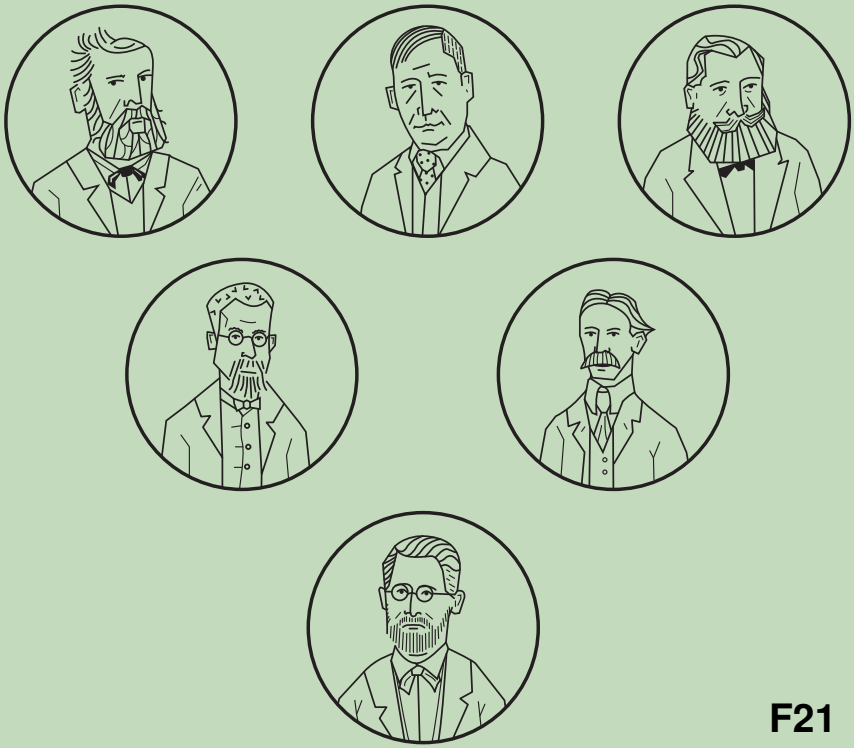
Set Theory

An Open Introduction



F21

Set Theory

The Open Logic Project

Instigator

Richard Zach, *University of Calgary*

Editorial Board

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

Contributors

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

Set Theory

An Open Introduction

Remixed by Tim Button

Fall 2021

The Open Logic Project would like to acknowledge the generous support of the Taylor Institute of Teaching and Learning of the University of Calgary, and the Alberta Open Educational Resources (ABOER) Initiative, which is made possible through an investment from the Alberta government.

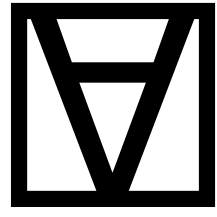
UNIVERSITY OF CALGARY
Taylor Institute for Teaching and Learning

Cover illustrations by Matthew Leadbeater, used under a Creative Commons Attribution-NonCommercial 4.0 International License.

Typeset in Baskervald X and Nimbus Sans by LaTeX.

This version of *Set Theory* is revision 6c40575 (2021-10-30), with content generated from *Open Logic Text* revision 6891b66 (2024-12-01). Free download at:

https://st.openlogicproject.org/



Contents

About this Book

This book is an Open Education Resource. It is written for students with a little background in logic, and some high school mathematics. It aims to scratch the tip of the surface of the philosophy of set theory. By the end of this book, students reading it might have a sense of:

1. why set theory came about;
2. how to embed large swathes of mathematics within set theory + arithmetic;
3. how to embed arithmetic itself within set theory;
4. what the cumulative iterative conception of set amounts to;
5. how one might try to justify the axioms of ZFC.

The book grew out of a short course that I taught in the Cambridge Philosophy department. Before me, it was lectured by Luca Incurvati and Michael Potter. In writing this book—and the course, more generally—I was hugely indebted to both Luca and Michael. I hope this comes through in the text of the book itself; but I also want to offer both of them my heartfelt thanks here.

Most of this book was originally released as *Open Set Theory*; this book is its successor. I have contributed all of the material

for this book to the *Open Logic Project*, where it is freely available. (Chapters 2 to 5 were drawn, with only tiny changes, from previously existing material in the OLP; these changes are now also part of the OLP.) Please see openlogicproject.org for more information.

Tim Button
University College London
October 2021

PART I

Prelude

CHAPTER 1

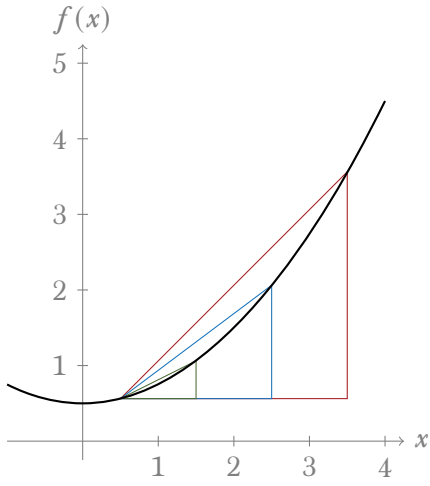
History and Mythology

To understand the philosophical significance of set theory, it will help to have some sense of why set theory arose at all. To understand that, it will help to think a little bit about the history and mythology of mathematics. So, before we get started on discussing set theory at all, we will start with a very brief "history". But we put this in scare-quotes, because it is very brief, extremely selective, and somewhat contestable.

1.1 Infinitesimals and Differentiation

Newton and Leibniz discovered the calculus (independently) at the end of the 17th century. A particularly important application of the calculus was *differentiation*. Roughly speaking, differentiation aims to give a notion of the "rate of change", or gradient, of a function at a point.

Here is a vivid way to illustrate the idea. Consider the function $f(x) = x^2/4 + 1/2$, depicted in black below:



Suppose we want to find the gradient of the function at $c = 1/2$. We start by drawing a triangle whose hypotenuse approximates the gradient at that point, perhaps the red triangle above. When β is the base length of our triangle, its height is $f(1/2+\beta) - f(1/2)$, so that the gradient of the hypotenuse is:

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

So the gradient of our red triangle, with base length 3, is exactly 1. The hypotenuse of a smaller triangle, the blue triangle with base length 2, gives a better approximation; its gradient is $3/4$. A yet smaller triangle, the green triangle with base length 1, gives a yet better approximation; with gradient $1/2$.

Ever-smaller triangles give us ever-better approximations. So we might say something like this: the hypotenuse of a triangle with an *infinitesimal* base length gives us the gradient at $c = 1/2$ itself. In this way, we would obtain a formula for the (first) derivative of the function f at the point c:

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

And, roughly, this is what Newton and Leibniz said.

However, since they have said this, we must ask them: what is an *infinitesimal*? A serious dilemma arises. If $\beta = 0$, then f' is ill-defined, for it involves dividing by 0. But if $\beta > 0$, then we just get an *approximation* to the gradient, and not the gradient itself.

This is not an anachronistic concern. Here is Berkeley, criticizing Newton's followers:

> I admit that signs may be made to denote either any thing or nothing: and consequently that in the original notation $c + \beta$, β might have signified either an increment or nothing. But then which of these soever you make it signify, you must argue consistently with such its signification, and not proceed upon a double meaning: Which to do were a manifest sophism. (Berkeley 1734, §XIII, variables changed to match preceding text)

To defend the infinitesimal calculus against Berkeley, one might reply that the talk of "infinitesimals" is merely figurative. One might say that, so long as we take a *really small* triangle, we will get a *good enough* approximation to the tangent. Berkeley had a reply to this too: whilst that might be good enough for engineering, it undermines the *status* of mathematics, for

> we are told that *in rebus mathematicis errores quàm minimi non sunt contemnendi.* [In the case of mathematics, the smallest errors are not to be neglected.] (Berkeley, 1734, §IX)

The italicised passage is a near-verbatim quote from Newton's own *Quadrature of Curves* (1704).

Berkeley's philosophical objections are deeply incisive. Nevertheless, the calculus was a massively successful enterprise, and mathematicians continued to use it without falling into error.

1.2 Rigorous Definition of Limits

These days, the standard solution to the foregoing problem is to get rid of the infinitesimals. Here is how.

We saw that, as β gets smaller, we get better approximations of the gradient. Indeed, as β gets arbitrarily close to 0, the value of $f'(c)$ "tends without limit" to the gradient we want. So, instead of considering what happens *at* $\beta = 0$, we need only consider the *trend* of $f'(c)$ as β approaches 0.

Put like this, the general challenge is to make sense of claims of this shape:

$$\text{As } x \text{ approaches } c,\ g(x) \text{ tends without limit to } \ell.$$

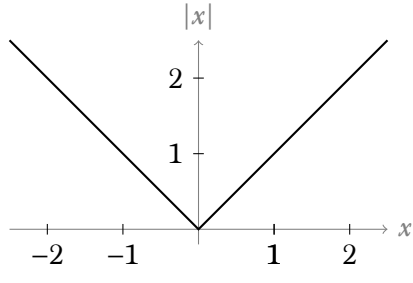
which we can write more compactly as follows:

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

In the 19th century, building upon earlier work by Cauchy, Weierstrass offered a perfectly rigorous definition of this expression. The idea is indeed that we can make $g(x)$ as close as we like to ℓ, by making x suitably close to c. More precisely, we stipulate that $\lim_{x \to c} g(x) = \ell$ will mean:

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

The vertical bars here indicate absolute magnitude. That is, $|x| = x$ when $x \geq 0$, and $|x| = -x$ when $x < 0$; you can depict *that* function as follows:



So the definition says roughly this: you can make your "error" less than ε (i.e., $|g(x) - \ell| < \varepsilon$) by choosing arguments which are no more than δ away from c (i.e., $|x - c| < \delta$).

Having defined the notion of a limit, we can use it to avoid infinitesimals altogether, stipulating that the gradient of f at c is given by:

$$f'(c) = \lim_{x \to 0} \left(\frac{f(c + x) - f(c)}{x} \right) \text{ where a limit exists.}$$

It is important, though, to realise why our definition needs the caveat "where a limit exists". To take a simple example, consider $f(x) = |x|$, whose graph we just saw. Evidently, $f'(0)$ is ill-defined: if we approach 0 "from the right", the gradient is always 1; if we approach 0 "from the left", the gradient is always -1; so the limit is undefined. As such, we might add that a function f is *differentiable* at x iff such a limit exists.

We have seen how to handle differentiation using the notion of a *limit*. We can use the same notion to define the idea of a *continuous* function. (Bolzano had, in effect, realised this by 1817.) The Cauchy–Weierstrass treatment of continuity is as follows. Roughly: a function f is continuous (at a point) provided that, if you demand a certain amount of precision concerning the output of the function, you can guarantee this by insisting upon a certain amount of precision concerning the input of the function. More precisely: f is continuous at c provided that, as x tends to zero, the difference between $f(c + x)$ and $f(c)$ itself tends to 0. Otherwise put: f is *continuous* at c iff $f(c) = \lim_{x \to c} f(x)$.

To go any further would just lead us off into real analysis, when our subject matter is set theory. So now we should pause, and state the moral. During the 19th century, mathematicians learnt how to do without infinitesimals, by invoking a rigorously defined notion of a *limit*.

1.3 Pathologies

However, the definition of a *limit* turned out to allow for some rather "pathological" constructions.

Around the 1830s, Bolzano discovered a function which was *continuous everywhere*, but *differentiable nowhere*. (Unfortunately, Bolzano never published this; the idea was first encountered by mathematicians in 1872, thanks to Weierstrass's independent discovery of the same idea.)[1] This was, to say the least, rather surprising. It is easy to find functions, such as $|x|$, which are continuous everywhere but not differentiable at a particular point. But a function which is continuous everywhere but differentiable *nowhere* is a very different beast. Consider, for a moment, how you might try to draw such a function. To ensure it is continuous, you must be able to draw it without ever removing your pen from the page; but to ensure it is differentiable nowhere, you would have to abruptly change the direction of your pen, constantly.

Further "pathologies" followed. In January 5 1874, Cantor wrote a letter to Dedekind, posing the problem:

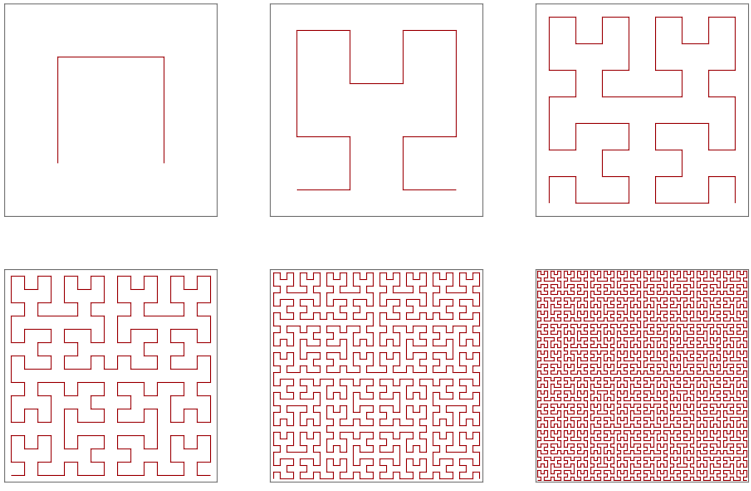
> Can a surface (say a square including its boundary) be one-to-one correlated to a line (say a straight line including its endpoints) so that to every point of the surface there corresponds a point of the line, and conversely to every point of the line there corresponds a point of the surface?
>
> It still seems to me at the moment that the answer to this question is very difficult—although here too one is so impelled to say *no* that one would like to hold the proof to be almost superfluous. [Quoted in Gouvêa 2011]

But, in 1877, Cantor proved that he had been wrong. In fact, a line and a square have exactly the same number of points. He

[1] The history is documented in extremely thorough footnotes to the Wikipedia article on the Weierstrass function.

wrote on 29 June 1877 to Dedekind "*je le vois, mais je ne le crois pas*"; that is, "I see it, but I don't believe it". In the "received history" of mathematics, this is often taken to indicate just how *literally incredible* these new results were to the mathematicians of the time. (The correspondence is presented in Gouvêa (2011), and we return to it in section 1.4. Cantor's proof is outlined in section 5.10.)

Inspired by Cantor's result, Peano started to consider whether it might be possible to map a line *smoothly* onto a plane. This would be a *curve which fills space*. In 1890, Peano constructed just such a curve. This is truly counter-intuitive: Euclid had defined a line as "breadthless length" (Book I, Definition 2), but Peano had shown that, by curling up a line appropriately, its length can be turned into breadth. In 1891, Hilbert described a slightly more intuitive space-filling curve, together with some pictures illustrating it. The curve is constructed in sequence, and here are the first six stages of the construction:



In the limit—a notion which had, by now, received rigorous definition—the entire square is filled in solid red. And, in passing, Hilbert's curve is continuous everywhere but differentiable nowhere; intuitively because, in the infinite limit, the function

abruptly changes direction at every moment. (We will outline Hilbert's construction in more detail in section 5.11.)

For better or worse, these "pathological" geometric constructions were treated as a reason to doubt appeals to geometric intuition. They became something approaching *propaganda* for a new way of doing mathematics, which would culminate in set theory. In the later myth-building of the subject, it was repeated, often, that these results were both perfectly rigorous and perfectly shocking. They therefore served a dual purpose: as a warning against relying upon geometric intuition, and as a demonstration of the fertility of new ways of thinking.

1.4 More Myth than History?

Looking back on these events with more than a century of hindsight, we must be careful not to take these verdicts on trust. The results were certainly novel, exciting, and surprising. But how truly shocking were they? And did they really demonstrate that we should not rely on geometric intuition?

On the question of shock, Gouvêa (2011) points out that Cantor's famous note to Dedekind, "*je le vois, mais je ne le crois pas*" is taken rather out of context. Here is more of that context (quoted from Gouvêa):

> Please excuse my zeal for the subject if I make so many demands upon your kindness and patience; the communications which I lately sent you are even for me so unexpected, so new, that I can have no peace of mind until I obtain from you, honoured friend, a decision about their correctness. So long as you have not agreed with me, I can only say: *je le vois, mais je ne le crois pas.*

Cantor knew his result was "so unexpected, so new". But it is doubtful that he ever found his result *unbelievable.* As Gouvêa

points out, he was simply asking Dedekind to check the proof he had offered.

On the question of geometric intuition: Peano published his space-filling curve without including any diagrams. But when Hilbert published his curve, he explained his purpose: he would provide readers with a clear way to understand Peano's result, if they "help themselves to the following geometric intuition"; whereupon he included a series of *diagrams* just like those provided in section 1.3.

More generally: whilst diagrams have fallen rather out of fashion in published proofs, there is no getting round the fact that mathematicians *frequently* use diagrams when proving things. (Roughly put: good mathematicians know when they can rely upon geometric intuition.)

In short: don't believe the hype; or at least, don't just take it on trust. For more on this, you could read Giaquinto (2007).

1.5 Roadmap

Part of the moral of the previous section is that the history of mathematics was largely written by the victors. They had axes to grind; philosophical and mathematical axes. Serious study of the history of mathematics is seriously difficult (and rewarding), and the Owl of Minerva takes flight only at dusk.

For all that, it's incontestable that the "pathological" results involved the development of fascinating new mathematical tools, and a re-thinking of the standards of mathematical rigour. For example, they required thinking of the continuum (the "real line") in a particular way, and thinking of functions as point-by-point maps. And, in the end, the full development of all of these tools required the rigorous development of set theory. The rest of this book will explain some of that development.

Part II will present a version of *naïve* set theory, which is easily sufficient to develop all of the mathematics just described. This will take a while. But, by end of part II we will be in a position to

understand how to treat real numbers as certain sets, and how to treat functions on them—including space-filling curves—as *further* sets.

But the *naïvety* of this set theory will emerge in part III, as we encounter set-theoretic paradoxes, and the felt need to describe things much more precisely. At this point, we will need to develop an axiomatic treatment of sets, which we can use to recapture all of our naïve results, whilst (hopefully) avoiding paradoxes. (The Owl of mathematical rigour takes flight only at dusk, too.)

PART II

Naïve Set Theory

Introduction to Part II

In part II, we will consider sets in a naïve, informal way. Chapter 2 will introduce the basic idea, that sets are collections considered extensionally, and will introduce some very basic operations. Then chapters 3 to 4 will explain how set theory allows us to speak about relations and (therefore) functions.

Chapters 5 to 7 will then consider some of the early achievements of naïve set theory. In chapter chapter 5, we explore how to compare sets with regard to their size. In chapter 6, we explore how one might reduce the integers, rationals, and reals to set theory plus basic arithmetic. In chapter 7, we consider how one might implement basic arithmetic within set theory.

To repeat, all of this will be done *naïvely*. But everything we do in part II can be done perfectly rigorously, in the formal set theory which we introduce in part III.

CHAPTER 2

Getting Started

2.1 Extensionality

A *set* is a collection of objects, considered as a single object. The objects making up the set are called *elements* or *members* of the set. If x is a member of a set A, we write $x \in A$; if not, we write $x \notin A$. The set which has no members is called the *empty* set and denoted "$\emptyset$".

It does not matter how we *specify* the set, or how we *order* its members, or indeed how *many times* we count its members. All that matters are what its members are. We codify this in the following principle.

Definition 2.1 (Extensionality). If A and B are sets, then $A = B$ iff every member of A is also a member of B, and vice versa.

Extensionality licenses some notation. In general, when we have some objects a_1, ..., a_n, then $\{a_1, \ldots, a_n\}$ is *the* set whose members are $a_1, \ldots, a_n$. We emphasise the word "*the*", since extensionality tells us that there can be only *one* such set. Indeed, extensionality also licenses the following:

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

This delivers on the point that, when we consider sets, we don't care about the order of their members, or how many times they are specified.

Example 2.2. Whenever you have a bunch of objects, you can collect them together in a set. The set of Richard's siblings, for instance, is a set that contains one person, and we could write it as $S = \{\text{Ruth}\}$. The set of positive integers less than 4 is $\{1,2,3\}$, but it can also be written as $\{3,2,1\}$ or even as $\{1,2,1,2,3\}$. These are all the same set, by extensionality. For every member of $\{1,2,3\}$ is also a member of $\{3,2,1\}$ (and of $\{1,2,1,2,3\}$), and vice versa.

