(m +_M m') = f(m) +_N f(m')$, ao passo que para $-_M$ e $-_N$ as operações que associam *vetores* aos seus *inversos aditivos*, (i) se traduz em $f(-_M m) = -_N f(m)$; no caso 0-ário, pode-se pensar na exigência típica de que morfismos de anéis levem a *unidade* de um anel *na unidade* do outro. Por sua vez, a condição (ii) encontra seus paralelos no contexto de ordens: uma função crescente $f: \mathbb{P} \rightarrow \mathbb{Q}$ entre ordens parciais $\langle \mathbb{P}, \leq_{\mathbb{P}} \rangle$ e $\langle \mathbb{Q}, \leq_{\mathbb{Q}} \rangle$, por exemplo, é um \mathcal{L}_{\leq} -morfismo da linguagem \mathcal{L}_{\leq} composta unicamente por um símbolo relacional de aridade 2, digamos \leq , já que se pede “ $a \leq_{\mathbb{P}} b \Rightarrow f(a) \leq_{\mathbb{Q}} f(b)$ ”, i.e., $f[\leq_{\mathbb{P}}] \subseteq \leq_{\mathbb{Q}}$. O mesmo fenômeno ocorre com funções compatíveis com uma relação de equivalência (Teorema B.1.16), posto que a igualdade é uma relação de equivalência e estas, por sua vez, são relações de aridade 2. ▲

Mesmo neste contexto introdutório, já é possível perceber alguns fenômenos bastante gerais.

Proposição F.1.7. Seja \mathcal{L} uma linguagem.

- (i) Se M é uma \mathcal{L} -estrutura, então Id_M é um \mathcal{L} -morfismo.
- (ii) Se M, N e O são \mathcal{L} -estruturas e $f: M \rightarrow N$ e $g: N \rightarrow O$ são \mathcal{L} -morfismos, então $g \circ f: M \rightarrow O$ é um \mathcal{L} -morfismo.

Demonstração. Os itens (i) e (ii) seguem, respectivamente, das identidades $\text{Id}_M^n = \text{Id}_{M^n}$ e $(g \circ f)^n = g^n \circ f^n$. O leitor pode cuidar dos detalhes². □

¹O que não é claro a partir da notação $\langle \{\ast\}, \{\ast\}, 2 \rangle$, já que \ast poderia indicar um símbolo relacional. Tal ambiguidade deliberada é fruto do meu receio de sobrecarregar ainda mais a notação. Em geral, as linguagens serão explicitadas *em prosa* e apenas abreviadas por símbolos, a fim de tornar a leitura mais simples para seres humanos, público alvo deste texto. O caso de \mathcal{L}_{mag} foi justamente uma ilustração de que mesmo de forma incompleta, o apego ao simbolismo tornaria a leitura ainda mais penosa.

²Cuidado para não confundir “ h^n ” com o resultado de n composições da função h . Aqui, a notação “ h^n ” está de acordo com o Exercício A.39. Em particular, como $M \neq N$, nem faria sentido pensar em tal expoente como indicativo de composições iteradas: como sempre, o contexto é importante.

Um \mathcal{L} -isomorfismo entre duas \mathcal{L} -estruturas M e N é um \mathcal{L} -morfismo $f: M \rightarrow N$ para o qual existe um \mathcal{L} -morfismo $g: N \rightarrow M$ com $g \circ f = \text{Id}_M$ e $f \circ g = \text{Id}_N$. Em tal situação, diz-se que M e N são \mathcal{L} -estruturas *isomorfas*.

Exemplo F.1.8. Em virtude do Exercício A.13, todo \mathcal{L} -isomorfismo é uma bijeção entre os universos subjacentes. Por sua vez, como qualquer bijeção $f: M \rightarrow N$ satisfaz a identidade $(f^{-1})^n = (f^n)^{-1}: M^n \rightarrow N^n$, resulta que para uma linguagem algébrica \mathcal{L} , i.e., sem símbolos relacionais, todo \mathcal{L} -morfismo bijetor é um \mathcal{L} -isomorfismo.

Proposição F.1.9. Se \mathcal{L} é uma linguagem algébrica, então um \mathcal{L} -morfismo $f: M \rightarrow N$ é um \mathcal{L} -isomorfismo se, e somente se, f é \mathcal{L} -morfismo bijetor.

Demonstração. Em vista da identidade observada acima, de $f \circ s_M = s_N \circ f^n$ resulta $f^{-1} \circ s_N \circ f^n = s_M$ e, por conseguinte, $f^{-1} \circ s_N = s_M \circ (f^{-1})^n$. \square

Contudo, o fenômeno acima não é regra. Ao aplicar o mesmo raciocínio para relações, obtém-se $s_M \subseteq (f^{-1})^n[s_N]$. Porém, seria preciso assegurar a inclusão oposta a fim de garantir a *compatibilidade* de f^{-1} com um símbolo de relação n -ário s . Há exemplos bastante imediatos de que a proposição acima, de fato, falha em linguagens não algébricas: ao se considerar \mathbb{P} como o conjunto $X := \{0, 1\}$ munido da ordem parcial $\preceq := \{\langle 0, 0 \rangle, \langle 1, 1 \rangle\}$ e \mathbb{P}' como o próprio conjunto X com a ordem usual³, a função $\text{Id}: \mathbb{P} \rightarrow \mathbb{P}'$ que faz $\text{Id}(x) := x$ é um \mathcal{L}_{\leq} -morfismo bijetor; no entanto, sua inversa $\text{Id}^{-1}: \mathbb{P}' \rightarrow \mathbb{P}$ não é um \mathcal{L}_{\leq} -morfismo, dado que $0 \leq 1$ em \mathbb{P}' enquanto não ocorre $0 \preceq 1$ em \mathbb{P} . \blacktriangle

Exemplo F.1.10 (Produtos). Dadas uma linguagem $\mathcal{L} := \langle \mathcal{S}, \Omega \rangle$ e duas \mathcal{L} -estruturas M e N , é muito natural definir uma \mathcal{L} -estrutura sobre o produto cartesiano $M \times N$: basta declarar $s_{M \times N} := s_M \times s_N$ para cada $s \in \mathcal{S}$, ou quase isso. Mais precisamente, para $s \in \mathcal{S}$ um símbolo operacional de aridade n , $s_{M \times N}$ denota a operação n -ária que faz

$$\underbrace{\langle \langle a_1, b_1 \rangle, \dots, \langle a_n, b_n \rangle \rangle}_{\text{em } (M \times N)^n} \mapsto \langle s_M(a_1, \dots, a_n), s_N(b_1, \dots, b_n) \rangle$$

para cada n -upla $\langle \langle a_1, b_1 \rangle, \dots, \langle a_n, b_n \rangle \rangle \in (M \times N)^n$. Na prática, tal correspondência é mais simples do que parece: trata-se da composição da bijeção *natural* que existe entre $M^n \times N^n$ e $(M \times N)^n$ com o produto cartesiano das funções $s_M: M^n \rightarrow M$ e $s_N: N^n \rightarrow N$, no sentido do que se definiu na Observação A.5.11, este sim denotado por $s_M \times s_N$,

$$\underbrace{\langle \langle a_1, \dots, a_n \rangle, \langle b_1, \dots, b_n \rangle \rangle}_{\text{em } M^n \times N^n} \xrightarrow{s_M \times s_N} \langle s_M(a_1, \dots, a_n), s_N(b_1, \dots, b_n) \rangle.$$

O caso de um símbolo relacional s de aridade n é análogo: $s_{M \times N} := \varphi[s_M \times s_N]$, onde $\varphi: M^n \times N^n \rightarrow (M \times N)^n$ é a bijeção *óvia*, o que faz sentido pois $s_M \times s_N \subseteq M^n \times N^n$. O leitor deve se sentir encorajado a reescrever tais definições para os casos $n := 0, 1$ e 2 , a fim de perceber que tudo se trata da boa e velha definição *coordenada a coordenada*. \blacktriangle

Classicamente, morfismos costumam ser xingados de *homomorfismos*, o que é etimologicamente razoável: “homos” e “morphē”, que vêm do grego, significariam algo como “mesma forma”. Embora essa tradução literal se aplique melhor aos *isomorfismos*, morfismos realmente preservam certas formas na passagem de uma estrutura para outra. Porém, discutir precisamente o que é preservado terá que esperar.

³Induzida de ω .

F.2 Subestruturas, núcleos e quocientes

Definição F.2.1. Sejam $\mathcal{L} := \langle \mathcal{S}, \mathcal{O} \rangle$ uma linguagem e $\langle M, \mathcal{I} \rangle$ uma \mathcal{L} -estrutura. Um subconjunto $N \subseteq M$ é dito uma **\mathcal{L} -subestrutura** se a restrição $\mathcal{I}|_N$ apropriada das operações e relações de \mathcal{I} a N faz de $\langle N, \mathcal{I}|_N \rangle$ uma \mathcal{L} -estrutura. ¶

Acima, por “restrição apropriada” entende-se a função $\mathcal{I}|_N$ que a cada símbolo $s \in \mathcal{S}$ cuja aridade é $\Omega(s) := n$ associa a restrição da operação s_M ao subconjunto N^n (no caso de um símbolo operacional s), i.e., $s_N := s_M|_{N^n}$, ou $s_N := s_M \cap N^n$ se s for um símbolo relacional. Explicitamente, a fim de que N seja subestrutura de $\langle M, \mathcal{I} \rangle$:

- ✓ as interpretações em M das constantes de \mathcal{L} devem pertencer a N ;
- ✓ deve-se ter $s_M(a_1, \dots, a_n) \in N$ sempre que $a_1, \dots, a_n \in N$ e s for um símbolo de operação com aridade n .

Observação F.2.2. A definição anterior deveria dizer que N é *subestruturável*, já que a princípio N é um subconjunto sem estrutura do universo M . Todavia, tal abuso é inofensivo, dado que a estrutura $\mathcal{I}|_N$ definida é a única a fazer da inclusão $i: N \rightarrow M$ um \mathcal{L} -morfismo – se, é claro, as condições observadas acima forem satisfeitas por N . △

Tais considerações sugerem uma pergunta ingênuas cuja resposta tem um alcance inimaginável: se M é uma \mathcal{L} -estrutura e $X \subseteq M$ não é uma subestrutura de M , como corrigir isso? Dada a aplicabilidade do próximo teorema em outros contextos, convém utilizar letras distintas das que tem sido usadas neste capítulo até aqui.

Teorema F.2.3 (Fecho indutivo). *Sejam A um conjunto, $X \subseteq A$ um subconjunto e \mathcal{F} uma família de operações finitárias em A . Então existe um (único) subconjunto $\mathcal{F}(X) \subseteq A$ com as seguintes propriedades:*

- (i) $X \subseteq \mathcal{F}(X)$;
- (ii) se $f \in \mathcal{F}$ tem aridade m e $c \in (\mathcal{F}(X))^m$, então $f(c) \in \mathcal{F}(X)$;
- (iii) se $B \subseteq A$ tem as duas propriedades acima, então $\mathcal{F}(X) \subseteq B$.

Demonstração. Diremos que $C \subseteq A$ é \mathcal{F} -indutivo se a condição (ii) acima for satisfeita para C em vez de $\mathcal{F}(X)$. Não é difícil perceber que A é \mathcal{F} -indutivo e, se \mathcal{J} for uma família de subconjuntos \mathcal{F} -indutivos, então $\bigcap \mathcal{J}$ é \mathcal{F} -indutivo. Logo, tomando-se a família $\mathcal{J} := \{C \subseteq A : X \subseteq C \text{ e } C \text{ é } \mathcal{F}\text{-indutivo}\}$, resulta que $\mathcal{F}(X) = \bigcap \mathcal{J}$. Note que a unicidade de $\mathcal{F}(X)$ segue por $\mathcal{F}(X)$ ser o ínfimo de \mathcal{J} (Exemplo B.2.5).

Alternativamente, e isto será importante, pode-se considerar $X_0 := X$ e, para $n \in \omega$,

$$X_{n+1} := X_n \cup \{f(c) : m \in \omega, f \in \mathcal{F} \text{ tem aridade } m \text{ e } c \in (X_n)^m\},$$

onde segue que $\mathcal{F}(X) = \bigcup_{n \in \omega} X_n$: por construção, $X_+ := \bigcup_{n \in \omega} X_n$ é \mathcal{F} -indutivo e contém X , o que prova $\mathcal{F}(X) \subseteq X_+$; por outro lado, não é difícil se convencer, por indução, que $X_n \subseteq \mathcal{F}(X)$ para todo n , donde se obtém $X_+ \subseteq \mathcal{F}(X)$. □

Exercício F.1. Compare a demonstração anterior com a *construção* da σ -álgebra gerada por \mathcal{E} no Exemplo C.5.9. ■

Observação F.2.4. Convém ressaltar que $|\mathcal{F}(X)| \leq |X| \cdot |\mathcal{F}| \cdot \aleph_0$, o que segue pois $|X_0| \leq |X| \cdot |\mathcal{F}| \cdot \aleph_0$ e, se $|X_n| \leq |X| \cdot |\mathcal{F}| \cdot \aleph_0$, então

$$|X_{n+1}| \leq |X_n| + |\mathcal{F}| \cdot \sup_{m \in \omega} |X_n|^m \leq |X| \cdot |\mathcal{F}| \cdot \aleph_0 + |\mathcal{F}| \cdot \sup_{m \in \omega} |X|^m |\mathcal{F}|^m \aleph_0^m,$$

o restante decorre dos resultados sobre aritmética cardinal discutidos no capítulo anterior⁴. Em particular, se $|X|, |\mathcal{F}| \leq \aleph_0$, então $|X| \leq |\mathcal{F}(X)| \leq \aleph_0$. \triangle

Observação F.2.5 (Indução sobre complexidade). Além de facilitar a estimativa da cardinalidade de $\mathcal{F}(X)$, a segunda construção tem a vantagem de permitir os chamados argumentos por *indução sobre complexidade*, que nada mais são do que induções clássicas envoltas numa *aura* de mistério computacional.

Teorema F.2.6 (Indução sobre complexidade). *Sejam Y um conjunto e \mathcal{P} uma família de subconjuntos de Y , bem ordenada por uma relação \prec , com $Y = \bigcup \mathcal{P}$. Se Φ for uma “propriedade” sobre elementos de Y tal que, para todo $P \in \mathcal{P}$ se tenha*

$$(\forall Q \in \mathcal{P} \quad Q \prec P \Rightarrow Q \subseteq \{y \in Y : \Phi(y)\}) \Rightarrow P \subseteq \{y \in Y : \Phi(y)\}, \quad (\text{F.1})$$

então todo $y \in Y$ tem a propriedade Φ .

Demonstração. Se não ocorresse o que se propõe, então a coleção

$$\mathcal{T} := \{P \in \mathcal{P} : \exists y \in P \text{ tal que } \neg \Phi(y)\}$$

seria não-vazia, o que permitiria tomar o menor $P \in \mathcal{T}$. A hipótese sobre Φ garantiria então que $\Phi(y)$ é verdadeira para todo $y \in P$, uma contradição. \square

O *modus operandi* anterior costuma ser chamado de **indução sobre a complexidade** pois, em geral, uma decomposição \mathcal{P} como acima surge em contextos nos quais o conjunto Y é descrito recursivamente a partir de conjuntos mais simples⁵. No caso, fazendo $Y := \mathcal{F}(X)$, segue que para demonstrar a validade de uma condição Φ acerca de todos os elementos de $\mathcal{F}(X)$, basta verificar que: cada $x \in X$ satisfaz Φ e $f(c)$ satisfaz Φ sempre que $f \in \mathcal{F}$ e c é uma upla do *tamanho certo* composta por elementos que satisfazem Φ . O leitor pode tratar dos detalhes. \triangle

Corolário F.2.7. *Sejam \mathcal{L} uma linguagem e M uma \mathcal{L} -estrutura. Para cada subconjunto $X \subseteq M$, a família das \mathcal{L} -subestruturas de M que contêm X tem um menor elemento. Mais precisamente: existe uma \mathcal{L} -subestrutura S de M tal que $X \subseteq S$ e $S \subseteq T$ para toda \mathcal{L} -subestrutura $T \subseteq M$ que satisfaz $X \subseteq T$.*

Definição F.2.8. Nas notações acima, denotaremos por $\text{sp}(X)$ a menor \mathcal{L} -subestrutura de M que contém X , que se diz ser **gerada** por X . \P

Em geral, quando \mathcal{L} é uma linguagem algébrica, as (sub)estruturas de tipo \mathcal{L} são chamadas de **(sub) álgebras** de tipo \mathcal{L} , ou apenas \mathcal{L} -(sub) álgebras. Claramente, trata-se de uma generalização dos *subgrupos*, *subanéis*, *submódulos*, etc., que o leitor já deve conhecer de suas experiências prévias com Álgebra Abstrata. Pelo restante desta seção, \mathcal{L} indicará uma linguagem algébrica.

⁴Especificamente: Observação E.2.4, Corolários E.2.14 e E.2.15 e Teorema E.2.22. Pode ser útil, ainda, observar que $\sup_{m \in \omega} |X|^m |\mathcal{F}|^m \aleph_0^m = \max\{|X|, |\mathcal{F}|, \aleph_0\}$.

⁵Pode-se inclusive definir a *complexidade* de $y \in Y$ como $\min\{P \in \mathcal{P} : y \in P\}$.

Proposição F.2.9. Sejam \mathcal{L} -álgebras A e B , bem como \mathcal{L} -morfismos $f, g: A \rightarrow B$. Se $A = \text{sp}(X)$ para algum $X \subseteq A$ e $f(x) = g(x)$ para todo $x \in X$, então $f = g$.

Demonstração. O conjunto $T := \{a \in A : f(a) = g(a)\}$ é uma \mathcal{L} -subálgebra de A , pois f e g são morfismos. Como $X \subseteq T$, segue que $T = A$ e, portanto, $f = g$. \square

Exemplo F.2.10 (Opcional (?): bases em Álgebra Linear). Explicitamente, o último resultado atesta a injetividade da correspondência $f \mapsto f|_X$ entre os \mathcal{L} -morfismos da forma $A \rightarrow B$ e as funções da forma $X \rightarrow B$. Isto lembra o fenômeno ocorrido com as *bases* em Álgebra Linear (de espaços vetoriais, módulos, etc.), mas há uma diferença: neste caso, garante-se (adicionalmente!) a sobrejetividade da correspondência $f \mapsto f|_X$. \blacktriangle

Uma vez em posse da definição de *estrutura algébrica*, a discussão dos quocientes e das *subestruturas* passa a ser realizável de modo bastante geral, o que pode revelar nuances importantes de tais construções, mas que costumam passar despercebidas quando feitas nos cenários típicos de *grupos*, *anéis* e *módulos*.

Definição F.2.11. Seja $f: A \rightarrow B$ um \mathcal{L} -morfismo. O **núcleo universal** de f é o subconjunto $\text{Ker}(f) := \{\langle a, a' \rangle \in A \times A : f(a) = f(a')\}$. \P

Embora pareça estranho para leitores já familiarizados com Álgebra Elementar, $\text{Ker}(f)$, definido como acima, tem duas propriedades interessantes:

- ✓ é uma relação de equivalência sobre A , i.e., para quaisquer $x, y, z \in A$ ocorre
 - $x \in \text{Ker}(f) x$,
 - $x \in \text{Ker}(f) y \Rightarrow y \in \text{Ker}(f) x$, e
 - $x \in \text{Ker}(f) y$ e $y \in \text{Ker}(f) z \Rightarrow x \in \text{Ker}(f) z$;
- ✓ é uma \mathcal{L} -subálgebra de $A \times A$, pois $\langle s_A, s_A \rangle \in \text{Ker}(f)$ para todo símbolo de constante s e, para as demais aridades,

$f(s_A(a_1, \dots, a_n)) = s_A(f(a_1), \dots, f(a_n)) = s_A(f(b_1), \dots, f(b_n)) = f(s_A(b_1, \dots, b_n)),$ sempre que $\langle a_1, b_1 \rangle, \dots, \langle a_n, b_n \rangle \in \text{Ker}(f)$, mostrando assim que

$$s_{A \times A}(\langle a_1, b_1 \rangle, \dots, \langle a_n, b_n \rangle) := \langle s_A(a_1, \dots, a_n), s_A(b_1, \dots, b_n) \rangle \in \text{Ker}(f).$$

Observação F.2.12. Acima, usou-se indução em complexidade. Se parecer muito confuso, reescreva tudo com $n := 2$. Acredite, vai melhorar. \triangle

Definição F.2.13. Uma **congruência** numa \mathcal{L} -álgebra A é uma relação de equivalência $C \subseteq A \times A$ que também é uma \mathcal{L} -subálgebra de $A \times A$. \P

Fixada uma congruência C em A , faz sentido considerar o conjunto $A/C := \{\bar{a} : a \in A\}$ das classes de equivalência da relação C . Como ocorre secretamente, com grupos, anéis, etc., é possível elevar A/C ao patamar de \mathcal{L} -álgebra, o que exige a definição de uma \mathcal{L} -estrutura $\mathcal{I}_{A/C}$ em A/C , o que se faz da maneira óbvia: para um símbolo $s \in \mathcal{S}$ de aridade $n \in \omega$, faz-se $s_{A/C}: (A/C)^n \rightarrow A/C$ por meio da correspondência $\langle \bar{a}_1, \dots, \bar{a}_n \rangle \mapsto \overline{p_A(a_1, \dots, a_n)}$.

A boa definição da operação $s_{A/C}$ segue da suposição de C ser \mathcal{L} -subálgebra: com efeito, se $\bar{a}_i = \bar{b}_i$ para cada $i \in \{1, \dots, n\}$, então $\langle a_i, b_i \rangle \in C$ e daí

$$p_{A \times A}(\langle a_1, b_1 \rangle, \dots, \langle a_n, b_n \rangle) := \langle p_A(a_1, \dots, a_n), p_A(b_1, \dots, b_n) \rangle,$$

que pertence a C pois este é fechado para a operação $s_{A \times A}$, acarretando a identidade $p_A(a_1, \dots, a_n) = p_A(b_1, \dots, b_n)$. A construção da estrutura de A/C é feita sob medida para que a projeção $\pi: A \rightarrow A/C$ seja um \mathcal{L} -morfismo. E não é só isso.

Teorema F.2.14. Nas notações acima, se $f: A \rightarrow B$ é um \mathcal{L} -morfismo com $C \subseteq \text{Ker}(f)$, então existe um único morfismo de \mathcal{L} -álgebras $\bar{f}: A/C \rightarrow B$ tal que $\bar{f} \circ \pi = f$.

Demonstração. O único modo *sensato* de definir \bar{f} é fazendo $\bar{f}(\bar{a}) = f(a)$, e daí precisa-se verificar que tal correspondência está bem definida (o que ocorre por conta da hipótese $C \subseteq \text{Ker}(f)$) e é o único morfismo de \mathcal{L} -álgebras satisfazendo $\bar{f} \circ \pi = f$ (o que segue da própria definição). \square

Exemplo F.2.15 (Núcleo de funções). As considerações acima a respeito de núcleos e congruências permanecem válidas nas situações em que $\mathcal{L} := \emptyset$. Em outras palavras, isto significa que núcleos (como acima) podem ser definidos para funções entre *conjuntos* desprovidos de *estruturas algébricas*. Em particular, se $f: X \rightarrow Y$ é uma função entre conjuntos e $R \subseteq X \times X$ é uma relação de equivalência com $R \subseteq \text{Ker}(f)$, então existe uma única função $\bar{f}: X/R \rightarrow Y$ com $\bar{f} \circ \pi = f$. Soa familiar? \blacktriangle

Os mecanismos acima permitem lidar muito bem com a *forma* das estruturas. É chegado o momento de discutir *conteúdo* ou, mais precisamente: *expressividade*.

F.3 Interpretação e satisfabilidade

Nesta seção, $\mathcal{L} := \langle \mathcal{S}, \Omega \rangle$ volta a denotar uma linguagem possivelmente não-algébrica, com $\mathcal{O}_{\mathcal{L}}$ e $\mathcal{R}_{\mathcal{L}}$ os conjuntos de símbolos operacionais e relacionais, respectivamente. Agora, o principal objetivo é descrever como os símbolos de \mathcal{S} e suas respectivas aridades permitem definir uma *gramática* capaz de expressar *termos*, *fórmulas* e *sentenças* acerca de elementos numa \mathcal{L} -estrutura.

Definição F.3.1. Sejam \mathcal{V} e \mathcal{A} conjuntos disjuntos, e $\partial: \mathcal{A} \rightarrow \omega$ uma função “aridade”. Considere também dois objetos adicionais, $\langle , \rangle \notin \mathcal{V} \cup \mathcal{A}$. Seja $\mathcal{C} := \mathcal{V} \cup \mathcal{A} \cup \{\langle , \rangle\}$ e chame por $\text{str}_{\mathcal{A}, \partial}(\mathcal{V}) := \bigcup_{n \in \omega} \mathcal{C}^n$. \P

Formalmente, $\text{str}_{\mathcal{A}, \partial}(\mathcal{V})$ é formado por todas as *strings* de *caracteres* pertencentes ao conjunto \mathcal{C} , i.e., são sequências finitas da forma $\langle c_1, \dots, c_n \rangle$ para certos $n \in \omega$ e $c_1, \dots, c_n \in \mathcal{C}$. No caso de $n := \emptyset$, tem-se a *string* vazia. Na prática, $\text{str}_{\mathcal{A}, \partial} \mathcal{V}$ é a coleção de todas as concatenações finitas de caracteres no conjunto \mathcal{C} .