Frequently we'll specify a set by some property that its members share. We'll use the following shorthand notation for that: $\{x : \varphi(x)\}$, where the $\varphi(x)$ stands for the property that x has to have in order to be counted among the members of the set.

Example 2.3. In our example, we could have specified S also as

$$S = \{x : x \text{ is a sibling of Richard}\}.$$

Example 2.4. A number is called *perfect* iff it is equal to the sum of its proper divisors (i.e., numbers that evenly divide it but aren't identical to the number). For instance, 6 is perfect because its proper divisors are 1, 2, and 3, and $6 = 1 + 2 + 3$. In fact, 6 is the only positive integer less than 10 that is perfect. So, using extensionality, we can say:

$$\{6\} = \{x : x \text{ is perfect and } 0 \leq x \leq 10\}$$

We read the notation on the right as "the set of x's such that x is perfect and $0 \leq x \leq 10$". The identity here confirms that, when we consider sets, we don't care about how they are specified. And, more generally, extensionality guarantees that there is always only one set of x's such that $\varphi(x)$. So, extensionality justifies calling $\{x : \varphi(x)\}$ *the* set of x's such that $\varphi(x)$.

Extensionality gives us a way for showing that sets are identical: to show that $A = B$, show that whenever $x \in A$ then also $x \in B$, and whenever $y \in B$ then also $y \in A$.

2.2 Subsets and Power Sets

We will often want to compare sets. And one obvious kind of comparison one might make is as follows: *everything in one set is in the other too.* This situation is sufficiently important for us to introduce some new notation.

Definition 2.5 (Subset). If every member of a set A is also a member of B, then we say that A is a *subset* of B, and write $A \subseteq B$. If A is not a subset of B we write $A \not\subseteq B$. If $A \subseteq B$ but $A \neq B$, we write $A \subsetneq B$ and say that A is a *proper subset* of B.

Example 2.6. Every set is a subset of itself, and $\emptyset$ is a subset of every set. The set of even numbers is a subset of the set of natural numbers. Also, $\{a, b\} \subseteq \{a, b, c\}$. But $\{a, b, e\}$ is not a subset of $\{a, b, c\}$.

Example 2.7. The number 2 is an member of the set of integers, whereas the set of even numbers is a subset of the set of integers. However, a set may happen to *both* be a member and a subset of some other set, e.g., $\{0\} \in \{0, \{0\}\}$ and also $\{0\} \subseteq \{0, \{0\}\}$.

Extensionality gives a criterion of identity for sets: $A = B$ iff every member of A is also a member of B and vice versa. The definition of "subset" defines $A \subseteq B$ precisely as the first half of this criterion: every member of A is also a member of B. Of course the definition also applies if we switch A and B: that is, $B \subseteq A$ iff every member of B is also a member of A. And that, in turn, is exactly the "vice versa" part of extensionality. In other words, extensionality entails that sets are equal iff they are subsets of one another.

Proposition 2.8. *$A = B$ iff both $A \subseteq B$ and $B \subseteq A$.*

Now is also a good opportunity to introduce some further bits of helpful notation. In defining when A is a subset of B we said that "every member of A is ...," and filled the "..." with

"a member of B". But this is such a common *shape* of expression that it will be helpful to introduce some formal notation for it.

Definition 2.9. $(\forall x \in A)\varphi$ abbreviates $\forall x(x \in A \to \varphi)$. Similarly, $(\exists x \in A)\varphi$ abbreviates $\exists x(x \in A \wedge \varphi)$.

Using this notation, we can say that $A \subseteq B$ iff $(\forall x \in A)x \in B$.

Now we move on to considering a certain kind of set: the set of all subsets of a given set.

Definition 2.10 (Power Set). The set consisting of all subsets of a set A is called the *power set of* A, written $\wp(A)$.

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

Example 2.11. What are all the possible subsets of $\{a, b, c\}$? They are: $\emptyset$, $\{a\}$, $\{b\}$, $\{c\}$, $\{a, b\}$, $\{a, c\}$, $\{b, c\}$, $\{a, b, c\}$. The set of all these subsets is $\wp(\{a, b, c\})$:

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

2.3 Some Important Sets

Example 2.12. We will mostly be dealing with sets whose members are mathematical objects. Four such sets are important enough to have specific names:

$\mathbb{N} = \{0, 1, 2, 3, \ldots\}$ the set of natural numbers

$\mathbb{Z} = \{\ldots, -2, -1, 0, 1, 2, \ldots\}$ the set of integers

$\mathbb{Q} = \{m/n : m, n \in \mathbb{Z} \text{ and } n \neq 0\}$ the set of rationals

$\mathbb{R} = (-\infty, \infty)$

the set of real numbers (the continuum)

These are all *infinite* sets, that is, they each have infinitely many members.

As we move through these sets, we are adding *more* numbers to our stock. Indeed, it should be clear that $\mathbb{N} \subseteq \mathbb{Z} \subseteq \mathbb{Q} \subseteq \mathbb{R}$: after all, every natural number is an integer; every integer is a rational; and every rational is a real. Equally, it should be clear that $\mathbb{N} \subsetneq \mathbb{Z} \subsetneq \mathbb{Q}$, since -1 is an integer but not a natural number, and $1/2$ is rational but not integer. It is less obvious that $\mathbb{Q} \subsetneq \mathbb{R}$, i.e., that there are some real numbers which are not rational.

We'll sometimes also use the set of positive integers $\mathbb{Z}^+ = \{1, 2, 3, \dots\}$ and the set containing just the first two natural numbers $\mathbb{B} = \{0, 1\}$.

Example 2.13 (Strings). Another interesting example is the set A^* of *finite strings* over an alphabet A: any finite sequence of elements of A is a string over A. We include the *empty string* Λ among the strings over A, for every alphabet A. For instance,

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

If $x = x_1 \dots x_n \in A^*$is a string consisting of n "letters" from A, then we say *length* of the string is n and write $\text{len}(x) = n$.

Example 2.14 (Infinite sequences). For any set A we may also consider the set A^ω of infinite sequences of members of A. An infinite sequence $a_1 a_2 a_3 a_4 \dots$ consists of a one-way infinite list of objects, each one of which is a member of A.

2.4 Unions and Intersections

In section 2.1, we introduced definitions of sets by abstraction, i.e., definitions of the form $\{x : \varphi(x)\}$. Here, we invoke some property φ, and this property can mention sets we've already

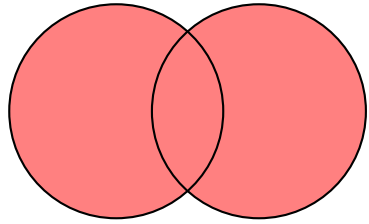


Figure 2.1: The union $A \cup B$ of two sets is set of members of A together with those of B.

defined. So for instance, if A and B are sets, the set $\{x : x \in A \vee x \in B\}$ consists of all those objects which are members of either A or B, i.e., it's the set that combines the members of A and B. We can visualize this as in Figure 2.1, where the highlighted area indicates the members of the two sets A and B together.

This operation on sets—combining them—is very useful and common, and so we give it a formal name and a symbol.

Definition 2.15 (Union). The *union* of two sets A and B, written $A \cup B$, is the set of all things which are members of A, B, or both.

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

Example 2.16. Since the multiplicity of members doesn't matter, the union of two sets which have a member in common contains that member only once, e.g., $\{a, b, c\} \cup \{a, 0, 1\} = \{a, b, c, 0, 1\}$.

The union of a set and one of its subsets is just the bigger set: $\{a, b, c\} \cup \{a\} = \{a, b, c\}$.

The union of a set with the empty set is identical to the set: $\{a, b, c\} \cup \emptyset = \{a, b, c\}$.

We can also consider a "dual" operation to union. This is the operation that forms the set of all members that are members of A and are also members of B. This operation is called *intersection*, and can be depicted as in Figure 2.2.

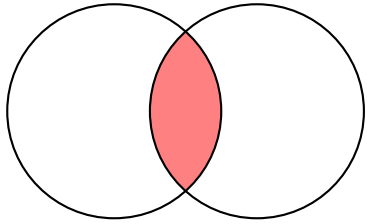


Figure 2.2: The intersection $A \cap B$ of two sets is the set of members they have in common.

Definition 2.17 (Intersection). The *intersection* of two sets A and B, written $A \cap B$, is the set of all things which are members of both A and B.

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

Two sets are called *disjoint* if their intersection is empty. This means they have no members in common.

Example 2.18. If two sets have no members in common, their intersection is empty: $\{a, b, c\} \cap \{0, 1\} = \emptyset$.

If two sets do have members in common, their intersection is the set of all those: $\{a, b, c\} \cap \{a, b, d\} = \{a, b\}$.

The intersection of a set with one of its subsets is just the smaller set: $\{a, b, c\} \cap \{a, b\} = \{a, b\}$.

The intersection of any set with the empty set is empty: $\{a, b, c\} \cap \emptyset = \emptyset$.

We can also form the union or intersection of more than two sets. An elegant way of dealing with this in general is the following: suppose you collect all the sets you want to form the union (or intersection) of into a single set. Then we can define the union of all our original sets as the set of all objects which belong to at least one member of the set, and the intersection as the set of all objects which belong to every member of the set.

Definition 2.19. If A is a set of sets, then $\bigcup A$ is the set of members of members of A:

$$\begin{aligned}\bigcup A &= \{x : x \text{ belongs to a member of } A\}, \text{ i.e.,}\\ &= \{x : \text{there is a } B \in A \text{ so that } x \in B\}\end{aligned}$$

Definition 2.20. If A is a set of sets, then $\bigcap A$ is the set of objects which all elements of A have in common:

$$\begin{aligned}\bigcap A &= \{x : x \text{ belongs to every member of } A\}, \text{ i.e.,}\\ &= \{x : \text{for all } B \in A, x \in B\}\end{aligned}$$

Example 2.21. Suppose $A = \{\{a,b\},\{a,d,e\},\{a,d\}\}$. Then $\bigcup A = \{a,b,d,e\}$ and $\bigcap A = \{a\}$.

We could also do the same for a sequence of sets A_1, A_2, ...

$$\begin{aligned}\bigcup_i A_i &= \{x : x \text{ belongs to one of the } A_i\}\\ \bigcap_i A_i &= \{x : x \text{ belongs to every } A_i\}.\end{aligned}$$

When we have an *index* of sets, i.e., some set I such that we are considering A_i for each $i \in I$, we may also use these abbreviations:

$$\begin{aligned}\bigcup_{i\in I} A_i &= \bigcup\{A_i : i \in I\}\\ \bigcap_{i\in I} A_i &= \bigcap\{A_i : i \in I\}\end{aligned}$$

Finally, we may want to think about the set of all members in A which are not in B. We can depict this as in Figure 2.3.

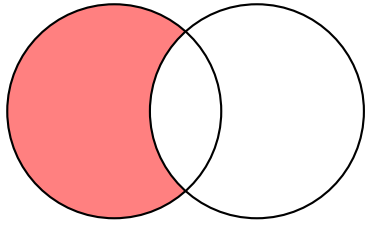


Figure 2.3: The difference $A \setminus B$ of two sets is the set of those members of A which are not also members of B.

Definition 2.22 (Difference). The *set difference* $A \setminus B$ is the set of all members of A which are not also members of B, i.e.,

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

2.5 Pairs, Tuples, Cartesian Products

It follows from extensionality that sets have no order to their elements. So if we want to represent order, we use *ordered pairs* $\langle x, y\rangle$. In an unordered pair $\{x, y\}$, the order does not matter: $\{x, y\} = \{y, x\}$. In an ordered pair, it does: if $x \neq y$, then $\langle x, y\rangle \neq \langle y, x\rangle$.

How should we think about ordered pairs in set theory? Crucially, we want to preserve the idea that ordered pairs are identical iff they share the same first element and share the same second element, i.e.:

$$\langle a, b\rangle = \langle c, d\rangle \text{ iff both } a = c \text{ and } b = d.$$

We can define ordered pairs in set theory using the Wiener–Kuratowski definition.

Definition 2.23 (Ordered pair). $\langle a, b\rangle = \{\{a\}, \{a, b\}\}$.

Having fixed a definition of an ordered pair, we can use it to define further sets. For example, sometimes we also want ordered sequences of more than two objects, e.g., *triples* $\langle x, y, z\rangle$,

quadruples $\langle x, y, z, u\rangle$, and so on. We can think of triples as special ordered pairs, where the first element is itself an ordered pair: $\langle x, y, z\rangle$ is $\langle\langle x, y\rangle, z\rangle$. The same is true for quadruples: $\langle x, y, z, u\rangle$ is $\langle\langle\langle x, y\rangle, z\rangle, u\rangle$, and so on. In general, we talk of *ordered n-tuples* $\langle x_1, \ldots, x_n\rangle$.

Certain sets of ordered pairs, or other ordered n-tuples, will be useful.

Definition 2.24 (Cartesian product). Given sets A and B, their *Cartesian product* $A \times B$ is defined by

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

Example 2.25. If $A = \{0, 1\}$, and $B = \{1, a, b\}$, then their product is

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

Example 2.26. If A is a set, the product of A with itself, $A \times A$, is also written A^2. It is the set of *all* pairs $\langle x, y\rangle$ with $x, y \in A$. The set of all triples $\langle x, y, z\rangle$ is A^3, and so on. We can give a recursive definition:

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

Proposition 2.27. *If A has n members and B has m members, then $A \times B$ has $n \cdot m$ elements.*

Proof. For every member x in A, there are m members of the form $\langle x, y\rangle \in A \times B$. Let $B_x = \{\langle x, y\rangle : y \in B\}$. Since whenever $x_1 \neq x_2$, $\langle x_1, y\rangle \neq \langle x_2, y\rangle$, $B_{x_1} \cap B_{x_2} = \emptyset$. But if $A = \{x_1, \ldots, x_n\}$, then $A \times B = B_{x_1} \cup \cdots \cup B_{x_n}$, and so has $n \cdot m$ members.

To visualize this, arrange the members of $A \times B$ in a grid:

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

Since the x_i are all different, and the y_j are all different, no two of the pairs in this grid are the same, and there are $n \cdot m$ of them.□

Example 2.28. If A is a set, a *word* over A is any sequence of members of A. A sequence can be thought of as an n-tuple of members of A. For instance, if $A = \{a, b, c\}$, then the sequence "bac" can be thought of as the triple $\langle b, a, c\rangle$. Words, i.e., sequences of symbols, are of crucial importance in computer science. By convention, we count members of A as sequences of length 1, and $\emptyset$ as the sequence of length 0. The set of *all* words over A then is

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

2.6 Russell's Paradox

Extensionality licenses the notation $\{x : \varphi(x)\}$, for *the* set of x's such that $\varphi(x)$. However, all that extensionality *really* licenses is the following thought. *If* there is a set whose members are all and only the φ's, *then* there is only one such set. Otherwise put: having fixed some φ, the set $\{x : \varphi(x)\}$ is unique, *if it exists.*

But this conditional is important! Crucially, not every property lends itself to *comprehension.* That is, some properties do *not* define sets. If they all did, then we would run into outright contradictions. The most famous example of this is Russell's Paradox.

Sets may be members of other sets—for instance, the power set of a set A is made up of sets. And so it makes sense to ask or investigate whether a set is a member of another set. Can a set be a member of itself? Nothing about the idea of a set seems to

rule this out. For instance, if *all* sets form a collection of objects, one might think that they can be collected into a single set—the set of all sets. And it, being a set, would be a member of the set of all sets.

Russell's Paradox arises when we consider the property of not having itself as a member, of being *non-self-membered*. What if we suppose that there is a set of all sets that do not have themselves as a member? Does

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

exist? It turns out that we can prove that it does not.

Theorem 2.29 (Russell's Paradox). *There is no set* $R = \{x : x \notin x\}$.

Proof. If $R = \{x : x \notin x\}$ exists, then $R \in R$ iff $R \notin R$, which is a contradiction. □

Let's run through this proof more slowly. If R exists, it makes sense to ask whether $R \in R$ or not. Suppose that indeed $R \in R$. Now, R was defined as the set of all sets that are not members of themselves. So, if $R \in R$, then R does not itself have R's defining property. But only sets that have this property are in R, hence, R cannot be a member of R, i.e., $R \notin R$. But R can't both be and not be a member of R, so we have a contradiction.

Since the assumption that $R \in R$ leads to a contradiction, we have $R \notin R$. But this also leads to a contradiction! For if $R \notin R$, then R itself does have R's defining property, and so R would be a member of R just like all the other non-self-membered sets. And again, it can't both not be and be a member of R.

How do we set up a set theory which avoids falling into Russell's Paradox, i.e., which avoids making the *inconsistent* claim that $R = \{x : x \notin x\}$ exists? Well, we would need to lay down axioms which give us very precise conditions for stating when sets exist (and when they don't).

The set theory sketched in this chapter doesn't do this. It's *genuinely naïve*. It tells you only that sets obey extensionality and

that, if you have some sets, you can form their union, intersection, etc. It is possible to develop set theory more rigorously than this.

Problems

Problem 2.1. Prove that there is at most one empty set, i.e., show that if A and B are sets without members, then $A = B$.

Problem 2.2. List all subsets of $\{a, b, c, d\}$.

Problem 2.3. Show that if A has n members, then $\wp(A)$ has 2^n members.

Problem 2.4. Prove that if $A \subseteq B$, then $A \cup B = B$.

Problem 2.5. Prove rigorously that if $A \subseteq B$, then $A \cap B = A$.

Problem 2.6. Show that if A is a set and $A \in B$, then $A \subseteq \bigcup B$.

Problem 2.7. Prove that if $A \subsetneq B$, then $B \setminus A \neq \emptyset$.

Problem 2.8. Using Definition 2.23, prove that $\langle a, b\rangle = \langle c, d\rangle$ iff both $a = c$ and $b = d$.

Problem 2.9. List all members of $\{1, 2, 3\}^3$.

Problem 2.10. Show, by induction on k, that for all $k \geq 1$, if A has n members, then A^k has n^k members.

CHAPTER 3

Relations

3.1 Relations as Sets

In section 2.3, we mentioned some important sets: $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$. You will no doubt remember some interesting relations between the members of some of these sets. For instance, each of these sets has a completely standard *order relation* on it. There is also the relation *is identical with* that every object bears to itself and to no other thing. There are many more interesting relations that we'll encounter, and even more possible relations. Before we review them, though, we will start by pointing out that we can look at relations as a special sort of set.

For this, recall two things from section 2.5. First, recall the notion of a *ordered pair*: given a and b, we can form $\langle a, b\rangle$. Importantly, the order of elements *does* matter here. So if $a \neq b$ then $\langle a, b\rangle \neq \langle b, a\rangle$. (Contrast this with unordered pairs, i.e., 2-element sets, where $\{a, b\} = \{b, a\}$.) Second, recall the notion of a *Cartesian product*: if A and B are sets, then we can form $A \times B$, the set of all pairs $\langle x, y\rangle$ with $x \in A$ and $y \in B$. In particular, $A^2 = A \times A$ is the set of all ordered pairs from A.

Now we will consider a particular relation on a set: the $<$-relation on the set $\mathbb{N}$ of natural numbers. Consider the set of all pairs of numbers $\langle n, m\rangle$ where $n < m$, i.e.,

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

There is a close connection between n being less than m, and the pair $\langle n, m\rangle$ being a member of R, namely:

$$n < m \text{ iff } \langle n, m\rangle \in R.$$

Indeed, without any loss of information, we can consider the set R to *be* the $<$-relation on $\mathbb{N}$.

In the same way we can construct a subset of $\mathbb{N}^2$ for any relation between numbers. Conversely, given any set of pairs of numbers $S \subseteq \mathbb{N}^2$, there is a corresponding relation between numbers, namely, the relationship n bears to m if and only if $\langle n, m\rangle \in S$. This justifies the following definition:

Definition 3.1 (Binary relation). A *binary relation* on a set A is a subset of A^2. If $R \subseteq A^2$ is a binary relation on A and $x, y \in A$, we sometimes write Rxy (or xRy) for $\langle x, y\rangle \in R$.