O propósito da definição acima é descrever o *ambiente* no qual as coisas que chamaremos de termos e fórmulas *habitam*. Mais precisamente, termos e fórmulas são definidos como subconjuntos apropriados de $\text{str}_{\mathcal{A}, \partial}(\mathcal{V})$, para certas escolhas de \mathcal{A} , ∂ e \mathcal{V} . Em tal processo, o Teorema F.2.3 (do fecho indutivo) pode ser usado pois, secretamente, $\text{str}_{\mathcal{A}, \partial}(\mathcal{V})$ está munido de uma família de operações finitárias induzidas pela função ∂ . Explicitamente:

- ✓ para $s \in \mathcal{A}$ com $\partial(s) := n > 0$, define-se $s_{\text{str}}: \text{str}_{\mathcal{A}, \partial}(\mathcal{V})^n \rightarrow \text{str}_{\mathcal{A}, \partial}(\mathcal{V})$ como uma operação n -ária, com

$$\langle \langle c_{1,1}, \dots, c_{1,j_1} \rangle, \dots, \langle c_{n,1}, \dots, c_{n,j_n} \rangle \rangle \mapsto \langle s, \langle \langle c_{1,1}, \dots, c_{1,j_1}, \dots, c_{n,1}, \dots, c_{n,j_n} \rangle \rangle \rangle;$$

- ✓ para $s \in \mathcal{A}$ com $\partial(s) := 0$, define-se $s_{\text{str}}: \text{str}_{\mathcal{A}, \partial}(\mathcal{V})^0 \rightarrow \text{str}_{\mathcal{A}, \partial}(\mathcal{V})$ como uma operação 0-ária, com $\emptyset \mapsto \langle s \rangle$, que pode ser escrito como $\emptyset \mapsto s$ por meio da bijeção óbvia entre $\text{str}_{\mathcal{A}, \partial}(\mathcal{V})$ e $\text{str}_{\mathcal{A}, \partial}(\mathcal{V})^1$.

A ideia por trás dessas definições é fazer das operações s_{str} meros concatenadores de caracteres: note que ao omitir parênteses e as vírgulas entre eles no primeiro caso, o resultado da operação seria $s(c_{1,1}, \dots, c_{n,j_n})$, de modo que os símbolos “(” e “)” funcionam apenas como delimitadores, ou símbolos de pontuação⁶ para indicar o *escopo* de s . Adotaremos a omissão de parênteses e vírgulas sugerida acima sempre que tratarmos de *termos* e *fórmulas*.

Proposição F.3.2. *Fixado um conjunto \mathcal{V} com $\mathcal{V} \cap \mathcal{S} = \emptyset$, existe um conjunto $\mathbb{T}_{\mathcal{L}}(\mathcal{V})$ de strings satisfazendo as seguintes condições:*

- (i) $\mathcal{V} \cup \mathcal{O}_{\mathcal{L},0} \subseteq \mathbb{T}_{\mathcal{L}}(\mathcal{V})$, onde $\mathcal{O}_{\mathcal{L},0} := \{s \in \mathcal{O}_{\mathcal{L}} : \Omega(s) = 0\}$;
- (ii) se $\tau_1, \dots, \tau_n \in \mathbb{T}_{\mathcal{L}}(\mathcal{V})$ e $s \in \mathcal{S}$ é um símbolo operacional de aridade $n > 0$, então $s(\tau_1, \dots, \tau_n) \in \mathbb{T}_{\mathcal{L}}(\mathcal{V})$;
- (iii) se $\tau \in \mathbb{T}_{\mathcal{L}}(\mathcal{V})$, então $\tau \in \mathcal{V} \cup \mathcal{O}_{\mathcal{L},0}$ ou existem $\tau_1, \dots, \tau_n \in \mathbb{T}_{\mathcal{L}}(\mathcal{V})$ para algum $n > 0$, bem como um símbolo operacional $s \in \mathcal{S}$ com $\Omega(s) := n$, tal que $\tau = s(\tau_1, \dots, \tau_n)$.

Demonstração. O resultado segue do Teorema F.2.3, especificamente com $A := \text{str}_{\mathcal{O}_{\mathcal{L}}, \Omega}(\mathcal{V})$, $\mathcal{F} := \{s_{\text{str}} : s \in \mathcal{O}_{\mathcal{L}}\}$ e $X := \mathcal{V}$. \square

Definição F.3.3. Nas condições acima, os elementos de $\mathbb{T}_{\mathcal{L}}(\mathcal{V})$ são chamados de *\mathcal{L} -termos nas variáveis* de \mathcal{V} .

Explicitamente, a *cláusula (i)* diz que tanto os símbolos de aridade 0 quanto as *variáveis* são *\mathcal{L} -termos*: íntimos costumam chamá-los de **\mathcal{L} -termos atômicos** de $\mathbb{T}_{\mathcal{L}}(\mathcal{V})$. Por sua vez, a cláusula *(ii)* determina de maneira recursiva como obter termos *novos* a partir de termos *previamente conhecidos*: se τ_1, \dots, τ_n são *\mathcal{L} -termos* (atômicos ou não) e $s \in \mathcal{O}_{\mathcal{L}}$ é um símbolo operacional de aridade n , então $s(\tau_1, \dots, \tau_n)$ é um *\mathcal{L} -termo*. A última cláusula é a tradução da minimalidade, no sentido da inclusão, da subestrutura $\mathcal{F}(X)$ construída no Teorema F.2.3: isso *exclui* a possibilidade de que $\mathbb{T}_{\mathcal{L}}(\mathcal{V})$ contenha elementos que não obedeçam as cláusulas anteriores. Naturalmente, *\mathcal{L} -termos* serão chamados apenas de *termos* quando a linguagem *\mathcal{L}* estiver clara pelo contexto.

Exemplo F.3.4 (Fundamental). O leitor que já tem familiaridade com polinômios deve usá-los como farol no entendimento dos termos: um polinômio *real* como $2x^3 + 5xy$, por exemplo, é tão somente uma expressão simbólica utilizando as *constantes* 2 e 5, as *variáveis* x e y e, por fim, os *símbolos de operação* $+$ e \cdot (este último implícito). Por si só, um polinômio não *pede* que suas variáveis sejam substituídas por *números* ou algo do tipo. Isso pode ser feito? Sim, pode.

Na verdade, seria *lícito* usar o polinômio $2x^3 + 5xy$ para definir uma função $\mathbb{R}^2 \rightarrow \mathbb{R}$, que associa cada par $\langle a, b \rangle \in \mathbb{R}^2$ ao *número* $2a^3 + 5ab$, i.e., substituindo as ocorrências das variáveis x e y pelos elementos a e b de \mathbb{R} , respectivamente. Porém, isto é problema de quem decide fazer tal *interpretação* do polinômio. Em outras palavras: polinômios são receitas multiuso para a realização de operações algébricas. Termos não são diferentes. \blacktriangle

Definição F.3.5. Fixado um conjunto de variáveis \mathcal{V} e um conjunto M , a notação $u: \mathcal{V} \rightarrow M$ indicará que u é uma função da forma $\mathcal{V}' \rightarrow M$ para algum subconjunto finito $\mathcal{V}' \subseteq \mathcal{V}$. Funções dessa forma serão xingadas de **\mathcal{V} -atribuições de valores** em M , e a coleção de todas as \mathcal{V} -atribuições em M será denotada por $[\mathcal{V} \rightarrow M]$. \P

⁶Optei por não usar parênteses ou colchetes para não gerar confusões com o uso *metalinguístico* que já fazemos desses símbolos.

Definição F.3.6. Para um \mathcal{L} -termo $\tau \in \mathbb{T}_{\mathcal{L}}(\mathcal{V})$ e uma \mathcal{L} -estrutura M , chamamos de **interpretação** de τ em M à função τ^M definida da seguinte forma:

- (i) (interpretação de variáveis) se $\tau := v \in \mathcal{V}$ e $u: \mathcal{V} \rightharpoonup M$ é tal que $v \in \text{dom}(u)$, então $\tau^M(u) := u(v)$; não se define $\tau^M(u)$ caso $v \notin \text{dom}(u)$;
- (ii) (interpretação de constantes) se $\tau := s \in \mathcal{O}_{\mathcal{L}}$ tem aridade 0, então $\tau^M(u) := s_M \in M$ para qualquer atribuição $u \in [\mathcal{V} \rightharpoonup M]$;
- (iii) (interpretação de termos) se $\tau := s(\tau_1, \dots, \tau_n)$, para $s \in \mathcal{O}_{\mathcal{L}}$ com aridade $n > 0$ e termos $\tau_1, \dots, \tau_n \in \mathbb{T}_{\mathcal{L}}(\mathcal{V})$, então $\tau^M(u) := s_M(\tau_1^M(u), \dots, \tau_n^M(u))$ sempre que $u \in [\mathcal{V} \rightharpoonup M]$ for uma atribuição cujas interpretações $\tau_1^M(u), \dots, \tau_n^M(u)$ estiverem definidas. ¶

Exemplo F.3.7. Ainda no caso de polinômios, note que a função $\mathbb{R}^2 \rightarrow \mathbb{R}$ induzida pelo polinômio y funciona como a projeção na segunda coordenada, enquanto a função induzida pelo polinômio x projeta pares $\langle a, b \rangle \in \mathbb{R}^2$ em a . Já as funções induzidas por polinômios constantes são constantes. Por fim, a função induzida pelo polinômio $2xy + y^2$ é a soma das funções induzidas pelos polinômios $2xy$ e y^2 . Esse tipo de analogia deve ajudar a aceitar a próxima proposição. ▲

Proposição F.3.8. Para um \mathcal{L} -termo $\tau \in \mathbb{T}_{\mathcal{L}}(\mathcal{V})$ e um \mathcal{L} -morfismo $g: M \rightarrow N$ entre \mathcal{L} -estruturas M e N , verifica-se $\tau^N(g \circ u) = g(\tau^M(u))$ para qualquer atribuição $u: \mathcal{V} \rightharpoonup M$ no domínio de τ^M . Menos precisamente: $\tau^N \circ g = g \circ \tau^M$.

Demonstração. Note que $g \circ u: \mathcal{V} \rightharpoonup N$ é uma atribuição de valores em N . Com isso, a ideia é usar indução na complexidade do termo a fim de *transportar* as identidades válidas em M para N :

- (i) se $\tau := v \in \mathcal{V}$, então $\tau^N(g \circ u) := (g \circ u)(v)$, enquanto $\tau^M(u) := u(v)$;
- (ii) se $\tau := s \in \mathcal{O}_{\mathcal{L}}$ tem aridade 0, então $\tau^N(g \circ u) := s_N$, enquanto $\tau^M(u) := s_M$, e daí $g(s_M) = s_N$;
- (iii) se $\tau := s(\tau_1, \dots, \tau_n)$ para $s \in \mathcal{O}_{\mathcal{L}}$ com aridade $n > 0$ e $\tau_1, \dots, \tau_n \in \mathbb{T}_{\mathcal{L}}(\mathcal{V})$ são tais que $\tau_i^N \circ g = g \circ \tau_i^M$ para cada $i \leq n$, então
$$\begin{aligned} \tau^N(g \circ u) &:= s_N(\tau_1^N(g \circ u), \dots, \tau_n^N(g \circ u)) = s_N(g(\tau_1^M(u)), \dots, g(\tau_n^M(u))) = \\ &= g(s_M(\tau_1^M(u), \dots, \tau_n^M(u))) = g(\tau^M(u)), \end{aligned}$$
como desejado. □

Observação F.3.9. A construção dos termos ignora *intencionalmente* os possíveis símbolos relacionais da linguagem \mathcal{L} : a ideia que se desdobra adiante consiste em usar os termos como *nomes* para objetos de uma estrutura M , de modo que M *satisfaz* uma *afirmação* φ sobre os termos τ_1, \dots, τ_n se a interpretação em M das relações e funções presentes na fórmula φ for verificada para os *objetos nomeados* pelos termos. Por exemplo, pode-se considerar a *expressão* “ $xy + 2x^2 < 5x - y$ ”: a depender das uplas $\langle a_x, a_y \rangle \in \mathbb{R}^2$ de números reais utilizadas, pode ser *verdadeiro* ou *falso* que $a_x a_y + 2a_x^2 < 5a_x - a_y$ em \mathbb{R} .

Dito isso, existem muitos resultados que podem ser desenvolvidos em linguagens algébricas, sem a intromissão de símbolos relacionais e *quantificadores*. Infelizmente, tratar de tais resultados, típicos da Álgebra Universal, ocuparia muito espaço num capítulo já bastante tumultuado⁷. Ou será que não? △

⁷Leitores interessados em mais discussões sobre Álgebra Universal podem conferir o cânones de Burris & Sankappanavar [6], além de Bergman [4] ou o sucinto Cohn [8].

F.3.1 Pausa dramática: estruturas livres

Em vez de considerar interpretações definidas em subconjuntos de $[\mathcal{V} \rightarrow M]$, seria lícito tratar de atribuições da forma $\mathcal{V} \rightarrow M$, de modo a fazer da interpretação τ^M de um termo $\tau \in \mathbb{T}_{\mathcal{L}}(\mathcal{V})$ uma função do tipo $M^{\mathcal{V}} \rightarrow M$. Neste caso:

- (i) (interpretação de variáveis) se $\tau := v \in \mathcal{V}$, então τ^M é a projeção na v -ésima coordenada do produto $M^{\mathcal{V}}$, i.e., $\tau^M(f) := f(v)$ para cada $f \in M^{\mathcal{V}}$;
- (ii) (interpretação de constantes) se $\tau := s \in \mathcal{O}_{\mathcal{L}}$ tem aridade 0, então $\tau^M(f) := s_M \in M$ para qualquer $f \in M^{\mathcal{V}}$;
- (iii) (interpretação de termos) se $\tau := s(\tau_1, \dots, \tau_n)$, para $s \in \mathcal{O}_{\mathcal{L}}$ com aridade $n > 0$ e termos $\tau_1, \dots, \tau_n \in \mathbb{T}_{\mathcal{L}}(\mathcal{V})$, então $\tau^M(f) := s_M(\tau_1^M(f), \dots, \tau_n^M(f))$, para qualquer função $f \in M^{\mathcal{V}}$.

Embora tais atribuições generalizadas sejam *desnecessárias* na prática, já que tanto termos quanto *fórmulas* utilizam apenas finitas variáveis, há vantagens no caso em que \mathcal{L} é algébrica.

Teorema F.3.10. *Suponha que \mathcal{L} seja uma linguagem algébrica. Para um conjunto de variáveis \mathcal{V} fixado, a família dos \mathcal{L} -termos $\mathbb{T}_{\mathcal{L}}(\mathcal{V})$ admite uma estrutura \mathcal{L} -algébrica tal que para qualquer função $f: \mathcal{V} \rightarrow A$, em que A é uma \mathcal{L} -álgebra, existe um único \mathcal{L} -morfismo $\tilde{f}: \mathbb{T}_{\mathcal{L}}(\mathcal{V}) \rightarrow A$ que torna o diagrama a seguir comutativo.*

$$\begin{array}{ccc} \mathbb{T}_{\mathcal{L}}(\mathcal{V}) & \xrightarrow{\tilde{f}} & A \\ i \uparrow & \nearrow f & \\ \mathcal{V} & & \end{array}$$

Demonstração. Por simplicidade, chamemos $\mathbb{T} := \mathbb{T}_{\mathcal{L}}(\mathcal{V})$. Agora, para cada $p \in \mathcal{O}$, definamos a função $p_{\mathbb{T}}: \mathbb{T}^{\Omega(p)} \rightarrow \mathbb{T}$ que faz uma $\Omega(p)$ -upla $\langle \tau_1, \dots, \tau_{\Omega(p)} \rangle$ de termos corresponder ao termo $p(\tau_1, \dots, \tau_{\Omega(p)})$ se $\Omega(p) > 0$, e $p_{\mathbb{T}}(\emptyset) := p$ se $\Omega(p) := 0$. Em outras palavras: são as operações herdadas da família de *strings* (confira a Proposição F.3.2).

Para a função $f: \mathcal{V} \rightarrow A$ dada, definamos \tilde{f} marotamente como a correspondência $\sigma \mapsto \sigma^A(f)$, i.e., \tilde{f} deve associar cada termo σ à sua interpretação σ^A com as variáveis de \mathcal{V} *valoradas* de acordo com a função f . A *boa definição* de \tilde{f} decorre da *unicidade de grafia* dos termos de \mathbb{T} : explicitamente, se $\tau := f(\tau_1, \dots, \tau_m)$, $\sigma := g(\sigma_1, \dots, \sigma_n)$ e $\tau = \sigma$, então $m = n$, $f = g$ e $\tau_i = \sigma_i$ para todo $i \in \{1, \dots, n\}$, o que decorre, essencialmente, da construção formal de \mathbb{T} (Proposição F.3.2).

Vejamos que \tilde{f} é um \mathcal{L} -morfismo entre \mathbb{T} e A que *comuta* o diagrama:

- ✓ se $p \in \mathcal{O}_{\mathcal{L}}$ tem aridade 0, então $(\tilde{f} \circ p_{\mathbb{T}})(\emptyset) = \tilde{f}(p_{\mathbb{T}}(\emptyset)) = \tilde{f}(p) := p^A(f) := p_A$;
- ✓ se $p \in \mathcal{O}_{\mathcal{L}}$ tem aridade $n > 0$ e $\tau_1, \dots, \tau_n \in \mathbb{T}$, então

$$\begin{aligned} (\tilde{f} \circ p_{\mathbb{T}})(\tau_1, \dots, \tau_n) &= \tilde{f}(p(\tau_1, \dots, \tau_n)) := p(\tau_1, \dots, \tau_n)^A(f) := p_A(\tau_1^A(f), \dots, \tau_n^A(f)) = \\ &= p_A(\tilde{f}(\tau_1), \dots, \tilde{f}(\tau_n)) = (p_A \circ (\tilde{f})^n)(\tau_1, \dots, \tau_n); \end{aligned}$$

- ✓ a comutatividade do diagrama segue pois $\tilde{f}(v) := v^A(f) := f(v)$.

Finalmente, se $\varphi: \mathbb{T} \rightarrow A$ é outro \mathcal{L} -morfismo que torna o diagrama comutativo, então ocorre $\varphi = \tilde{f}$. A verificação deve ser feita por indução na *complexidade* de $\tau \in \mathbb{T}$:

- ✓ se τ é uma variável de \mathcal{V} , então $\varphi(\tau) = f(\tau) = \tilde{f}(\tau)$;
- ✓ se τ é uma constante, então por φ ser um morfismo vale $\varphi \circ \tau_{\mathbb{T}} = \tau_A$, resultando em $\varphi(\tau) = \varphi(\tau_{\mathbb{T}}(\emptyset)) = \tau_A(\emptyset) = \tau^A(f) = \tilde{f}(\tau)$;
- ✓ se $\tau := p(\tau_1, \dots, \tau_n)$ com $\tau_1, \dots, \tau_n \in \mathbb{T}$, $p \in \mathcal{O}_{\mathcal{L}}$, $\Omega(p) := n$ e $\varphi(\tau_i) = \tilde{f}(\tau_i)$ para todo $i \in \{1, \dots, n\}$, então por φ ser um morfismo se infere $\varphi \circ p_{\mathbb{T}} = p_A \circ \varphi^n$ e, por conseguinte

$$\begin{aligned}\varphi(\tau) &= (\varphi \circ p_{\mathbb{T}})(\tau_1, \dots, \tau_n) = p_A \circ \varphi^n(\tau_1, \dots, \tau_n) = p_A(\varphi(\tau_1), \dots, \varphi(\tau_n)) = \\ &= p_A(\tilde{f}(\tau_1), \dots, \tilde{f}(\tau_n)) = p(\tilde{f}(\tau_1), \dots, \tilde{f}(\tau_n))^A(f) = \tilde{f}(\tau).\end{aligned}\quad \square$$

A próxima definição é um subcaso do que será o critério de satisfabilidade de \mathcal{L} -fórmulas: trata-se da generalização do que significa, para uma álgebra, satisfazer uma *identidade polinomial*. Em tempo: pelo restante desta subseção, \mathcal{L} volta a indicar uma linguagem algébrica.

Definição F.3.11. Dados termos $\sigma, \tau \in \mathbb{T}_{\mathcal{L}}(\mathcal{V})$ e uma \mathcal{L} -álgebra A , diremos que A **satisfaz**⁸ a *relação de identidade* $\sigma \approx \tau$, simbolicamente indicado por $A \models \sigma \approx \tau$, se ocorrer $\sigma^A = \tau^A$. ¶

Exercício F.2. Nas condições anteriores, mostre que são equivalentes:

- (i) $A \models \sigma \approx \tau$;
- (ii) $\varphi(\sigma) = \varphi(\tau)$ para todo \mathcal{L} -morfismo $\varphi: \mathbb{T}_{\mathcal{L}}(\mathcal{V}) \rightarrow A$.

Dica: se $\sigma^A = \tau^A$, então $\sigma^A(f) = \tau^A(f)$ para toda função $f: \mathcal{V} \rightarrow A$, e tais funções caracterizam os morfismos $\mathbb{T}_{\mathcal{L}}(\mathcal{V}) \rightarrow A$, em virtude do último teorema. ■

Tanto a nomenclatura “*relação de identidade*” quanto a simbologia “ $\sigma \approx \tau$ ” são afagos psicológicos: formalmente, “ $\sigma \approx \tau$ ” indica o par ordenado $\langle \sigma, \tau \rangle \in \mathbb{T}_{\mathcal{L}}(\mathcal{V}) \times \mathbb{T}_{\mathcal{L}}(\mathcal{V})$. De qualquer forma, passa a fazer sentido xingar um subconjunto $\Sigma \subseteq \mathbb{T}_{\mathcal{L}}(\mathcal{V}) \times \mathbb{T}_{\mathcal{L}}(\mathcal{V})$ de *família de identidades*.

Definição F.3.12. Nas notações acima, diremos que uma \mathcal{L} -álgebra A **satisfaz** Σ (ou é de **tipo** Σ), simbolizado por $A \models \Sigma$, se ocorrer $A \models \sigma \approx \tau$ para todo par $\langle \sigma, \tau \rangle \in \Sigma$. ¶

Ao se fixar uma família de identidades Σ , pode-se considerar a *classe* de todas as \mathcal{L} -álgebras que satisfazem Σ : a **classe equacional**⁹ das \mathcal{L} -álgebras de tipo Σ . O leitor não deve ter problemas em cozinhá-las: famílias de identidades que descrevam a classe dos monoides, dos grupos, dos anéis e dos módulos sobre um anel fixado: são apenas os *axiomas* que descrevem tais animais, escritos na linguagem da Álgebra Universal.

Finalmente, pode-se provar de maneira cabal o velho mantra de que “estruturas algébricas isomórficas são indistinguíveis algebricamente”.

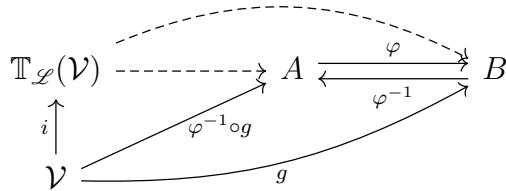
⁸Ou para os mais afoitos: A é *modelo* da identidade $\sigma \approx \tau$.

⁹Textos especializados no assunto costumam chamar tais classes de *varieties*. O leitor não deve confundir isso com as variedades da Geometria: trata-se de um falso cognato com a palavra *manifold*.

Teorema F.3.13. Sejam A e B duas \mathcal{L} -álgebras isomórfas. Se $\sigma, \tau \in \mathbb{T}_{\mathcal{L}}(\mathcal{V})$ são \mathcal{L} -termos para algum conjunto \mathcal{V} de variáveis, então $A \models \sigma \approx \tau \Leftrightarrow B \models \sigma \approx \tau$.

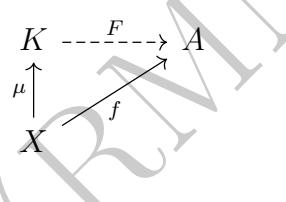
Dado que tal resultado será demonstrado para \mathcal{L} -fórmulas quaisquer na próxima subseção, convém deixar a demonstração do presente subcaso a cargo do leitor.

Exercício F.3. Demonstre o teorema anterior. Dica: encare o diagrama



até que ele te encare de volta. ■

Definição F.3.14. Seja \mathcal{K} uma classe de \mathcal{L} -álgebras¹⁰. Diremos que uma \mathcal{L} -álgebra $K \in \mathcal{K}$ é uma \mathcal{K} -álgebra livre sobre um conjunto X se existir uma função $\mu: X \rightarrow K$ tal que para toda \mathcal{L} -álgebra $A \in \mathcal{K}$ dotada de uma função $f: X \rightarrow A$ existe um único \mathcal{L} -morfismo $F: K \rightarrow A$ que torna comutativo o diagrama



i.e., tal que $F \circ \mu = f$. ¶

A definição acima é a generalização (algébrica) dos grupos livres, módulos livres, anéis de polinômios, etc. Moralmente, ela se comporta como a menor álgebra na classe \mathcal{K} que contém X como subconjunto. De fato, na maioria dos casos, a função $\mu: X \rightarrow K$ é injetora, o que permite assumir $X \subseteq K$.