Example 3.2. The set $\mathbb{N}^2$ of pairs of natural numbers can be listed in a 2-dimensional matrix like this:

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

We have put the diagonal, here, in bold, since the subset of $\mathbb{N}^2$ consisting of the pairs lying on the diagonal, i.e.,

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

is the *identity relation on* $\mathbb{N}$. (Since the identity relation is popular, let's define $\mathrm{Id}_A = \{\langle x, x\rangle : x \in A\}$ for any set A.) The subset of all pairs lying above the diagonal, i.e.,

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

is the *less than* relation, i.e., Lnm iff $n < m$. The subset of pairs below the diagonal, i.e.,

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

is the *greater than* relation, i.e., Gnm iff $n > m$. The union of L with I, which we might call $K = L \cup I$, is the *less than or equal to* relation: Knm iff $n \leq m$. Similarly, $H = G \cup I$ is the *greater than or equal to relation.* These relations L, G, K, and H are special kinds of relations called *orders.* L and G have the property that no number bears L or G to itself (i.e., for all n, neither Lnn nor Gnn). Relations with this property are called *irreflexive*, and, if they also happen to be orders, they are called *strict orders.*

Although orders and identity are important and natural relations, it should be emphasized that according to our definition *any* subset of A^2 is a relation on A, regardless of how unnatural or contrived it seems. In particular, $\emptyset$ is a relation on any set (the *empty relation*, which no pair of elements bears), and A^2 itself is a relation on A as well (one which every pair bears), called the *universal relation.* But also something like $E = \{\langle n,m\rangle : n > 5 \text{ or } m \times n \geq 34\}$ counts as a relation.

3.2 Philosophical Reflections

In section 3.1, we defined relations as certain sets. We should pause and ask a quick philosophical question: what is such a definition *doing*? It is extremely doubtful that we should want to say that we have *discovered* some metaphysical identity facts; that, for example, the order relation on $\mathbb{N}$ *turned out* to be the set $R = \{\langle n,m\rangle : n,m \in \mathbb{N} \text{ and } n < m\}$ that we defined in section 3.1. Here are three reasons why.

First: in Definition 2.23, we defined $\langle a,b\rangle = \{\{a\},\{a,b\}\}$. Consider instead the definition $\|a,b\| = \{\{b\},\{a,b\}\} = \langle b,a\rangle$. When $a \neq b$, we have that $\langle a,b\rangle \neq \|a,b\|$. But we could equally have regarded $\|a,b\|$ as our definition of an ordered pair, rather

than $\langle a, b\rangle$. Both definitions would have worked equally well. So now we have two equally good candidates to "be" the order relation on the natural numbers, namely:

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

Since $R \neq S$, by extensionality, it is clear that they cannot *both* be identical to the order relation on $\mathbb{N}$. But it would just be arbitrary, and hence a bit embarrassing, to claim that R rather than S (or vice versa) *is* the ordering relation, as a matter of fact. (This is a very simple instance of an argument against set-theoretic reductionism which Benacerraf made famous in 1965. We will revisit it several times.)

Second: if we think that *every* relation should be identified with a set, then the relation of set-membership itself, $\in$, should be a particular set. Indeed, it would have to be the set $\{\langle x, y\rangle : x \in y\}$. But does this set exist? Given Russell's Paradox, it is a non-trivial claim that such a set exists. In fact, it is possible to develop set theory in a rigorous way as an axiomatic theory, and that theory will indeed deny the existence of this set. So, even if some relations can be treated as sets, the relation of set-membership will have to be a special case.

Third: when we "identify" relations with sets, we said that we would allow ourselves to write Rxy for $\langle x, y\rangle \in R$. This is fine, provided that the membership relation, "$\in$", is treated *as* a predicate. But if we think that "$\in$" stands for a certain kind of set, then the expression "$\langle x, y\rangle \in R$" just consists of three singular terms which stand for sets: "$\langle x, y\rangle$", "$\in$", and "R". And such a list of names is no more capable of expressing a proposition than the nonsense string: "the cup penholder the table". Again, even if some relations can be treated as sets, the relation of set-membership must be a special case. (This rolls together a simple version of Frege's concept *horse* paradox, and a famous objection that Wittgenstein once raised against Russell.)

So where does this leave us? Well, there is nothing *wrong* with our saying that the relations on the numbers are sets. We

just have to understand the spirit in which that remark is made. We are not stating a metaphysical identity fact. We are simply noting that, in certain contexts, we can (and will) *treat* (certain) relations as certain sets.

3.3 Special Properties of Relations

Some kinds of relations turn out to be so common that they have been given special names. For instance, $\leq$ and $\subseteq$ both relate their respective domains (say, $\mathbb{N}$ in the case of $\leq$ and $\wp(A)$ in the case of $\subseteq$) in similar ways. To get at exactly how these relations are similar, and how they differ, we categorize them according to some special properties that relations can have. It turns out that (combinations of) some of these special properties are especially important: orders and equivalence relations.

Definition 3.3 (Reflexivity). A relation $R \subseteq A^2$ is *reflexive* iff, for every $x \in A$, Rxx.

Definition 3.4 (Transitivity). A relation $R \subseteq A^2$ is *transitive* iff, whenever Rxy and Ryz, then also Rxz.

Definition 3.5 (Symmetry). A relation $R \subseteq A^2$ is *symmetric* iff, whenever Rxy, then also Ryx.

Definition 3.6 (Anti-symmetry). A relation $R \subseteq A^2$ is *anti-symmetric* iff, whenever both Rxy and Ryx, then $x = y$ (or, in other words: if $x \neq y$ then either $\neg Rxy$ or $\neg Ryx$).

In a symmetric relation, Rxy and Ryx always hold together, or neither holds. In an anti-symmetric relation, the only way for

Rxy and Ryx to hold together is if $x = y$. Note that this does not *require* that Rxy and Ryx holds when $x = y$, only that it isn't ruled out. So an anti-symmetric relation can be reflexive, but it is not the case that every anti-symmetric relation is reflexive. Also note that being anti-symmetric and merely not being symmetric are different conditions. In fact, a relation can be both symmetric and anti-symmetric at the same time (e.g., the identity relation is).

Definition 3.7 (Connectivity). A relation $R \subseteq A^2$ is *connected* if for all $x, y \in A$, if $x \neq y$, then either Rxy or Ryx.

Definition 3.8 (Irreflexivity). A relation $R \subseteq A^2$ is called *irreflexive* if, for all $x \in A$, not Rxx.

Definition 3.9 (Asymmetry). A relation $R \subseteq A^2$ is called *asymmetric* if for no pair $x, y \in A$ we have both Rxy and Ryx.

Note that if $A \neq \emptyset$, then no irreflexive relation on A is reflexive and every asymmetric relation on A is also anti-symmetric. However, there are $R \subseteq A^2$ that are not reflexive and also not irreflexive, and there are anti-symmetric relations that are not asymmetric.

3.4 Equivalence Relations

The identity relation on a set is reflexive, symmetric, and transitive. Relations R that have all three of these properties are very common.

Definition 3.10 (Equivalence relation). A relation $R \subseteq A^2$

that is reflexive, symmetric, and transitive is called an *equivalence relation*. Members x and y of A are said to be *R-equivalent* if Rxy.

Equivalence relations give rise to the notion of an *equivalence class*. An equivalence relation "chunks up" the domain into different partitions. Within each partition, all the objects are related to one another; and no objects from different partitions relate to one another. Sometimes, it's helpful just to talk about these partitions *directly*. To that end, we introduce a definition:

Definition 3.11. Let $R \subseteq A^2$ be an equivalence relation. For each $x \in A$, the *equivalence class* of x in A is the set $[x]_R = \{y \in A : Rxy\}$. The *quotient* of A under R is $A/_R = \{[x]_R : x \in A\}$, i.e., the set of these equivalence classes.

The next result vindicates the definition of an equivalence class, in proving that the equivalence classes are indeed the partitions of A:

Proposition 3.12. *If $R \subseteq A^2$ is an equivalence relation, then Rxy iff $[x]_R = [y]_R$.*

Proof. For the left-to-right direction, suppose Rxy, and let $z \in [x]_R$. By definition, then, Rxz. Since R is an equivalence relation, Ryz. (Spelling this out: as Rxy and R is symmetric we have Ryx, and as Rxz and R is transitive we have Ryz.) So $z \in [y]_R$. Generalising, $[x]_R \subseteq [y]_R$. But exactly similarly, $[y]_R \subseteq [x]_R$. So $[x]_R = [y]_R$, by extensionality.

For the right-to-left direction, suppose $[x]_R = [y]_R$. Since R is reflexive, Ryy, so $y \in [y]_R$. Thus also $y \in [x]_R$ by the assumption that $[x]_R = [y]_R$. So Rxy. □

Example 3.13. A nice example of equivalence relations comes from modular arithmetic. For any a, b, and $n \in \mathbb{Z}^+$, say that $a \equiv_n b$ iff dividing a by n gives the same remainder as dividing b by n. (Somewhat more symbolically: $a \equiv_n b$ iff, for some $k \in \mathbb{Z}$,

$a - b = kn$.) Now, $\equiv_n$ is an equivalence relation, for any n. And there are exactly n distinct equivalence classes generated by $\equiv_n$; that is, $\mathbb{N}/_{\equiv_n}$ has n members. These are: the set of numbers divisible by n without remainder, i.e., $[0]_{\equiv_n}$; the set of numbers divisible by n with remainder 1, i.e., $[1]_{\equiv_n}$; ...; and the set of numbers divisible by n with remainder $n-1$, i.e., $[n-1]_{\equiv_n}$.

3.5 Orders

Many of our comparisons involve describing some objects as being "less than", "equal to", or "greater than" other objects, in a certain respect. These involve *order* relations. But there are different kinds of order relations. For instance, some require that any two objects be comparable, others don't. Some include identity (like $\leq$) and some exclude it (like $<$). It will help us to have a taxonomy here.

Definition 3.14 (Preorder). A relation which is both reflexive and transitive is called a *preorder*.

Definition 3.15 (Partial order). A preorder which is also anti-symmetric is called a *partial order*.

Definition 3.16 (Linear order). A partial order which is also connected is called a *total order* or *linear order*.

Example 3.17. Every linear order is also a partial order, and every partial order is also a preorder, but the converses don't hold. The universal relation on A is a preorder, since it is reflexive and transitive. But, if A has more than one member, the universal relation is not anti-symmetric, and so not a partial order.

Example 3.18. Consider the *no longer than* relation $\preccurlyeq$ on $\mathbb{B}^*$: $x \preccurlyeq y$ iff $\text{len}(x) \leq \text{len}(y)$. This is a preorder (reflexive and transitive),

and even connected, but not a partial order, since it is not anti-symmetric. For instance, $01 \preccurlyeq 10$ and $10 \preccurlyeq 01$, but $01 \neq 10$.

Example 3.19. An important partial order is the relation $\subseteq$ on a set of sets. This is not in general a linear order, since if $a \neq b$ and we consider $\wp(\{a, b\}) = \{\emptyset, \{a\}, \{b\}, \{a, b\}\}$, we see that $\{a\} \not\subseteq \{b\}$ and $\{a\} \neq \{b\}$ and $\{b\} \not\subseteq \{a\}$.

Example 3.20. The relation of *divisibility without remainder* gives us a partial order which isn't a linear order. For integers n, m, we write $n \mid m$ to mean n (evenly) divides m, i.e., iff there is some integer k so that $m = kn$. On $\mathbb{N}$, this is a partial order, but not a linear order: for instance, $2 \nmid 3$ and also $3 \nmid 2$. Considered as a relation on $\mathbb{Z}$, divisibility is only a preorder since it is not anti-symmetric: $1 \mid -1$ and $-1 \mid 1$ but $1 \neq -1$.

Definition 3.21 (Strict order). A *strict order* is a relation which is irreflexive, asymmetric, and transitive.

Definition 3.22 (Strict linear order). A strict order which is also connected is called a *strict total order* or *strict linear order.*

Example 3.23. $\leq$ is the linear order corresponding to the strict linear order $<$. $\subseteq$ is the partial order corresponding to the strict order $\subsetneq$.

Any strict order R on A can be turned into a partial order by adding the diagonal Id_A, i.e., adding all the pairs $\langle x, x\rangle$. (This is called the *reflexive closure* of R.) Conversely, starting from a partial order, one can get a strict order by removing Id_A. These next two results make this precise.

Proposition 3.24. *If R is a strict order on A, then $R^+ = R \cup \mathrm{Id}_A$*

is a partial order. Moreover, if R is a strict linear order, then R^+ is a linear order.

Proof. Suppose R is a strict order, i.e., $R \subseteq A^2$ and R is irreflexive, asymmetric, and transitive. Let $R^+ = R \cup \mathrm{Id}_A$. We have to show that R^+ is reflexive, anti-symmetric, and transitive.

R^+ is clearly reflexive, since $\langle x, x\rangle \in \mathrm{Id}_A \subseteq R^+$ for all $x \in A$.

To show R^+ is anti-symmetric, suppose for reductio that R^+xy and R^+yx but $x \neq y$. Since $\langle x, y\rangle \in R \cup \mathrm{Id}_A$, but $\langle x, y\rangle \notin \mathrm{Id}_A$, we must have $\langle x, y\rangle \in R$, i.e., Rxy. Similarly, Ryx. But this contradicts the assumption that R is asymmetric.

To establish transitivity, suppose that R^+xy and R^+yz. If both $\langle x, y\rangle \in R$ and $\langle y, z\rangle \in R$, then $\langle x, z\rangle \in R$ since R is transitive. Otherwise, either $\langle x, y\rangle \in \mathrm{Id}_A$, i.e., $x = y$, or $\langle y, z\rangle \in \mathrm{Id}_A$, i.e., $y = z$. In the first case, we have that R^+yz by assumption, $x = y$, hence R^+xz. Similarly in the second case. In either case, R^+xz, thus, R^+ is also transitive.

Concerning the "moreover" clause, suppose that R is also connected. So for all $x \neq y$, either Rxy or Ryx, i.e., either $\langle x, y\rangle \in R$ or $\langle y, x\rangle \in R$. Since $R \subseteq R^+$, this remains true of R^+, so R^+ is connected as well. □

Proposition 3.25. *If R is a partial order on A, then $R^- = R \setminus \mathrm{Id}_A$ is a strict order. Moreover, if R is a linear order, then R^- is a strict linear order.*

Proof. This is left as an exercise. □

The following simple result establishes that strict linear orders satisfy an extensionality-like property:

Proposition 3.26. *If $<$ is a strict linear order on A, then:*

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

Proof. Suppose $(\forall x \in A)(x < a \leftrightarrow x < b)$. If $a < b$, then $a < a$, contradicting the fact that $<$ is irreflexive; so $a \not< b$. Exactly similarly, $b \not< a$. So $a = b$, as $<$ is connected. □

3.6 Operations on Relations

It is often useful to modify or combine relations. In Proposition 3.24, we considered the *union* of relations, which is just the union of two relations considered as sets of pairs. Similarly, in Proposition 3.25, we considered the relative difference of relations. Here are some other operations we can perform on relations.

Definition 3.27. Let R, S be relations, and A be any set.

The *inverse* of R is $R^{-1} = \{\langle y, x\rangle : \langle x, y\rangle \in R\}$.

The *relative product* of R and S is $(R \mid S) = \{\langle x, z\rangle : \exists y(Rxy \wedge Syz)\}$.

The *restriction* of R to A is $R\restriction_A = R \cap A^2$.

The *application* of R to A is $R[A] = \{y : (\exists x \in A)Rxy\}$

Example 3.28. Let $S \subseteq \mathbb{Z}^2$ be the successor relation on $\mathbb{Z}$, i.e., $S = \{\langle x, y\rangle \in \mathbb{Z}^2 : x + 1 = y\}$, so that Sxy iff $x + 1 = y$.

S^{-1} is the predecessor relation on $\mathbb{Z}$, i.e., $\{\langle x, y\rangle \in \mathbb{Z}^2 : x - 1 = y\}$.

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

$S\restriction_{\mathbb{N}}$ is the successor relation on $\mathbb{N}$.

$S[\{1, 2, 3\}]$ is $\{2, 3, 4\}$.

Definition 3.29 (Transitive closure). Let $R \subseteq A^2$ be a binary relation.

The *transitive closure* of R is $R^+ = \bigcup_{0<n\in\mathbb{N}} R^n$, where we recursively define $R^1 = R$ and $R^{n+1} = R^n \mid R$.

The *reflexive transitive closure* of R is $R^* = R^+ \cup \mathrm{Id}_A$.

Example 3.30. Take the successor relation $S \subseteq \mathbb{Z}^2$. S^2xy iff $x + 2 = y$, S^3xy iff $x + 3 = y$, etc. So S^+xy iff $x + n = y$ for some $n \geq 1$. In other words, S^+xy iff $x < y$, and S^*xy iff $x \leq y$.

Problems

Problem 3.1. List the members of the relation $\subseteq$ on the set $\wp(\{a, b, c\})$.

Problem 3.2. Give examples of relations that are (a) reflexive and symmetric but not transitive, (b) reflexive and anti-symmetric, (c) anti-symmetric, transitive, but not reflexive, and (d) reflexive, symmetric, and transitive. Do not use relations on numbers or sets.

Problem 3.3. Show that $\equiv_n$ is an equivalence relation, for any $n \in \mathbb{Z}^+$, and that $\mathbb{N}/_{\equiv_n}$ has exactly n members.

Problem 3.4. Give a proof of Proposition 3.25.

Problem 3.5. Show that the transitive closure of R is in fact transitive.

CHAPTER 4

Functions

4.1 Basics

A *function* is a map which sends each member of a given set to a specific member in some (other) given set. For instance, the operation of adding 1 defines a function: each number n is mapped to a unique number $n + 1$.

More generally, functions may take pairs, triples, etc., as inputs and return some kind of output. Many functions are familiar to us from basic arithmetic. For instance, addition and multiplication are functions. They take in two numbers and return a third.

In this mathematical, abstract sense, a function is a *black box*: what matters is only what output is paired with what input, not the method for calculating the output.

Definition 4.1 (Function). A *function* $f : A \to B$ is a mapping of each member of A to an member of B.

We call A the *domain* of f and B the *codomain* of f. The members of A are called inputs or *arguments* of f, and the member of B that is paired with an argument x by f is called the *value of* f for argument x, written $f(x)$.

The *range* $\text{ran}(f)$ of f is the subset of the codomain consisting of the values of f for some argument; $\text{ran}(f) = \{f(x) : x \in A\}$.

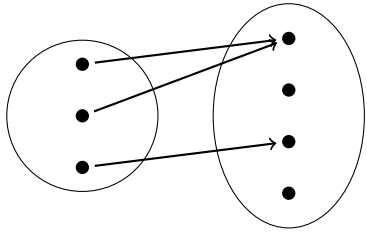


Figure 4.1: A function is a mapping of each member of one set to a member of another. An arrow points from an argument in the domain to the corresponding value in the codomain.

The diagram in Figure 4.1 may help to think about functions. The ellipse on the left represents the function's *domain*; the ellipse on the right represents the function's *codomain*; and an arrow points from an *argument* in the domain to the corresponding *value* in the codomain.

Example 4.2. Multiplication takes pairs of natural numbers as inputs and maps them to natural numbers as outputs, so goes from $\mathbb{N} \times \mathbb{N}$ (the domain) to $\mathbb{N}$ (the codomain). As it turns out, the range is also $\mathbb{N}$, since every $n \in \mathbb{N}$ is $n \times 1$.

Example 4.3. Multiplication is a function because it pairs each input—each pair of natural numbers—with a single output: $\times : \mathbb{N}^2 \to \mathbb{N}$. By contrast, the square root operation applied to the domain $\mathbb{N}$ is not functional, since each positive integer n has two square roots: $\sqrt{n}$ and $-\sqrt{n}$. We can make it functional by only returning the positive square root: $\sqrt{\ } : \mathbb{N} \to \mathbb{R}$.

Example 4.4. The relation that pairs each student in a class with their final grade is a function—no student can get two different final grades in the same class. The relation that pairs each student in a class with their parents is not a function: students can have zero, or two, or more parents.