Lema F.3.15. Seja $K \in \mathcal{K}$ uma \mathcal{K} -álgebra livre sobre um conjunto X . Se existir $C \in \mathcal{K}$ com $|C| \geq 2$, então μ é injetora.

Demonastração. De fato, se μ não fosse injetora, ocorreria $\mu(x) = \mu(x')$ para certos $x, x' \in X$ distintos. Daí, tomando-se qualquer $f: X \rightarrow C$ com $f(x) \neq f(x')$, o que é possível pela suposição sobre C , a hipótese sobre K garante um único \mathcal{L} -morfismo $F: K \rightarrow C$ com $F \circ \mu = f$, o que levaria a concluir $f(x) = f(x')$. □

Como a maioria das classes equacionais contém álgebras *não-triviais*, i.e., com pelo menos dois elementos distintos, será seguro assumir que X é um subconjunto de qualquer \mathcal{K} -álgebra livre sobre si. Aliás, antes de discutir a existência, convém observar a unicidade das \mathcal{K} -álgebras livres – a menos de isomorfismos.

Proposição F.3.16. Nas notações acima, se K e K' forem \mathcal{K} -álgebras livres sobre um conjunto X , então existe um único \mathcal{L} -isomorfismo $K \rightarrow K'$.

¹⁰No sentido usual de classe que tem sido utilizado ao longo do texto.

Demonstração. Se $\mu: X \rightarrow K$ e $\mu': X \rightarrow K'$ são \mathcal{K} -álgebras livres sobre o conjunto X , então existem únicos \mathcal{L} -morfismos $F: K \rightarrow K'$ e $G: K' \rightarrow K$ satisfazendo $F \circ \mu = \mu'$ e $G \circ \mu' = \mu$; daí, por ocorrer $(G \circ F) \circ \mu = \mu$ e $\text{Id}_K \circ \mu = \mu$, resulta $G \circ F = \text{Id}_K$. Analogamente, $F \circ G = \text{Id}_{K'}$. \square

Teorema F.3.17. *Sejam \mathcal{V} um conjunto de variáveis e \mathcal{C} uma classe equacional de \mathcal{L} -álgebras para alguma família de identidades $\Sigma \subseteq \mathbb{T}_{\mathcal{L}}(\mathcal{V}) \times \mathbb{T}_{\mathcal{L}}(\mathcal{V})$. Então existe uma \mathcal{C} -álgebra livre sobre \mathcal{V} .*

Demonstração. Seja $\tilde{\Sigma} := \{\langle \sigma, \tau \rangle \in \mathbb{T}_{\mathcal{L}}(\mathcal{V}) \times \mathbb{T}_{\mathcal{L}}(\mathcal{V}) : \forall A \in \mathcal{C} (A \models \sigma \approx \tau)\}$, i.e., a família das identidades satisfeitas por todas as álgebras da classe \mathcal{C} , para a qual evidentemente vale $\Sigma \subseteq \tilde{\Sigma}$. O grande truque consiste em tomar $K := \mathbb{T}_{\mathcal{L}}(\mathcal{V})/\tilde{\Sigma}$.

Afirmiação. $\tilde{\Sigma}$ é uma relação de equivalência sobre $\mathbb{T}_{\mathcal{L}}(\mathcal{V})$.

Prova da afirmação. Isto segue pois $\langle \sigma, \tau \rangle \in \tilde{\Sigma}$ se, e somente se, $\sigma^A = \tau^A$ para toda \mathcal{L} -álgebra $A \in \mathcal{C}$, donde as propriedades de reflexividade, simetria e transitividade seguem do fato de a relação de igualdade satisfazer todas essas condições. \square

Afirmiação. $\tilde{\Sigma}$ é uma subálgebra de $\mathbb{T}_{\mathcal{L}}(\mathcal{V}) \times \mathbb{T}_{\mathcal{L}}(\mathcal{V})$.

Prova da afirmação. Dado um símbolo $p \in \mathcal{O}_{\mathcal{L}}$ de aridade 0, toda \mathcal{L} -álgebra A satisfaz $p^A = p^A$, atestando $p_{\mathbb{T}_{\mathcal{L}}(\mathcal{V}) \times \mathbb{T}_{\mathcal{L}}(\mathcal{V})} := \langle p, p \rangle \in \tilde{\Sigma}$; se $\langle \sigma_1, \tau_1 \rangle, \dots, \langle \sigma_n, \tau_n \rangle \in \tilde{\Sigma}$ e $p \in \mathcal{O}_{\mathcal{L}}$ tem aridade $n > 0$, então

$$p_{\mathbb{T}_{\mathcal{L}}(\mathcal{V}) \times \mathbb{T}_{\mathcal{L}}(\mathcal{V})}(\langle \sigma_1, \tau_1 \rangle, \dots, \langle \sigma_n, \tau_n \rangle) = \langle p(\sigma_1, \dots, \sigma_n), p(\tau_1, \dots, \tau_n) \rangle \in \tilde{\Sigma},$$

pois se $A \in \mathcal{C}$, então $\sigma_i^A = \tau_i^A$ para cada i e, consequentemente,

$$p(\sigma_1, \dots, \sigma_n)^A(u) = p_A(\sigma_1^A(u), \dots, \sigma_n^A(u)) = p_A(\tau_1^A(u), \dots, \tau_n^A(u)) = p(\tau_1, \dots, \tau_n)^A(u)$$

para qualquer \mathcal{V} -upla $u \in A^{\mathcal{V}}$. \square

Em outras palavras, $\tilde{\Sigma}$ é uma congruência em $\mathbb{T}_{\mathcal{L}}(\mathcal{V})$, donde segue que $K := \mathbb{T}_{\mathcal{L}}(\mathcal{V})/\tilde{\Sigma}$ é uma \mathcal{L} -álgebra. Mostraremos que $K \in \mathcal{C}$ e que para toda \mathcal{L} -álgebra $A \in \mathcal{C}$ dotada de uma função $f: \mathcal{V} \rightarrow A$ existe um único \mathcal{L} -morfismo $F: K \rightarrow A$ satisfazendo $F \circ \mu = f$, onde $\mu: \mathcal{V} \rightarrow K$ é a composição da inclusão $i: \mathcal{V} \rightarrow \mathbb{T}_{\mathcal{L}}(\mathcal{V})$ com a projeção $\pi: \mathbb{T}_{\mathcal{L}}(\mathcal{V}) \rightarrow K$.

Afirmiação. *Se $\gamma: \mathbb{T}_{\mathcal{L}}(\mathcal{V}) \rightarrow \mathbb{T}_{\mathcal{L}}(\mathcal{V})$ é um \mathcal{L} -morfismo e $\langle \sigma, \tau \rangle \in \Sigma$, então ocorre $\langle \gamma(\sigma), \gamma(\tau) \rangle \in \tilde{\Sigma}$.*

Em posse desta afirmação, que será provada adiante, mostraremos que se $\langle \sigma, \tau \rangle \in \Sigma$ e $\alpha: \mathbb{T}_{\mathcal{L}}(\mathcal{V}) \rightarrow K$ é um \mathcal{L} -morfismo, então $\alpha(\sigma) = \alpha(\tau)$, o que equivale a $K \models \sigma \approx \tau$ em virtude do Exercício F.2. Ao se tomar $f: \mathcal{V} \rightarrow \mathbb{T}$ satisfazendo $\alpha(v) = \overline{f(v)} \in K$ para cada $v \in \mathcal{V}$, o Teorema F.3.10 garante que existe um único \mathcal{L} -morfismo $\gamma: \mathbb{T}_{\mathcal{L}}(\mathcal{V}) \rightarrow \mathbb{T}_{\mathcal{L}}(\mathcal{V})$ que estende f . Agora, os detalhes implícitos de toda essa construção permitem mostrar que $B := \{t \in \mathbb{T}_{\mathcal{L}}(\mathcal{V}) : \alpha(t) = \overline{\gamma(t)}\}$ é uma \mathcal{L} -subálgebra de $\mathbb{T}_{\mathcal{L}}(\mathcal{V})$ que contém \mathcal{V} e, por este último ser gerador de $\mathbb{T}_{\mathcal{L}}(\mathcal{V})$, segue-se $\mathbb{T}_{\mathcal{L}}(\mathcal{V}) = B$. Finalmente, como $\langle \gamma(\sigma), \gamma(\tau) \rangle \in \tilde{\Sigma}$ pela afirmação feita, verifica-se $\gamma(\sigma) = \overline{\gamma(\tau)}$ e, consequentemente, $\alpha(\sigma) = \alpha(\tau)$, como queríamos.

Prova da afirmação. Para provar a afirmação, note que se $C \in \mathcal{C}$ e $\beta: \mathbb{T}_{\mathcal{L}}(\mathcal{V}) \rightarrow C$ é um \mathcal{L} -morfismo, então $\beta \circ \gamma: \mathbb{T}_{\mathcal{L}}(\mathcal{V}) \rightarrow C$ é um \mathcal{L} -morfismo. Daí, por valer $C \models \sigma \approx \tau$, deve-se ter a identidade $(\beta \circ \gamma)(\sigma) = (\beta \circ \gamma)(\tau)$, donde a arbitrariedade de C acarreta $\langle \gamma(\sigma), \gamma(\tau) \rangle \in \tilde{\Sigma}$, como afirmado. \square

Por fim, provemos a existência e unicidade do \mathcal{L} -morfismo F . Pela *propriedade universal* de $\mathbb{T}_{\mathcal{L}}(\mathcal{V})$, existe um único \mathcal{L} -morfismo $g: \mathbb{T}_{\mathcal{L}}(\mathcal{V}) \rightarrow A$ tal que $g(v) = f(v)$ para cada $v \in \mathcal{V}$. Notemos então que $\tilde{\Sigma} \subseteq \text{Ker}(g)$: se $B \models \sigma \approx \tau$ para toda \mathcal{L} -álgebra $B \in \mathcal{C}$, então em particular $\sigma^A = \tau^A$ e, consequentemente, $g(\sigma) := \sigma^A(f) = \tau^A(f) := g(\tau)$. Logo, existe um único \mathcal{L} -morfismo $\bar{g}: K \rightarrow A$ satisfazendo $\bar{g} \circ \pi = g$ e, consequentemente, $\bar{g} \circ \mu = f$. Daí, não é difícil concluir que \bar{g} é o \mathcal{L} -morfismo F procurado. \square

Proposição F.3.18. *Se a classe de \mathcal{L} -álgebras \mathcal{K} é fechada por subálgebras¹¹ e existe uma \mathcal{K} -álgebra livre $\mu: X \rightarrow K$ sobre um conjunto X , então para cada $\overline{Y} \subseteq X$, a subálgebra $\text{sp}(\mu[Y]) \subseteq K$ é uma \mathcal{K} -álgebra livre sobre Y .*

Demonstração. De fato, basta tomar $\mu': Y \rightarrow \text{sp}(\mu[Y])$ como a restrição da função μ : uma função $g: Y \rightarrow A$ pode ser arbitrariamente estendida a uma função $f: X \rightarrow A$, que por sua vez se estende de forma única a um \mathcal{L} -morfismo $F: K \rightarrow A$, cuja restrição $G := F|_{\text{sp}(\mu[Y])}$ é o único \mathcal{L} -morfismo que estende g , em virtude da Proposição F.2.9. \square

Corolário F.3.19. *Seja \mathcal{C} uma classe equacional de \mathcal{L} -álgebras. Então para cada conjunto X existe uma \mathcal{C} -álgebra livre sobre X .*

Demonstração. Sejam \mathcal{V} um conjunto de variáveis e $\Sigma \subseteq \mathbb{T}_{\mathcal{L}}(\mathcal{V}) \times \mathbb{T}_{\mathcal{L}}(\mathcal{V})$ a família de identidades que caracteriza a classe \mathcal{C} , tudo isso com $X \cap \mathcal{V} = \emptyset$. Considerando-se a família auxiliar $\Sigma' := \Sigma \cup \{\langle x, x \rangle : x \in X\} \subseteq \mathbb{T}_{\mathcal{L}}(\mathcal{V} \cup X) \times \mathbb{T}_{\mathcal{L}}(\mathcal{V} \cup X)$ e definindo \mathcal{C}' a classe das \mathcal{L} -álgebras que satisfazem Σ' , o teorema anterior garante uma \mathcal{C}' -álgebra livre sobre $\mathcal{V} \cup X$, donde a proposição anterior provê uma \mathcal{C}' -álgebra livre sobre X . Por fim, como $\mathcal{C} = \mathcal{C}'$, o resultado segue. \square

F.3.2 De volta ao itinerário: fórmulas e modelos

Proposição F.3.20. *Fixado um conjunto \mathcal{V} com $\mathcal{V} \cap \mathcal{S} = \emptyset$, existe um conjunto $\mathbb{F}_{\mathcal{L}}(\mathcal{V})$ de strings satisfazendo as seguintes condições:*

- (i) se $\sigma, \tau \in \mathbb{T}_{\mathcal{L}}(\mathcal{V})$, então $\sigma \approx \tau \in \mathbb{F}_{\mathcal{L}}(\mathcal{V})$;
- (ii) se $\sigma_1, \dots, \sigma_n \in \mathbb{T}_{\mathcal{L}}(\mathcal{V})$ e $R \in \mathcal{R}_{\mathcal{L}}$ tem aridade $n > 0$, então $R \langle \sigma_1, \dots, \sigma_n \rangle \in \mathbb{F}_{\mathcal{L}}(\mathcal{V})$;
- (iii) se $\varphi \in \mathbb{F}_{\mathcal{L}}(\mathcal{V})$, então $\neg \varphi \in \mathbb{F}_{\mathcal{L}}(\mathcal{V})$;
- (iv) se $\varphi, \psi \in \mathbb{F}_{\mathcal{L}}(\mathcal{V})$, então $(\varphi \wedge \psi), (\varphi \vee \psi), (\varphi \rightarrow \psi)$ e $(\varphi \leftrightarrow \psi)$ pertencem a $\mathbb{F}_{\mathcal{L}}(\mathcal{V})$;
- (v) se $\varphi \in \mathbb{F}_{\mathcal{L}}(\mathcal{V})$ e $x \in \mathcal{V}$, então $\exists x \varphi$ e $\forall x \varphi$ pertencem a $\mathbb{F}_{\mathcal{L}}(\mathcal{V})$;
- (vi) se $\varphi \in \mathbb{F}_{\mathcal{L}}(\mathcal{V})$, então φ obedece a alguma das cláusulas anteriores.

Demonstração. Repita a demonstração da Proposição F.3.2, mutatis mutandis¹². \square

¹¹É o que o nome sugere: se $C \in \mathcal{K}$ e $D \subseteq C$ é subálgebra de C , então $D \in \mathcal{K}$. Note que classes equacionais são fechadas por subálgebras.

¹²Grosso modo, \wedge , \vee , \rightarrow e \leftrightarrow devem ser tratados como operações binárias, enquanto \neg , $\exists x$ e $\forall x$ são operações unárias (para cada $x \in \mathcal{V}$) que agem nas strings descritas pelas duas primeiras cláusulas.

Definição F.3.21 (Sintaxe de primeira ordem). Nas condições acima, os elementos de $\mathbb{F}_{\mathcal{L}}(\mathcal{V})$ são chamados de **\mathcal{L} -fórmulas** (de primeira ordem)¹³ de \mathcal{V} .

Explicitamente, as cláusulas (i) e (ii) dizem que tanto as *expressões* do tipo $\sigma \approx \tau$ quanto as expressões do tipo $R(\sigma_1, \dots, \sigma_n)$ são \mathcal{L} -fórmulas, desde que $\tau, \sigma, \sigma_1, \dots, \sigma_n$ sejam \mathcal{L} -termos e R seja um símbolo de relação apropriado, i.e., com aridade igual ao número de termos. Dada a simplicidade de tais fórmulas, elas costumam ser chamadas de **\mathcal{L} -fórmulas atômicas**. As outras cláusulas determinam como obter fórmulas *válidas* a partir de fórmulas válidas já conhecidas por meio dos *símbolos lógicos* e dos *quantificadores*¹⁴. A última exigência exclui a possibilidade de que $\mathbb{F}_{\mathcal{L}}(\mathcal{V})$ contenha “fórmulas” que não obedecem às cláusulas anteriores.

É importante destacar que os **símbolos lógicos** $\approx, \neg, \wedge, \vee, \rightarrow, \leftrightarrow, \exists$ e \forall são necessariamente distintos dos possíveis símbolos já presentes na linguagem \mathcal{L} : o primeiro, \approx , será utilizado na *interpretação* da relação de igualdade¹⁵; \exists e \forall serão usados para expressar *quantificação*; os demais serão usados como *operadores* ou *conectivos* entre as fórmulas. Os parênteses usuais “(” e “)” serão usados como símbolos de pontuação entre fórmulas, enquanto o papel de “(...)” será designar o escopo dos possíveis símbolos operacionais¹⁶.

Observação F.3.22 (Honestidade notacional – ou falta dela). Ao longo do texto, muitos parênteses serão omitidos sem maiores menções, tendo como base as regras implícitas de abreviação usuais que o leitor certamente já aprendeu *nas ruas*. Além disso, seguindo a tradição ocidental, daqui em diante, concatenações do tipo “ $*(*x, *(*y, z))$ ”, em que $*$ é um símbolo de aridade 2, serão transcritas como “ $x * (y * z)$ ”, e assim por diante. △

Uma vez definidas as regras de *sintaxe* para as fórmulas de uma linguagem, é chegado o momento de definir como *interpretar* essas fórmulas numa dada estrutura.

Definição F.3.23 (Semântica de Tarski). Nas condições acima, sejam M uma \mathcal{L} -estrutura e $u: \mathcal{V} \rightarrow M$ uma \mathcal{V} -atribuição em M . Para uma fórmula $\varphi \in \mathbb{F}_{\mathcal{L}}(\mathcal{V})$ cujas variáveis pertençam ao domínio de u , o que será abreviado dizendo-se que u é φ -compatível, escrevere-se $M \models \varphi[u]$ a fim de indicar que M , juntamente com a estrutura implícita \mathcal{I} , **satisfaz** φ com a atribuição u , o que se define de acordo com a complexidade da fórmula.

(i) Para φ atômica, há dois casos:

Forma da φ	Critério para $M \models \varphi[u]$
$\tau \approx \sigma$	$\tau^M(u) = \sigma^M(u)$
$R(\tau_1, \dots, \tau_n)$	$\langle \tau_1^M(u), \dots, \tau_n^M(u) \rangle \in R_M$

onde $\tau, \sigma, \tau_1, \dots, \tau_n \in \mathbb{T}_{\mathcal{L}}(\mathcal{V})$ e R é simbolo relacional com aridade $n > 0$.

¹³Grosso modo, a expressão “primeira ordem” indica que as variáveis serão usadas apenas para interpretar *elementos* dos modelos, e não *subconjuntos*. O leitor interessado em sintaxes de *segunda ordem* pode conferir [9].

¹⁴Isso fica explícito na construção recursiva de $\mathbb{F}_{\mathcal{L}}(\mathcal{V})$ por meio do Teorema F.2.3.

¹⁵E é bem mais comum escrever “=” em vez de “ \approx ”, mas eu prefiro reservar o símbolo “=” para a igualdade *entre conjuntos* ou – peço perdão pela expressão – para a igualdade *verdadeira*.

¹⁶É possível expressar tanto fórmulas quanto termos sem ambiguidades *e sem parênteses*. O leitor interessado deve pesquisar pela chamada “notação *prefixa*”.

(ii) Para os demais casos sem quantificadores:

Forma da φ	Critério para $M \models \varphi[u]$
$\neg\psi$	não ocorre $M \models \psi[u]$
$\psi_1 \vee \psi_2$	ocorre $M \models \psi_1[u]$ ou ocorre $M \models \psi_2[u]$
$\psi_1 \wedge \psi_2$	$M \models \psi_1[u]$ e $M \models \psi_2[u]$ ocorrem
$\psi_1 \rightarrow \psi_2$	não ocorre $M \models \psi_1[u]$ sem que $M \models \psi_2[u]$ ocorra
$\psi_1 \leftrightarrow \psi_2$	$M \models \psi_1[u]$ e $M \models \psi_2[u]$ ocorrem (ou falham) simultaneamente

onde $\psi, \psi_1, \psi_2 \in \mathbb{F}_{\mathcal{L}}(\mathcal{V})$.

(iii) Para os casos com quantificadores, convém introduzir a seguinte notação: para um elemento $a \in M$ e uma variável $x \in \mathcal{V}$, indica-se por $u_{x \mapsto a}$ à atribuição $v: \mathcal{V} \rightarrow M$ que faz $v(y) := u(y)$ para todo $y \in \text{dom}(u) \setminus \{x\}$, enquanto $v(x) := a$. Com isso dito:

Forma da φ	Critério para $M \models \varphi[u]$
$\exists x\psi$	ocorre $M \models \psi[u_{x \mapsto a}]$ para algum $a \in M$
$\forall x\psi$	ocorre $M \models \psi[u_{x \mapsto a}]$ para qualquer $a \in M$

onde $\psi \in \mathbb{F}_{\mathcal{L}}(\mathcal{V})$ e $x \in \mathcal{V}$ é uma variável. ¶

Desse modo, enquanto a interpretação de um termo $\tau \in \mathbb{T}_{\mathcal{L}}(\mathcal{V})$ é uma *função parcial* $\tau^M: [\mathcal{V} \rightarrow M] \rightarrow M$, a *interpretação* de uma fórmula $\varphi \in \mathbb{F}_{\mathcal{L}}(\mathcal{V})$ não *toma valores* em M , mas usa as interpretações em M dos eventuais termos que ocorrem em φ a fim de decidir acerca da ocorrência de $M \models \varphi[u]$: nesse sentido, é mais coerente dizer que “ $M \models \varphi[u]$ ” define uma relação de *satisfabilidade*, que determina se M satisfaz, ou não, uma determinada fórmula com os valores das variáveis dados por uma certa atribuição¹⁷. Por sua vez, os critérios utilizados para realizar tal decisão tão somente sistematizam o modo pelo qual argumentamos corriqueiramente com “não”, “e”, “ou”, “se..., então”, “se, e somente se”, “existe” e “para todo”.

Observação F.3.24 (\Rightarrow vs. \rightarrow). Ao longo do texto, como de costume, o símbolo “ \Rightarrow ” é usado como abreviação para “se..., então”. Nesse sentido, o leitor pode pensar em “ \Rightarrow ” como a versão *metalingüística* da noção subjacente de *implicação*, como em

$$M \models (\varphi \rightarrow \psi)[u] \text{ se, e somente se, } M \models \varphi[u] \Rightarrow M \models \psi[u].$$

Não faria sentido escrever “ $M \models \varphi[u]$ se, e somente se, $M \models \varphi[u] \rightarrow M \models \psi[u]$ ”, já que “ $M \models \varphi[u]$ ” e “ $M \models \psi[u]$ ” não são \mathcal{L} -fórmulas. Analogamente, o “se, e somente se” poderia ser trocado por “ \Leftrightarrow ”, mas não por “ \leftrightarrow ”. △

Definição F.3.25. Diz-se que uma \mathcal{L} -estrutura M é um **modelo** para uma fórmula $\varphi \in \mathbb{F}_{\mathcal{L}}(\mathcal{V})$, o que será indicado por $M \models \varphi$, se ocorrer $M \models \varphi[u]$ para qualquer atribuição de valores $u: \mathcal{V} \rightarrow M$ que for φ -compatível. Por extensão, diz-se que M é um **modelo** para um conjunto de fórmulas $T \subseteq \mathbb{F}_{\mathcal{L}}(\mathcal{V})$ se $M \models \varphi$ para toda fórmula $\varphi \in T$, o que se abrevia com $M \models T$. ¶

¹⁷Isto define uma função $\|\cdot\|_u: \mathbb{F}_{\mathcal{L}}(\mathcal{V}) \rightarrow \{0, 1\}$ em que $\|\varphi\|_u := 1$ se, e somente se, $M \models \varphi[u]$.

Na prática, a definição acima se restringe a um tipo particular de fórmula chamado de **sentença**, que por sua vez é uma fórmula sem *variáveis livres*. O que é uma variável livre? Simples: é uma variável da fórmula que não ocorre “presa” a algum quantificador. Por exemplo: na fórmula “ $\exists x \forall y x + y \approx z$ ”, tanto x quanto y ocorrem presas aos quantificadores \exists e \forall , respectivamente, enquanto a variável z é *livre*. Na prática, a satisfabilidade de tal fórmula numa estrutura apropriada depende do valor atribuído a z . Nesse sentido, é comum escrever $\varphi(x_1, \dots, x_n)$ para indicar que φ é uma fórmula em que as variáveis x_1, \dots, x_n ocorrem livres.

De volta à definição de modelo: note que se $M \models \varphi$, então $M \models \varphi[u]$ para qualquer atribuição u , donde segue que $M \models \forall x_1 \dots \forall x_n \varphi$, onde x_1, \dots, x_n são as possíveis variáveis livres de φ . A *sentença* $\forall x_1 \dots \forall x_n \varphi$ pode ser chamada de *fecho universal de φ* .