We can define functions by specifying in some precise way what the value of the function is for every possible argument.

Different ways of doing this are by giving a formula, describing a method for computing the value, or listing the values for each argument. However functions are defined, we must make sure that for each argument we specify one, and only one, value.

Example 4.5. Let $f : \mathbb{N} \to \mathbb{N}$ be defined such that $f(x) = x + 1$. This is a definition that specifies f as a function which takes in natural numbers and outputs natural numbers. It tells us that, given a natural number x, f will output its successor $x + 1$. In this case, the codomain $\mathbb{N}$ is not the range of f, since the natural number 0 is not the successor of any natural number. The range of f is the set of all positive integers, $\mathbb{Z}^+$.

Example 4.6. Let $g : \mathbb{N} \to \mathbb{N}$ be defined such that $g(x) = x+2-1$. This tells us that g is a function which takes in natural numbers and outputs natural numbers. Given a natural number n, g will output the predecessor of the successor of the successor of x, i.e., $x + 1$.

We just considered two functions, f and g, with different *definitions*. However, these are the *same function*. After all, for any natural number n, we have that $f(n) = n + 1 = n + 2 - 1 = g(n)$. Otherwise put: our definitions for f and g specify the same mapping by means of different equations. Implicitly, then, we are relying upon a principle of extensionality for functions,

$$\text{if } \forall x\, f(x) = g(x), \text{ then } f = g$$

provided that f and g share the same domain and codomain.

Example 4.7. We can also define functions by cases. For instance, we could define $h : \mathbb{N} \to \mathbb{N}$ by

$$h(x) = \begin{cases} \frac{x}{2} & \text{if } x \text{ is even} \\ \frac{x+1}{2} & \text{if } x \text{ is odd.} \end{cases}$$

Since every natural number is either even or odd, the output of this function will always be a natural number. Just remember that

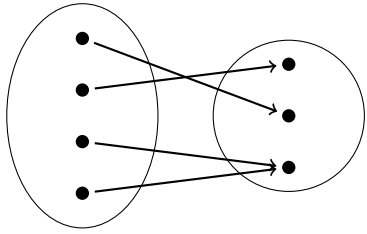


Figure 4.2: A surjective function has every member of the codomain as a value.

if you define a function by cases, every possible input must fall into exactly one case. In some cases, this will require a proof that the cases are exhaustive and exclusive.

4.2 Kinds of Functions

It will be useful to introduce a kind of taxonomy for some of the kinds of functions which we encounter most frequently.

To start, we might want to consider functions which have the property that every member of the codomain is a value of the function. Such functions are called surjective, and can be pictured as in Figure 4.2.

Definition 4.8 (Surjective function). A function $f : A \to B$ is *surjective* iff B is also the range of f, i.e., for every $y \in B$ there is at least one $x \in A$ such that $f(x) = y$, or in symbols:

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

We call such a function a surjection from A to B.

If you want to show that f is a surjection, then you need to show that every object in f's codomain is the value of $f(x)$ for some input x.

Note that any function *induces* a surjection. After all, given a function $f : A \to B$, let $f' : A \to \text{ran}(f)$ be defined by $f'(x) =$

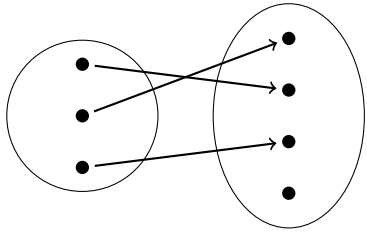


Figure 4.3: An injective function never maps two different arguments to the same value.

$f(x)$. Since $\text{ran}(f)$ is *defined* as $\{f(x) \in B : x \in A\}$, this function f' is guaranteed to be a surjection

Now, any function maps each possible input to a unique output. But there are also functions which never map different inputs to the same outputs. Such functions are called injective, and can be pictured as in Figure 4.3.

Definition 4.9 (Injective function). A function $f : A \to B$ is *injective* iff for each $y \in B$ there is at most one $x \in A$ such that $f(x) = y$. We call such a function an injection from A to B.

If you want to show that f is an injection, you need to show that for any members x and y of f's domain, if $f(x) = f(y)$, then $x = y$.

Example 4.10. The constant function $f : \mathbb{N} \to \mathbb{N}$ given by $f(x) = 1$ is neither injective, nor surjective.

The identity function $f : \mathbb{N} \to \mathbb{N}$ given by $f(x) = x$ is both injective and surjective.

The successor function $f : \mathbb{N} \to \mathbb{N}$ given by $f(x) = x + 1$ is injective but not surjective.

The function $f : \mathbb{N} \to \mathbb{N}$ defined by:

$$f(x) = \begin{cases} \frac{x}{2} & \text{if } x \text{ is even} \\ \frac{x+1}{2} & \text{if } x \text{ is odd.} \end{cases}$$

is surjective, but not injective.

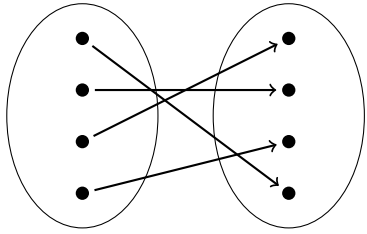


Figure 4.4: A bijective function uniquely pairs the elements of the codomain with those of the domain.

Often enough, we want to consider functions which are both injective and surjective. We call such functions bijective. They look like the function pictured in Figure 4.4. Bijections are also sometimes called *one-to-one correspondences*, since they uniquely pair elements of the codomain with elements of the domain.

Definition 4.11 (Bijection). A function $f : A \to B$ is *bijective* iff it is both surjective and injective. We call such a function a bijection from A to B (or between A and B).

4.3 Functions as Relations

A function which maps members of A to members of B obviously defines a relation between A and B, namely the relation which holds between x and y iff $f(x) = y$. In fact, we might even—if we are interested in reducing the building blocks of mathematics for instance—*identify* the function f with this relation, i.e., with a set of pairs. This then raises the question: which relations define functions in this way?

Definition 4.12 (Graph of a function). Let $f : A \to B$ be a function. The *graph* of f is the relation $R_f \subseteq A \times B$ defined by

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

The graph of a function is uniquely determined, by extensionality. Moreover, extensionality (on sets) will immediately vindicate the implicit principle of extensionality for functions, whereby if f and g share a domain and codomain then they are identical if they agree on all values.

Similarly, if a relation is "functional", then it is the graph of a function.

Proposition 4.13. *Let $R \subseteq A \times B$ be such that:*

1. *If Rxy and Rxz then $y = z$; and*
2. *for every $x \in A$ there is some $y \in B$ such that $\langle x, y \rangle \in R$.*

Then R is the graph of the function $f : A \to B$ defined by $f(x) = y$ iff Rxy.

Proof. Suppose there is a y such that Rxy. If there were another $z \neq y$ such that Rxz, the condition on R would be violated. Hence, if there is a y such that Rxy, this y is unique, and so f is well-defined. Obviously, $R_f = R$. □

Every function $f : A \to B$ has a graph, i.e., a relation on $A \times B$ defined by $f(x) = y$. On the other hand, every relation $R \subseteq A \times B$ with the properties given in Proposition 4.13 is the graph of a function $f : A \to B$. Because of this close connection between functions and their graphs, we can think of a function simply as its graph. In other words, functions can be identified with certain relations, i.e., with certain sets of tuples. Note, though, that the spirit of this "identification" is as in section 3.2: it is not a claim about the metaphysics of functions, but an observation that it is convenient to *treat* functions as certain sets. One reason that this is so convenient, is that we can now consider performing similar operations on functions as we performed on relations (see section 3.6). In particular:

Definition 4.14. Let $f : A \to B$ be a function with $C \subseteq A$.

The *restriction* of f to C is the function $f \restriction_C : C \to B$ defined by $(f \restriction_C)(x) = f(x)$ for all $x \in C$. In other words, $f \restriction_C = \{\langle x, y\rangle \in R_f : x \in C\}$.

The *application* of f to C is $f[C] = \{f(x) : x \in C\}$. We also call this the *image* of C under f.

It follows from these definitions that $\mathrm{ran}(f) = f[\mathrm{dom}(f)]$, for any function f. These notions are exactly as one would expect, given the definitions in section 3.6 and our identification of functions with relations. But two other operations—inverses and relative products—require a little more detail. We will provide that in section 4.4 and section 4.5.

4.4 Inverses of Functions

We think of functions as maps. An obvious question to ask about functions, then, is whether the mapping can be "reversed." For instance, the successor function $f(x) = x + 1$ can be reversed, in the sense that the function $g(y) = y - 1$ "undoes" what f does.

But we must be careful. Although the definition of g defines a function $\mathbb{Z} \to \mathbb{Z}$, it does not define a *function* $\mathbb{N} \to \mathbb{N}$, since $g(0) \notin \mathbb{N}$. So even in simple cases, it is not quite obvious whether a function can be reversed; it may depend on the domain and codomain.

This is made more precise by the notion of an inverse of a function.

Definition 4.15. A function $g : B \to A$ is an *inverse* of a function $f : A \to B$ if $f(g(y)) = y$ and $g(f(x)) = x$ for all $x \in A$ and $y \in B$.

If f has an inverse g, we often write f^{-1} instead of g.

Now we will determine when functions have inverses. A good candidate for an inverse of $f : A \to B$ is $g : B \to A$ "defined by"

$$g(y) = \text{"the" } x \text{ such that } f(x) = y.$$

But the scare quotes around "defined by" (and "the") suggest that this is not a definition. At least, it will not always work, with complete generality. For, in order for this definition to specify a function, there has to be one and only one x such that $f(x) = y$—the output of g has to be uniquely specified. Moreover, it has to be specified for every $y \in B$. If there are x_1 and $x_2 \in A$ with $x_1 \neq x_2$ but $f(x_1) = f(x_2)$, then $g(y)$ would not be uniquely specified for $y = f(x_1) = f(x_2)$. And if there is no x at all such that $f(x) = y$, then $g(y)$ is not specified at all. In other words, for g to be defined, f must be both injective and surjective.

Let's go slowly. We'll divide the question into two: Given a function $f : A \to B$, when is there a function $g : B \to A$ so that $g(f(x)) = x$? Such a g "undoes" what f does, and is called a *left inverse* of f. Secondly, when is there a function $h : B \to A$ so that $f(h(y)) = y$? Such an h is called a *right inverse* of f—f "undoes" what h does.

Proposition 4.16. *If $f : A \to B$ is injective, then there is a* left inverse *$g : B \to A$ of f so that $g(f(x)) = x$ for all $x \in A$.*

Proof. Suppose that $f : A \to B$ is injective. Consider a $y \in B$. If $y \in \operatorname{ran}(f)$, there is an $x \in A$ so that $f(x) = y$. Because f is injective, there is only one such $x \in A$. Then we can define: $g(y) = x$, i.e., $g(y)$ is "the" $x \in A$ such that $f(x) = y$. If $y \notin \operatorname{ran}(f)$, we can map it to any $a \in A$. So, we can pick an $a \in A$ and define $g : B \to A$ by:

$$g(y) = \begin{cases} x & \text{if } f(x) = y \\ a & \text{if } y \notin \operatorname{ran}(f). \end{cases}$$

It is defined for all $y \in B$, since for each such $y \in \operatorname{ran}(f)$ there is exactly one $x \in A$ such that $f(x) = y$. By definition, if $y = f(x)$, then $g(y) = x$, i.e., $g(f(x)) = x$. □

Proposition 4.17. *If $f : A \to B$ is surjective, then there is a* right inverse *$h : B \to A$ of f so that $f(h(y)) = y$ for all $y \in B$.*

Proof. Suppose that $f : A \to B$ is surjective. Consider a $y \in B$. Since f is surjective, there is an $x_y \in A$ with $f(x_y) = y$. Then we can define: $h(y) = x_y$, i.e., for each $y \in B$ we choose some $x \in A$ so that $f(x) = y$; since f is surjective there is always at least one to choose from.[1] By definition, if $x = h(y)$, then $f(x) = y$, i.e., for any $y \in B$, $f(h(y)) = y$. □

By combining the ideas in the previous proof, we now get that every bijection has an inverse, i.e., there is a single function which is both a left and right inverse of f.

Proposition 4.18. *If $f : A \to B$ is bijective, there is a function $f^{-1} : B \to A$ so that for all $x \in A$, $f^{-1}(f(x)) = x$ and for all $y \in B$, $f(f^{-1}(y)) = y$.*

Proof. Exercise. □

There is a slightly more general way to extract inverses. We saw in section 4.2 that every function f induces a surjection $f' : A \to \mathrm{ran}(f)$ by letting $f'(x) = f(x)$ for all $x \in A$. Clearly, if f is injective, then f' is bijective, so that it has a unique inverse by Proposition 4.18. By a very minor abuse of notation, we sometimes call the inverse of f' simply "the inverse of f."

[1] Since f is surjective, for every $y \in B$ the set $\{x : f(x) = y\}$ is nonempty. Our definition of h requires that we choose a single x from each of these sets. That this is always possible is actually not obvious—the possibility of making these choices is simply assumed as an axiom. In other words, this proposition assumes the so-called Axiom of Choice, an issue we will revisit in chapter 16. However, in many specific cases, e.g., when $A = \mathbb{N}$ or is finite, or when f is bijective, the Axiom of Choice is not required. (In the particular case when f is bijective, for each $y \in B$ the set $\{x : f(x) = y\}$ has exactly one member, so that there is no choice to make.)

Proposition 4.19. *Show that if $f : A \to B$ has a left inverse g and a right inverse h, then $h = g$.*

Proof. Exercise. □

Proposition 4.20. *Every function f has at most one inverse.*

Proof. Suppose g and h are both inverses of f. Then in particular g is a left inverse of f and h is a right inverse. By Proposition 4.19, $g = h$. □

4.5 Composition of Functions

We saw in section 4.4 that the inverse f^{-1} of a bijection f is itself a function. Another operation on functions is composition: we can define a new function by composing two functions, f and g, i.e., by first applying f and then g. Of course, this is only possible if the ranges and domains match, i.e., the range of f must be a subset of the domain of g. This operation on functions is the analogue of the operation of relative product on relations from section 3.6.

A diagram might help to explain the idea of composition. In Figure 4.5, we depict two functions $f : A \to B$ and $g : B \to C$ and their composition $(g \circ f)$. The function $(g \circ f) : A \to C$ pairs each member of A with a member of C. We specify which member of C a member of A is paired with as follows: given an input $x \in A$, first apply the function f to x, which will output some $f(x) = y \in B$, then apply the function g to y, which will output some $g(f(x)) = g(y) = z \in C$.

Definition 4.21 (Composition). Let $f : A \to B$ and $g : B \to C$ be functions. The *composition* of f with g is $g \circ f : A \to C$, where $(g \circ f)(x) = g(f(x))$.

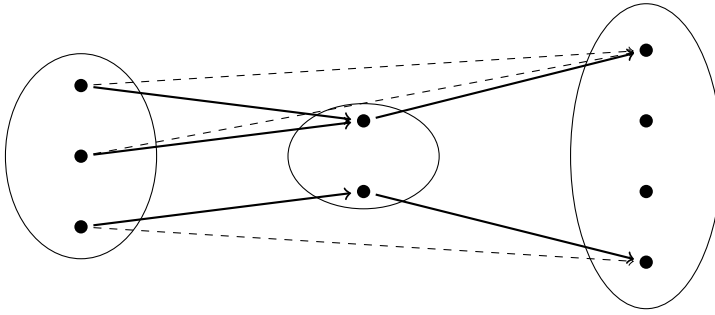


Figure 4.5: The composition $g \circ f$ of two functions f and g.

Example 4.22. Consider the functions $f(x) = x+1$, and $g(x) = 2x$. Since $(g \circ f)(x) = g(f(x))$, for each input x you must first take its successor, then multiply the result by two. So their composition is given by $(g \circ f)(x) = 2(x+1)$.

Problems

Problem 4.1. Show that if $f : A \to B$ has a left inverse g, then f is injective.

Problem 4.2. Show that if $f : A \to B$ has a right inverse h, then f is surjective.

Problem 4.3. Prove Proposition 4.18. You have to define f^{-1}, show that it is a function, and show that it is an inverse of f, i.e., $f^{-1}(f(x)) = x$ and $f(f^{-1}(y)) = y$ for all $x \in A$ and $y \in B$.

Problem 4.4. Prove Proposition 4.19.

Problem 4.5. Show that if $f : A \to B$ and $g : B \to C$ are both injective, then $g \circ f : A \to C$ is injective.

Problem 4.6. Show that if $f : A \to B$ and $g : B \to C$ are both surjective, then $g \circ f : A \to C$ is surjective.

Problem 4.7. Suppose $f : A \to B$ and $g : B \to C$. Show that the graph of $g \circ f$ is $R_f \mid R_g$.

CHAPTER 5

The Size of Sets

5.1 Introduction

When Georg Cantor developed set theory in the 1870s, one of his aims was to make palatable the idea of an infinite collection—an actual infinity, as the medievals would say. A key part of this was his treatment of the *size* of different sets. If a, b and c are all distinct, then the set $\{a, b, c\}$ is intuitively *larger* than $\{a, b\}$. But what about infinite sets? Are they all as large as each other? It turns out that they are not.

The first important idea here is that of an enumeration. We can list every finite set by listing all its members. For some infinite sets, we can also list all their members if we allow the list itself to be infinite. Such sets are called countable. Cantor's surprising result, which we will fully understand by the end of this chapter, was that some infinite sets are not countable.

5.2 Enumerations and Countable Sets

We can specify finite set is by simply enumerating its members. We do this when we define a set like so:

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

Assuming that the members $a_1, \ldots, a_n$ are all distinct, this gives us a bijection between A and the first n natural numbers $0, \ldots,$

$n-1$. Conversely, since every finite set has only finitely many members, every finite set can be put into such a correspondence. In other words, if A is finite, there is a bijection between A and $\{0,\ldots,n-1\}$, where n is the number of members of A.

If we allow for certain kinds of infinite sets, then we will also allow some infinite sets to be enumerated. We can make this precise by saying that an infinite set is enumerated by a bijection between it and all of $\mathbb{N}$.

Definition 5.1 (Enumeration, set-theoretic). An *enumeration* of a set A is a bijection whose range is A and whose domain is either an initial set of natural numbers $\{0,1,\ldots,n\}$ or the entire set of natural numbers $\mathbb{N}$.

There is an intuitive underpinning to this use of the word *enumeration*. For to say that we have enumerated a set A is to say that there is a bijection f which allows us to count out the elements of the set A. The 0th element is $f(0)$, the 1st is $f(1)$, ...the nth is $f(n)$....[1] The rationale for this may be made even clearer by adding the following:

Definition 5.2. A set A is countable iff either $A = \emptyset$ or there is an enumeration of A. We say that A is uncountable iff A is not countable.

So a set is countable iff it is empty or you can use an enumeration to count out its members.

Example 5.3. A function enumerating the natural numbers is simply the identity function $\mathrm{Id}_{\mathbb{N}}\colon \mathbb{N} \to \mathbb{N}$ given by $\mathrm{Id}_{\mathbb{N}}(n) = n$. A function enumerating the *positive* natural numbers, $\mathbb{N}^+ = \mathbb{N} \setminus \{0\}$, is the function $g(n) = n + 1$, i.e., the successor function.

[1] Yes, we count from 0. Of course we could also start with 1. This would make no big difference. We would just have to replace $\mathbb{N}$ by $\mathbb{Z}^+$.

Example 5.4. The functions $f : \mathbb{N} \to \mathbb{N}$ and $g : \mathbb{N} \to \mathbb{N}$ given by

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

respectively enumerate the even natural numbers and the odd natural numbers. But neither is surjective, so neither is an enumeration of $\mathbb{N}$.