Observação F.3.26 (Relações de identidade vs. fórmulas). Embora seja quase explícito, convém reforçar: no caso em que a fórmula φ é da forma $\tau \approx \sigma$ para certos termos $\tau, \sigma \in \mathbb{T}_{\mathcal{L}}(\mathcal{V})$, verifica-se $M \models \tau \approx \sigma$ no sentido anterior se, e somente se, $M \models \tau \approx \sigma$ como na Definição F.3.11, i.e., no contexto da Álgebra Universal.

Com efeito, de acordo com a definição da subseção anterior, $M \models \tau \approx \sigma$ indica que $\tau^M(u) = \sigma^M(u)$ para qualquer função $u: \mathcal{V} \rightarrow M$. Uma vez que a definição da interpretação de termos para funções apenas estende as regras da interpretação de termos para atribuições, é imediato que $M \models (\tau \approx \sigma)[u]$ para qualquer atribuição $u: \mathcal{V} \rightarrow M$ e, portanto, $M \models \tau \approx \sigma$ de acordo com o *criterio de Tarski*. Reciprocamente, se $M \models (\tau \approx \sigma)[u]$ para qualquer atribuição φ -compatível $u: \mathcal{V} \rightarrow M$, então $\tau^M(u') = \sigma^M(u')$ para qualquer função $u': \mathcal{V} \rightarrow M$, já que as variáveis fora do *escopo* de τ e σ não interferem nas interpretações; pode-se argumentar de forma mais honesta por indução na complexidade dos termos, o que por sua vez é um edificante exercício para o leitor interessado. \triangle

Exemplo F.3.27 (Grupos, anéis e módulos). Seja $\mathcal{L} := \mathcal{L}_{\text{group}}$ a linguagem dos grupos: composta apenas por símbolos operacionais de aridades 0, 1 e 2, digamos e , i e $*$, respectivamente. Apesar do nome atribuído a tal linguagem, *qualquer* conjunto X munido de uma operação binária $*_X$, uma operação unária i_X e uma constante fixada e_X merece a alcunha de \mathcal{L} -estrutura. Dito isso, grupos são as \mathcal{L} -estruturas que modelam as fórmulas

- (i) $\xi := \forall x \forall y \forall z ((x * y) * z) \approx (x * (y * z))$,
- (ii) $\psi := \forall x ((x * e) \approx x) \wedge ((e * x) \approx x)$, e
- (iii) $\zeta := \forall x ((x * i(x)) \approx e) \wedge ((i(x) * x) \approx e)$,

onde x, y e z são variáveis num conjunto \mathcal{V} infinito enumerável¹⁸. Note que se G é uma \mathcal{L} -estrutura, então as interpretações dos termos na primeira fórmula correspondem às funções $G \times G \times G \rightarrow G$ que fazem $\langle a, b, c \rangle \mapsto (a *_G b) *_G c$ e $\langle a, b, c \rangle \mapsto a *_G (b *_G c)$, de modo que $G \models \xi$ se, e somente se, a operação $*_G: G \times G \rightarrow G$ é *associativa*, no sentido clássico. Da mesma forma, se $G \models \psi$ e $G \models \zeta$, então e_G é o *elemento neutro* da operação $*_G$, enquanto i_G se revela como a função que associa cada elemento de G ao seu $*_G$ -*inverso*. Em outras palavras: G é um *grupo* ou, caso se queira enfatizar o papel das operações, $\langle G, *_G, i_G, e_G \rangle$ é um grupo.

¹⁸Nada impede que se considerem outros conjuntos de variáveis. Porém, a fim de *capturar* todas as \mathcal{L} -estruturas merecedoras da alcunha de grupo, precisa-se de pelo menos três variáveis. Na prática, o mais comum é considerar $|\mathcal{V}| = \aleph_0$ e seguir com a vida.

Com o exemplo acima em mente, torna-se uma tarefa simples definir apropriadamente uma linguagem $\mathcal{L}' := \mathcal{L}_{\text{ring}}$ que permita descrever anéis como as \mathcal{L}' -estruturas que modelam um certo conjunto de fórmulas parecidas com as fórmulas ξ , ψ e ζ utilizadas na descrição dos grupos. Todavia, o caso dos *módulos* é menos imediato. Com um anel comutativo A fixado, diz-se que um grupo abeliano M é um **A -módulo** se existir uma função $*: A \times M \rightarrow M$ satisfazendo as seguintes condições:

- (i) $a * (b * m) = (a \cdot_A b) * m$ para quaisquer $a, b \in A$ e $m \in M$;
- (ii) $1_A * m = m$ para qualquer $m \in M$;
- (iii) $(a +_A b) * m = (a * m) +_M (b * m)$ para quaisquer $a, b \in A$ e $m \in M$;
- (iv) $a * (m +_M n) = (a * m) +_M (a * n)$ para quaisquer $a \in A$ e $m, n \in M$.

Num primeiro momento, a função $*: A \times M \rightarrow M$ parece fugir do poder de expressividade das operações, já que operações n -árias não tem esse formato. Porém, com $a \in A$ fixado, obtém-se uma operação unária $*_a: M \rightarrow M$ que faz $m \mapsto a * m$. Desse modo, a primeira exigência, por exemplo, se traduz na identidade $*_a(*_b(m)) = *_a \cdot_A b(m)$. Com isso, fica menos difícil perceber como definir uma linguagem apropriada para descrever os A -módulos: além dos símbolos que regem o comportamento do anel A e da estrutura de grupo abeliano de M , para cada $a \in A$ se acrescenta um símbolo $*_a$ de operação unária. As fórmulas, por sua vez, apenas traduzem as condições acima nesta linguagem. ▲

Observação F.3.28 (Magmas e morfismos). Leitores com alguma bagagem em Álgebra Abstrata podem ter se incomodado com as definições anteriores por uma questão de aparente excesso de hipóteses. Por exemplo, é comum definir *grupos* como sendo conjuntos dotados de apenas uma operação binária associativa para a qual *existe um* elemento neutro e em que todo elemento tem *um* inverso¹⁹. Daí, em virtude do Lema B.1.27, conclui-se que, de fato, existe *a* função que associa cada elemento ao seu *único* inverso, bem como a função que escolhe o *único* elemento neutro como constante.

Assim, é lícito considerar $\mathcal{L}'' := \mathcal{L}_{\text{mag}}$, a **linguagem dos magmas**²⁰, que tem um único símbolo operacional binário, e dizer que grupos são precisamente as \mathcal{L}'' -estruturas que satisfazem as fórmulas apropriadas. Naturalmente, o mesmo pode ser feito para anéis (daí com duas operações binárias) e módulos. Porém, isso traz pelo menos duas consequências.

- (i) Diferente do que se fez no exemplo anterior, a descrição dos grupos enquanto estruturas da linguagem dos magmas requer fórmulas mais complexas do que as utilizadas para descrevê-los enquanto $\mathcal{L} := \mathcal{L}_{\text{group}}$ -estruturas, especificamente no que se refere a quantificadores existenciais. A desvantagem é que com fórmulas mais complexas, os teoremas da Álgebra Universal acerca de estruturas algébricas não se aplicam²¹.
- (ii) Embora a mudança da linguagem possa não afetar os grupos, anéis e módulos (a menos de alterar as fórmulas que devem ser satisfeitas), linguagens diferentes induzem morfismos diferentes!

¹⁹Por sinal, confira a Definição B.1.28.

²⁰Um **magma** é apenas um conjunto dotado de uma operação binária.

²¹Lembre-se: em Álgebra Universal, as fórmulas *permitidas* são do tipo $\forall x \forall y \forall z \dots \tau \approx \sigma$ (confira a Observação F.3.26).

Com efeito, um \mathcal{L}'' -morfismo $f: M \rightarrow N$ entre duas \mathcal{L}'' -estruturas deve apenas satisfazer $f(a \star_M b) = f(a) \star_N f(b)$ para quaisquer $a, b \in M$, enquanto um \mathcal{L} -morfismo $f: G \rightarrow H$ entre \mathcal{L} -estruturas deve ser tal que $f(a *_G b) = f(a) *_H f(b)$, $f(i_G(a)) = i_H(f(a))$ e $f(e_G) = e_H$ para quaisquer $a, b \in M$. No caso dos grupos propriamente ditos, a distinção inexiste por conta da existência de inversos. Porém, ao se fazer isso para anéis, chega-se a duas definições de morfismo: uma em que se pede $f(1_A) = 1_B$ e outra em que isso não precisa ocorrer. \triangle

Exemplo F.3.29 (Corpos). Apesar de corpos serem $\mathcal{L}_{\text{ring}}$ -estruturas, as fórmulas utilizadas para descrevê-los são fundamentalmente mais complexas: a começar pela exigência $0_A \neq 1_A$, que consiste na negação de uma identidade entre termos constantes. Embora pareça um detalhe sutil, isto faz com que corpos não sejam *tão algébricos* quanto grupos e anéis: não é coincidência que produtos de grupos sejam grupos, ao passo que produtos de corpos, em geral, não sejam corpos (confira o Exercício F.13). \blacktriangle

Exemplo F.3.30. *Corpos ordenados* (definidos *en passant* na demonstração do Teorema C.3.2), são um tipo de estrutura cuja linguagem emprega símbolos operacionais e relacionais. Recordemo-nos então de que o corpo \mathbb{R} dos números reais costuma ser definido como (o único, a menos de isomorfismo) corpo ordenado e *completo*, onde a *completude* exprime o seguinte: todo subconjunto não-vazio de \mathbb{R} e limitado superiormente admite supremo. Tal propriedade não é exprimível como uma fórmula de primeira ordem por um motivo muito simples: a expressão “todo subconjunto... admite supremo” está quantificada sobre subconjuntos da estrutura, e não apenas sobre seus elementos. O que permite detectar isso é o fato de que quaisquer dois corpos ordenados completos têm a mesma cardinalidade, o que em geral não ocorre com o tipo de estrutura tratada nesta seção. Isso ficará mais claro adiante. \blacktriangle

Antes de iniciar a parte mais delicada desta subseção, convém retornar ao problema de entender o que morfismos preservam. A Proposição F.3.8 já deu uma dica, no sentido de mostrar que um morfismo entre \mathcal{L} -estruturas preserva identidades entre termos interpretados. Assim, seria natural que \mathcal{L} -estruturas isomórficas satisfizessem as mesmas fórmulas/sentenças. De fato:

Teorema F.3.31. *Se $f: M \rightarrow N$ é um \mathcal{L} -isomorfismo e $T \subseteq \mathbb{F}_{\mathcal{L}}(\mathcal{V})$ é um conjunto de fórmulas, então M é modelo para T se, e somente se, N é modelo para T .*

Demonstração. Basta mostrar que $M \models \varphi[f^{-1} \circ v] \Leftrightarrow N \models \varphi[v]$ para quaisquer \mathcal{L} -fórmula φ e atribuição $v: \mathcal{V} \rightarrow N$ φ -compatível, o que se faz por indução na complexidade de φ . Note que $u := f^{-1} \circ v: \mathcal{V} \rightarrow M$ é uma \mathcal{V} -atribuição φ -compatível em N .

- (i) Para $\varphi := \tau \approx \sigma$, se $\tau^M(u) = \sigma^M(u)$, então a identidade $\tau^N(v) = \sigma^N(v)$ segue da Proposição F.3.8. Analogamente, $N \models \varphi[v] \Rightarrow M \models \varphi[u]$.
- (ii) Para $\varphi := R(\tau_1, \dots, \tau_n)$, onde R é símbolo relacional n -ário e τ_1, \dots, τ_n são termos, se valer $M \models \varphi$, então $\langle \tau_1^M(u), \dots, \tau_n^M(u) \rangle \in R_M$, donde novamente a Proposição F.3.8 aliada às hipóteses sobre f acarretam $\langle \tau_1^N(v), \dots, \tau_n^N(v) \rangle \in R_N$. A outra implicação é análoga.
- (iii) Os demais casos não atômicos são imediatos. \square

A recíproca, i.e., “se duas \mathcal{L} -estruturas satisfazem as mesmas sentenças de primeira ordem, então ambas são isomórficas”, é falsa.

Exemplo F.3.32. Considere uma linguagem \mathcal{L} composta por um símbolo de relação binária R e duas constantes, digamos c' e c'' . Tomando-se ω com a relação de ordem usual, $c' := 0$ e $c'' := 2$, tem-se ω uma \mathcal{L} -estrutura, ao passo que o subconjunto P dos números naturais pares é, evidentemente, uma subestrutura. Porém, existem fórmulas na linguagem \mathcal{L} que interpretadas em P são verdadeiras em ω mas falsas em P : basta tomar $\exists x (c' R x \wedge x R c'')$. Ainda assim, note o seguinte: a existência de uma *testemunha* para (a interpretação da) fórmula fora de P foi o que permitiu a ocorrência de tal fenômeno. ▲

Definição F.3.33. Sejam M uma \mathcal{L} -estrutura e N uma subestrutura. Diz-se que N é **subestrutura elementar** de M se para toda \mathcal{L} -fórmula φ e toda atribuição de valores $u: \mathcal{V} \rightarrow N$ φ -compatível, ocorrer $N \models \varphi[u]$ se, e somente se, $M \models \varphi[u]$. ¶

Na prática, dizer que N é subestrutura elementar de M significa que toda fórmula interpretada em M com *parâmetros* em N é verdadeira em M se, e somente se, for verdadeira em N . Assim, o exemplo que antecede a definição acima exibe uma \mathcal{L} -subestrutura não-elementar de ω .

Verbalmente: se existem $a_1, \dots, a_n \in N$ e $b \in M$ tornando a interpretação de φ em M verdadeira, então pode-se trocar a *testemunha* b em M por outra testemunha $a \in N$. Vamos chamar tal propriedade de “substituição de testemunhas”. O importante a destacar: vale a volta.

Lema F.3.34 (Tarski-Vaught). *Se a subestrutura N tem a propriedade de substituição de testemunhas acima, então N é subestrutura elementar de M .*

Demonstração. Basta argumentar por indução na complexidade da fórmula φ presente na definição de subestrutura elementar. Com raciocínio essencialmente algébrico, a afirmação procurada valerá para todas as fórmulas sem quantificadores. No caso do quantificador existencial, é justamente a hipótese que permite provar $M \models \exists x \psi[u] \Rightarrow N \models \exists x \psi[u]$. Os detalhes ficam a cargo do leitor. □

Teorema F.3.35 (Löwenheim-Skolem, versão enumerável). *Sejam \mathcal{L} uma linguagem e \mathcal{V} um conjunto de variáveis, ambas enumeráveis. Se $T \subseteq \mathbb{F}_{\mathcal{L}}(\mathcal{V})$ é um conjunto de sentenças que admite um modelo infinito, então T admite um modelo infinito enumerável.*

Demonstração. Na prática, basta mostrar que se M é uma \mathcal{L} -estrutura com M infinito, então existe uma subestrutura elementar $N \subseteq M$ com $|N| = \aleph_0$: de fato, se isso for feito, basta tomar M como modelo de T , pois daí a elementaridade de N garantirá que $N \models T$. Para construir N , vamos obter uma cadeia $\langle N_j \rangle_{j \in \omega}$ de subestruturas de M tal que as fórmulas existenciais satisfeitas por M com parâmetros em N_j tenham testemunhas em N_{j+1} , para todo j ; com isso feito basta tomar $N := \bigcup_{j \in \omega} N_j$ com as interpretações “óbvias”.

O *Katzensprung* da prova está contido na Observação F.2.4: se $|X| = \aleph_0$ e $|\mathcal{F}| \leq \aleph_0$, então $|\mathcal{F}(X)| = \aleph_0$, onde $\mathcal{F}(X)$ é a subestrutura gerada por X . Assim, fixando-se $N'_0 \subseteq M$ um subconjunto infinito enumerável:

- ✓ toma-se $N_0 := \text{sp}(N'_0)$, i.e., fecha-se N'_0 por constantes e operações algébricas;
- ✓ faz-se $N'_1 := N_0 \cup \{a_{\varphi,u} : \varphi \in \mathbb{F}_{\mathcal{L}}(\mathcal{V}), u: \mathcal{V} \rightarrow N_0 \text{ é } \varphi\text{-compatível e } M \models \exists x \varphi[u]\}$, onde $a_{\varphi,u} \in M$ é escolhido de tal forma que $M \models \varphi[u_{x \mapsto a_{\varphi,u}}]$; como cada atribuição u percorre apenas finitas variáveis (entre as quais se incluem as possíveis variáveis livres de φ), não é difícil ver²² que N'_1 é um subconjunto enumerável de M , de modo que $N_1 := \text{sp}(N'_1)$ é ainda uma subestrutura de M com $N_0 \subseteq N_1$.

²²Talvez apesar lembrar que $|[\mathcal{V}]^{<\aleph_0}| \leq \aleph_0$.

Procedendo recursivamente dessa forma, o *teste de Tarski-Vaught* garante que N é uma subestrutura elementar de M , como desejado. \square

Como isso mostra a falsidade da recíproca do Teorema F.3.31? Muito simples: fixada uma \mathcal{L} -estrutura infinita, é formalmente lícito (embora imoral) considerar o conjunto de sentenças $T := \{\varphi \in \mathbb{F}_{\mathcal{L}}(\mathcal{V}) : \varphi \text{ é sentença e } M \models \varphi\}$; como $M \models T$, Löwenheim-Skolem garante um subestrutura N de M com $N \models T$ e $|N| = \aleph_0$. Em particular, se $|M| > \aleph_0$, então M e N não podem ser isomorfas!

Exemplo F.3.36 (A completude Revisitada). O argumento acima permite mostrar, sem grandes dificuldades, que não existe *axiomática de primeira-ordem* capaz de capturar *todos* os fatos acerca de corpos ordenados completos. Mais precisamente, uma vez que o *axioma de completude* trata de *subconjuntos* da estrutura, segue que não se pode utilizará-lo *ipsis litteris* como fórmula de *primeira ordem*: a sintaxe de tais fórmulas versa apenas sobre *elementos* da estrutura.

Porém, isto não significa, *a priori*, que não existam outras fórmulas de primeira ordem capazes de fazer o trabalho. Por exemplo: alguém poderia propor que considerássemos, para cada fórmula φ na linguagem dos corpos ordenados, a fórmula $\sup \varphi$ dada por

$$(\exists x\varphi \wedge \exists y\forall z (\varphi(z) \rightarrow z \leq y)) \rightarrow \exists a\forall z (\varphi(z) \rightarrow z \leq a) \wedge \forall b\forall z (\varphi(z) \rightarrow z \leq b) \rightarrow a \leq b,$$

que em certo sentido expressa que um subconjunto da estrutura descrito pela fórmula φ admitirá supremo, desde que seja não-vazio ($\exists x\varphi$) e limitado superiormente ($\exists y\dots$). Certamente, $\mathbb{R} \models \sup \varphi$ para qualquer fórmula φ ; no entanto, como qualquer conjunto infinito tem mais subconjuntos do que fórmulas para descrevê-los (Observação E.2.21), não parece crível que tal abordagem seja capaz de caracterizar a completude de corpos ordenados.

De fato, fixada *qualquer* linguagem enumerável \mathcal{L} tal que \mathbb{R} seja uma \mathcal{L} -estrutura²³, para um conjunto de variáveis \mathcal{V} com $|\mathcal{V}| = \aleph_0$ e $\mathcal{R} := \{\varphi \in \mathbb{F}_{\mathcal{L}}(\mathcal{V}) : \mathbb{R} \models \varphi\}$, o Teorema de Löwenheim-Skolem garante uma \mathcal{L} -estrutura R com $R \models \mathcal{R}$ e $|R| = \aleph_0$. Uma vez que todo corpo ordenado completo tem cardinalidade 2^{\aleph_0} , resulta que o modelo R não pode ser completo – embora satisfaça todas as sentenças de primeira ordem satisfeitas por \mathbb{R} ! \blacktriangleleft

F.4 Verdade ou consequência

Definição F.4.1. Sejam \mathcal{L} uma linguagem e \mathcal{V} um conjunto infinito enumerável de variáveis²⁴. Uma *sentença* $\varphi \in \mathbb{F}_{\mathcal{L}}(\mathcal{V})$ é uma **consequência (semântica)** de um conjunto de sentenças $T \subseteq \mathbb{F}_{\mathcal{L}}(\mathcal{V})$ se todo modelo M de T também for modelo de φ . Em outras palavras: se para toda \mathcal{L} -estrutura M valer que $M \models \varphi$ sempre que $M \models T$. Como de costume nos textos da área, isto será abreviado com $T \models \varphi$. \P

A definição acima poderia ser feita para um conjunto de fórmulas T em vez de um conjunto de sentenças. No entanto, se $T \models \varphi$, então $T' \models \varphi$, onde T' é o conjunto dos fechos universais de cada fórmula $\psi \in T$. Neste caso, conjuntos de sentenças, como T' , ganham um nome especial: são chamados de **teorias de primeira ordem na linguagem \mathcal{L}** ou, de maneira mais concisa, **\mathcal{L} -teorias**²⁵.

²³A linguagem dos corpos ordenados, por exemplo.

²⁴Os elementos de \mathcal{V} serão denotados por letras minúsculas com ou sem subíndices, como de costume.

²⁵Algumas referências chamam de *teoria* apenas os conjuntos de sentenças *fechados* por *consequência*, no sentido da Definição F.4.1.

Embora modelos estejam intrinsecamente ligados à definição de “ $T \models \varphi$ ”, o que se obtém é uma relação (de consequência) entre sentenças, que na prática coincide com o que já se faz no dia a dia: a fim de provar, por exemplo, a afirmação “num corpo ordenado \mathbb{K} ocorre $0_{\mathbb{K}} < 1_{\mathbb{K}}$ ”, toma-se um corpo ordenado \mathbb{K} qualquer (i.e., um modelo para o conjunto de sentenças \mathcal{A}_{OF} que descreve os corpos ordenados) no qual se verifica que, de fato, $0_{\mathbb{K}} < 1_{\mathbb{K}}$. Da arbitrariedade do corpo ordenado tomado, a única conclusão razoável é a de que a afirmação é *consequência* das sentenças em \mathcal{A}_{OF} .

Dado que costumamos pensar em modelos como os lugares em que a matemática “acontece”, é razoável refrasear “ $T \models \varphi$ ” como “ φ é verdade sempre que T for verdade”, onde **verdade** expressa apenas a ocorrência em modelos. Essa é uma das razões pelas quais profissionais chamam tal abordagem de *semântica*: ela depende dos significados dados pelas interpretações em modelos. Por mais que esta seja a minha abordagem preferida, ela não é única: é possível determinar noções de consequência baseadas, tão somente, na estrutura (sintaxe) das fórmulas.

Definição F.4.2. Sejam \mathcal{L} uma linguagem e \mathcal{V} um conjunto infinito enumerável de variáveis.

- (i) Um **sistema de dedução** $\mathcal{D} := \langle \mathcal{A}, I \rangle$ para $\mathbb{F}_{\mathcal{L}}(\mathcal{V})$ consiste de um conjunto \mathcal{A} de fórmulas, chamadas **axiomas lógicos**, e de uma função $I: D_I \rightarrow \mathbb{F}_{\mathcal{L}}(\mathcal{V})$ chamada **regra de inferência**, onde D_I é uma coleção de sequências finitas de \mathcal{L} -fórmulas.
- (ii) Seja $\varphi \in \mathbb{F}_{\mathcal{L}}(\mathcal{V})$ uma \mathcal{L} -fórmula e $\Sigma \subseteq \mathbb{F}_{\mathcal{L}}(\mathcal{V})$ um conjunto de sentenças, chamados de **axiomas não-lógicos**. Diremos que uma sequência finita $\langle \varphi_1, \dots, \varphi_n \rangle$ de \mathcal{L} -fórmulas é uma \mathcal{D} -prova para φ com respeito a Σ se:
 - para cada $i < n$, $\varphi_i \in \mathcal{A} \cup \Sigma$ ou existem $j_1, \dots, j_m < i$ com $I(\varphi_{j_1}, \dots, \varphi_{j_m}) = \varphi_i$;
 - $\varphi_n := \varphi$.

Em tal situação, escreveremos $\Sigma \vdash_{\mathcal{D}} \varphi$ para indicar que existe uma \mathcal{D} -prova para φ com respeito a Σ , i.e., que φ é **demonstrável** (ou **dedutível**²⁶) segundo o sistema \mathcal{D} com os axiomas não-lógicos Σ . ¶

A Definição F.4.2, adaptada do texto de Ben-Ari [3] e Manin [25], é tremendamente geral e, para os propósitos modestos deste texto, abrangente demais, por englobar inclusive sistemas dedutivos *inúteis*: basta tomar $\mathcal{A} := D_I := \emptyset$ ou $\mathcal{A} := \mathbb{F}_{\mathcal{L}}(\mathcal{V})$, por exemplo²⁷. Um sistema de dedução *útil* deveria, pelo menos, concordar com a *prática matemática*, no sentido de demonstrar *verdades*. Posto de outra forma:

Definição F.4.3. Diremos que um sistema de dedução \mathcal{D} é **correto** se a ocorrência de $\Sigma \vdash_{\mathcal{D}} \varphi$ implicar em $\mathcal{A} \cup \Sigma \models \varphi$ para qualquer coleção $\Sigma \cup \{\varphi\}$ de \mathcal{L} -sentenças. ¶

Na contrapositiva, a definição acima diz que se ao menos um modelo M para os axiomas em $\mathcal{A} \cup \Sigma$ não satisfaz uma sentença φ , então o sistema de dedução \mathcal{D} não deve demonstrar φ . Agora, se um sistema de dedução \mathcal{D} satisfaz a recíproca dessa condição, então sempre que uma sentença φ se verificar em *todos* os modelos de $\mathcal{A} \cup \Sigma$ deverá ser possível encontrar uma \mathcal{D} -prova para φ : em outras palavras, \mathcal{D} é *completo*, no sentido de que prova todas as *verdades* de $\mathcal{A} \cup \Sigma$.

²⁶E ainda há quem *insista* em **consequência sintática**.

²⁷Leitores interessados numa abordagem ainda mais generalista para sistemas dedutivos podem conferir a obra de Anita Wasilewska [34].

Definição F.4.4. Diremos que um sistema de dedução \mathcal{D} é **completo** se a ocorrência de $\mathcal{A} \cup \Sigma \models \varphi$ implicar em $\Sigma \vdash_{\mathcal{D}} \varphi$ para qualquer coleção $\Sigma \cup \{\varphi\}$ de \mathcal{L} -sentenças. ¶

Analisar e discutir os diversos sistemas de dedução razoáveis (via \vdash) é uma das atribuições da *Teoria da Prova* que, como o nome sugere, transforma as boas e velhas *demonstrações* em objetos matemáticos passíveis de análise. Por sua vez, a *Teoria dos Modelos* se ocupa do estudo das noções de consequência (via \models) e suas aplicações. Embora pareçam afastadas da prática matemática usual, ambas fazem com o *modus operandi* matemático o que a Álgebra fez com a Aritmética, por exemplo: abstrair seus métodos por meio de abordagens axiomáticas. É por tal razão que costuma-se pensar em tais subáreas como pertencentes à *Metamatemática*.

Em virtude da minha predileção pela abordagem semântica (além da extensão do texto), não serão apresentados sistemas dedutivos específicos: o leitor interessado encontrará fartas discussões na obra de Negri & von Plato [27] ou, se preferir uma abordagem mais *natural*, o recente livro de Halbeisen & Kraft [15] – além do (já) clássico Dirk van Dalen [33]. Por ora, basta saber que os sistemas dedutivos usuais são simultaneamente corretos e completos. Uma *consequência* (aparentemente) inócuia disso é o

Teorema F.4.5 (Compacidade). *Uma \mathcal{L} -teoria $\Sigma \subseteq \mathbb{F}_{\mathcal{L}}(\mathcal{V})$ tem modelo se, e somente se, todo subconjunto finito de Σ tem modelo.*

Esboço da demonstração. É claro que se $M \models \Sigma$, então $M \models F$ para todo subconjunto finito $F \subseteq \Sigma$. Agora, se Σ não tem modelo, então ao fixar o seu sistema de dedução *correto e completo* favorito para $\mathbb{F}_{\mathcal{L}}(\mathcal{V})$, digamos $\mathcal{D} := \langle \mathcal{A}, I \rangle$, seria possível encontrar $\varphi \in \Sigma$ juntamente com uma \mathcal{D} -prova para $\neg\varphi$ com respeito a Σ (por completude!); uma vez que tal \mathcal{D} -prova envolve apenas um subconjunto finito de fórmulas $\Sigma' \subseteq \Sigma$, chega-se a $\Sigma' \vdash_{\mathcal{D}} \neg\varphi$ e, pela completude do sistema, $\mathcal{A} \cup \Sigma' \models \neg\varphi$. Logo, qualquer modelo de Σ' satisfaz $\neg\varphi$, donde segue que não existe modelo para $\Sigma' \cup \{\varphi\} \subseteq \Sigma$: o contrário daria um modelo em que φ e $\neg\varphi$ seriam satisfeitas simultaneamente. □

O esboço acima esconde detalhes sintáticos importantes: a rigor, a argumentação depende de alguns mecanismos de sistemas de dedução específicos, como provas por contraposição, absurdo, etc.; além disso, as sentenças que compõem os axiomas lógicos da família \mathcal{A} costumam ser satisfeitas por qualquer estrutura da linguagem \mathcal{L} , o que permite omitir sua última ocorrência a fim de obter a (realmente desejada) expressão “ $\Sigma \models \neg\varphi$ ”. Não obstante, é possível demonstrar o Teorema da Compacidade sem apelar para os recursos da Teoria da Prova, por argumentos totalmente semânticos²⁸.

Observação F.4.6. Antes de prosseguir, é altamente recomendado revisar as definições de filtro e ultrafiltro no Exemplo D.3.9. △

Exemplo F.4.7. Suponha que para cada $n \in \omega$ seja dado um grupo G_n , com G_0 seja não-abeliano e G_n abeliano para todo $n > 0$. Mesmo com a *maioria* dos grupos sendo abelianos, ao se considerar a estrutura usual em $\prod_{n \in \omega} G_n$, o resultado é um grupo não-abeliano. Não seria interessante corrigir isso por meio de alguma relação de equivalência que respeitasse a voz da maioria? Entram em cena os *ultraprodutos!* ▲

²⁸Cabe destacar que tal desvio não se justifica apenas para evitar as pequenas tecnicidades mencionadas. Com efeito, embora a falta de ênfase possa ter passado a impressão contrária, não é trivial provar que os sistemas dedutivos usuais são completos! Na prática, provar a completude de um sistema consiste em mostrar que teorias que não demonstram contradições (fórmulas do tipo $\neg\varphi \wedge \varphi$) têm modelos! Em outras palavras: a partir de uma teoria *consistente*, deve-se “construir” um modelo para a teoria.

Sejam $\mathcal{I} \neq \emptyset$ um conjunto de índices e $\{M_i : i \in \mathcal{I}\}$ uma família de conjuntos. Dado um filtro \mathcal{F} em \mathcal{I} , considere $\sim_{\mathcal{F}}$ a relação binária em $M := \prod_{i \in \mathcal{I}} M_i$ definida por $f \sim_{\mathcal{F}} g \Leftrightarrow \{i \in \mathcal{I} : f(i) = g(i)\} \in \mathcal{F}$.

Lema F.4.8. A relação $\sim_{\mathcal{F}}$ é uma equivalência sobre M .

Demonstração. Embora a prova seja simples, convém apresentá-la para aquecer os ânimos do leitor.

✓ $\sim_{\mathcal{F}}$ é reflexiva: para $f \in M$, tem-se $f(i) = f(i)$ para todo $i \in \mathcal{I}$, i.e.,

$$\mathcal{I} = \{i \in \mathcal{I} : f(i) = f(i)\};$$

uma vez que $\mathcal{I} \in \mathcal{F}$ por \mathcal{F} ser filtro²⁹, tem-se $f \sim_{\mathcal{F}} f$.

✓ $\sim_{\mathcal{F}}$ é simétrica: se $f \sim_{\mathcal{F}} g$, então $\{i \in \mathcal{I} : f(i) = g(i)\} \in \mathcal{I}$; como se tem

$$\{i \in \mathcal{I} : f(i) = g(i)\} = \{i \in \mathcal{I} : g(i) = f(i)\},$$

segue que $g \sim_{\mathcal{F}} f$;

✓ $\sim_{\mathcal{F}}$ é transitiva: para $f, g, h \in M$ com $f \sim_{\mathcal{F}} g$ e $g \sim_{\mathcal{F}} h$, a definição da relação impõe que

$$I := \{i \in \mathcal{I} : f(i) = g(i)\} \text{ e } J := \{i \in \mathcal{I} : g(i) = h(i)\}$$

pertencem ao filtro \mathcal{F} ; posto que $I \cap J \subseteq \{i \in \mathcal{I} : f(i) = h(i)\} := K$ com $I \cap J \in \mathcal{F}$ (pois \mathcal{F} é filtro!), resulta que $K \in \mathcal{F}$ e, portanto, $f \sim_{\mathcal{F}} h$. \square

Exercício F.4. No último item, mostre que a inclusão $I \cap J \subseteq K$ pode ser própria. ■

Vamos denotar por $\prod_{\mathcal{F}} M_i$ o quociente $(\prod_{i \in \mathcal{I}} M_i) / \sim_{\mathcal{F}}$, usualmente chamado de **produto reduzido** de $\langle M_i \rangle_{i \in \mathcal{I}}$ com respeito ao filtro \mathcal{F} . Por simplicidade, indicaremos por \bar{f} a classe de equivalência de $f \in \prod_{i \in \mathcal{I}} M_i$ com respeito à relação $\sim_{\mathcal{F}}$. Nas felizes situações em que \mathcal{F} for um ultrafiltro em \mathcal{I} , $\prod_{\mathcal{F}} M_i$ será xingado apenas de **ultraproduto**.

Exemplo F.4.9 (Opcional (?): medidas e ultrafiltros). Fixado um ultrafiltro \mathfrak{u} num conjunto \mathcal{I} , é lícito definir a função $m_{\mathfrak{u}}: \wp(\mathcal{I}) \rightarrow \{0, 1\}$ que a cada subconjunto $A \subseteq \mathcal{I}$ faz $m_{\mathfrak{u}}(A) := 1$ se $A \in \mathfrak{u}$, e $m_{\mathfrak{u}}(A) := 0$ caso contrário. Agora, em virtude da Proposição D.3.13, não é difícil perceber que $m_{\mathfrak{u}}$ tem as seguintes propriedades:

- (i) $m_{\mathfrak{u}}(\emptyset) = 0$ e $m_{\mathfrak{u}}(\mathcal{I}) = 1$;
- (ii) se $A_0, \dots, A_n \subseteq \mathcal{I}$ são subconjuntos dois a dois disjuntos, então

$$m_{\mathfrak{u}} \left(\bigcup_{j \leq n} A_j \right) = \sum_{j \leq n} m_{\mathfrak{u}}(A_j),$$

o que se resume em dizer que $m_{\mathfrak{u}}$ é uma *medida finitamente aditiva*.

Com isso em mente, ao considerar todos os conjuntos M_i iguais entre si, digamos $M_i := M$ para todo $i \in \mathcal{I}$, o produtório $\prod_{i \in \mathcal{I}} M_i$ coincide com as funções da forma $\mathcal{I} \rightarrow M$. Nesse sentido, $f \sim_{\mathfrak{u}} g$ se revela ser a afirmação de que as duas funções $f, g: \mathcal{I} \rightarrow M$ são *iguais a menos de um conjunto de medida $m_{\mathfrak{u}}$ nula*: de fato, se $f \sim_{\mathfrak{u}} g$, então $C := \{i \in \mathcal{I} : f(i) \neq g(i)\} \in \mathfrak{u}$ e, por conseguinte, o conjunto dos pontos de \mathcal{I} em que f e g diferem não pertence a \mathfrak{u} (pelo item (ii) da Proposição D.3.13) e, portanto, sua medida $m_{\mathfrak{u}}$ deve ser nula; a recíproca é análoga. ▲

²⁹Filtros são não-vazios e “fechados para cima”. Logo, existe $\mathcal{J} \in \mathcal{F}$ e $\mathcal{J} \subseteq \mathcal{I}$, acarretando $\mathcal{I} \in \mathcal{F}$!

Nas situações em que cada M_i é \mathcal{L} -estrutura de uma linguagem \mathcal{L} fixada, ainda é possível promover o produto reduzido $M_{\mathcal{F}} := \prod_{\mathcal{F}} M_i$ ao patamar de \mathcal{L} -estrutura. Explicitamente:

- (i) para um símbolo de constante $c \in \mathcal{O}_{\mathcal{L}}$, faz-se $c_{M_{\mathcal{F}}} := \overline{\langle c^{M_i} \rangle_{i \in \mathcal{I}}}$;
- (ii) para um símbolo operacional $s \in \mathcal{O}_{\mathcal{L}}$ com aridade $n > 0$, $s_{M_{\mathcal{F}}}$ é a operação que faz

$$s_{M_{\mathcal{F}}} (\langle \overline{f_1}, \dots, \overline{f_n} \rangle) := \overline{\langle s_{M_i}(f_1(i), \dots, f_n(i)) \rangle_{i \in \mathcal{I}}}$$

para cada n -upla $\langle \overline{f_1}, \dots, \overline{f_n} \rangle \in M_{\mathcal{F}}^n$ (note que para cada $j \leq n$, f_j é uma \mathcal{I} -upla $\langle f_j(i) \rangle_{i \in \mathcal{I}} \in \prod_{i \in \mathcal{I}} M_i$);

- (iii) enfim, para um símbolo relacional $R \in \mathcal{R}_{\mathcal{L}}$ com aridade $n > 0$,

$$R_{M_{\mathcal{F}}} := \left\{ \langle \overline{f_1}, \dots, \overline{f_n} \rangle \in M_{\mathcal{F}}^n : \underbrace{\{i \in \mathcal{I} : \langle f_1(i), \dots, f_n(i) \rangle \in R_{M_i}\}}_{T_f} \in \mathcal{F} \right\}.$$

As coisas todas ficam bem definidas por \mathcal{F} ser filtro:

- (i) nada precisa ser mostrado para a definição da constante $c_{M_{\mathcal{F}}}$;
- (ii) no caso de $s_{M_{\mathcal{F}}}$, note que se $\overline{f_j} = \overline{g_j}$ para cada $j \leq n$, então o conjunto de índices $C_j := \{i \in \mathcal{I} : f_j(i) = g_j(i)\}$ deve pertencer ao filtro \mathcal{F} para cada j , enquanto

$$\bigcap_{j \leq n} C_j \subseteq \{i \in \mathcal{I} : s_{M_i}(f_1(i), \dots, f_n(i)) = s_{M_i}(g_1(i), \dots, g_n(i))\},$$

com $\bigcap_{j \leq n} C_j \in \mathcal{F}$ pois \mathcal{F} é filtro;

- (iii) a relação $R_{M_{\mathcal{F}}}$ fica bem definida no sentido de satisfazer a equivalência

$$\langle \overline{f_1}, \dots, \overline{f_n} \rangle \in R_{M_{\mathcal{F}}} \Leftrightarrow \underbrace{\{i \in \mathcal{I} : \langle f_1(i), \dots, f_n(i) \rangle \in R_{M_i}\}}_{T_f} \in \mathcal{I} \quad (\text{F.2})$$

independentemente da escolha dos representantes de cada classe $\overline{f_j}$, o que segue pois se $\overline{g_j} = \overline{f_j}$ para cada j , então, com as notações do item anterior,

$$T_f \cap \bigcap_{j \leq n} C_j \subseteq T_g \text{ e } T_g \cap \bigcap_{j \leq n} C_j \subseteq T_f,$$

onde a equivalência desejada segue.

Exercício F.5. Mostre que para um termo $\tau \in \mathbb{T}_{\mathcal{L}}(\mathcal{V})$, verifica-se $\tau^{M_{\mathcal{F}}} = \overline{\tau^M}$, em que M indica, por simplicidade, o produto $\prod_{i \in \mathcal{I}} M_i$ dotado de sua \mathcal{L} -estrutura natural³⁰. Dica: indução na complexidade. ■

Até aqui, todas as considerações feitas foram meramente estruturais, sem qualquer relação aparente com satisfabilidade de fórmulas – o que dizer então sobre o Teorema da Compacidade? Se estivéssemos numa partida de RPG, o mestre poderia dizer: role um D6 para fazer um *check* de sanidade antes de prosseguir.

³⁰Confira o Exemplo F.1.10 para se lembrar da estrutura produto (ou encare a estrutura no produto reduzido anterior, com $\mathcal{F} := \{\mathcal{I}\}$, até que ela te encare de volta)

Para simplificar as notações, vamos escrever M para indicar $\prod_{i \in \mathcal{I}} M_i$, enquanto n -uplas em M^n serão denotadas por $\vec{\alpha}$: assim $\vec{\alpha} := \langle \alpha_1, \dots, \alpha_n \rangle$, onde cada α_j é uma \mathcal{I} -upla; em particular, $\vec{\alpha}(i) := \langle \alpha_1(i), \dots, \alpha_n(i) \rangle$ é uma \mathcal{V} -atribuição em M_i , enquanto $\vec{\alpha}$ é uma atribuição em M . Agora, para uma fórmula $\varphi \in \mathbb{F}_{\mathcal{L}}(\mathcal{V})$ com n variáveis livres, digamos x_1, \dots, x_n , e uma n -upla $\vec{\alpha} \in M^n$, defina

$$\|\varphi[\vec{\alpha}]\| := \{i \in \mathcal{I} : M_i \models \varphi[\vec{\alpha}(i)]\},$$

onde a expressão “ $M_i \models \varphi[\vec{\alpha}(i)]$ ” abrevia que M_i satisfaz φ com respeito a $\vec{\alpha}(i) \in M_i^n$.

Lema F.4.10. *Para fórmulas $\varphi, \psi \in \mathbb{F}_{\mathcal{L}}(\mathcal{V})$ com n variáveis livres e $\vec{\alpha} \in M^n$, valem as identidades:*

- i) $\|(\varphi \wedge \psi)[\vec{\alpha}]\| = \|\varphi[\vec{\alpha}]\| \cap \|\psi[\vec{\alpha}]\|$;
- ii) $\|(\varphi \vee \psi)[\vec{\alpha}]\| = \|\varphi[\vec{\alpha}]\| \cup \|\psi[\vec{\alpha}]\|$;
- iii) $\|\neg\varphi[\vec{\alpha}]\| = \mathcal{I} \setminus \|\varphi[\vec{\alpha}]\|$.

Demonstração. Basta “abrir” a definição de “ $M_i \models \varphi[\vec{\alpha}(i)]$ ” de acordo com a semântica de Tarski (Definição F.3.23). O leitor pode cuidar dos detalhes. \square

Em certo sentido, $\|\varphi[\vec{\alpha}]\|$ codifica os índices em \mathcal{I} cujos fatores M_i satisfazem $\varphi[\vec{\alpha}(i)]$. Por outro lado, ultrafiltros capturam o significado de ser *maioria* em \mathcal{I} . A junção dessas duas observações resulta no

Teorema F.4.11 (Łoś). *Sejam \mathcal{L} , M , φ e $\vec{\alpha}$ como antes, e suponha que \mathfrak{u} seja um ultrafiltro em \mathcal{I} . Em tais condições,*

$$M_{\mathfrak{u}} \models \varphi[\vec{\alpha}] \Leftrightarrow \|\varphi[\vec{\alpha}]\| \in \mathfrak{u}$$

Demonstração. Como já era de se esperar, a prova será por indução na complexidade da \mathcal{L} -fórmula φ .

- (i) Para φ atômica, pode-se ter $\varphi := \tau \approx \sigma$ ou $\varphi := R(\tau_1, \dots, \tau_n)$: para o primeiro caso, a semântica de Tarski estipula, por um lado, $M_{\mathfrak{u}} \models \tau \approx \sigma$ como sinônimo de $\tau^{M_{\mathfrak{u}}}(v) = \sigma^{M_{\mathfrak{u}}}(v)$ para qualquer atribuição φ -compatível $v : \mathcal{V} \rightharpoonup M_{\mathfrak{u}}$, ao passo que o exercício anterior revela que $\tau^{M_{\mathfrak{u}}}(v) = \tau^M(v)$ e $\sigma^{M_{\mathfrak{u}}}(v) = \sigma^M(v)$, donde a equivalência desejada segue pela definição de $\sim_{\mathfrak{u}}$; o segundo caso é análogo (e se vale da equivalência destacada em (F.2)).
- (ii) Supondo $\varphi := \neg\psi$, com o resultado válido para ψ , note que $M_{\mathfrak{u}} \models \varphi[\vec{\alpha}]$ se, e somente se, $M_{\mathfrak{u}} \not\models \psi[\vec{\alpha}]$, o que por hipótese equivale a $\|\psi[\vec{\alpha}]\| \notin \mathfrak{u}$. Ora, por \mathfrak{u} ser ultrafiltro, isto equivale (pelo lema anterior) a $\mathcal{I} \setminus \|\psi[\vec{\alpha}]\| = \|\neg\psi[\vec{\alpha}]\| \in \mathfrak{u}$.
- (iii) O caso em que $\varphi := \psi_1 \Box \psi_2$, com $\Box \in \{\wedge, \vee, \rightarrow, \leftrightarrow\}$ e o resultado válido para ψ_1 e ψ_2 , é análogo: por exemplo, utilizando o lema anterior juntamente com o fato de \mathfrak{u} ser fechado por interseções finitas, segue sem grandes dificuldades que $M_{\mathfrak{u}} \models (\psi_1 \wedge \psi_2)[\vec{\alpha}]$ se, e somente se, $\|\psi_1[\vec{\alpha}]\| \cap \|\psi_2[\vec{\alpha}]\| = \|(\psi_1 \wedge \psi_2)[\vec{\alpha}]\| \in \mathfrak{u}$; daí os demais casos saem aliados ao anterior por meio das identidades booleanas usuais³¹, embora nada impeça que o leitor demonstre tais instâncias, como exercício.

³¹ $\neg\neg\psi_1 = \psi_1$, $\psi_1 \vee \psi_2 = \neg(\neg\psi_1 \wedge \neg\psi_2)$, $\psi_1 \rightarrow \psi_2 = \neg\psi_1 \vee \psi_2$ e $\psi_1 \leftrightarrow \psi_2 = (\psi_1 \rightarrow \psi_2) \wedge (\psi_2 \rightarrow \psi_1)$.

- (iv) Enfim, os casos quantificados. Para $\varphi := \exists x\psi$, com a equivalência válida para ψ , note que se $M_u \models \varphi[\vec{\alpha}]$, então existe $\vec{b} \in M_u$ com $M \models \psi[\vec{\alpha}, \vec{b}]$, donde a hipótese acarreta $\{i \in \mathcal{I} : M_i \models \psi[\alpha(i), b(i)]\} \in u$, com tal família de índices contida em $\|\varphi[\vec{\alpha}]\|$, o que assegura a pertinência desejada. Para a recíproca, se $\|\varphi[\vec{\alpha}]\| \in u$, então para cada $i \in \|\varphi[\vec{\alpha}]\|$ é lícito escolher $b_i \in M_i$ com $M_i \models \psi[\vec{\alpha}(i), b_i]$; tomando-se $b_i \in M_i$ arbitrariamente nos casos em que $i \notin \|\varphi[\vec{\alpha}]\|$, resulta que $\{i \in \mathcal{I} : M_i \models \psi[\vec{\beta}]\} \in u$, onde $\vec{\beta} := \langle \alpha_1, \dots, \alpha_n, b \rangle$; logo, pela hipótese acerca de ψ , deve-se ter $M_u \models \psi[\vec{\beta}]$, i.e., $M_u \models \psi[\vec{\alpha}, \vec{b}]$, que significa, precisamente, $M_u \models \varphi[\vec{\alpha}]$. O caso do quantificador \forall segue pois $\forall x\varphi = \neg\exists x\neg\varphi$. \square

No caso particular em que φ não tem variáveis livres, o Teorema de Łoś se transforma na equivalência $M_u \models \varphi \Leftrightarrow \|\varphi\| \in u$, em que $\|\varphi\| := \{i \in \mathcal{I} : M_i \models \varphi\}$. Pode não parecer, mas acabamos de provar o Teorema da Compacidade.

Demonstração do Teorema da Compacidade. Já que, por hipótese, para cada subconjunto finito $F \subseteq \Sigma$ existe uma \mathcal{L} -estrutura M_F com $M_F \models F$, é razoável chamar de \mathcal{I} a família dos subconjuntos finitos e não-vazios de Σ . Agora, $\mathcal{H} := \{\{F \in \mathcal{I} : G \subseteq F\} : G \in \mathcal{I}\}$ é uma família de subconjuntos de \mathcal{I} com a *propriedade da interseção finita*: sempre que $\mathcal{F}_0, \dots, \mathcal{F}_n \in \mathcal{H}$, assegura-se $\bigcap_{j \leq n} \mathcal{F}_j \neq \emptyset$. No caso, isto vale pois cada \mathcal{F}_j é da forma $\{F \in \mathcal{I} : G_j \subseteq J\}$ para algum $G_j \in \mathcal{I}$, donde segue que $\bigcup_{j \leq n} G_j \in \bigcap_{j \leq n} \mathcal{F}_j$. Tal observação é relevante pois, na presença dessa propriedade, a família

$$\mathcal{H}^\uparrow := \left\{ \mathcal{F} \subseteq \mathcal{I} : \exists n \in \omega \text{ e } \mathcal{F}_0, \dots, \mathcal{F}_n \in \mathcal{H} \text{ com } \bigcap_{j \leq n} \mathcal{F}_j \subseteq \mathcal{F} \right\}$$

se revela um filtro próprio, que deve estar contido num ultrafiltro u de \mathcal{I} (pelo Teorema D.3.15). Agora, basta mostrar que a \mathcal{L} -estrutura $M_u := \prod_u M_F$ é um modelo para Σ . Pelo teorema anterior, a fim de mostrar que $M_u \models \varphi$ para algum \mathcal{L} -sentença $\varphi \in \Sigma$ fixada, basta verificar $\|\varphi\| \in u$. Ora, como $G := \{\varphi\} \in \mathcal{I}$, temos $\{F \in \mathcal{I} : \varphi \in F\} \in \mathcal{H} \subseteq u$ e, por definição, se $\varphi \in F$, então $M_F \models \varphi$, acarretando $M_F \models \varphi$, OU SEJA: $\{F \in \mathcal{I} : \varphi \in F\} \subseteq \|\varphi\|$ e, por fim, $\|\varphi\| \in u$. \square

Tanto trabalho para nada. Será?