Example 5.5. Let $\lceil x \rceil$ be the *ceiling* function, which rounds x up to the nearest integer. Then the function $f : \mathbb{N} \to \mathbb{Z}$ given by:

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

enumerates the set of integers $\mathbb{Z}$ as follows:

$f(0)$	$f(1)$	$f(2)$	$f(3)$	$f(4)$	$f(5)$	$f(6)$	$\ldots$
$\left\lceil \frac{0}{2} \right\rceil$	$-\left\lceil \frac{1}{2} \right\rceil$	$\left\lceil \frac{2}{2} \right\rceil$	$-\left\lceil \frac{3}{2} \right\rceil$	$\left\lceil \frac{4}{2} \right\rceil$	$-\left\lceil \frac{5}{2} \right\rceil$	$\left\lceil \frac{6}{2} \right\rceil$	$\ldots$
0	-1	1	-2	2	-3	3	$\ldots$

Notice how f generates the values of $\mathbb{Z}$ by "hopping" back and forth between positive and negative integers. You can also think of f as defined by cases as follows:

$$f(n) = \begin{cases} \frac{n}{2} & \text{if } n \text{ is even} \\ -\frac{n+1}{2} & \text{if } n \text{ is odd} \end{cases}$$

5.3 Cantor's Zig-Zag Method

We've already considered some "easy" enumerations. Now we will consider something a bit harder. Consider the set of pairs of natural numbers, which we defined in section 2.5 thus:

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

We can organize these ordered pairs into an *array*, like so:

	0	**1**	**2**	**3**	...
0	$\langle 0,0\rangle$	$\langle 0,1\rangle$	$\langle 0,2\rangle$	$\langle 0,3\rangle$	...
1	$\langle 1,0\rangle$	$\langle 1,1\rangle$	$\langle 1,2\rangle$	$\langle 1,3\rangle$	...
2	$\langle 2,0\rangle$	$\langle 2,1\rangle$	$\langle 2,2\rangle$	$\langle 2,3\rangle$	...
3	$\langle 3,0\rangle$	$\langle 3,1\rangle$	$\langle 3,2\rangle$	$\langle 3,3\rangle$	...
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\ddots$

Clearly, every ordered pair in $\mathbb{N} \times \mathbb{N}$ will appear exactly once in the array. In particular, $\langle n, m\rangle$ will appear in the nth row and mth column. But how do we organize the elements of such an array into a "one-dimensional" list? The pattern in the array below demonstrates one way to do this (although of course there are many other options):

	0	**1**	**2**	**3**	**4**	...
0	0	1	3	6	10	...
1	2	4	7	11	...	...
2	5	8	12	...	...	...
3	9	13	...	...	...	...
4	14	...	...	...	...	...
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	...	$\ddots$

This pattern is called *Cantor's zig-zag method*. It enumerates $\mathbb{N} \times \mathbb{N}$ as follows:

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

And this establishes the following:

Proposition 5.6. *$\mathbb{N} \times \mathbb{N}$ is countable.*

Proof. Let $f : \mathbb{N} \to \mathbb{N} \times \mathbb{N}$ take each $k \in \mathbb{N}$ to the tuple $\langle n, m\rangle \in \mathbb{N} \times \mathbb{N}$ such that k is the value of the nth row and mth column in Cantor's zig-zag array. □

This technique also generalises rather nicely. For example, we can use it to enumerate the set of ordered triples of natural numbers, i.e.:

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

We think of $\mathbb{N} \times \mathbb{N} \times \mathbb{N}$ as the Cartesian product of $\mathbb{N} \times \mathbb{N}$ with $\mathbb{N}$, that is,

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

and thus we can enumerate $\mathbb{N}^3$ with an array by labelling one axis with the enumeration of $\mathbb{N}$, and the other axis with the enumeration of $\mathbb{N}^2$:

	0	**1**	**2**	**3**	...
$\langle \mathbf{0}, \mathbf{0}\rangle$	$\langle 0,0,0\rangle$	$\langle 0,0,1\rangle$	$\langle 0,0,2\rangle$	$\langle 0,0,3\rangle$	...
$\langle \mathbf{0}, \mathbf{1}\rangle$	$\langle 0,1,0\rangle$	$\langle 0,1,1\rangle$	$\langle 0,1,2\rangle$	$\langle 0,1,3\rangle$	...
$\langle \mathbf{1}, \mathbf{0}\rangle$	$\langle 1,0,0\rangle$	$\langle 1,0,1\rangle$	$\langle 1,0,2\rangle$	$\langle 1,0,3\rangle$	...
$\langle \mathbf{0}, \mathbf{2}\rangle$	$\langle 0,2,0\rangle$	$\langle 0,2,1\rangle$	$\langle 0,2,2\rangle$	$\langle 0,2,3\rangle$	...
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\ddots$

Thus, by using a method like Cantor's zig-zag method, we may similarly obtain an enumeration of $\mathbb{N}^3$. And we can keep going, obtaining enumerations of $\mathbb{N}^n$ for any natural number n. So, we have:

Proposition 5.7. *$\mathbb{N}^n$ is countable, for every $n \in \mathbb{N}$.*

5.4 Pairing Functions and Codes

Cantor's zig-zag method makes the enumerability of $\mathbb{N}^n$ visually evident. But let us focus on our array depicting $\mathbb{N}^2$. Following the zig-zag line in the array and counting the places, we can check that $\langle 1,2\rangle$ is associated with the number 7. However, it would be nice if we could compute this more directly. That is, it would

be nice to have to hand the *inverse* of the zig-zag enumeration, $g\colon \mathbb{N}^2 \to \mathbb{N}$, such that

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

This would enable us to calculate exactly where $\langle n,m\rangle$ will occur in our enumeration.

In fact, we can define g directly by making two observations. First: if the nth row and mth column contains value v, then the $(n+1)$st row and $(m-1)$st column contains value $v+1$. Second: the first row of our enumeration consists of the triangular numbers, starting with 0, 1, 3, 6, etc. The kth triangular number is the sum of the natural numbers $< k$, which can be computed as $k(k+1)/2$. Putting these two observations together, consider this function:

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

We often just write $g(n,m)$ rather that $g(\langle n,m\rangle)$, since it is easier on the eyes. This tells you first to determine the $(n+m)^{\text{th}}$ triangle number, and then add n to it. And it populates the array in exactly the way we would like. So in particular, the pair $\langle 1,2\rangle$ is sent to $\frac{4\times 3}{2} + 1 = 7$.

This function g is the *inverse* of an enumeration of a set of pairs. Such functions are called *pairing functions*.

Definition 5.8 (Pairing function). A function $f\colon A\times B \to \mathbb{N}$ is an arithmetical *pairing function* if f is injective. We also say that f *encodes* $A\times B$, and that $f(x,y)$ is the *code* for $\langle x,y\rangle$.

We can use pairing functions to encode, e.g., pairs of natural numbers; or, in other words, we can represent each *pair* of elements using a *single* number. Using the inverse of the pairing function, we can *decode* the number, i.e., find out which pair it represents.

5.5 Uncountable Sets

The set $\mathbb{N}$ of natural numbers is infinite. It is also trivially countable. But the remarkable fact is that there are *uncountable* sets, i.e., sets which are not countable (see Definition 5.2).

This might be surprising. After all, to say that A is uncountable is to say that there is *no* bijection $f : \mathbb{N} \to A$; that is, no function mapping the infinitely many members of $\mathbb{N}$ to A exhausts all of A. So if A is uncountable, there are "more" members of A than there are natural numbers.

To prove that a set is uncountable, you have to show that no appropriate bijection can exist. The best way to do this is to show that every attempt to enumerate members of A must leave at least one member out; this shows that no function $f : \mathbb{N} \to A$ is surjective. And a general strategy for establishing this is to use Cantor's *diagonal method*. Given a list of members of A, say, x_1, x_2, ..., we construct another member of A which, by its construction, cannot possibly be on that list.

But all of this is best understood by example. So, our first example is the set $\mathbb{B}^\omega$ of all infinite strings of 0's and 1's. (The '$\mathbb{B}$' stands for binary, and we can just think of it as the two-element set $\{0,1\}$.)

Theorem 5.9. *$\mathbb{B}^\omega$ is uncountable.*

Proof. Consider any enumeration of a subset of $\mathbb{B}^\omega$. So we have some list s_0, s_1, s_2, ... where every s_n is an infinite string of 0's and 1's. Let $s_n(m)$ be the nth digit of the mth string in this list. So we can now think of our list as an array, where $s_n(m)$ is placed at the nth row and mth column:

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

We will now construct an infinite string, d, of 0's and 1's which is not on this list. We will do this by specifying each of its entries, i.e., we specify $d(n)$ for all $n \in \mathbb{N}$. Intuitively, we do this by reading down the diagonal of the array above (hence the name "diagonal method") and then changing every 1 to a 0 and every 1 to a 0. More abstractly, we define $d(n)$ to be 0 or 1 according to whether the n-th member of the diagonal, $s_n(n)$, is 1 or 0, that is:

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

Clearly $d \in \mathbb{B}^\omega$, since it is an infinite string of 0's and 1's. But we have constructed d so that $d(n) \neq s_n(n)$ for any $n \in \mathbb{N}$. That is, d differs from s_n in its nth entry. So $d \neq s_n$ for any $n \in \mathbb{N}$. So d cannot be on the list $s_0, s_1, s_2, \ldots$

We have shown, given an arbitrary enumeration of some subset of $\mathbb{B}^\omega$, that it will omit some member of $\mathbb{B}^\omega$. So there is no enumeration of the set $\mathbb{B}^\omega$, i.e., $\mathbb{B}^\omega$ is uncountable. □

This proof method is called "diagonalization" because it uses the diagonal of the array to define d. However, diagonalization need not involve the presence of an array. Indeed, we can show that some set is uncountable by using a similar idea, even when no array and no actual diagonal is involved. The following result illustrates how.

Theorem 5.10. *$\wp(\mathbb{N})$ is not countable.*

Proof. We proceed in the same way, by showing that every list of subsets of $\mathbb{N}$ omits some subset of $\mathbb{N}$. So, suppose that we have some list $N_0, N_1, N_2, \ldots$ of subsets of $\mathbb{N}$. We define a set D as follows: $n \in D$ iff $n \notin N_n$:

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

Clearly $D \subseteq \mathbb{N}$. But D cannot be on the list. After all, by construction $n \in D$ iff $n \notin N_n$, so that $D \neq N_n$ for any $n \in \mathbb{N}$. □

The preceding proof did not mention a diagonal. Still, you can think of it as involving a diagonal if you picture it this way: Imagine the sets N_0, N_1, ..., written in an array, where we write N_n on the nth row by writing m in the mth column iff if $m \in N_n$. For example, say the first four sets on that list are $\{0,1,2,\dots\}$, $\{1,3,5,\dots\}$, $\{0,1,4\}$, and $\{2,3,4,\dots\}$; then our array would begin with

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

Then D is the set obtained by going down the diagonal, placing $n \in D$ iff n is *not* on the diagonal. So in the above case, we would leave out 0 and 1, we would include 2, we would leave out 3, etc.

5.6 Reduction

We proved that $\mathbb{B}^\omega$ is uncountable by a diagonalization argument. We used a similar diagonalization argument to show that $\wp(\mathbb{N})$ is uncountable. But here's another way we can prove that $\wp(\mathbb{N})$ is uncountable: show that *if $\wp(\mathbb{N})$ is countable then $\mathbb{B}^\omega$ is also countable.* Since we know $\mathbb{B}^\omega$ is uncountable, it will follow that $\wp(\mathbb{N})$ is too.

This is called *reducing* one problem to another. In this case, we reduce the problem of enumerating $\mathbb{B}^\omega$ to the problem of enumerating $\wp(\mathbb{N})$. A solution to the latter—an enumeration of $\wp(\mathbb{N})$—would yield a solution to the former—an enumeration of $\mathbb{B}^\omega$.

To reduce the problem of enumerating a set B to that of enumerating a set A, we provide a way of turning an enumeration of A into an enumeration of B. The easiest way to do that is to define a surjection $f : A \to B$. If $x_1, x_2, \dots$ enumerates A, then $f(x_1), f(x_2), \dots$ would enumerate B. In our case, we are looking for a surjection $f : \wp(\mathbb{N}) \to \mathbb{B}^\omega$.

Proof of Theorem 5.10 by reduction. For a reduction, suppose that $\wp(\mathbb{N})$ is countable, and thus that there is an enumeration of it, N_1, N_2, N_3, ...

Define the function $f : \wp(\mathbb{N}) \to \mathbb{B}^\omega$ by letting $f(N)$ be the string s_k such that $s_k(n) = 1$ iff $n \in N$, and $s_k(n) = 0$ otherwise.

This clearly defines a function, since whenever $N \subseteq \mathbb{N}$, any $n \in \mathbb{N}$ either is a member of N or isn't. For instance, the set $2\mathbb{N} = \{2n : n \in \mathbb{N}\} = \{0, 2, 4, 6, \dots\}$ of even naturals gets mapped to the string $1010101\dots$; $\emptyset$ gets mapped to $0000\dots$; $\mathbb{N}$ gets mapped to $1111\dots$.

It is also surjective: every string of 0s and 1s corresponds to some set of natural numbers, namely the one which has as its members those natural numbers corresponding to the places where the string contains a 1s. More precisely, if $s \in \mathbb{B}^\omega$, then define $N \subseteq \mathbb{N}$ by:

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

Then $f(N) = s$, as can be verified by consulting the definition of f.

Now consider the list

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

Since f is surjective, every member of $\mathbb{B}^\omega$ must appear as a value of f for some argument, and so must appear on the list. This list must therefore enumerate all of $\mathbb{B}^\omega$.

So if $\wp(\mathbb{N})$ were countable, $\mathbb{B}^\omega$ would be countable. But $\mathbb{B}^\omega$ is uncountable (Theorem 5.9). Hence $\wp(\mathbb{N})$ is uncountable. □

5.7 Equinumerosity

We have an intuitive notion of "size" of sets, which works fine for finite sets. But what about infinite sets? If we want to come up with a formal way of comparing the sizes of two sets of *any* size, it is a good idea to start by defining when sets are the same size. Here is Frege:

> If a waiter wants to be sure that he has laid exactly as many knives as plates on the table, he does not need to count either of them, if he simply lays a knife to the right of each plate, so that every knife on the table lies to the right of some plate. The plates and knives are thus uniquely correlated to each other, and indeed through that same spatial relationship. (Frege, 1884, §70)

The insight of this passage can be brought out through a formal definition:

Definition 5.11. A is *equinumerous* with B, written $A \approx B$, iff there is a bijection $f : A \to B$.

Proposition 5.12. *Equinumerosity is an equivalence relation.*

Proof. We must show that equinumerosity is reflexive, symmetric, and transitive. Let A, B, and C be sets.

Reflexivity. The identity map $\mathrm{Id}_A : A \to A$, where $\mathrm{Id}_A(x) = x$ for all $x \in A$, is a bijection. So $A \approx A$.

Symmetry. Suppose $A \approx B$, i.e., there is a bijection $f : A \to B$. Since f is bijective, its inverse f^{-1} exists and is also bijective. Hence, $f^{-1} : B \to A$ is a bijection, so $B \approx A$.

Transitivity. Suppose that $A \approx B$ and $B \approx C$, i.e., there are bijections $f : A \to B$ and $g : B \to C$. Then the composition $g \circ f : A \to C$ is bijective, so that $A \approx C$. □

Proposition 5.13. *If $A \approx B$, then A is countable if and only if B is.*

Proof. Suppose $A \approx B$, so there is some bijection $f : A \to B$, and suppose that A is countable. Then either $A = \emptyset$ or there is a bijection g whose range is A and whose domain is either $\mathbb{N}$ or an initial sequence of natural numbers. If $A = \emptyset$, then $B = \emptyset$ also (otherwise there would be some $y \in B$ with no $x \in A$ such that

$g(x) = y$). So suppose we have our bijection g. Then $f \circ g$ is a bijection with range B and domain the same as that of g (i.e., either $\mathbb{N}$ or an initial segment of it), so that B is countable.

If B is countable, we obtain that A is countable by repeating the argument with the bijection $f^{-1}: B \to A$ instead of f. □

5.8 Sets of Different Sizes, and Cantor's Theorem

We have offered a precise statement of the idea that two sets have the same size. We can also offer a precise statement of the idea that one set is smaller than another. Our definition of "is smaller than (or equinumerous)" will require, instead of a bijection between the sets, an injection from the first set to the second. If such a function exists, the size of the first set is less than or equal to the size of the second. Intuitively, an injection from one set to another guarantees that the range of the function has at least as many members as the domain, since no two members of the domain map to the same member of the range.

Definition 5.14. A is *no larger than* B, written $A \preceq B$, iff there is an injection $f: A \to B$.

It is clear that this is a reflexive and transitive relation, but that it is not symmetric (this is left as an exercise). We can also introduce a notion, which states that one set is (strictly) smaller than another.

Definition 5.15. A is *smaller than* B, written $A \prec B$, iff there is an injection $f: A \to B$ but no bijection $g: A \to B$, i.e., $A \preceq B$ and $A \not\approx B$.

It is clear that this relation is irreflexive and transitive. (This is left as an exercise.) Using this notation, we can say that a set A is countable iff $A \preceq \mathbb{N}$, and that A is uncountable iff $\mathbb{N} \prec A$.

This allows us to restate Theorem 5.10 as the observation that $\mathbb{N} \prec \wp(\mathbb{N})$. In fact, Cantor (1892) proved that this last point is *perfectly general*:

Theorem 5.16 (Cantor). *$A \prec \wp(A)$, for any set A.*

Proof. The map $f(x) = \{x\}$ is an injection $f : A \to \wp(A)$, since if $x \neq y$, then also $\{x\} \neq \{y\}$ by extensionality, and so $f(x) \neq f(y)$. So we have that $A \preceq \wp(A)$.

It remains to show that $A \not\approx \wp(A)$. For reductio, suppose $A \approx \wp(A)$, i.e., there is some bijection $g : A \to \wp(A)$. Now consider:

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

Note that $D \subseteq A$, so that $D \in \wp(A)$. Since g is a bijection, there is some $y \in A$ such that $g(y) = D$. But now we have:

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

This is a contradiction; so $A \not\approx \wp(A)$. □

It's instructive to compare the proof of Theorem 5.16 to that of Theorem 5.10. There we showed that for any list N_0, N_1, N_2, ..., of subsets of $\mathbb{N}$ we can construct a set D of numbers guaranteed not to be on the list. It was guaranteed not to be on the list because $n \in N_n$ iff $n \notin D$, for every $n \in \mathbb{N}$. We follow the same idea here, except the indices n are now members of A rather than of $\mathbb{N}$. The set D is defined so that it is different from $g(x)$ for each $x \in A$, because $x \in g(x)$ iff $x \notin D$.

The proof is also worth comparing with the proof of Russell's Paradox, Theorem 2.29. Indeed, Cantor's Theorem was the inspiration for Russell's own paradox.

5.9 The Notion of Size, and Schröder-Bernstein

Here is an intuitive thought: if A is no larger than B and B is no larger than A, then A and B are equinumerous. To be honest, if

this thought were *wrong*, then we could scarcely justify the thought that our defined notion of equinumerosity has anything to do with comparisons of "sizes" between sets! Fortunately, though, the intuitive thought is correct. This is justified by the Schröder-Bernstein Theorem.

Theorem 5.17 (Schröder-Bernstein). *If $A \preceq B$ and $B \preceq A$, then $A \approx B$.*

In other words, if there is an injection from A to B, and an injection from B to A, then there is a bijection from A to B.

This result, however, is really rather *difficult* to prove. Indeed, although Cantor stated the result, others proved it.[2] For now, you can (and must) take it on trust.

Fortunately, Schröder-Bernstein is *correct*, and it vindicates our thinking of the relations we defined, i.e., $A \approx B$ and $A \preceq B$, as having something to do with "size". Moreover, Schröder-Bernstein is very *useful*. It can be difficult to think of a bijection between two equinumerous sets. The Schröder-Bernstein Theorem allows us to break the comparison down into cases so we only have to think of an injection from the first to the second, and vice-versa.

5.10 Cantor on the Line and the Plane

Some of the circumstances surrounding the proof of Schröder-Bernstein tie in with the history we discussed in section 1.3. Recall that, in 1877, Cantor proved that there are exactly as many points on a square as on one of its sides. Here, we will present his (first attempted) proof.