Exemplo F.4.12 (Corpos algébricamente fechados). Um corpo L é dito **algebricamente fechado** se todo polinômio não-constante em L tem raiz em L . Corpos ordenados não têm tal propriedade (pois $x^2 - 1$ não tem raiz), assim como corpos finitos (se $L := \{l_1, \dots, l_n\}$, então $1 - \prod_{j \leq n} (x - l_j)$ não tem raiz). Um modo típico de mostrar que *todo corpo é subcorpo de um corpo algebricamente fechado* consiste em seguir os seguintes passos:

- i. prova-se que existe uma extensão K' de K tal que todo polinômio $f \in K[x] \setminus K$ tem (pelo menos) uma raiz em K' ;
- ii. por meio da etapa anterior, considera-se $L_0 := K$, $L_1 = (L_0)'$ e, mais geralmente, $L_{n+1} := (L_n)'$ para cada $n \in \omega$, para daí se definir $L := \bigcup_{n \in \omega} L_n$;
- iii. como L_{n+1} estende L_n para cada $n \in \omega$, resulta que L é um corpo que contém K ;
- iv. finalmente, se $p \in L[x]$, então existe $n \in \omega$ tal que $p \in L_n[x]$ e, pelo modo como se tomaram os corpos intermediários, segue que p tem pelo menos uma raiz em L_{n+1} e, por conseguinte, em L .

Das etapas acima, a menos trivial costuma ser a primeira, usualmente realizada por meio de um quociente esperto de um anel de polinômios em infinitas indeterminadas por um ideal maximal (garantido por AC). No presente contexto, pode-se fazer a mesma coisa de modo muito mais mirabolante. Fixando-se o corpo K e considerando a linguagem \mathcal{L}_K composta pelos símbolos usuais da linguagem de anéis juntamente com uma constante para cada elemento de K , consideram-se três *grupos* de sentenças, que serão tomadas como axiomas de uma teoria \mathcal{T} :

- (i) axiomas usuais de *um* corpo;
- (ii) axiomas estipulando que as constantes atreladas a K fazem de K um (sub) corpo; e
- (iii) para cada $p \in K[x]$, um axioma estipulando que p tem raiz³².

Ocorre que para qualquer polinômio não-constante $p \in K[x]$, *sempre* existe uma extensão L de K em que p tem raiz: basta tomar um fator irredutível q de p e considerar $L := K[x]/\text{sp}(q)$. Repetindo-se tal observação finitas vezes, mostra-se que qualquer subconjunto finito de \mathcal{L}_K tem modelo. Logo, existe um modelo de \mathcal{L}_K , i.e., uma *extensão* L/K em que todo polinômio de K tem raiz em L , como desejado. ▲

Exemplo F.4.13 (Ordens totais). O Teorema da Compacidade permite mostrar que todo conjunto admite uma ordem total: fixado um conjunto A , considera-se a linguagem \mathcal{L}_A , com o símbolo de relação binária \leq e, para cada $a \in A$, uma constante c_a ; como axiomas, impõem-se as regras usuais para ordens totais e, para cada par de elementos distintos $a, b \in A$, considera-se o axioma $\neg(c_a \approx c_b)$. Como a ordem $\langle \omega, \leq \rangle$ nos ensina, toda subcoleção finita desses axiomas admite um modelo, donde segue que existe uma ordem total M que satisfaz, simultaneamente, $c_a^M \neq c_b^M$ sempre que $a, b \in A$ são distintos. Logo, basta importar a ordem total de M por meio da injeção $a \mapsto c_a^M$. ▲

Observação F.4.14. É claro que a conclusão acima é redundante sob ZFC, já que AC garante a boa ordenação de todos os conjuntos. Contudo, acima utilizou-se apenas o Teorema da Compacidade, que é *estritamente mais fraco* do que AC. △

Exemplo F.4.15 (Análise *Não-Standard*). Considere $R_n := \mathbb{R}$ para todo $n \in \omega$, bem como um ultrafiltro \mathfrak{u} sobre ω . Que tipo de curiosidades poderia esconder o ultraproduto $\prod_{\mathfrak{u}} \mathbb{R}$? Ao se fixar a linguagem \mathcal{L}_{OF} dos corpos ordenados, segue que a \mathcal{L}_{OF} -estrutura natural de $\prod_{\mathfrak{u}} \mathbb{R}$ o torna um corpo ordenado: ora, se $\varphi \in \mathbb{F}_{\mathcal{L}_{OF}}(\mathcal{V})$ é qualquer uma das sentenças envolvidas na definição de corpo ordenado, então por termos $\omega = \{n \in \omega : \mathbb{R} \models \varphi\}$ e $\omega \in \mathfrak{u}$, resulta (pelo Teorema de Łoś) que $\prod_{\mathfrak{u}} \mathbb{R} \models \varphi$. Mais geralmente, pelo mesmo argumento, toda \mathcal{L}_{OF} -sentença satisfeita por \mathbb{R} também será satisfeita por $\prod_{\mathfrak{u}} \mathbb{R}$. Será que tal animal é a própria reta? Não necessariamente.

Com efeito, se o ultrafiltro \mathfrak{u} tomado não for *principal*³³, então $\prod_{\mathfrak{u}} \mathbb{R}$ é um corpo não-arquimediano! Com efeito, neste caso, o elemento $\varepsilon := \overline{\left\langle \frac{1}{2^n} \right\rangle_{n \in \omega}}$ é, em $\prod_{\mathfrak{u}} \mathbb{R}$, tanto maior do que 0 quanto menor do que qualquer $r \in \mathbb{R}$ fixado com $r > 0$: a primeira desigualdade segue pois $\omega = \{n \in \omega : 0 < \frac{1}{2^n}\}$, enquanto a segunda decorre da finitude do conjunto $\{n \in \omega : \frac{1}{2^n} \leq r\}$, obrigando que seu complementar pertença a \mathfrak{u} (por propriedades de ultrafiltros não-principais).

³²Um polinômio $p := \alpha_0 + \dots + \alpha_n x^n \in K[x]$ é um \mathcal{L}_K -termo τ_p com apenas uma variável livre x , o que permite escrever “ $\exists x (\tau_p \approx 0)$ ” como uma fórmula cuja interpretação *diz* que p tem raiz no modelo.

³³Isto é, \mathfrak{u} não é da forma $\{A \subseteq \mathbb{R} : x \in A\}$ para algum $x \in \mathbb{R}$ fixado.

As considerações acima são pinceladas muito modestas da (usualmente xingada de) Análise *Não-Standard*, que usa o Teorema de Łoś para formalizar a intuição Leibniziana do princípio da transferência, que permite empregar rigorosamente infinitésimos no estudo das noções de Análise. Obviamente, trata-se de um assunto que não cabe nas páginas finais de um livro introdutório de *teoria dos conjuntos*, apresentado aqui apenas como *propaganda*: o leitor interessado em aprender mais sobre ultrafiltros com ênfase em Análise Não-Standard pode conferir o recente texto de Goldbring [14]. ▲

Observação F.4.16. Para mais Teoria dos Modelos, confira as obras de Chang & Keisler [7] e Rothmaler [29], não necessariamente nesta ordem. △

F.5 Dobrando a meta

Aplicações como as que foram apresentadas ao longo do capítulo não costumam gerar incômodo por tratarem de teorias cujo escopo é, em algum sentido, *local*. A coisa começa a ficar estranha quando se pensa no seguinte: fixando-se uma linguagem dotada de um único símbolo relacional de aridade 2, digamos “ ε ”, seria possível escrever como axiomas de primeira ordem todas as sentenças que compõem a axiomática de ZFC? Se sim, não seria lícito tratar a própria *Teoria dos Conjuntos* como uma das *teorias* discutidas nas seções anteriores?

Por exemplo: podemos considerar a fórmula

$$\text{Ext} := \forall x \forall y \forall z ((z \varepsilon x \leftrightarrow z \varepsilon y) \rightarrow x \approx y).$$

Agora, um conjunto M dotado de uma relação binária $E \subseteq M \times M$ é tal que $M \models \text{Ext}$ se, e somente se, para quaisquer $a, b \in M$, a ocorrência de $a E b$ for equivalente à igualdade entre os conjuntos $\{c \in M : c E a\}$ e $\{c \in M : c E b\}$. Naturalmente, se M é um número ordinal e $E := \{\langle \alpha, \beta \rangle \in M \times M : \alpha \in \beta\}$, então $M \models \text{Ext}$, o que vale mais geralmente para qualquer conjunto transitivo. Como a abreviação sugere, Ext é o Axioma da Extensão expresso como sentença de primeira ordem – e, com alguma paciência, não é difícil perceber que todos os outros axiomas de ZFC podem ser transcritos da mesma forma.

Exercício F.6. Traduza o Axioma do Infinito como uma ε -fórmula, abreviada Inf. Mostre que ao considerar ω como uma ε -estrutura (com $\varepsilon_\omega := \{\langle m, n \rangle \in \omega \times \omega : m \in n\}$), deve-se ter $\omega \not\models \text{Inf}$. Dica: algum $n \in \omega$ é infinito? ■

Exercício F.7. Exiba um ordinal α tal que $\alpha \models \text{Inf}$. Dica: é mais fácil do que parece. ■

Exercício F.8. Traduza outros axiomas de ZFC como ε -sentenças. Depois, investigue se suas ε -estruturas favoritas (i.e., conjuntos dotados de relações binárias!) satisfazem ou não os axiomas que você traduziu³⁴. ■

Depois de dedicar algum tempo às atividades recreativas propostas acima, é inevitável se fazer a

Pergunta: existe um modelo para ZFC (onde, agora, ZFC indica os axiomas usuais de ZFC escritos como sentenças de primeira ordem)?

³⁴E o leitor deve se atentar aos sutis axiomas da separação e substituição: a rigor, tratam-se de *esquemas* de axiomas, um para cada fórmula. O contrário implicaria em escrever sentenças com fórmulas quantificadas, o que não faz parte da sintaxe de primeira ordem.

A resposta é mais delicada do que aparenta. Num primeiro momento, é tentador dizer que o universo \mathbb{V} é um modelo para ZFC. Porém, pela definição adotada, modelos *deveriam* ser conjuntos, o que impede tratar a classe própria \mathbb{V} como um modelo. Por um lado, isto é tranquilizador, já que parece deixar *nossa universo* fora do escopo da Teoria de Modelos. Por outro lado, é inquietante: será que ZFC não tem modelos?

Definição F.5.1. Diremos que Σ é (sintaticamente) **consistente**, abreviado $\text{Con}_{\vdash}(\Sigma)$, se não existir sentença $\psi \in \mathbb{F}_{\mathcal{L}}(\mathcal{V})$ tal que $\Sigma \vdash \psi \wedge \neg\psi$ (e coisas como “ $\neg\psi \wedge \psi$ ” são chamadas de *contradições*). Diremos que Σ é (semanticamente) **consistente**, abreviado $\text{Con}_{\models}(\Sigma)$, se existir uma \mathcal{L} -estrutura $\langle M, \mathcal{I} \rangle$ com $M \models \Sigma$.

As duas noções de consistência acima podem chamar a atenção por sua aparente assimetria. Entretanto, o contexto em que nossos modelos são tomados torna irrelevante definir a consistência semântica de modo análogo ao caso sintático: com efeito, não existem sentença $\psi \in \mathbb{F}_{\mathcal{L}}(\mathcal{V})$ e \mathcal{L} -estrutura M com $M \models \psi \wedge \neg\psi$ (Exercício F.19). Logo, se $M \models \Sigma$, i.e., se $\text{Con}_{\models}(\Sigma)$, então $\text{Con}_{\vdash}(\Sigma)$: o contrário levaria a concluir, pela corretude do sistema dedutivo, que $\Sigma \models \psi \wedge \neg\psi$ para alguma sentença ψ . É a completude (assumida) do sistema dedutivo que manifesta a recíproca desta implicação, no seguinte sentido:

Proposição F.5.2. *Para uma \mathcal{L} -teoria Σ , as seguintes afirmações são equivalentes:*

- (i) $\text{Con}_{\vdash}(\Sigma) \Rightarrow \text{Con}_{\models}(\Sigma)$;
- (ii) *o sistema dedutivo é completo, i.e., se $\Sigma \models \varphi$, então $\Sigma \vdash \varphi$.*

Esboço da prova. Uma argumentação honesta depende do sistema dedutivo considerado, razão pela qual se apresenta apenas um esboço. Primeiro, se vale a afirmação (i) e $\Sigma \models \varphi$, então $\Sigma \cup \{\neg\varphi\}$ não tem modelo, i.e., é semanticamente *inconsistente*. Logo, por (i), $\Sigma \cup \{\neg\varphi\}$ é sintaticamente inconsistente, o que significa (via “provas por absurdo”), que $\Sigma \vdash \varphi$. Reciprocamente, se valem (ii) e $\text{Con}_{\vdash}(\Sigma)$, então Σ tem modelo: se $\Sigma = \emptyset$, nada há a fazer; se não, então para $\varphi \in \Sigma$ tem-se $\Gamma \cup \{\varphi\}$ sintaticamente consistente, onde $\Gamma := \Sigma \setminus \{\varphi\}$; logo, $\Gamma \not\models \neg\varphi$, donde (ii) acarreta que nem todo modelo de Γ satisfaz $\neg\varphi$ ou, em outras palavras, existe um modelo para $\Gamma \cup \{\varphi\} = \Sigma$. \square

Corolário F.5.3 (Compacidade, versão sintática). *Uma \mathcal{L} -teoria Σ é sintaticamente consistente se, e somente se, todo subconjunto finito de Σ é sintaticamente consistente.*

Demonstração. Pelo Teorema F.4.5 (da compacidade)³⁵, $\text{Con}_{\models}(\Sigma)$ ocorre se, e somente se, $\text{Con}_{\models}(F)$ ocorre para cada subconjunto finito $F \subseteq \Sigma$. Logo, o resultado desejado segue da proposição anterior, segundo a qual $\text{Con}_{\vdash}(\Sigma)$ é equivalente a $\text{Con}_{\models}(\Sigma)$. \square

Podemos voltar à última pergunta: será que ZFC não tem modelos?

Um olhar desatento para a última proposição poderia trazer certo pavor diante da possibilidade de uma resposta “sim”: *se ZFC não tem modelos, então ZFC tem contradições?!* No entanto, trata-se apenas de uma interpretação equivocada: a rigor, tanto as definições de consistência quanto a última proposição foram *implementadas* em \mathbb{V} , onde *tudo acontece* e, por conseguinte, o escopo de seus resultados comprehende apenas estruturas e teorias definidas em \mathbb{V} , i.e., por meio de conjuntos. Em outras palavras: o universo \mathbb{V} , e tudo o que sabemos acerca de \mathbb{V} , constituem a *metateoria* em que as discussões anteriores foram realizadas.

³⁵Esta versão sintática costuma ser provada diretamente na presença de um sistema dedutivo explícito, por meio de um argumento bastante simples: se Σ é sintaticamente inconsistente, então as sentenças utilizadas na dedução de uma contradição constituem o subconjunto finito que será inconsistente!

O sufixo “meta”, frequentemente utilizado com as expressões “linguagem” e “teoria”, é usado para designar, respectivamente:

- ✓ a linguagem utilizada na *discussão* de *outra* linguagem (a *linguagem objeto*);
- ✓ a teoria utilizada para tirar conclusões acerca de *outra* teoria (a *teoria objeto*).

Em contextos elementares de Álgebra, Análise e Geometria, por exemplo, conjuntos costumam ser utilizados como a *base ontológica* da discussão, no seguinte sentido: os objetos estudados por tais Teorias são conjuntos com certas propriedades e estruturas adicionais, que por sua vez também são conjuntos. Dessa forma, conjuntos constituem a metalinguagem formal³⁶, enquanto os resultados sobre conjuntos utilizados nas argumentações (e.g., Lema de Zorn, Teorema da Recursão, etc.) constituem a metateoria.

Exemplo F.5.4. Ao construir o corpo de frações de \mathbb{Z} como um quociente do conjunto $\mathbb{Z} \times (\mathbb{Z} \setminus \{0\})$, apela-se para diversos *aparatos* oriundos da Teoria dos Conjuntos. Assim, enquanto a afirmação “existe um corpo de frações de \mathbb{Z} ” pode ser considerada como um resultado da teoria (objeto) dos anéis, uma argumentação como a sugerida pode ser considerada metateórica por usar recursos da metateoria. Por outro lado, a afirmação “todo elemento de um anel tem um único inverso aditivo” se faz tão somente por meio dos axiomas de anéis, i.e., na linguagem (objeto) dos anéis. ▲

Nada disso causa dor de cabeça, pois é bastante clara a distinção entre teoria/linguagem objeto e metateoria/metalinguagem nos casos anteriores. A situação muda, porém, quando o foco da investigação passa a ser a metateoria; no caso, quando os objetos de estudo são conjuntos. A rigor, no *momento* em que a *metateoria* passa a ser *objeto* de investigação, ela automaticamente se torna uma *teoria objeto* analisada por alguma *metametateoria*!

Nessa perspectiva particular, há duas posturas relativamente comuns que merecem ser mencionadas.

- ✓ A postura *platonista* supõe que existe um universo *real* e independente, mas possivelmente intangível, de objetos matemáticos sobre os quais as discussões matemáticas ocorrem, de modo que coisas como ZFC (ou outras axiomatizações para conjuntos) apenas tentam descrever os objetos desse universo abstrato. Isto resolve o problema da metateoria de modo fulminante, ao supor um substrato quase palpável para as ponderações abstratas realizadas: \mathbb{V} é o universo matemático platônico.
- ✓ A postura *formalista* supõe que tudo se resume a símbolos escritos, sem qualquer comprometimento com alguma realidade *adjacente*. Em certo sentido, ZFC (ou outra axiomatização para conjuntos) consiste num jogo de manipulação simbólica. Não existem conjuntos, nem funções, nem cardinais ou coisas do tipo: existem apenas símbolos que xingamos com tais nomes a depender de como eles são descritos pelas regras do jogo.

As duas posturas acima não encerram o leque de atitudes *defensáveis* dentro da Filosofia da Matemática para as questões ontológicas – tampouco são imunes a críticas. Porém, para um texto que não tem por objetivo último discutir tais problemas, é suficiente considerá-las como eixos norteadores³⁷.

³⁶Enquanto a língua de quem escreve (português, inglês, francês, etc.) poderia ser entendida como a *metametalinguagem* ou, ao ser misturada com os jargões informais de conjuntos, apenas como a *metalinguagem informal*.

³⁷O leitor com inclinações filosóficas leves pode se aprofundar com [17]. Para o caso de inclinações mais graves, talvez seja melhor usar [32].

Com os ânimos acalmados, convém reformular a pergunta original.

Pergunta: existem um conjunto M e uma relação binária E em M tal que $M \models \text{ZFC}$?

Note que nenhuma das posturas anteriores inviabiliza a pergunta *a priori*, bem como não trivializa suas respostas: no caso platonista, a riqueza inesgotável de \mathbb{V} poderia muito bem prover um objeto $\mathbb{U} \in \mathbb{V}$ que imitasse \mathbb{V} , no sentido de satisfazer ZFC, mas também poderia ser o caso de tal riqueza ser irreplicável, no sentido de não existir objeto como \mathbb{U} ; no caso formalista, *nada existe* no sentido usual, de modo que a existência de um modelo para ZFC se traduz na *dedução* de que um símbolo particular satisfaz os critérios para ser xingado de modelo para ZFC, enquanto a inexistência consiste numa prova de que não existe tal símbolo.

Antes de trazer mais ingredientes para a discussão, vamos supor que existe um modelo para ZFC. Neste caso, do ponto de vista dos habitantes de M , conjuntos são todos os elementos de M . Em particular, “ $M \notin M$ ” apenas reflete o já conhecido resultado de que o universo não é um conjunto. Mas M é um conjunto, não? Para *nós*, sim, mas para os habitantes de M , não: pode ser elucidativo dizer que os elementos de M são M -conjuntos, e que ao repetir em M a prova de que não existe conjunto universo, prova-se (do nosso ponto de vista) que M não é um M -conjunto. Tudo parece inofensivo... até agora.

Exemplo F.5.5. Ao tomar \mathcal{L} como a linguagem dos anéis e Γ como os axiomas que *descrevem* os corpos, considere φ a \mathcal{L} -sentença que expressa a asserção “existe x tal que $x \cdot x = -1$ ”. Pergunta-se: $\Gamma \vdash \varphi$? Ora, com a dobradinha “corretude+completude” dos sistemas dedutivos, a pergunta equivale a investigar o *status* de $\Gamma \models \varphi$.

Por um lado, $\Gamma \not\models \varphi$, já que existem corpos (e.g., \mathbb{Q}) que não verificam φ , i.e., nos quais não existe elemento x satisfazendo $x^2 = -1$. Todavia, isto não significa que $\Gamma \models \neg\varphi$! Com efeito, o corpo \mathbb{C} dos **números complexos** possui um elemento, usualmente xingado de i , cujo quadrado é -1 . Logo, $\Gamma \not\models \neg\varphi$. ▲

Definição F.5.6. Uma \mathcal{L} -teoria Σ é **incompleta** se existe ao menos uma \mathcal{L} -sentença φ tal que $\Sigma \not\models \varphi$ e $\Sigma \not\models \neg\varphi$ (equivalentemente: $\Sigma \not\models \varphi$ e $\Sigma \not\models \neg\varphi$), caso em que se diz que a sentença φ é **independente** de Σ . Teorias **completas** são aquelas que não são incompletas. ¶

Assim, o exemplo anterior mostra que Γ é uma teoria incompleta. Por sua vez, fixada uma \mathcal{L} -estrutura M , a teoria $\Sigma_M := \{\varphi \in \mathbb{F}_{\mathcal{L}}(\mathcal{V}) : M \models \varphi\}$ é completa (Exercício F.20). Finalmente, entram em cena as últimas peças de nossa discussão, enunciadas de forma bem pouco rigorosa em razão do *adiantado da hora*.

Primeira Incompletude de Gödel³⁸. *Uma teoria rica o suficiente, procedural e sintaticamente consistente é incompleta.*

Segunda Incompletude de Gödel. *Uma teoria nas condições anteriores, e capaz de expressar uma fórmula Con_\vdash para a consistência da teoria, não demonstra Con_\vdash .*

Acima, o “rica o suficiente” joga para debaixo do tapete diversas exigências técnicas que fogem do escopo deste texto, mas que, grosso modo, podem ser pensadas como “desenvolver aritmética e um pouco de recursão”. Já o “procedural” se refere à *descritibilidade* da teoria: em certo sentido, seus axiomas devem ser descritíveis de alguma forma algorítmica (diferente da teoria Σ_M anterior, por exemplo, descrita de forma intencional). E, porém, na segunda noção de incompletude que se revela o grande problema.

³⁸Cabe destacar: o *Teorema da Completude de Gödel* estabelece a completude de sistemas dedutivos; já os teoremas de *Incompletude* de Gödel tratam da (não) completude de teorias.

Um pouco mais detalhadamente, mas nem tanto, a grande sacada de Gödel foi utilizar a aritmética da teoria para descrever/codificar as fórmulas da teoria (por meio de uma bijeção explícita entre ω e a coleção de fórmulas) e, assim, permitir que a teoria discursasse sobre suas próprias fórmulas. No caso da segunda noção de incompletude, em sua formulação original (ou quase), Con_{\models} codifica a afirmação “não existe número que seja a codificação da prova de uma contradição”, de modo que o restante da demonstração se torna um argumento de ponto fixo.

Tais meandros são necessários por conta das poucas exigências impostas sobre a teoria, que essencialmente dá conta de descrever aritmética e mais nada. No caso de uma metateoria robusta em que consistência significa existência de modelos (como a nossa, por exemplo, em vista da Proposição F.5.2!), é possível argumentar diretamente e sem (muitos) rodeios.

Teorema F.5.7 (Segunda Incompletude de Gödel, para conjuntos). *Se $\text{ZFC} \vdash \text{Con}_{\models}(\text{ZFC})$, então ZFC não admite um modelo.*

Demonstração (com abanação das mãos). Seja $\langle V_0, E_0 \rangle$ um modelo para ZFC. Note que se $\text{ZFC} \vdash \text{Con}_{\models}(\text{ZFC})$, então por $V_0 \models \text{ZFC}$, segue que existem $V_1 \in V_0$ e uma relação binária E_1 sobre V_1 tal que $\langle V_1, E_1 \rangle \models \text{ZFC}$, bem como existem $V_2 \in V_1$ com uma relação binária... recursivamente, isto resultaria numa sequência $\langle V_n \rangle_{n \in \omega}$ que violaria o Exercício A.35, mostrando que $\langle V_0, E_0 \rangle$ não modela ZFC. \square

Exercício F.9. A “demonstração” acima, adaptada de [16], tem alguns erros. Encontre pelo menos um deles. ■

Demonstração. Primeiro, observe que se ZFC demonstra $\text{Con}_{\models}(\text{ZFC})$, então existe um subconjunto finito $\Sigma \subsetneq \text{ZFC}$ capaz de definir as noções de modelo e satisfabilidade, bem como conter as eventuais (e finitas!) instâncias dos esquemas de axiomas (substituição e separação) necessários para os próximos argumentos e, mais importante de tudo, com $\Sigma \vdash \text{Con}_{\models}(\text{ZFC})$.