Let L be the unit line, i.e., the set of points $[0,1]$. Let S be the unit square, i.e., the set of points $\mathrm{L} \times \mathrm{L}$. In these terms, Cantor proved that $\mathrm{L} \approx \mathrm{S}$. He wrote a note to Dedekind, essentially containing the following argument.

[2] For more on the history, see e.g., Potter (2004, pp. 165–6).

Theorem 5.18. $L \approx S$

Proof: first part.. Fix $a, b \in \mathrm{L}$. Write them in binary notation, so that we have infinite sequences of 0s and 1s, a_1, a_2, ..., and b_1, b_2, ..., such that:

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

Now consider the function $f : \mathrm{S} \to \mathrm{L}$ given by

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

Now f is an injection, since if $f(a,b) = f(c,d)$, then $a_n = c_n$ and $b_n = d_n$ for all $n \in \mathbb{N}$, so that $a = c$ and $b = d$. □

Unfortunately, as Dedekind pointed out to Cantor, this does not answer the original question. Consider $0.\dot{1}\dot{0} = 0.1010101010\ldots$. We need that $f(a,b) = 0.\dot{1}\dot{0}$, where:

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

But $a = 0.\dot{1}\dot{1} = 1$. So, when we say "write a and b in binary notation", we have to choose *which* notation to use; and, since f is to be a *function*, we can use only *one* of the two possible notations. But if, for example, we use the simple notation, and write a as "$1.000\ldots$", then we have no pair $\langle a,b\rangle$ such that $f(a,b) = 0.\dot{1}\dot{0}$.

To summarise: Dedekind pointed out that, given the possibility of certain recurring decimal expansions, Cantor's function f is an injection but *not* a surjection. So Cantor has shown only that $\mathrm{S} \preceq \mathrm{L}$ and *not* that $\mathrm{S} \approx \mathrm{L}$.

Cantor wrote back to Dedekind almost immediately, essentially suggesting that the proof could be completed as follows:

Proof: completed.. So, we have shown that $\mathrm{S} \preceq \mathrm{L}$. But there is obviously an injection from L to S: just lay the line flat along one side of the square. So $\mathrm{L} \preceq \mathrm{S}$ and $\mathrm{S} \preceq \mathrm{L}$. By Schröder–Bernstein (Theorem 5.17), $\mathrm{L} \approx \mathrm{S}$. □

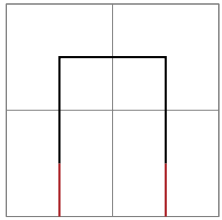
But of course, Cantor could not complete the last line in these terms, for the Schröder-Bernstein Theorem was not yet proved. Indeed, although Cantor would subsequently formulate this as a general conjecture, it was not satisfactorily proved until 1897. (And so, later in 1877, Cantor offered a different proof of Theorem 5.18, which did not go via Schröder–Bernstein.)

5.11 Appendix: Hilbert's Space-filling Curves

In chapter section 1.3, we mentioned that Cantor's proof that a line and a square have exactly the same number of points (Theorem 5.18) prompted Peano to ask whether there might be a space-filling *curve*. He obtained a positive answer in 1890. In this section, we explain (in a hand-wavy way) how to construct Hilbert's space-filling curve (with a tiny tweak).[3]

We must define a function, h, as the limit of a sequence of functions h_1, h_2, h_3, ... We first describe the construction. Then we show it is space-filling. Then we show it is a curve.

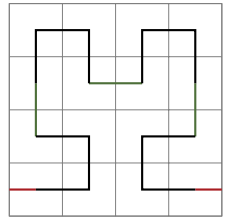
We will take h's range to be the unit square, S. Here is our first approximation to h, i.e., h_1:



To keep track of things, we have imposed a 2×2 grid on the square. We can think of the curve starting in the bottom left quarter, moving to the top left, then to the top right, then finally

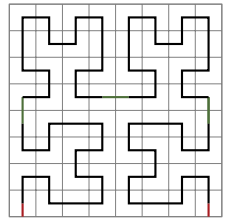
[3]For a more rigorous explanation, see Rose (2010). The tweak amounts to the inclusion of the red parts of the curves below. This makes it slightly easier to check that the curve is continuous.

to the bottom right. Here is the second stage in the construction, i.e., h_2:



The different colours will help explain how h_2 was constructed. We first place scaled-down copies of the non-red bit of h_1 into the bottom left, top left, top right, and bottom right of our square (drawn in black). We then connect these four figures (with green lines). Finally, we connect our figure to the boundary of the square (with red lines).

Now to h_3. Just as h_2 was made from four connected, scaled-down copies of the non-red bit of h_1, so h_3 is made up of four scaled-down copies of the non-red bit of h_2 (drawn in black), which are then joined together (with green lines) and finally connected to the boundary of the square (with red lines).



And now we see the general pattern for defining h_{n+1} from h_n. At last we define the curve h *itself* by considering the point-by-point limit of these successive functions h_1, h_2, ... That is, for each $x \in S$:

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

We now show that this curve fills space. When we draw the curve h_n, we impose a $2^n \times 2^n$ grid onto S. By Pythagoras's Theorem, the diagonal of each grid-location is of length:

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

and evidently h_n passes through every grid-location. So each point in S is *at most* $2^{(\frac{1}{2}-n)}$ distance away from some point on h_n. Now, h is defined as the limit of the functions h_1, h_2, h_3, ... So the maximum distance of any point from h is given by:

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

That is: every point in S is 0 distance from h. In other words, every point of S lies *on* the curve. So h fills space!

It remains to show that h is, indeed, a *curve*. To show this, we must define the notion. The modern definition builds on one given by Jordan in 1887 (i.e., only a few years before the first space-filling curve was provided):

Definition 5.19. A curve is a continuous map from L to $\mathbb{R}^2$.

This is fairly intuitive: a curve is, intuitively, a "smooth" map which takes a canonical line onto the plane $\mathbb{R}^2$. Our function, h, is indeed a map from L to $\mathbb{R}^2$. So, we just need to show that h is continuous. We defined continuity in section 1.2 using ε/δ notation. In the vernacular, we want to establish the following: *If you specify a point p in S, together with any desired level of precision ε, we can find an open section of L such that, given any x in that open section, $h(x)$ is within ε of p.*

So: assume that you have specified p and ε. This is, in effect, to draw a circle with centre p and radius ε on S. (The circle might spill off the edge of S, but that doesn't matter.) Now, recall that, when describing the function h_n, we drew a $2^n \times 2^n$ grid upon S. It is obvious that, no matter how small ε is, there is some n such

that some individual grid-location of the $2^n \times 2^n$ grid on S lies wholly within the circle with centre p and radius ε.

So, take that n, and let I be the largest open part of L which h_n maps wholly into the relevant grid location. (It is clear that (a, b) exists, since we already noted that h_n passes through every grid-location in the $2^n \times 2^n$ grid.) It now suffices to show to show that, whenever $x \in I$ the point $h(x)$ lies in that same grid-location. And to do *this*, it suffices to show that $h_m(x)$ lies in that same grid location, for any $m > n$. But this is obvious. If we consider what happens with h_m for $m > n$, we see that exactly the "same part" of the unit interval is mapped into the same grid-location; we just map it into that region in an increasingly stretched-out, wiggly fashion.

Problems

Problem 5.1. Show that a set A is countable iff either $A = \emptyset$ or there is a surjection $f : \mathbb{N} \to A$. Show that A is countable iff there is an injection $g : A \to \mathbb{N}$.

Problem 5.2. Define an enumeration of the square numbers 1, 4, 9, 16, ...

Problem 5.3. Show that if A and B are countable, so is $A \cup B$.

Problem 5.4. Show by induction on n that if $A_1, A_2, \ldots, A_n$ are all countable, so is $A_1 \cup \cdots \cup A_n$.

Problem 5.5. Show that $(\mathbb{Z}^+)^n$ is countable, for every $n \in \mathbb{N}$.

Problem 5.6. Show that $(\mathbb{Z}^+)^*$ is countable. You may assume problem 5.5.

Problem 5.7. Give an enumeration of the set of all non-negative rational numbers.

Problem 5.8. Show that $\mathbb{Q}$ is countable. Recall that any rational number can be written as a fraction z/m with $z \in \mathbb{Z}$, $m \in \mathbb{N}^+$.

Problem 5.9. Define an enumeration of $\mathbb{B}^*$.

Problem 5.10. Recall from your introductory logic course that each possible truth table expresses a truth function. In other words, the truth functions are all functions from $\mathbb{B}^k \to \mathbb{B}$ for some k. Prove that the set of all truth functions is enumerable.

Problem 5.11. Show that the set of all finite subsets of an arbitrary infinite countable set is countable.

Problem 5.12. A subset of $\mathbb{N}$ is said to be *cofinite* iff it is the complement of a finite set $\mathbb{N}$; that is, $A \subseteq \mathbb{N}$ is cofinite iff $\mathbb{N} \setminus A$ is finite. Let I be the set whose members are exactly the finite and cofinite subsets of $\mathbb{N}$. Show that I is countable.

Problem 5.13. Show that the countable union of countable sets is countable. That is, whenever $A_1, A_2, \ldots$ are sets, and each A_i is countable, then the union $\bigcup_{i=1}^{\infty} A_i$ of all of them is also countable. [NB: this is hard!]

Problem 5.14. Let $f : A \times B \to \mathbb{N}$ be an arbitrary pairing function. Show that the inverse of f is an enumeration of $A \times B$.

Problem 5.15. Specify a function that encodes $\mathbb{N}^3$.

Problem 5.16. Show that the set of all functions $f : \mathbb{N} \to \mathbb{N}$ is uncountable by an explicit diagonal argument. That is, show that if $f_1, f_2, \ldots$, is a list of functions and each $f_i : \mathbb{N} \to \mathbb{N}$, then there is some $g : \mathbb{N} \to \mathbb{N}$ not on this list.

Problem 5.17. Show that if there is an injective function $g : B \to A$, and B is uncountable, then so is A. Do this by showing how you can use g to turn an enumeration of A into one of B.

Problem 5.18. Show that the set X of all functions $f : \mathbb{N} \to \mathbb{N}$ is uncountable by a reduction argument (Hint: give a surjective function from X to $\mathbb{B}^\omega$.)

Problem 5.19. Show that the set of all *sets of* pairs of natural numbers, i.e., $\wp(\mathbb{N} \times \mathbb{N})$, is uncountable by a reduction argument.

Problem 5.20. Show that $\mathbb{N}^\omega$, the set of infinite sequences of natural numbers, is uncountable by a reduction argument.

Problem 5.21. Let S be the set of all surjections from $\mathbb{N}$ to the set $\{0,1\}$, i.e., S consists of all surjections $f : \mathbb{N} \to \mathbb{B}$. Show that S is uncountable.

Problem 5.22. Show that the set $\mathbb{R}$ of all real numbers is uncountable.

Problem 5.23. Show that if $A \approx C$ and $B \approx D$, and $A \cap B = C \cap D = \emptyset$, then $A \cup B \approx C \cup D$.

Problem 5.24. Show that if A is infinite and countable, then $A \approx \mathbb{N}$.

Problem 5.25. Show that there cannot be an injection $g : \wp(A) \to A$, for any set A. Hint: Suppose $g : \wp(A) \to A$ is injective. Consider $D = \{g(B) : B \subseteq A \text{ and } g(B) \notin B\}$. Let $x = g(D)$. Use the fact that g is injective to derive a contradiction.

CHAPTER 6

Arithmetization

In chapter 1, we considered some of the historical background, as to *why* we even have set theory. Chapters 2 to 5 then worked through through some principles of naïve set theory. So we now understand, naïvely, how to construct relations and functions and compare the sizes of sets, and *things like that.*

With this under our belts, we can approach some of the early achievements of set theory, in reducing (in some sense) large chunks of mathematics to set theory and arithmetic. That is the aim of this chapter.

6.1 From $\mathbb{N}$ to $\mathbb{Z}$

Here are two basic realisations:

1. Every integer can be written in the form $n - m$, with $n, m \in \mathbb{N}$.

2. The information encoded in an expression $n-m$ can equally be encoded by an ordered pair $\langle n, m \rangle$.

We already know that the ordered pairs of natural numbers are the members of $\mathbb{N}^2$. And we are assuming that we understand $\mathbb{N}$. So here is a naïve suggestion, based on the two realisations we have had: *let's treat integers as ordered pairs of natural numbers.*

In fact, this suggestion is too naïve. Obviously we want it to be the case that $0 - 2 = 4 - 6$. But evidently $\langle 0,2\rangle \neq \langle 4,6\rangle$. So we cannot simply say that $\mathbb{N}^2$ is the set of integers.

Generalising from the preceding problem, what we want is the following:

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

(It should be obvious that this is how integers are *meant* to behave: just add b and d to both sides.) And the easy way to guarantee this behaviour is just to define an equivalence relation between ordered pairs, $\sim$, as follows:

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

We now have to show that this is an equivalence relation.

Proposition 6.1. *$\sim$ is an equivalence relation.*

Proof. We must show that $\sim$ is reflexive, symmetric, and transitive.

Reflexivity: Evidently $\langle a,b\rangle \sim \langle a,b\rangle$, since $a + b = b + a$.

Symmetry: Suppose $\langle a,b\rangle \sim \langle c,d\rangle$, so $a + d = c + b$. Then $c + b = a + d$, so that $\langle c,d\rangle \sim \langle a,b\rangle$.

Transitivity: Suppose $\langle a,b\rangle \sim \langle c,d\rangle \sim \langle m,n\rangle$. So $a + d = c + b$ and $c + n = m + d$. So $a + d + c + n = c + b + m + d$, and so $a + n = m + b$. Hence $\langle a,b\rangle \sim \langle m,n\rangle$. □

Now we can use this equivalence relation to take equivalence classes:

Definition 6.2. The integers are the equivalence classes, under $\sim$, of ordered pairs of natural numbers; that is, $\mathbb{Z} = \mathbb{N}^2/_\sim$.

Now, one might have plenty of different *philosophical* reactions to this stipulative definition. Before we consider those reactions, though, it is worth continuing with some of the technicalities.

Having said what the integers are, we shall need to define basic functions and relations on them. Let's write $[m,n]_\sim$ for the equivalence class under $\sim$ with $\langle m,n\rangle$ as a member.[1] That is:

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

So now we offer some definitions:

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

(As is common, I'm using 'ab' stand for '$(a \times b)$', just to make the axioms easier to read.) Now, we need to make sure that these definitions behave as they *ought* to. Spelling out what this means, and checking it through, is rather laborious; we relegate the details to section 6.6. But the short point is: everything works!

One final thing remains. We have constructed the integers using natural numbers. But this will mean that the natural numbers *are not themselves integers.* We will return to the philosophical significance of this in section 6.5. On a purely technical front, though, we will need some way to be able to treat natural numbers *as* integers. The idea is quite easy: for each $n \in \mathbb{N}$, we just stipulate that $n_\mathbb{Z} = [n,0]_\sim$. We need to confirm that this definition is well-behaved, i.e., that for any $m,n \in \mathbb{N}$

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

But this is all pretty straightforward. For example, to show that the second of these obtains, we can simply help ourselves to the behaviour of the natural numbers and reason as follows:

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

[1] Note: using the notation introduced in Definition 3.11, we would have written $[\langle m,n\rangle]_\sim$ for the same thing. But that's just a bit harder to read.

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

We leave it as an exercise to confirm that the other two conditions hold.

6.2 From $\mathbb{Z}$ to $\mathbb{Q}$

We just saw how to construct the integers from the natural numbers, using some naïve set theory. We shall now see how to construct the rationals from the integers in a very similar way. Our initial realisations are:

1. Every rational can be written in the form i/j, where both i and j are integers but j is non-zero.

2. The information encoded in an expression i/j can equally be encoded in an ordered pair $\langle i, j \rangle$.

The obvious approach would be to think of the rationals *as* ordered pairs drawn from $\mathbb{Z} \times (\mathbb{Z} \setminus \{0_\mathbb{Z}\})$. As before, though, that would be a bit too naïve, since we want $3/2 = 6/4$, but $\langle 3, 2 \rangle \neq \langle 6, 4 \rangle$. More generally, we will want the following:

$$a/b = c/d \text{ iff } a \times d = b \times c$$

To get this, we define an equivalence relation on $\mathbb{Z} \times (\mathbb{Z} \setminus \{0_\mathbb{Z}\})$ thus:

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

We must check that this is an equivalence relation. This is very much like the case of $\sim$, and we will leave it as an exercise. But it allows us to say:

Definition 6.3. The rationals are the equivalence classes, under $\backsim$, of pairs of integers (whose second element is non-zero). That is, $\mathbb{Q} = (\mathbb{Z} \times (\mathbb{Z} \setminus \{0_{\mathbb{Z}}\}))/\backsim$.

As with the integers, we also want to define some basic operations. Where $[i, j]_\backsim$ is the equivalence class under $\backsim$ with $\langle i, j\rangle$ as a member, we say:

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

To define $r \leq s$ on these rationals, we use the fact that $r \leq s$ iff $s - r$ is not negative, i.e., $r - s$ can be written as i/j with i non-negative and j positive:

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

for some $i \in \mathbb{N}$ and $0 \neq j \in \mathbb{N}$.

We then need to check that these definitions behave as they *ought* to; and we relegate this to section 6.6. But they indeed do! Finally, we want some way to treat integers *as* rationals; so for each $i \in \mathbb{Z}$, we stipulate that $i_{\mathbb{Q}} = [i, 1_{\mathbb{Z}}]_\backsim$. Again, we check that all of this behaves correctly in section 6.6.

6.3 The Real Line

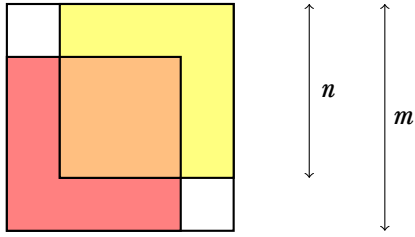
The next step is to show how to construct the reals from the rationals. Before that, we need to understand what is *distinctive* about the reals.

The reals behave very much like the rationals. (Technically, both are examples of *ordered fields*; for the definition of this, see Definition 6.9.) Now, if you worked through the exercises to chapter 5, you will know that there are strictly more reals than rationals, i.e., that $\mathbb{Q} \prec \mathbb{R}$. This was first proved by Cantor. But it's been known for about two and a half millennia that there are irrational numbers, i.e., reals which are not rational. Indeed:

Theorem 6.4. *$\sqrt{2}$ is not rational, i.e., $\sqrt{2} \notin \mathbb{Q}$*

Proof. Suppose, for reductio, that $\sqrt{2}$ is rational. So $\sqrt{2} = {}^{m}/_{n}$ for some natural numbers m and n. Indeed, we can choose m and n so that the fraction cannot be reduced any further. Re-organising, $m^2 = 2n^2$. From here, we can complete the proof in two ways:

First, geometrically (following Tennenbaum).[2] Consider these squares:



□

Since $m^2 = 2n^2$, the region where the two squares of side n overlap has the same area as the region which neither of the two squares cover; i.e., the area of the orange square equals the sum of the area of the two unshaded squares. So where the orange square has side p, and each unshaded square has side q, $p^2 = 2q^2$. But now $\sqrt{2} = {}^{p}/_{q}$, with $p < m$ and $q < n$ and $p, q \in \mathbb{N}$. This contradicts the fact that m and n were chosen to be as small as possible.

Second, formally. Since $m^2 = 2n^2$, it follows that m is even. (It is easy to show that, if x is odd, then x^2 is odd.) So $m = 2r$, for some $r \in \mathbb{N}$. Rearranging, $2r^2 = n^2$, so n is also even. So both m and n are even, and hence the fraction ${}^{m}/_{n}$ *can* be reduced further. Contradiction!

In passing, this diagrammatic proof allows us to revisit the material from section 1.4. Tennenbaum (1927–2006) was a thoroughly modern mathematician; but the proof is undeniably lovely, completely rigorous, and appeals to geometric intuition!

[2] This proof is reported by Conway (2006).