Agora, para $\langle M, \varepsilon^M \rangle$ e $\langle N, \varepsilon^N \rangle$ modelos de Σ , vamos escrever “ $M < N$ ” para indicar a existência de $\langle X, \varepsilon^X \rangle \in N$ satisfazendo $\varepsilon^M = (\varepsilon^X)^N$, onde $(\varepsilon^X)^N := \{\langle x, y \rangle : N \models x \varepsilon^X y\}$, com ε^X uma relação binária em X . Intuitivamente, “ $M < N$ ” busca expressar que M é aquilo que N pensa que X é. Convém destacar algumas coisas com calma:

- (i) se $M < N$, então para uma ε -sentença σ qualquer, deve-se ter $M \models \sigma$ se, e somente se, $N \models (X \models \sigma)$, o que segue diretamente da definição de “ $<$ ”;
- (ii) em particular, como $M \models \Sigma$ e Σ é finito, resulta que se $M < N$ e X é como no item anterior, então $N \models (X \models \Sigma)$, o que se pode abbreviar com $N \models (M \models \Sigma)$;
- (iii) vale a recíproca do item anterior, no sentido de que se existe $X \in N$ tal que $N \models (X \models \Sigma)$, então com $X := M$, $\langle M, (\varepsilon^X)^N \rangle$ é um modelo para Σ tal que $M < N$;
- (iv) consequentemente, se $M_0 < M_1$ e $M_1 < M_2$, então $M_0 < M_2$.

Fixemos então uma bijeção G entre ω e a coleção \mathcal{S} de todas as fórmulas da teoria dos conjuntos, o que pode ser feito já que o conjunto de tais sentenças é infinito enumerável. Para $n \in \omega$, seja $S_n := \{m \in \omega : \Sigma \models G(n)[m]\}$, i.e., o conjunto dos números naturais que satisfazem a n -ésima fórmula segundo a bijeção G . A penúltima peça do mecanismo é o conjunto $S := \{n \in \omega : \exists M (M \models n \notin S_n)\}$.

Note que a menos de expandir todas as definições, a fórmula usada para descrever o conjunto S pertence à coleção \mathcal{S} . Chamando-a de ψ , segue que existe $k \in \omega$ com $G(k) = \psi$. O detalhe final da *mise en place* é perceber que existe uma fórmula τ em \mathcal{S} que expressa “ $k \in S$ ”. Prossigamos com a lista de observações.

- (v) Para qualquer modelo N de Σ , deve-se ter

$$N \models \tau \Leftrightarrow \exists M (M < N \wedge M \models \neg\tau).$$

Por um lado, se $N \models \tau$, então N “acredita que k pertence a S ”, ou seja, existe $\langle X, \varepsilon^X \rangle \in N$ com $\langle X, \varepsilon^X \rangle \models \neg\tau$, o que permite obter um modelo $\langle M, \varepsilon^M \rangle$ para Σ com $M < N$ e $M \models \neg\tau$ (como em (iii)); por outro lado, se existe tal M , então a definição de “ $<$ ” assegura que existe $\langle X, \varepsilon^X \rangle \in N$ com $X \models \neg\tau$, justamente o que significa $k \in S$ (em N).

- (vi) Por fim, xingando de *positivos* os modelos que satisfazem τ e de *negativos* os demais, resulta que se M for negativo, então todo N com $N < M$ é positivo (basta encarar a equivalência anterior até que ela te encare de volta).

Destacadas todas as afirmações acima, a demonstração pode começar. Suponha que ZFC tenha um modelo, digamos M_0 , que em particular é modelo para Σ . O *Katzen-prung* consiste em definir um modelo negativo M_1 *esperto* para Σ :

- (+) se M_0 for positivo, então a afirmação (v) assegura algum modelo $M_1 < M_0$ negativo;
- (-) se M_0 for negativo, façamos $M_1 := M_0$.

Dado que $\Sigma \vdash \text{Con}_{\models}(\text{ZFC})$ e $M_1 \models \Sigma$, existe M_2 com $M_2 < M_1$ (como em (iii)) e, por conseguinte, a afirmação (vi) garante que M_2 é positivo. Ao aplicar (v) novamente, desta vez com respeito a M_2 , obtém-se M_3 negativo com $M_3 < M_2$. No entanto, por (iv), ocorre $M_3 < M_0$, contrariando a negatividade de M_0 (em virtude de (vi)). \square

Observação F.5.8. A delicadeza na definição da relação $M < N$ se deve ao fato de que ao expressar algo como “ $N \models (M \models \Sigma)$ ”, não há razões para supor que a relação binária em M seja o que N acredita ser a pertinência “real”. Nesse sentido, talvez fique mais fácil perceber o que havia de errado com a primeira demonstração apresentada. O leitor interessado nos diversos detalhes omitidos na última demonstração pode conferir os textos de Jech [20] e Bagaria [1], que foram as referências utilizadas para compor a exposição anterior. \triangle

Corolário F.5.9. $\text{Con}_{\models}(\text{ZFC}) \Rightarrow \text{ZFC} \not\models \text{Con}_{\models}(\text{ZFC})$.

De forma mais verbal: se ZFC é sintaticamente consistente, então ZFC *não pode* provar isso! Portanto, na *prática*, fica respondida a primeira pergunta: se ZFC for consistente, então ZFC não prova que existe um modelo para ZFC³⁹. Apesar de parecer uma constatação desesperadora, ela se revela libertadora: ao assumir que ZFC é consistente, torna-se lícito *supor* que existe um modelo para ZFC, ou quase isso.

³⁹E se ZFC for inconsistente... ora, neste caso, teríamos problemas bem mais sérios do que modelos de ZFC para nos preocuparmos – mesmo assim, como toda boa teoria inconsistente, ZFC provaria qualquer coisa, inclusive a existência de modelos.

A rigor, acrescenta-se um símbolo de constante à linguagem $\langle \varepsilon, \langle \varepsilon, 2 \rangle \rangle$ de ZFC, digamos \mathbb{M} e, além dos axiomas usuais de ZFC, adicionam-se sentenças que expressem a validade dos axiomas “originais” quando *interpretados* em \mathbb{M} . Por exemplo: além da sentença

$$\text{Pair} := \forall x \forall y \exists z \forall w (w \in z \rightarrow (w \approx x \vee w \approx y))$$

conhecida como Axioma do Par, acrescenta-se

$$\text{Pair}' := \forall x \forall y ((x \in \mathbb{M} \wedge y \in \mathbb{M}) \rightarrow \exists z (z \in \mathbb{M} \wedge \forall w (w \in z \rightarrow (w \approx x \vee w \approx y)))),$$

que ao ser interpretada num modelo V , *diz* essencialmente que o Axioma do Par se verifica quando restrito aos “elementos” do elemento $\mathbb{M}_V \in V$ que interpreta a constante \mathbb{M} .

Ao xingar de ZFC^+ o resultado dessa brincadeira, chega-se a $\text{ZFC}^+ \vdash \text{Con}_{\models}(\text{ZFC})$, pois os axiomas de ZFC^+ impõem, explicitamente, que a constante \mathbb{M} se comporte como um modelo para ZFC. Porém, isto não significa que ZFC^+ provou sua própria consistência, já que os “axiomas” satisfeitos por \mathbb{M} não *prescrevem* que o próprio tenha um modelo para ZFC! Mais importante: se assumirmos que ZFC é consistente (sintaticamente, digamos), então ZFC^+ também será, já que o contrário daria uma demonstração em ZFC^+ para uma contradição que, por sua vez, poderia ser convertida numa demonstração em ZFC para uma contradição. Para mais detalhes, o leitor pode conferir [35].

Portanto, ao usar ZFC^+ como metateoria, os resultados de Teoria dos Modelos ficam disponíveis para serem aplicados sobre o modelo \mathbb{M} de ZFC – e de maneira completamente honesta. Consequentemente, afirmações do tipo “se \mathbb{M} é um modelo para ZFC, então existe um modelo \mathbb{M}' para $\text{ZFC}+\neg\text{CH}$ ” passam a fazer mais sentido do que antes, quando ingenuamente tratávamos “ser modelo para ZFC” como sinônimo para “ser o universo”: ainda hoje, não é incomum encontrar pessoas que consideram resultados sobre consistência relativa de teorias como epifanias, revelações de uma realidade transcendental grandiosa⁴⁰. No entanto, quem comprehende o conteúdo das afirmações sabe que se tratam de miudezas tão corriqueiras quanto, digamos, o fato de \mathbb{C} poder ser obtido a partir de \mathbb{R} por meio da adjunção de uma indeterminada e posterior quociente por um ideal apropriado.

Exemplo F.5.10 (*Forcing* e CH). Grosso modo, pode-se dizer que o *forcing* consiste numa técnica que converte modelos de $\text{ZF}(\mathbb{C})$ em modelos de $\text{ZF}(\mathbb{C})+\varphi$, para certas sentenças φ na linguagem da teoria dos conjuntos⁴¹. Em particular, nas situações em que se consegue fazer isso tanto para φ quanto para $\neg\varphi$, chega-se a uma conclusão inevitável: a fórmula φ é independente de ZFC (lembre-se: tanto $\text{ZFC} \not\models \varphi$ quanto $\text{ZFC} \not\models \neg\varphi$ ocorrem). Dessa forma, a menos do escopo de aplicabilidade das teorias consideradas, resultados de independência não deveriam ser mais chocantes do que o Exemplo F.5.5. ▲

⁴⁰É digno de nota que tal comportamento, frequentemente, vem acompanhado de uma pseudo-reverência aliada à autodepreciação, como em “Isto aí é Filosofia, belíssimo, mas profundo demais! É melhor manter os pés no chão e fazer...”, onde as reticências costumam ser preenchidas por áreas típicas da Matemática Abstrata contemporânea. Porém, na prática (no fim do dia, no contar das bolsas, etc.), não passa de uma justificativa perfumada para desestimular o estudo da Lógica-Matemática como se esta tratasse apenas de assuntos incompreensíveis ou inúteis.

⁴¹Um dos primeiros usos de tal técnica, desenvolvida por Paul Cohen no século passado, obteve modelos para $\text{ZF}+\neg\text{AC}$ e $\text{ZFC}+\neg\text{CH}$ a partir de modelos para ZF e ZFC, respectivamente, mostrando assim que $\text{ZF} \not\models \text{AC}$ e $\text{ZFC} \not\models \text{CH}$. Algumas décadas antes, Gödel já havia mostrado (por meio de outro tipo de argumentação) como obter um modelo para $\text{ZFC}+\text{CH}$ a partir de um modelo para ZF, mostrando assim que $\text{ZF} \not\models \neg\text{AC}$ e $\text{ZF} \not\models \neg\text{CH}$. Concluiu-se, assim, que AC e CH são independentes de ZF e ZFC, respectivamente.

Uma abordagem bastante semelhante é discutida na edição mais recente do *Mathematical Logic* [9], de Ebbinghaus, Flum & Thomas. Após *defenderem* a eficácia da *lógica de primeira ordem*⁴² para abstrair as diversas *metodologias matemáticas* por meio de descrições conjuntistas, os autores sugerem que a utilização de (alguma lista de) axiomas para conjuntos seja feita em dois níveis explicitamente distintos.

- (i) No nível metateórico, tais axiomas descrevem o *verdadeiro* universo do discurso, que os autores chamam de *teoria dos conjuntos de fundo* (*background set theory*), cujo propósito é servir como fundamento ontológico para qualquer discussão teórica *objetiva* posterior. Explicitamente: linguagens, sentenças, conjuntos, funções, estruturas para linguagens, etc. são, invariavelmente, objetos desse universo verdadeiro que não discutimos, mas apenas usamos de acordo com os axiomas fixados.
- (ii) No nível teórico, os mesmos axiomas podem ser usados na análise de estruturas particulares, assim como os axiomas para grupos e anéis em Álgebra. Em tal situação, os axiomas descrevem a *teoria-objeto dos conjuntos*: nesse sentido, um modelo para tais axiomas é apenas um objeto do universo do discurso (que existe e é descrito pela teoria dos conjuntos de fundo!) que *interpreta* os axiomas considerados para conjuntos; em particular, se M é um tal modelo que verifica $a \in_M b$, isto não significa que a seja *verdadeiramente* elemento de b (isto é, no universo do discurso), mas quer dizer apenas que a interpretação que M dá para o símbolo \in faz com que M acredite em tal pertinência (confira os Exercícios F.21 e F.22).

Outra alternativa, sugerida por Kunen [23], é a de “fundamentar” a Matemática *duas vezes*: primeiro, desenvolve-se ZFC como o que se fez nos primeiros capítulos; depois, realizam-se as discussões sobre Teoria de Modelos e Teoria da Prova; por fim, repete-se o primeiro passo, numa nova camada. Em certo sentido, desenvolve-se a teoria objeto até o ponto em que ela se torna robusta o suficiente para servir como metateoria para a discussão de uma nova teoria objeto que espelha a anterior⁴³. Numa linha parecida, utilizada por Halbeisen & Krapf [15], desenvolve-se primeiro uma teoria da prova e uma teoria de modelos “finitista”, para que só depois se discutam ZFC e todo o resto: neste caso, as menções a conjuntos em Teoria da Prova e Teoria dos Modelos ficam restritas a coleções finitas (ou, na pior das hipóteses, potencialmente infinitas), quase como se fossem meras figuras de linguagem para sequências de símbolos. Para aprofundar as discussões, o leitor pode conferir, além das referências supracitadas, os textos de Kunen [22, 24] e Fraenkel et al. [12].

Em última análise, cabe o alerta típico que se faz em *Modelagem Matemática* quando se analisam certos fenômenos naturais: os modelos de tais fenômenos apenas *modelam* os fenômenos – eles não *são* os fenômenos. Analogamente, o que se fez acima não descreveu a Matemática e seu universo ontológico como eles são, mas apenas como *poderiam ser*: são *modelos* que buscam descrever a Matemática e suas metodologias, que existem fora de tais modelos, sejam como entidades num universo platônico ou como fruto da engenhosidade humana. Nesse sentido, a impressão de circularidade sempre estará presente quando tentarmos justificar as bases escolhidas para fundamentar nossas justificativas: invariavelmente, chega um momento *na vida* em que a única resposta para um eventual “por quê?” é “sim”. Equivalentemente:

ex nihilo nihil fit.

⁴²Isto é, o combo “sentenças de primeira ordem” + “um bom sistema dedutivo”.

⁴³É quase como se a Matemática decidisse fazer terapia.

Exercícios adicionais

Exercício F.10. Um **reticulado** consiste de uma ordem parcial \mathbb{P} em que para quaisquer $a, b \in \mathbb{P}$ existem $\sup\{a, b\}$ e $\inf\{a, b\}$.

1. Chamando $a \wedge b := \inf\{a, b\}$ e $a \vee b := \sup\{a, b\}$, mostre que para quaisquer $a, b \in A$ valem as identidades $a \vee (a \wedge b) = a$ e $a \wedge (a \vee b) = a$.
2. Seja L um conjunto dotado de duas operações binárias, associativas e comutativas, \vee e \wedge , satisfazendo as identidades do item anterior. Escrevendo $l \leq p$ para indicar $l = p \wedge l$, mostre que $\langle L, \leq \rangle$ é um reticulado.
3. Seja $\Omega_{lat}: \{\wedge, \vee\} \rightarrow \omega$ a assinatura que faz $\Omega(\wedge) := \Omega(\vee) := 2$. Para reticulados \mathbb{P} e \mathbb{Q} , mostre que todo Ω_{lat} -morfismo $\mathbb{P} \rightarrow \mathbb{Q}$ é uma função crescente. Vale a volta? ■

Exercício F.11. Sejam \mathcal{L} uma linguagem algébrica e A um \mathcal{L} -álgebra. Mostre que A^A , i.e., o conjunto das funções da forma $A \rightarrow A$, pode ser munido de uma \mathcal{L} -estrutura que faz dele uma \mathcal{L} -álgebra. ■

Exercício F.12. Sejam \mathcal{L}_{group} a linguagem dos grupos (Exemplo F.3.27), $A := \omega$, $e_\omega := 0$, $\text{inv}_\omega := -$ e $*_\omega := +$. Para $\mathcal{V} := \{x, y\}$, mostre que os termos $*(x, y)$ e $*(y, x)$ têm interpretações idênticas em A . ■

Exercício F.13. Mostre que se A e B são \mathcal{L} -álgebras pertencentes a uma classe equacional \mathcal{K} , então $A \times B$ (com a estrutura do Exemplo F.1.10) também pertence a \mathcal{K} . Conclua que a classe dos corpos não é equacional. Caso se lembre da definição de *domínio* (de integridade), mostre que a classe composta por eles também não é equacional. ■

Exercício F.14. Sejam G um grupo, $C \subseteq G \times G$ e $H \subseteq G$ subconjuntos.

- a) Mostre que C é um congruência (relação de equivalência em G e subgrupo de G) se, e somente se, $N_C := \{ab^{-1} : (a, b) \in C\}$ é subgrupo normal de G .
- b) Mostre que H é subgrupo normal de G se, e somente se, $\sim_H := \{(a, b) \in G \times G : ab^{-1} \in H\}$ é uma congruência.
- c) Conclua que os quocientes de álgebra universal generalizam os quocientes usuais em teoria de grupos, anéis, etc. ■

Exercício F.15. Adapte o exercício anterior, trocando G por um anel e subgrupos por ideais. ■

Exercício F.16. Seja \mathcal{L} uma linguagem algébrica e Σ uma família de identidades. Mostre que se \mathcal{K} é a classe das álgebras que satisfazem Σ , então \mathcal{K} é fechada por subálgebras. ■

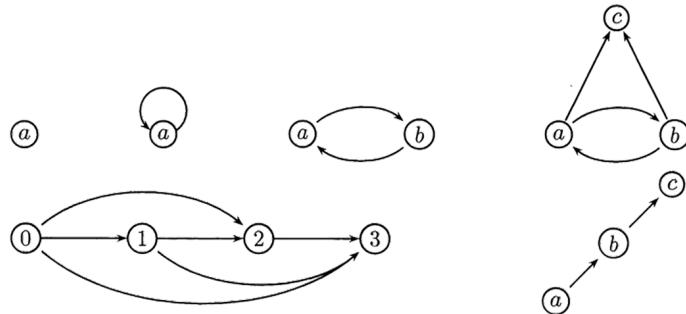
Exercício F.17. Mostre que subálgebras de grupos, anéis, módulos, etc. são também grupos, anéis e módulos, respectivamente. ■

Exercício F.18. Mostre que todo grupo é o quociente de um grupo livre por algum subgrupo. Generalize para álgebras em classes equacionais. Dica: para todo conjunto X , existe o grupo livre $G(X)$ sobre X caracterizado pela propriedade universal; no caso em que o próprio X for um grupo, tente encontrar um morfismo de grupos sobrejetor de $G(X)$ em X . ■

Exercício F.19. Para uma \mathcal{L} -estrutura M e uma \mathcal{L} -fórmula φ , mostre que $M \not\models \varphi \wedge \neg\varphi$. Dica: indução na complexidade de φ . ■

Exercício F.20. Mostre que se M é uma \mathcal{L} -estrutura, então $\Sigma_M := \{\varphi \in \mathbb{F}_{\mathcal{L}}(\mathcal{V}) : M \models \varphi\}$ é uma teoria completa. ■

Exercício F.21. Cada grafo a seguir representa uma relação binária: uma seta do tipo $x \rightarrow y$ indica $x R y$.



Considere então os ε -axiomas:

$$\begin{aligned}\varphi_0 &:= \forall x \forall y \forall z ((z \in x \leftrightarrow z \in y) \rightarrow x \approx y); \\ \varphi_1 &:= \forall x \exists y (y \in x) \rightarrow \exists z (z \in x \wedge \neg (\exists w (w \in x \wedge w \in z))); \\ \varphi_2 &:= \forall x \forall y \exists z \forall w (w \in z \rightarrow (w \approx x \vee w \approx y)); \\ \varphi_3 &:= \forall x \exists a \forall y \forall z (z \in y \wedge y \in x \rightarrow z \in a).\end{aligned}$$

Quais axiomas acima são satisfeitos pelos grafos anteriores? Quais não são? Por quê? Você reconhece tais sentenças? Dica: leia " $x \rightarrow y$ " como " $x \in y$ ". ■

Exercício F.22. Seja $V := \omega$.

- Para $m, n \in \omega$, escreva $m \varepsilon_V n$ se m for divisor de n . Quais axiomas de ZFC são satisfeitos pela estrutura $\langle V, \varepsilon_V \rangle$? Dica: a ideia é que os *elementos* de um *conjunto* n são os seus divisores.
- Para $m, n \in \omega$, escreva $m \varepsilon_V n$ se na representação binária de n ocorrer o número 1 na m -ésima posição, contando-se a partir da 0-ésima posição e pela direita. Quais axiomas de ZFC são satisfeitos pela estrutura $\langle V, \varepsilon_V \rangle$? Dica: como a representação de 13 em base 2 é 1101, tem-se $0 \varepsilon_V 13$, $1 \not\varepsilon_V 13$, $2 \varepsilon_V 13$, $3 \varepsilon_V 13$ e $4 \not\varepsilon_V 13$; na prática, pode ser mais interessante substituir V por $\{0, 1\}^{<\omega}$. ■

Exercício F.23 (Axiomas de Peano, de primeira ordem). Considere a linguagem $\mathcal{P} := \langle +, \cdot, \sigma, 0 \rangle$ composta exclusivamente por símbolos de operação, com aridades 2, 2, 1 e 0, respectivamente.

- Transcreva sentenças de primeira ordem cujas interpretações numa \mathcal{P} -estrutura \mathcal{N} sejam:
 - $\sigma_{\mathcal{N}}$ é injetiva;
 - $0_{\mathcal{N}} \notin \text{im}(\sigma_{\mathcal{N}})$.
- Para uma \mathcal{P} -fórmula⁴⁴ φ com uma variável livre x , transcreva uma sentença de primeira ordem cuja interpretação numa \mathcal{P} -estrutura seja “se \mathcal{N} satisfaz $\varphi(0_{\mathcal{N}})$ e \mathcal{N} satisfaz $\varphi(\sigma_{\mathcal{N}}(n))$ sempre que \mathcal{N} satisfaz $\varphi(n)$, então \mathcal{N} satisfaz $\varphi(x)$ para todo x em \mathcal{N} ”.
- As sentenças do item (a), juntamente com o esquema de sentenças do item (b), constituem a parte *principal*⁴⁵ da *Axiomática (de 1ª ordem) de Peano para Aritmética*. Convença-se de que tais axiomas admitem um modelo. Dica: considere ω com as operações usuais.

⁴⁴Com respeito a um conjunto de variáveis \mathcal{V} infinito enumerável.

⁴⁵Além de outras quatro que ditam o comportamento operatório: “ $\forall x (x + 0 \simeq x)$ ”, “ $\forall x (x \cdot 0 \simeq 0)$ ”, “ $\forall x \forall y (x + \sigma(y) \simeq \sigma(x + y))$ ” e “ $\forall x \forall y (x \cdot \sigma(y) \simeq (x \cdot y) + x)$ ”.

- d) Prove que existe uma \mathcal{P} -estrutura *elementarmente equivalente*⁴⁶ a ω mas que não é isomorfa a ω . Dica: basta considerar⁴⁷ Ψ o conjunto de todas as sentenças satisfeitas por ω , e daí utilizar $\Psi' := \Psi \cup \{\neg(x \approx n) : n \in \omega\}$; como cada subconjunto finito de Ψ' tem modelo, o Teorema da Compacidade garante um modelo para Ψ' . ■

Exercício F.24. Com as notações anteriores, seja \mathcal{N} uma $\{\sigma, 0\}$ -estrutura satisfazendo as sentenças do item (a), além da seguinte *afirmação*: *para todo subconjunto $X \subseteq \mathcal{N}$, se $0_{\mathcal{N}} \in X$ e $\sigma_{\mathcal{N}}(x) \in X$ sempre que $x \in X$, então $X = \mathcal{N}$.*

- a) Mostre que \mathcal{N} admite uma boa ordem com *order type* ω .
- b) Mostre que se \mathcal{N}' é outra $\{\sigma, 0\}$ -estrutura com as mesmas propriedades, então existe um único $\{\sigma, 0\}$ -isomorfismo entre \mathcal{N} e \mathcal{N}' .
- c) Reflita: o que mudou com respeito ao exercício anterior? ■

Exercício F.25. Mostre que $\text{Con}_{\vdash}(\text{ZF}) \Rightarrow \text{ZF} \not\models \text{Con}_{\models}(\text{ZF})$. Dica: imite a demonstração do Teorema F.5.7 ou, como Gödel, observe que $\text{Con}_{\models}(\text{ZF}) \Rightarrow \text{Con}_{\models}(\text{ZFC})$, o que está longe de ser trivial⁴⁸. ■

Exercício F.26 (Paradoxo de Löwenheim-Skolem). Seja $\langle M, \varepsilon_M \rangle$ uma ε -estrutura infinita que modele ZFC.

- a) Mostre que existe $N \subseteq M$, com $|N| = \aleph_0$, tal que $N \models \text{ZFC}$.
- b) Como conciliar o item anterior com o fato de que $N \models \psi$, onde ψ é a ε -sentença que expressa a existência de conjuntos não-enumeráveis? Dica: o que significa dizer, explicitamente, que um conjunto é não-enumerável? ■

Exercício F.27. Diremos que um cardinal não-enumerável κ é **fortemente inacessível** se $2^\lambda < \kappa$ para todo cardinal $\lambda < \kappa$.

- a) Mostre que cardinais fortemente inacessíveis são fracamente inacessíveis.
- b) Mostre que se κ é fortemente inacessível, então $\mathbb{V}_\kappa \models \text{ZFC}$.
- c) Mostre que se ZFC é consistente, então ZFC não demonstra a afirmação “existe cardinal fortemente inacessível”. Dica: apele para a incompletude de Gödel; alternativamente, suponha que exista tal cardinal para daí considerar o menor deles.
- d) Mostre que se ZFC é consistente, então $\text{ZFC} + \text{“não existe cardinal fortemente inacessível”}$ é consistente.
- e) Suponha que M seja um modelo para ZFC satisfazendo GCH ($2^{\aleph_\alpha} = \aleph_{\alpha+1}$ para todo $\alpha \in \text{ORD}$). Com isso, mostre que ZFC não demonstra a afirmação “existe cardinal fracamente inacessível”. ■

⁴⁶No sentido implícito do tipo de equivalência satisfeita pelas subestruturas elementares: mais precisamente, M e N são estruturas *elementarmente equivalentes* se, e somente se, M e N satisfazem precisamente as mesmas sentenças da linguagem.

⁴⁷Não é tão simples quanto a dica sugere: convém estender a linguagem e considerar cada $n \in \omega$ como uma constante da linguagem.

⁴⁸O leitor interessado no método deve, literalmente, fazer o \mathbb{L} : a classe dos construtíveis de Gödel.

Lista de símbolos e siglas

\in	símbolo de pertinência, 9
$A \subseteq B$	A está contido em/é subconjunto de B , 10
$A \not\subseteq B$	A não é subconjunto de B , 10
$A \subsetneq B$	A é subconjunto próprio de B , 10
$A = B$	igualdade entre os conjuntos A e B , 10
$\{x : \mathcal{P}(x)\}$	conjunto dos x 's com a propriedade \mathcal{P} , 10
ZFC	axiomática Zermelo-Fraenkel-Choice, 11
$\{x \in A : \mathcal{P}(x)\}$	subconjunto de A no qual $\mathcal{P}(x)$ é satisfeita, 11
$A := B$	igualdade por definição, 11
\emptyset	conjunto vazio, 12
$\bigcap \mathcal{F}$ ou $\bigcap_{F \in \mathcal{F}} F$	interseção da família não-vazia \mathcal{F} , 12
$A \setminus B$	complementar de B em A , 13
$\{a, b\}$	par não-ordenado, 13
$\bigcup \mathcal{F}$ ou $\bigcup_{F \in \mathcal{F}} F$	reunião da família \mathcal{F} , 14
$A \cap B$	interseção entre A e B , 14
$A \cup B$	(re)união dos conjuntos A e B , 14
$\wp(X)$	conjunto das partes, 16
$X \times Y$	produto cartesiano, 17
$x R y$	R -relação; x e y estão R -relacionados, 17
$x \not R y$	negação de $x R y$, 17
$\text{dom}(R)$	domínio de R , 17
$\text{im}(R)$	imagem de R , 17
$f(x)$	valor de f em x , 18
$x \xrightarrow{f} y$	$f(x) = y$, 18
$f: X \rightarrow Y$ ou $X \xrightarrow{f} Y$	função f de X em Y , 18
$\text{graf}(f)$	gráfico da função f , 18
Id_X	função identidade de X , 19
R^{-1}	relação inversa de R , 19
$g \circ f$	composição das funções g e f , 21

$f[A]$	imagem direta de A por f , 21
$f^{-1}[B]$	pré-imagem de B por f , 21
Y^X	conjunto das funções de X em Y , 21
X_+	sucessor de X , 23
ω	conjunto dos números naturais, 25
$\prod_{i \in \mathcal{I}} X_i$	produto cartesiano generalizado, 27
$\langle f_i : i \in \mathcal{I} \rangle$ ou $\langle f_i \rangle_{i \in \mathcal{I}}$	\mathcal{I} -upla, 27
$\{x : \mathcal{P}(x)\}$	classe dos conjuntos x tais que vale $\mathcal{P}(x)$, 30
\mathbb{V}	universo (dos conjuntos), 31
$\mathcal{F}(x)$	imagem de x por \mathcal{F} , 31
$\mathcal{F}[X]$ ou $\{\mathcal{F}(x) : x \in X\}$	imagem de X por \mathcal{F} , 31
$\langle f_i \rangle_{i \in \mathcal{I}}$	produto diagonal das funções f_i , com $i \in \mathcal{I}$, 33
$\prod_{i \in \mathcal{I}} f_i$ ou $f_0 \times \dots \times f_n$	produto cartesiano de funções, 33
$\Delta_{i \in \mathcal{I}}$	função diagonal, 34
\bar{x}	classe de equivalência de x , 37
ZF	axiomática Zermelo-Fraenkel, 39
$\langle G, * \rangle$	conjunto G munido de operação $*$, 42
$\langle G, *, e \rangle$	conjunto G munido de operação $*$ com elemento e fixado, 42
$\mathbb{S}(X)$ ou \mathbb{S}_κ	conjunto das bijeções de X sobre X , com $ X = \kappa$, 42
\mathbb{N}	conjunto dos números naturais maiores do que 0, 43
\mathbb{Z}	conjunto dos números inteiros, 45
\mathbb{Q}	conjunto dos números racionais, 47
$\langle X, \preceq \rangle$	conjunto X munido de relação binária \preceq , em geral uma ordem parcial, 47
$\min A$, $\min_{a \in A} a$ ou $\min \leq A$	o menor elemento de A com respeito à ordem \leq , 49
$\max A$, $\max_{a \in A} a$ ou $\max \leq A$	o maior elemento de A com respeito à ordem \leq , 49
$\inf A$, $\inf_{a \in A} a$ ou apenas $\inf \leq A$	ínfimo de A com respeito à ordem \leq , 49
$\sup A$, $\sup_{a \in A} a$ ou $\sup \leq A$	supremo de A com respeito à ordem \leq , 49
$a.k.a.$	“também conhecido como”, 52
$\overline{\text{seq}}(X)$	sequências finitas de X , 55
$n!$	fatorial de n , 56
$\#X$	(classe) cardinalidade de X , 60
$ X = Y $	cardinalidade de X igual à cardinalidade de Y , 63
$ X \neq Y $	cardinalidade de X diferente da cardinalidade de Y , 63
$ X \leq Y $	cardinalidade de X menor do que a cardinalidade de Y , 63
$ X < Y $	cardinalidade de X estritamente menor do que a cardinalidade de Y , 63

\mathbb{R}	conjunto dos números reais (<i>a.k.a.</i> reta real), 68
\mathbb{I}	conjunto dos números irracionais, 68
$A[x]$	anel de polinômios com coeficientes em A e indeterminada x , 68
$\alpha < \beta$	$\alpha \in \beta$ para ordinais α e β , 70
ORD	classe de todos os ordinais, 73
$\langle \mathbb{P}, \leq \rangle \cong \langle \mathbb{P}', \preceq \rangle$	isomorfismo entre as ordens $\langle \mathbb{P}, \leq \rangle$ e $\langle \mathbb{P}', \preceq \rangle$, 74
$\text{ord}(\mathbb{W})$	tipo de ordem de \mathbb{W} , 75
$H(X)$	número de Hartogs de X , 76
ω_1	primeiro ordinal não-enumerável, 76
ω_α	α -ésimo ordinal limite, 79
<i>ex nihilo nihil fit</i>	nada vem do vazio, 80
$\sup \mathcal{X}$	supremo de uma família \mathcal{X} de ordinais, 82
<i>Katzensprung</i>	“pulo do gato”, 86
$\text{sp}(X)$	subestrutura/subálgebra (subgrupo, subespaço vetorial, etc.) gerada por X , 87
AC	Axioma da Escolha, 96
$ X $	número cardinal de X , 99
κ^+	cardinal sucessor de κ , 100
\aleph_α	α -ésimo número cardinal transfinito, 102
$\kappa + \lambda$	soma dos cardinais κ e λ , 104
$\kappa \cdot \lambda$ ou $\kappa\lambda$	produto entre os cardinais κ e λ , 104
κ^λ	κ elevado à λ -ésima potência, 104
$\sum_{i \in \mathcal{I}} \kappa_i$	soma dos cardinais κ_i para $i \in \mathcal{I}$, 104
$\prod_{i \in \mathcal{I}} \kappa_i$	produto dos cardinais κ_i para $i \in \mathcal{I}$, 104
$X^{<\alpha}$	conjunto das funções da forma $\gamma \rightarrow X$ para $\gamma < \alpha$, 108
$X^{\leq\alpha}$	conjunto das funções da forma $\gamma \rightarrow X$ para $\gamma \leq \alpha$, 108
$[X]^\kappa$	subconjuntos de X com cardinalidade κ , 108
$[X]^{<\kappa}$	subconjuntos de X com cardinalidade $< \kappa$, 108
$[X]^{\leq\kappa}$	subconjuntos de X com cardinalidade $\leq \kappa$, 108
\mathfrak{c}	contínuo ou continuum, 110
CH	hipótese do contínuo, 110
PCP	Princípio da Casa dos Pombos, 114
$\text{cof}(\mathbb{P})$	cofinalidade de \mathbb{P} , 115
$\text{Ker}(f)$	núcleo universal do morfismo f , 131
$\mathbb{T}_{\mathcal{L}}(\mathcal{V})$	\mathcal{L} -termos nas variáveis de \mathcal{V} , 133
$u: \mathcal{V} \rightarrow M$	\mathcal{V} -atribuição em M , 133

$[\mathcal{V} \rightarrow M]$	conjunto das \mathcal{V} -atribuições, 133
$M \models \Gamma$	a álgebra/modelo satisfaz as identidades/relações/axiomas de Γ , 136
$\mathbb{F}_{\mathcal{L}}(\mathcal{V})$	\mathcal{L} -fórmula (de primeira ordem) nas variáveis de \mathcal{V} , 140
$M \models \varphi[u]$	M satisfaz φ com a atribuição u , 140
$T \models \varphi$	φ é consequência semântica de T , 146
$\Sigma \vdash \varphi$	Σ demonstra φ , 147
$\prod_{\mathcal{F}} M_i$	produto reduzido/ultraproduto de $\langle M_i \rangle_{i \in \mathcal{I}}$ pelo (ultra) filtro \mathcal{F} , 149
$\text{Con}_{\vdash}(\Sigma)$	Σ é sintaticamente consistente, 155
$\text{Con}_{\models}(\Sigma)$	Σ é semanticamente consistente, 155
\mathbb{C}	corpo dos complexos, 157

Referências Bibliográficas

- [1] J. Bagaria. A Short Guide to Gödel's Second Incompleteness Theorem. *Teorema: Revista Internacional de Filosofía*, 22(3):5–15, 2003.
- [2] B. Banaschewski. A new proof that “Krull implies Zorn”. *Mathematical Logic Quarterly*, 40:478–480, 1944.
- [3] M. Ben-Ari. *Mathematical Logic for Computer Science*. Springer, 2012.
- [4] G. M. Bergman. *An invitation to General Algebra and Universal Constructions*. Springer, 2017.
- [5] A. Blass. Existence of bases implies the Axiom of Choice. *American Journal of Mathematics*, 31:31–33, 1984.
- [6] S. Burris and H. P. Sankappanavar. *A Course in Universal Algebra*. Burris and Sankappanavar, 2012.
- [7] C. C. Chang and H. J. Keisler. *Model Theory*. Studies in Logic and the Foundations of Mathematics. Elsevier, 1990.
- [8] P. M. Cohn. *Universal Algebra*. Mathematics and Its Applications 6. Springer, 1 edition, 1981.
- [9] H.-D. Ebbinghaus, J. Flum, and W. Thomas. *Mathematical Logic*. Graduate Texts in Mathematics. Springer, 3 edition, 2021.
- [10] J. Ferreirós. *Labyrinth of thought: A history of set theory and its role in modern mathematics*. Birkhäuser Basel, 2^a edition, 2007.
- [11] G. B. Folland. *Real analysis: modern techniques and their applications*. Wiley, 2 edition, 1999.
- [12] A. A. Fraenkel, Y. Bar-Hillel, and A. Levy. *Foundations of set theory*. Elsevier, 2 edition, 1973.
- [13] J. H. Gallier. *Logic for Computer Science: Foundations of Automatic Theorem Proving*. Dover, 2 edition, 2015.
- [14] I. Goldbring. *Ultrafilters Throughout Mathematics*. Graduate Studies in Mathematics, 220. American Mathematical Society, 2022.
- [15] L. Halbeisen and R. Krapf. *Gödel's Theorems and Zermelo's Axioms: A Firm Foundation of Mathematics*. Springer, 2020.
- [16] L. J. Halbeisen. *Combinatorial Set Theory: With a Gentle Introduction to Forcing*. Springer Monographs in Mathematics. Springer, 2 edition, 2017.
- [17] J. D. Hamkins. *Lectures on the Philosophy of Mathematics*. MIT Press, 2021.
- [18] H. Herrlich. *Axiom of Choice*. Lecture Notes in Mathematics 1876. Springer, 1 edition, 2006.
- [19] K. Hrbacek and T. Jech. *Introduction to set theory*. Monographs and Textbooks in Pure and Applied Mathematics 220. M. Dekker, New York, 3 edition, 1999.
- [20] T. Jech. On Gödel's Second Incompleteness Theorem. *Proceedings of the American Mathematical Society*, 121(1):311–313, 1994.
- [21] T. Jech. *Set Theory: The Third Millennium Edition, revised and expanded*. Springer Monographs in Mathematics. Springer-Verlag, Berlin, 2006.
- [22] K. Kunen. *Set Theory: An Introduction to Independence Proofs*. Studies in Logic and the Foundations of Mathematics 102. North Holland, 1983.

- [23] K. Kunen. *The Foundations of Mathematics*. Studies in Logic: Mathematical Logic and Foundations 19. College Publications, 2009.
- [24] K. Kunen. *Set Theory*. College Publications, London, 2011.
- [25] Y. I. Manin. *A Course in Mathematical Logic for Mathematicians*. Graduate Texts in Mathematics. Springer, 2 edition, 2010.
- [26] R. M. Mezabarba. Fundamentos de Topologia Geral, 2023. Manuscrito disponível em https://github.com/mezabarbare/Fund_Top_Geral.
- [27] S. Negri and J. V. Plato. *Structural Proof Theory*. Cambridge, 2008.
- [28] B. Rin. Transfinite recursion and computation in the iterative conception of set. *Synthese*, 192(8):2437–2462, 2015.
- [29] P. Rothmaler. *Introduction to model theory*. CRC Press, 2000.
- [30] B. Russell. *Introduction to Mathematical Philosophy*. Dover, 2 edition, 1993.
- [31] E. Schechter. *Handbook of Analysis and Its Foundations*. Academic Press, 1996.
- [32] S. Shapiro, editor. *The Oxford Handbook of Philosophy of Mathematics and Logic*. Oxford Handbooks in Philosophy. Oxford University, 2005.
- [33] D. van Dalen. *Logic and Structure*. Universitext. Springer, 3rd edition, 1994.
- [34] A. Wasilewska. *Logics for Computer Science: Classical and Non-Classical*. Springer, 2018.
- [35] N. Weaver. *Forcing for Mathematicians*. World Scientific, 2014.

Índice Remissivo

- σ -álgebra, 77, 116
 - de Borel, 121
- álgebra
 - congruência, 131
 - de tipo \mathcal{L} , 130
 - livre, 137
 - subálgebra de uma, 130
- a.k.a.*, 52
- anel, 46
 - com unidade, 46
 - comutativo, 46
 - de polinômios, 68
- aplicação (*a.k.a.* função), 18
- aridade, 126
- assinatura
 - de uma linguagem, 126
- atribuição de valores, 133
- axioma, 10, 147
 - lógico, 147
 - não-lógico, 147
- Axioma (de ZFC)
 - da Escolha, 28, 83
 - da Extensão, 10
 - da Fundação, 29
 - da Separação, 11
 - da Substituição, 30
 - da União, 14
 - das Escolhas Múltiplas, 97
 - das Partes, 16
 - do Infinito, 24
 - do Par, 13
- boa definição, 23
- cadeia, 88
- cardinal
 - a.k.a.* “ordinal inicial”, 71
 - adição, 104
 - exponenciação, 104
 - fortemente inacessível, 164
 - fracamente inacessível, 120
 - limite, 119
 - multiplicação, 104
 - regular, 118
 - singular, 118
 - sucessor, 100, 119
- cardinalidade
 - como classe de equivalência, 60
 - como relação, 59
- classe
 - de equivalência, 36
 - de representantes, 38
 - equacional, 136
 - imprópria (*a.k.a.* conjunto), 30
 - própria, 30
 - subclasse, 31
 - universo (dos conjuntos), 31
- codomínio (de uma função), 18
- cofinalidade, 115
- coleção (ver conjunto), 13
- complemento, 13
- conjunto, 9
 - das partes, 16
 - dos números naturais, 25
 - enumerável, 67
 - finito, 23
 - indutivo, 25
 - infinito, 23
 - limitado, 48
 - não-enumerável, 67
 - parcialmente ordenado, 47
 - quociente, 37
 - sucessor, 23
 - transitivo, 69
 - unitário, 13
 - vazio, 12
- conjunto dos números
 - complexos, 157
 - inteiros, 45
 - irracionais, 68
 - racionais, 47
 - reais, 68
- conjuntos disjuntos, 14
- consequência
 - semântica, 146
 - sintática (*a.k.a.* dedução), 147
- contínuo (ou *continuum*, 110)
- corpo, 67
 - algebricamente fechado, 152
 - arquimediano, 67
 - ordenado, 67
 - ordenado completo, 67
- Critério ou Princípio
 - (da semântica) de Tarski, 140

- de Hume, 99
- de Tarski-Vaught, 145
- dedução (*a.k.a.* prova), 147
- demonstração (*a.k.a.* prova), 147
- diagrama comutativo, 41
- domínio (de uma relação/função), 17
- elemento, 9
 - (elementos) equivalentes, 35
 - inverso, 43
 - inverso à direita, 43
 - inverso à esquerda, 43
 - neutro da operação, 41
- estrutura
 - associada a uma linguagem, 126
 - morfismo de, 127
 - subestrutura de uma, 129
 - universo da, 126
- fórmula (ou \mathcal{L} -fórmula)
 - atômica, 140
 - com variável livre, 142
 - de primeira ordem, 140
 - funcional em x , 29
 - parcialmente funcional, 34
 - sentença, 142
- família (ver conjunto), 13
- filtro, 94
- função (como conjunto de pares), 18
 - \mathcal{F} -recursiva, 53
 - \mathcal{G} -recursiva (de classes), 73
 - \mathcal{O} -recursiva, 55, 74
 - bijetora, 20
 - codomínio de uma, 18
 - composição, 21
 - crescente, 57
 - de X em Y , 18
 - de classe, 31
 - decrescente, 57
 - escolha, 27
 - estritamente crescente, 57
 - estritamente decrescente, 57
 - funcional linear, 92
 - gráfico de uma, 18
 - identidade, 19
 - imagem de um elemento, 18
 - imagem direta, 21
 - inclusão, 19
 - injetora, 19, 20
 - monótona, 57
 - oráculo, 55
 - ponto fixo de uma, 63
 - pré-imagem, 21
 - produto diagonal entre, 33
 - projeção, 33
 - restrição, 21
 - sobrejetora, 20
 - funcional linear, 92
- grupo, 43
 - abeliano (*a.k.a.* comutativo), 42
 - de Grothendieck, 45
- Hipótese do Contínuo, 110
- I-upla, 27
- ideal
 - de um anel, 93
 - maximal (num anel), 93
 - sobre um conjunto, 94
- imagem
 - de um conjunto por uma função, 21, 31
 - de um elemento por uma função, 18
 - de uma relação/função, 17
 - direta, 21
- Incompletude de Gödel
 - Primeira, 157
 - Segunda, 157
- indução
 - finita, 25
 - numa boa ordem, 51
 - sobre complexidade, 130
 - transfinita, 80
- interseção, 12
- isomorfismo
 - de ordens, 74
 - entre estruturas, 128
- Katzensprung, 86
- kernel universal, 131
- lei do cancelamento, 44
- Leis de De Morgan, 16
- Lema
 - de Steinitz, 110
 - de Zorn, 89
 - do ultrafiltro, 95
- linguagem, 126
 - algébrica, 126, 128
 - assinatura de uma, 126
- módulo, 143
- magma, 143
- mapa (*a.k.a.* função), 18
- medida, 121
 - completa, 121
 - finitamente aditiva, 149
- modelo, 141
- monoide, 42
 - abeliano, 42
 - comutativo (ver abeliano), 42
- morfismo
 - entre estruturas, 127
 - núcleo universal do, 131
- núcleo, 131
- número
 - cardinal, 71
 - cardinal de X , 99
 - cardinal finito, 61

- cardinal regular, 118
- cardinal singular, 118
- de Hartogs, 76
- fatorial, 56
- inteiro, 45
- natural, 25
- ordinal, 69
- ordinal inicial, 71
- ordinal limite, 71
- ordinal sucessor, 71
- norma, 92
- operação
 - n -ária, 126
 - associativa, 41
 - binária, 41, 126
 - comutativa, 41
 - elemento neutro da, 41
 - únaria, 126
- ordem
 - ínfimo, 48
 - boa ordem, 50
 - elemento máximo, 48
 - elemento mínimo, 48
 - elemento maximal, 48
 - elemento minimal, 48
 - estrita, 48
 - isomorfa a outra, 74
 - limitante inferior, 48
 - limitante superior, 48
 - maior elemento, 48
 - menor elemento, 48
 - parcial, 48
 - supremo, 48
 - tipo, 75
 - total, 50
- par
 - coordenadas do, 14
 - não-ordenado, 13
 - ordenado, 14
- Paradoxo
 - de Burali-Forti, 73
 - de Cantor, 102
 - de Löwenheim-Skolem, 164
 - de Russell, 11
- partição, 37
- permutação, 42
- polinômio, 68
 - grau do, 68
- ponto
 - aderente, 28
 - fixo, 63
- pré-imagem, 21
- Princípio
 - da Casa dos Pombos, 114
 - da indução, 25
 - da indução (segunda forma), 51
- produto
 - cartesiano, 17
 - cartesiano generalizado, 27
 - reduzido, 149
 - ultra, 149
- projeção (ou projeção canônica), 40
- propriedade universal
 - do produto de conjuntos, 33
 - do quociente de conjuntos, 40
 - dos \mathcal{L} -termos, 135
- prova (*a.k.a.* demonstração), 147
- recursão
 - em boas ordens, 54
 - finita (em ω), 55
 - transfinita (em ordinais), 73
- regra de inferência, 147
- relação
 - n -ária, 126
 - antissimétrica, 47
 - assimétrica, 48
 - binária, 17
 - de equivalência, 35
 - de identidade, 136
 - de ordem estrita, 47
 - de ordem parcial, 47
 - domínio de, 17
 - imagem de, 17
 - inversa, 19
 - irreflexiva, 47
 - simétrica, 35
 - transitiva, 35
- reunião, 14
- símbolo
 - operacional, 126
 - relacional, 126
- satisfabilidade
 - (algébrica) de identidades, 136
 - de fórmulas, 140
- semigrupo, 42
- sentença, 142
 - independente, 157
- sequência
 - finita, 55
 - infinita, 28
- sistema de dedução, 147
 - completo, 148
 - correto, 147
- subálgebra
 - de uma álgebra, 130
 - gerada, 130
- subconjunto, 10
 - cofinal, 114
 - próprio, 10
- subestrutura, 129
 - elementar, 145
- Teorema
 - da indução finita, 25

- da indução finita (segunda forma), 51
de Bolzano-Weierstrass, 116
de Cantor, 62
de Cantor-Bernstein, 64, 82
de Hahn-Banach, 92
de König, 106
de Löwenheim-Skolem, 145
de Tarski-Vaught, 145
de Zermelo, 86
de Łoś, 151
do ponto fixo de Tarski, 63
teoria, 146
- completa, 157
incompleta, 157
semanticamente consistente, 155
sintaticamente consistente, 155
termo (ou \mathcal{L} -termo), 133
atômico, 133
- ultraproduto, 149
- união, 14
- universo
de uma estrutura, 126
dos conjuntos, 31