In any case: the reals are "more expansive" than the rationals. In some sense, there are "gaps" in the rationals, and these are filled by the reals. Weierstrass realised that this describes a single property of the real numbers, which distinguishes them from the rationals, namely the Completeness Property: *Every non-empty set of real numbers with an upper bound has a least upper bound.*

It is easy to see that the rationals do not have the Completeness Property. For example, consider the set of rationals less than $\sqrt{2}$, i.e.:

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

This has an upper bound in the rationals; its elements are all smaller than 3, for example. But what is its least upper bound? We want to say '$\sqrt{2}$'; but we have just seen that $\sqrt{2}$ is *not* rational. And there is no *least* rational number greater than $\sqrt{2}$. So the set has an upper bound but no least upper bound. Hence the rationals lack the Completeness Property.

By contrast, the continuum "morally ought" to have the Completeness Property. We do not just want $\sqrt{2}$ to be a real number; we want to fill all the "gaps" in the rational line. Indeed, we want the continuum itself to have no "gaps" in it. That is just what we will get via Completeness.

6.4 From $\mathbb{Q}$ to $\mathbb{R}$

In essence, the Completeness Property shows that any point α of the real line divides that line into two halves perfectly: those for which α is the least upper bound, and those for which α is the greatest lower bound. To *construct* the real numbers from the rational numbers, Dedekind suggested that we simply think of the reals as the *cuts* that partition the rationals. That is, we identify $\sqrt{2}$ with the *cut* which separates the rationals $< \sqrt{2}$ from the rationals $> \sqrt{2}$.

Let's tidy this up. If we cut the rational numbers into two halves, we can uniquely identify the partition we made just by

considering its *bottom* half. So, getting precise, we offer the following definition:

Definition 6.5 (Cut). A *cut* α is any non-empty proper initial segment of the rationals with no greatest element. That is, α is a cut iff:

1. *non-empty, proper*: $\emptyset \neq \alpha \subsetneq \mathbb{Q}$
2. *initial*: for all $p, q \in \mathbb{Q}$: if $p < q \in \alpha$ then $p \in \alpha$
3. *no maximum*: for all $p \in \alpha$ there is a $q \in \alpha$ such that $p < q$

Then $\mathbb{R}$ is the set of cuts.

So now we can say that $\sqrt{2} = \{p \in \mathbb{Q} : p^2 < 2 \text{ or } p < 0\}$. Of course, we need to check that this *is* a cut, but we relegate that to section 6.6.

As before, having defined some entities, we next need to define basic functions and relations upon them. We begin with an easy one:

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

This definition of an order allows to *state* the central result, that the set of cuts has the Completeness Property. Spelled out fully, the statement has this shape. If S is a non-empty set of cuts with an upper bound, then S has a least upper bound. In more detail: there is a cut, λ, which is an upper bound for S, i.e. $(\forall \alpha \in S)\alpha \subseteq \lambda$, and λ is the least such cut, i.e. $(\forall \beta \in \mathbb{R})((\forall \alpha \in S)\alpha \subseteq \beta \rightarrow \lambda \subseteq \beta)$. Now here is the proof of the result:

Theorem 6.6. *The set of cuts has the Completeness Property.*

Proof. Let S be any non-empty set of cuts with an upper bound. Let $\lambda = \bigcup S$. We first claim that λ is a cut:

1. Since S has an upper bound, at least one cut is in S, so $\emptyset \neq \lambda$. Since S is a set of cuts, $\lambda \subseteq \mathbb{Q}$. Since S has an

upper bound, some $p \in \mathbb{Q}$ is absent from every cut $\alpha \in S$. So $p \notin \lambda$, and hence $\lambda \subsetneq \mathbb{Q}$.

2. Suppose $p < q \in \lambda$. So there is some $\alpha \in S$ such that $q \in \alpha$. Since α is a cut, $p \in \alpha$. So $p \in \lambda$.

3. Suppose $p \in \lambda$. So there is some $\alpha \in S$ such that $p \in \alpha$. Since α is a cut, there is some $q \in \alpha$ such that $p < q$. So $q \in \lambda$.

This proves the claim. Moreover, clearly $(\forall \alpha \in S)\alpha \subseteq \bigcup S = \lambda$, i.e. λ is an upper bound on S. So now suppose $\beta \in \mathbb{R}$ is also an upper bound, i.e. $(\forall \alpha \in S)\alpha \subseteq \beta$. For any $p \in \mathbb{Q}$, if $p \in \lambda$, then there is $\alpha \in S$ such that $p \in \alpha$, so that $p \in \beta$. Generalizing, $\lambda \subseteq \beta$. So λ is the *least* upper bound on S. □

So we have a bunch of entities which satisfy the Completeness Property. And one way to put this is: there are no "gaps" in our cuts. (So: taking further "cuts" of reals, rather than rationals, would yield no interesting new objects.)

Next, we must define some operations on the reals. We start by embedding the rationals into the reals by stipulating that $p_{\mathbb{R}} = \{q \in \mathbb{Q} : q < p\}$ for each $p \in \mathbb{Q}$. We then define:

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

To handle the other multiplication cases, first let:

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

and then stipulate:

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

We then need to check that each of these definitions always yields a cut. And finally, we need to go through an easy (but long-winded) demonstration that the cuts, so defined, behave exactly as they should. But we relegate all of this to section 6.6.

6.5 Some Philosophical Reflections

So much for the technicalities. But what did they achieve?

Well, pretty uncontestably, they gave us some lovely pure mathematics. Moreover, there were some deep conceptual achievements. It was a profound insight, to see that the Completeness Property expresses the crucial difference between the reals and the rationals. Moreover, the explicit construction of reals, as Dedekind cuts, puts the subject matter of analysis on a firm footing. We know that the notion of a *complete ordered field* is coherent, for the cuts form just such a field.

For all that, we should air a few reservations about these achievements.

First, it is not clear that thinking of reals in terms of cuts is any *more* rigorous than thinking of reals in terms of their familiar (possibly infinite) decimal expansions. This latter "construction" of the reals has some resemblance to the construction of the reals via Cauchy sequence; but in fact, it was essentially known to mathematicians from the early 17th century onwards (see section 6.7). The real increase in rigour came from the realisation that the reals have the Completeness Property; the ability to construct real numbers as particular sets is perhaps not, by itself, so very interesting.

It is even less clear that the (much easier) arithmetization of the integers, or of the rationals, increases rigour in those areas. Here, it is worth making a simple observation. Having *constructed* the integers as equivalence classes of ordered pairs of naturals, and then constructed the rationals as equivalence classes of ordered pairs of integers, and then constructed the reals as sets of rationals, we immediately *forget about* the constructions. In

particular: no one would ever want to *invoke* these constructions during a mathematical proof (excepting, of course, a proof that the constructions behaved as they were supposed to). It's much easier to speak about a real, directly, than to speak about some set of sets of sets of sets of sets of sets of sets of naturals.

It is most doubtful of all that these definitions tell us what the integers, rationals, or reals *are*, *metaphysically speaking*. That is, it is doubtful that the reals (say) *are* certain sets (of sets of sets...). The main barrier to such a view is that the construction could have been done in many different ways. In the case of the reals, there are some genuinely interestingly different constructions (see section 6.7). But here is a really trivial way to obtain some different constructions: as in section 3.2, we could have defined ordered pairs slightly differently; if we had used this alternative notion of an ordered pair, then our constructions would have worked precisely as well as they did, but we would have ended up with different objects. As such, there are many rival set-theoretic constructions of the integers, the rationals, and the reals. And now it would just be arbitrary (and embarrassing) to claim that the integers (say) are *these* sets, rather than *those*. (As in section 3.2, this is an instance of an argument made famous by Benacerraf 1965.)

A further point is worth raising: there is something quite *odd* about our constructions. We started with the natural numbers. We then construct the integers, and construct "the 0 of the integers", i.e., $[0,0]_\sim$. But $0 \neq [0,0]_\sim$. Indeed, given our constructions, *no* natural number is an integer. But that seems extremely counter-intuitive. Indeed, in section 2.3, we claimed without much argument that $\mathbb{N} \subseteq \mathbb{Q}$. If the constructions tell us exactly *what* the numbers are, this claim was trivially false.

Standing back, then, where do we get to? Working in a naïve set theory, and helping ourselves to the naturals, we are able to *treat* integers, rationals, and reals as certain sets. In that sense, we can *embed* the theories of these entities within a set theory. But the philosophical import of this embedding is just not that straightforward.

Of course, none of this is the last word! The point is only this. Showing that the arithmetization of the reals *is* of deep philosophical significance would require some additional *philosophical* argument.

6.6 Ordered Rings and Fields

Throughout this chapter, we claimed that certain definitions behave "as they ought". In this technical appendix, we will spell out what we mean, and (sketch how to) show that the definitions do behave "correctly".

In section 6.1, we defined addition and multiplication on $\mathbb{Z}$. We want to show that, as defined, they endow $\mathbb{Z}$ with the structure we "would want" it to have. In particular, the structure in question is that of a commutative ring.

Definition 6.7. A *commutative ring* is a set S, equipped with specific elements 0 and 1 and operations $+$ and $\times$, satisfying these eight formulas:

Associativity	$a + (b + c) = (a + b) + c$
	$(a \times b) \times c = a \times (b \times c)$
Commutativity	$a + b = b + a$
	$a \times b = b \times a$
Identities	$a + 0 = a$
	$a \times 1 = a$
Additive Inverse	$(\exists b \in S) 0 = a + b$
Distributivity	$a \times (b + c) = (a \times b) + (a \times c)$

Implicitly, these are all bound with universal quantifiers restricted to S. And note that the elements 0 and 1 here need not be the natural numbers with the same name.

So, to check that the integers form a commutative ring, we just need to check that we meet these eight conditions. None of

the conditions is difficult to establish, but this is a bit laborious. For example, here is how to prove *Associativity*, in the case of addition:

Proof. Fix $i, j, k \in \mathbb{Z}$. So there are $a_1, b_1, a_2, b_2, a_3, b_3 \in \mathbb{N}$ such that $i = [a_1, b_1]$ and $j = [a_2, b_2]$ and $k = [a_3, b_3]$. (For legibility, we write "$[x, y]$" rather than "$[x, y]_\sim$"; we'll do this throughout this section.) Now:

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

helping ourselves freely to the behavior of addition on $\mathbb{N}$. □

Equally, here is how to prove *Additive Inverse*:

Proof. Fix $i \in \mathbb{Z}$, so that $i = [a, b]$ for some $a, b \in \mathbb{N}$. Let $j = [b, a] \in \mathbb{Z}$. Helping ourselves to the behaviour of the naturals, $(a + b) + 0 = 0 + (a + b)$, so that $\langle a + b, b + a\rangle \sim_{\mathbb{Z}} \langle 0, 0\rangle$ by definition, and hence $[a + b, b + a] = [0, 0] = 0_{\mathbb{Z}}$. So now $i + j = [a, b] + [b, a] = [a + b, b + a] = [0, 0] = 0_{\mathbb{Z}}$. □

And here is a proof of *Distributivity*:

Proof. As above, fix $i = [a_1, b_1]$ and $j = [a_2, b_2]$ and $k = [a_3, b_3]$. Now:

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

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

□

We leave it as an exercise to prove the remaining five conditions. Having done that, we have shown that $\mathbb{Z}$ constitutes a commutative ring, i.e., that addition and multiplication (as defined) behave as they should.

But our task is not over. As well as defining addition and multiplication over $\mathbb{Z}$, we defined an ordering relation, $\leq$, and we must check that this behaves as it should. In more detail, we must show that $\mathbb{Z}$ constitutes an *ordered* ring.[3]

Definition 6.8. An *ordered ring* is a commutative ring which is also equipped with a total order relation, $\leq$, such that:

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

As before, it is laborious but routine to show that $\mathbb{Z}$, as constructed, is an ordered ring. We will leave that to you.

This takes care of the integers. But now we need to show very similar things of the rationals. In particular, we now need to show that the rationals form an ordered *field*, under our given definitions of $+$, $\times$, and $\leq$:

Definition 6.9. An *ordered field* is an ordered ring which also satisfies:

Multiplicative Inverse $\quad (\forall a \in S \setminus \{0\})(\exists b \in S)\, a \times b = 1$

[3] Recall from Definition 3.16 that a total order is a relation which is reflexive, transitive, anti-symmetric, and connected. In the context of order relations, connectedness is sometimes called *trichotomy*, since for any a and b we have $a \leq b \vee a = b \vee a \geq b$.

Once you have shown that $\mathbb{Z}$ constitutes an ordered ring, it is easy but laborious to show that $\mathbb{Q}$ constitutes an ordered field.

Having dealt with the integers and the rationals, it only remains to deal with the reals. In particular, we need to show that $\mathbb{R}$ constitutes a *complete* ordered field, i.e., an ordered field with the Completeness Property. Now, Theorem 6.6 established that $\mathbb{R}$ has the Completeness Property. However, it remains to run through the (tedious) of checking that $\mathbb{R}$ is an ordered field.

Before tearing off into *that* laborious exercise, we need to check some more "immediate" things. For example, we need a guarantee that $\alpha + \beta$, as defined, is indeed a *cut*, for any cuts α and β. Here is a proof of that fact:

Proof. Since α and β are both cuts, $\alpha+\beta = \{p+q : p \in \alpha \wedge q \in \beta\}$ is a non-empty proper subset of $\mathbb{Q}$. Now suppose $x < p + q$ for some $p \in \alpha$ and $q \in \beta$. Then $x - p < q$, so $x - p \in \beta$, and $x = p + (x - p) \in \alpha + \beta$. So $\alpha + \beta$ is an initial segment of $\mathbb{Q}$. Finally, for any $p + q \in \alpha + \beta$, since α and β are both cuts, there are $p_1 \in \alpha$ and $q_1 \in \beta$ such that $p < p_1$ and $q < q_1$; so $p + q < p_1 + q_1 \in \alpha + \beta$; so $\alpha + \beta$ has no maximum. □

Similar efforts will allow you to check that $\alpha - \beta$ and $\alpha \times \beta$ and $\alpha \div \beta$ are cuts (in the last case, ignoring the case where β is the zero-cut). Again, though, we will simply leave this to you.

But here is a small loose end to tidy up. In section 6.4, we suggest that we can take $\sqrt{2} = \{p \in \mathbb{Q} : p < 0 \text{ or } p^2 < 2\}$. But we do need to show that this set is a *cut*. Here is a proof of that fact:

Proof. Clearly this is a nonempty proper initial segment of the rationals; so it suffices to show that it has no maximum. In particular, it suffices to show that, where p is a positive rational with $p^2 < 2$ and $q = \frac{2p+2}{p+2}$, both $p < q$ and $q^2 < 2$. To see that $p < q$, just note:

$$p^2 < 2$$
$$p^2 + 2p < 2 + 2p$$

$$p(p+2) < 2 + 2p$$
$$p < \tfrac{2+2p}{p+2} = q$$

To see that $q^2 < 2$, just note:

$$p^2 < 2$$
$$2p^2 + 4p + 2 < p^2 + 4p + 4$$
$$4p^2 + 8p + 4 < 2(p^2 + 4p + 4)$$
$$(2p+2)^2 < 2(p+2)^2$$
$$\tfrac{(2p+2)^2}{(p+2)^2} < 2$$
$$q^2 < 2$$

□

6.7 Appendix: the Reals as Cauchy Sequences

In section 6.4, we constructed the reals as Dedekind cuts. In this section, we explain an alternative construction. It builds on Cauchy's definition of (what we now call) a Cauchy sequence; but the use of this definition to *construct* the reals is due to other nineteenth-century authors, notably Weierstrass, Heine, Méray and Cantor. (For a nice history, see O'Connor and Robertson 2005.)

Before we get to the nineteenth century, it's worth considering Simon Stevin (1548–1620). In brief, Stevin realised that we can think of each real in terms of its decimal expansion. Thus even an irrational number, like $\sqrt{2}$, has a nice decimal expansion, beginning:

$$1.41421356237\ldots$$

It is very easy to model decimal expansions in set theory: simply consider them as functions $d\colon \mathbb{N} \to \mathbb{N}$, where $d(n)$ is the nth decimal place that we are interested in. We will then need a bit of tweak, to handle the bit of the real number that comes

before the decimal point (here, just 1). We will also need a further tweak (an equivalence relation) to guarantee that, for example, $0.999\ldots = 1$. But it is not difficult to offer a perfectly rigorous construction of the real numbers, in the manner of Stevin, within set theory.

Stevin is not our focus. (For more on Stevin, see Katz and Katz 2012.) But here is a closely related thought. Instead of treating $\sqrt{2}$'s decimal expansion directly, we can instead consider a *sequence* of increasingly accurate rational approximations to $\sqrt{2}$, by considering the increasingly precise expansions:

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

The idea that reals can be considered via "increasingly good approximations" provides us with the basis for another sequence of insights (akin to the realisations that we used when constructing $\mathbb{Q}$ from $\mathbb{Z}$, or $\mathbb{Z}$ from $\mathbb{N}$). The basic insights are these:

1. Every real can be written as a (perhaps infinite) decimal expansion.

2. The information encoded by a (perhaps infinite) decimal expansion can be equally be encoded by a sequence of rational numbers.

3. A sequence of rational numbers can be thought of as a function from $\mathbb{N}$ to $\mathbb{Q}$; just let $f(n)$ be the nth rational in the sequence.

Of course, not just *any* function from $\mathbb{N}$ to $\mathbb{Q}$ will give us a real number. For instance, consider this function:

$$f(n) = \begin{cases} 1 & \text{if } n \text{ is odd} \\ 0 & \text{if } n \text{ is even} \end{cases}$$

Essentially the worry here is that the sequence $0, 1, 0, 1, 0, 1, 0, \ldots$ doesn't seem to "hone in" on any real. So: to ensure that we

consider sequences which do hone in on some real, we need to restrict our attention to sequences which have some *limit*.

We have already encountered the idea of a limit, in section 1.2. But we cannot use *quite* the same definition as we used there. The expression "$(\forall \varepsilon > 0)$" there tacitly involved quantification over the real numbers; and we were considering the limits of functions on the real numbers; so invoking that definition would be to help ourselves to the real numbers; and they are exactly what we were aiming to *construct*. Fortunately, we can work with a closely related idea of a limit.

Definition 6.10. A function $f : \mathbb{N} \to \mathbb{Q}$ is a *Cauchy sequence* iff for any positive $\varepsilon \in \mathbb{Q}$ we have that $(\exists \ell \in \mathbb{N})(\forall m, n > \ell)|f(m) - f(n)| < \varepsilon$.

The general idea of a limit is the same as before: if you want a certain level of precision (measured by ε), there is a "region" to look in (any input greater than ℓ). And it is easy to see that our sequence 1, 1.4, 1.414, 1.4142, 1.41421... has a limit: if you want to approximate $\sqrt{2}$ to within an error of $1/10^n$, then just look to any entry after the nth.

The obvious thought, then, would be to say that a real number just *is* any Cauchy sequence. But, as in the constructions of $\mathbb{Z}$ and $\mathbb{Q}$, this would be too naïve: for any given real number, multiple different Cauchy sequences indicate that real number. A simple way to see this as follows. Given a Cauchy sequence f, define g to be exactly the same function as f, except that $g(0) \neq f(0)$. Since the two sequences agree everywhere after the first number, we will (ultimately) want to say that they have the same limit, in the sense employed in Definition 6.10, and so should be thought of "defining" the same real. So, we should really think of these Cauchy sequences as the same real number.

Consequently, we again need to define an equivalence relation on the Cauchy sequences, and identify real numbers with equivalence relations. First we need the idea of a function which tends to 0 in the limit. For any function $h : \mathbb{N} \to \mathbb{Q}$,

say that h *tends to* 0 iff for any positive $\varepsilon \in \mathbb{Q}$ we have that $(\exists \ell \in \mathbb{N})(\forall n > \ell)|f(n)| < \varepsilon$.[4] Further, where f and g are functions $\mathbb{N} \to \mathbb{Q}$, let $(f - g)(n) = f(n) - g(n)$. Now define:

$$f \eqsim g \text{ iff } (f - g) \text{ tends to } 0.$$

We need to check that $\eqsim$ is an equivalence relation; and it is. We can then, if we like, define the reals as the equivalence classes, under $\eqsim$, of all Cauchy sequences from $\mathbb{N} \to \mathbb{Q}$.

Having done this, we shall as usual write $[f]_{\eqsim}$ for the equivalence class with f as a member. However, to keep things readable, in what follows we will drop the subscript and write just $[f]$. We also stipulate that, for each $q \in \mathbb{Q}$, we have $q_{\mathbb{R}} = [c_q]$, where c_q is the constant function $c_q(n) = q$ for all $n \in \mathbb{N}$. We then define basic relations and operations on the reals, e.g.:

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

where $(f + g)(n) = f(n) + g(n)$ and $(f \times g)(n) = f(n) \times g(n)$. Of course, we also need to check that each of $(f + g)$, $(f - g)$ and $(f \times g)$ are Cauchy sequences when f and g are; but they are, and we leave this to you.

Finally, we define we a notion of order. Say $[f]$ is *positive* iff both $[f] \neq 0_{\mathbb{Q}}$ and $(\exists \ell \in \mathbb{N})(\forall n > \ell)0 < f(n)$. Then say $[f] < [g]$ iff $[(g - f)]$ is positive. We have to check that this is well-defined (i.e., that it does not depend upon choice of "representative" function from the equivalence class). But having done this, it is quite easy to show that these yield the right algebraic properties; that is:

Theorem 6.11. *The Cauchy sequences constitute an ordered field.*

Proof. Exercise. □

It is harder to prove that the reals, so constructed, have the Completeness Property, so we will give the proof.

[4] Compare this with the definition of $\lim_{x \to \infty} f(x) = 0$ in section 1.2.

Theorem 6.12. *Every non-empty set of Cauchy sequences with an upper bound has a least upper bound.*

Proof sketch. Let S be any non-empty set of Cauchy sequences with an upper bound. So there is some $p \in \mathbb{Q}$ such that $p_{\mathbb{R}}$ is an upper bound for S. Let $r \in S$; then there is some $q \in \mathbb{Q}$ such that $q_{\mathbb{R}} < r$. So if a least upper bound on S exists, it is between $q_{\mathbb{R}}$ and $p_{\mathbb{R}}$ (inclusive).

We will hone in on the l.u.b., by approaching it simultaneously from below and above. In particular, we define two functions, $f, g\colon \mathbb{N} \to \mathbb{Q}$, with the aim that f will hone in on the l.u.b. from above, and g will hone on in it from below. We start by defining:

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

Then, where $a_n = \frac{f(n)+g(n)}{2}$, let:[5]

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

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

Both f and g are Cauchy sequences. (This can be checked fairly easily; but we leave it as an exercise.) Note that the function $(f - g)$ tends to 0, since the difference between f and g halves at every step. Hence $[f] = [g]$.

We will show that $(\forall h \in S)[h] \leq [f]$, invoking Theorem 6.11 as we go. Let $h \in S$ and suppose, for reductio, that $[f] < [h]$, so that $0_{\mathbb{R}} < [(h-f)]$. Since f is a monotonically decreasing Cauchy sequence, there is some $n \in \mathbb{N}$ such that $[(c_{f(n)} - f)] < [(h-f)]$. So:

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

[5] This is a recursive definition. But we have not *yet* given any reason to think that recursive definitions are ok.

contradicting the fact that, by construction, $[h] \leq (f(k))_{\mathbb{R}}$.

In an exactly similar way, we can show that $(\forall [h]_{\in} S)[g] \leq [h]$. So $[f] = [g]$ is the *least* upper bound for S. □

Problems

Problem 6.1. Show that $(m+n)_{\mathbb{Z}} = m_{\mathbb{Z}} + n_{\mathbb{Z}}$ and $m \leq n \leftrightarrow m_{\mathbb{Z}} \leq n_{\mathbb{Z}}$, for any $m, n \in \mathbb{N}$.

Problem 6.2. Show that $\backsim$ is an equivalence relation.

Problem 6.3. Show that $(i + j)_{\mathbb{Q}} = i_{\mathbb{Q}} + j_{\mathbb{Q}}$ and $(i \times j)_{\mathbb{Q}} = i_{\mathbb{Q}} \times j_{\mathbb{Q}}$ and $i \leq j \leftrightarrow i_{\mathbb{Q}} \leq j_{\mathbb{Q}}$, for any $i, j \in \mathbb{Z}$.

Problem 6.4. Prove that $\mathbb{Z}$ is a commutative ring.

Problem 6.5. Prove that $\mathbb{Z}$ is an ordered ring.

Problem 6.6. Prove that $\mathbb{Q}$ is an ordered field.

Problem 6.7. Prove that $\mathbb{R}$ is an ordered field.

Problem 6.8. Let $f(n) = 0$ for every n. Let $g(n) = \frac{1}{(n+1)^2}$. Show that both are Cauchy sequences, and indeed that the limit of both functions is 0, so that also $f \sim_{\mathbb{R}} g$.

Problem 6.9. Prove that the Cauchy sequences constitute an ordered field.

CHAPTER 7

Infinite Sets

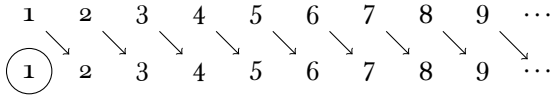
In the previous chapter, we showed how to construct a bunch of things—integers, rationals, and reals—assuming some naïve set theory and the natural numbers. The question for this chapter is: Can we construct the set of natural numbers *itself* using set theory?

7.1 Hilbert's Hotel

The set of the natural numbers is obviously infinite. So, if we do not want to *help ourselves* to the natural numbers, our first step must be characterize an infinite set in terms that do not require mentioning the natural numbers themselves. Here is a nice approach, presented by Hilbert in a lecture from 1924. He asks us to imagine

> [...] a hotel with a finite number of rooms. All of these rooms should be occupied by exactly one guest. If the guests now swap their rooms somehow, [but] so that each room still contains no more than one person, then no rooms will become free, and the hotel-owner cannot in this way create a new place for a newly arriving guest [...¶...]
>
> Now we stipulate that the hotel shall have infinitely many numbered rooms 1, 2, 3, 4, 5, ..., each of which

> is occupied by exactly one guest. As soon as a new guest comes along, the owner only needs to move each of the old guests into the room associated with the number one higher, and room 1 will be free for the newly-arriving guest.



> (published in Hilbert 2013, 730; our translation)

The crucial point is that Hilbert's Hotel has infinitely many rooms; and we can take his explanation to define what it means to say this. Indeed, this was Dedekind's approach (presented here, of course, with massive anachronism; Dedekind's definition is from 1888):

Definition 7.1. A set A is *Dedekind infinite* iff there is an injection from A to a proper subset of A. That is, there is some $o \in A$ and an injection $f : A \to A$ such that $o \notin \mathrm{ran}(f)$.

7.2 Dedekind Algebras

We not only want natural numbers to be infinite; we want them to have certain (algebraic) properties: they need to behave well under addition, multiplication, and so forth.

Dedekind's idea was to take the idea of the *successor function* as basic, and then characterise the numbers as those with the following properties:

1. There is a number, 0, which is not the successor of any number
   i.e., $0 \notin \mathrm{ran}(s)$
   i.e., $\forall x\ s(x) \neq 0$

2. Distinct numbers have distinct successors
   i.e., s is an injection
   i.e., $\forall x \forall y (s(x) = s(y) \rightarrow x = y)$

3. Every number is obtained from 0 by repeated applications of the successor function.

The first two conditions are easy to deal with using first-order logic (see above). But we cannot deal with (3) just using first-order logic. Dedekind's breakthrough was to reformulate condition (3), set-theoretically, as follows:

3′. The natural numbers are the smallest set that is *closed under the successor function*: that is, if we apply s to any member of the set, we obtain another member of the set.

But we shall need to spell this out slowly.

Definition 7.2. For any function f, the set X is *f-closed* iff $(\forall x \in X) f(x) \in X$. Now define, for any o:

$$\mathrm{clo}_f(o) = \bigcap \{X : o \in X \text{ and } X \text{ is } f\text{-closed}\}$$

So $\mathrm{clo}_f(o)$ is the intersection of all the f-closed sets with o as a member. Intuitively, then, $\mathrm{clo}_f(o)$ is the *smallest* f-closed set with o as a member. This next result makes that intuitive thought precise;

Lemma 7.3. *For any function f and any $o \in A$:*

1. $o \in \mathrm{clo}_f(o)$*; and*
2. $\mathrm{clo}_f(o)$ *is f-closed; and*
3. *if X is f-closed and $o \in X$, then* $\mathrm{clo}_f(o) \subseteq X$

Proof. Note that there is at least one f-closed set with o as a member, namely $\mathrm{ran}(f) \cup \{o\}$. So $\mathrm{clo}_f(o)$, the intersection of *all* such sets, exists. We must now check (1)–(3).

Concerning (1): $o \in \mathrm{clo}_f(o)$ as it is an intersection of sets which all have o as a member.

Concerning (2): suppose $x \in \mathrm{clo}_f(o)$. So if $o \in X$ and X is f-closed, then $x \in X$, and now $f(x) \in X$ as X is f-closed. So $f(x) \in \mathrm{clo}_f(o)$.

Concerning (3): quite generally, if $X \in C$ then $\bigcap C \subseteq X$. □

Using this, we can say:

Definition 7.4. A *Dedekind algebra* is a set A together with a function $f : A \to A$ and some $o \in A$ such that:

1. $o \notin \mathrm{ran}(f)$
2. f is an injection
3. $A = \mathrm{clo}_f(o)$

Since $A = \mathrm{clo}_f(o)$, our earlier result tells us that A is the smallest f-closed set with o as a member. Clearly a Dedekind algebra is Dedekind infinite; just look at clauses (1) and (2) of the definition. But the more exciting fact is that any Dedekind infinite set can be turned into a Dedekind algebra.

Theorem 7.5. *If there is a Dedekind infinite set, then there is a Dedekind algebra.*

Proof. Let D be Dedekind infinite. So there is an injection $g : D \to D$ and an element $o \in D \setminus \mathrm{ran}(g)$. Now let $A = \mathrm{clo}_g(o)$; by Lemma 7.3, A exists and $o \in A$. Let $f = g \restriction_A$. We will show that A, f, o comprise a Dedekind algebra.

Concerning (1): $o \notin \mathrm{ran}(g)$ and $\mathrm{ran}(f) \subseteq \mathrm{ran}(g)$ so $o \notin \mathrm{ran}(f)$.

Concerning (2): g is an injection on D; so $f \subseteq g$ must be an injection.

Concerning (3): by Lemma 7.3, A is g-closed; a fortiori, A is f-closed. So $\mathrm{clo}_f(o) \subseteq A$ by Lemma 7.3. Since also $\mathrm{clo}_f(o)$

is f-closed and $f = g\restriction_A$, it follows that $\mathrm{clo}_f(o)$ is g-closed. So $A \subseteq \mathrm{clo}_f(o)$ by Lemma 7.3. □

7.3 Dedekind Algebras and Arithmetical Induction

Crucially, now, a Dedekind algebra—indeed, *any* Dedekind algebra—will serve as a surrogate for the natural numbers. This is thanks to the following trivial consequence:

Theorem 7.6 (Arithmetical induction). *Let N, s, o comprise a Dedekind algebra. Then for any set X:*

$$\textit{if } o \in X \textit{ and } (\forall n \in N \cap X)s(n) \in X, \textit{ then } N \subseteq X.$$

Proof. By the definition of a Dedekind algebra, $N = \mathrm{clo}_s(o)$. Now if both $o \in X$ and $(\forall n \in N)(n \in X \to s(n) \in X)$, then $N = \mathrm{clo}_s(o) \subseteq X$. □

Since induction is characteristic of the natural numbers, the point is this. Given any Dedekind infinite set, we can form a Dedekind algebra, and use that algebra as our surrogate for the natural numbers.

Admittedly, Theorem 7.6 formulates induction in *set-theoretic* terms. But we can easily put the principle in terms which might be more familiar:

Corollary 7.7. *Let N, s, o comprise a Dedekind algebra. Then for any formula $\varphi(x)$, which may have parameters:*

$$\textit{if } \varphi(o) \textit{ and } (\forall n \in N)(\varphi(n) \to \varphi(s(n))), \textit{ then } (\forall n \in N)\varphi(n)$$

Proof. Let $X = \{n \in N : \varphi(n)\}$, and now use Theorem 7.6 □

In this result, we spoke of a formula "having parameters". What this means, roughly, is that for any objects $c_1, \ldots, c_k$, we can work with $\varphi(x, c_1, \ldots, c_k)$. More precisely, we can state the result

without mentioning "parameters" as follows. For any formula $\varphi(x, v_1, \ldots, v_k)$, whose free variables are all displayed, we have:

$$\forall v_1 \ldots \forall v_k((\varphi(o, v_1, \ldots, v_k) \wedge (\forall x \in N)(\varphi(x, v_1, \ldots, v_k) \rightarrow \varphi(s(x), v_1, \ldots, v_k))) \rightarrow (\forall x \in N)\varphi(x, v_1, \ldots, v_k))$$

Evidently, speaking of "having parameters" can make things much easier to read. (In part III, we will use this device rather frequently.)

Returning to Dedekind algebras: given any Dedekind algebra, we can also define the usual arithmetical functions of addition, multiplication and exponentiation. This is non-trivial, however, and it involves the technique of *recursive definition*. That is a technique which we shall introduce and justify much later, and in a much more general context. (Enthusiasts might want to revisit this after chapter 13, or perhaps read an alternative treatment, such as Potter 2004, pp. 95–8.) But, where N, s, o comprise a Dedekind algebra, we will ultimately be able to stipulate the following:

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

and show that these behave as one would hope.

7.4 Dedekind's "Proof" of the Existence of an Infinite Set

In this chapter, we have offered a set-theoretic treatment of the natural numbers, in terms of Dedekind algebras. In section 6.5, we reflected on the philosophical significance of the arithmetisation of analysis (among other things). Now we should reflect on the significance of what we have achieved here.

Throughout chapter 6, we took the natural numbers as given, and used them to construct the integers, rationals, and reals,

explicitly. In this chapter, we have not given an explicit construction of the natural numbers. We have just shown that, *given any Dedekind infinite set*, we can define a set which will behave just like we want $\mathbb{N}$ to behave.

Obviously, then, we cannot claim to have answered a metaphysical question, such as *which objects are the natural numbers*. But that's a good thing. After all, in section 6.5, we emphasized that we would be wrong to think of the definition of $\mathbb{R}$ as the set of Dedekind cuts as a *discovery*, rather than a convenient stipulation. The crucial observation is that the Dedekind cuts exemplify the key mathematical properties of the real numbers. So too here: the crucial observation is that *any* Dedekind algebra exemplifies the key mathematical properties of the natural numbers. (Indeed, Dedekind pushed this point home by proving that all Dedekind algebras are *isomorphic* (1888, Theorems 132–3). It is no surprise, then, that many contemporary "structuralists" cite Dedekind as a forerunner.)

Moreover, we have shown how to embed the theory of the natural numbers into a naïve simple set theory, which itself still remains rather informal, but which doesn't (apparently) assume the natural numbers as given. So, we may be on the way to realising Dedekind's own ambitious project, which he explained thus:

> In science nothing capable of proof ought to be believed without proof. Though this demand seems reasonable, I cannot regard it as having been met even in the most recent methods of laying the foundations of the simplest science; viz., that part of logic which deals with the theory of numbers. In speaking of arithmetic (algebra, analysis) as merely a part of logic I mean to imply that I consider the number-concept entirely independent of the notions or intuitions of space and time—that I rather consider it an immediate product of the pure laws of thought. (Dedekind, 1888, preface)

Dedekind's bold idea is this. We have just shown how to build

the natural numbers using (naïve) set theory alone. In chapter 6, we saw how to construct the reals given the natural numbers and some set theory. So, perhaps, "arithmetic (algebra, analysis)" turn out to be "merely a part of logic" (in Dedekind's extended sense of the word "logic").

That's the idea. But hold on for a moment. Our construction of a Dedekind algebra (our surrogate for the natural numbers) is conditional on the existence of a Dedekind infinite set. (Just look back to Theorem 7.5.) Unless the existence of a Dedekind infinite set can be established via "logic" or "the pure laws of thought", the project stalls.

So, *can* the existence of a Dedekind infinite set be established by "the pure laws of thought"? Here was Dedekind's effort:

> My own realm of thoughts, i.e., the totality S of all things which can be objects of my thought, is infinite. For if s signifies an element of S, then the thought s' that s can be an object of my thought, is itself an element of S. If we regard this as an image $\varphi(s)$ of the element s, then ... S is [Dedekind] infinite, which was to be proved. (Dedekind, 1888, §66)

This is quite an astonishing thing to find in the middle of a book which largely consists of highly rigorous mathematical proofs. Two remarks are worth making.

First: this "proof" scarcely has what we would now recognize as a "mathematical" character. It speaks of psychological objects (thoughts), and merely *possible* ones at that.

Second: at least as we have presented Dedekind algebras, this "proof" has a straightforward technical shortcoming. If Dedekind's argument is successful, it establishes only that there are infinitely many things (specifically, infinitely many thoughts). But Dedekind also needs to give us a reason to regard S as a single *set*, with infinitely many members, rather than thinking of S as *some things* (in the plural).

The fact that Dedekind did not see a gap here might suggest that his use of the word "totality" does not precisely track

our use of the word "set".[1] But this would not be too surprising. The project we have pursued in the last two chapters—a "construction" of the naturals, and from them a "construction" of the integers, reals and rationals—has all been carried out naïvely. We have helped ourselves to this set, or that set, as and when we have needed them, without laying down many general principles concerning exactly which sets exist, and when. But we know that we need *some* general principles, for otherwise we will fall into Russell's Paradox.

The time has come for us to outgrow our naïvety.

7.5 Appendix: Proving Schröder-Bernstein

Before we depart from naïve set theory, we have one last naïve (but sophisticated!) proof to consider. This is a proof of Schröder-Bernstein (Theorem 5.17): if $A \preceq B$ and $B \preceq A$ then $A \approx B$; i.e., given injections $f : A \to B$ and $g : B \to A$ there is a bijection $h : A \to B$.

In this chapter, we followed Dedekind's notion of *closures*. In fact, Dedekind provided a lovely proof of Schröder-Bernstein using this notion, and we will present it here. The proof closely follows Potter (2004, pp. 157–8), if you want a slightly different but essentially similar treatment. A little googling will also convince you that this is a theorem—rather like the irrationality of $\sqrt{2}$—for which *many* interesting and different proofs exist.

Using similar notation as Definition 7.2, let

$$\mathrm{Clo}_f(B) = \bigcap \{X : B \subseteq X \text{ and } X \text{ is } f\text{-closed}\}$$

for each set B and function f. Defined thus, $\mathrm{Clo}_f(B)$ is the smallest f-closed set containing B, in that:

[1] Indeed, we have other reasons to think it did not; see Potter (2004, p. 23).

Lemma 7.8. *For any function f, and any B:*

1. $B \subseteq \mathrm{Clo}_f(B)$; *and*
2. $\mathrm{Clo}_f(B)$ *is f-closed; and*
3. *if X is f-closed and $B \subseteq X$, then* $\mathrm{Clo}_f(B) \subseteq X$.

Proof. Exactly as in Lemma 7.3. □

We need one last fact to get to Schröder-Bernstein:

Proposition 7.9. *If $A \subseteq B \subseteq C$ and $A \approx C$, then $A \approx B \approx C$.*

Proof. Given a bijection $f : C \to A$, let $F = \mathrm{Clo}_f(C \setminus B)$ and define a function g with domain C as follows:

$$g(x) = \begin{cases} f(x) & \text{if } x \in F \\ x & \text{otherwise} \end{cases}$$

We'll show that g is a bijection from $C \to B$, from which it will follow that $g \circ f^{-1} : A \to B$ is a bijection, completing the proof.

First we claim that if $x \in F$ but $y \notin F$ then $g(x) \neq g(y)$. For reductio suppose otherwise, so that $y = g(y) = g(x) = f(x)$. Since $x \in F$ and F is f-closed by Lemma 7.8, we have $y = f(x) \in F$, a contradiction.

Now suppose $g(x) = g(y)$. So, by the above, $x \in F$ iff $y \in F$. If $x, y \in F$, then $f(x) = g(x) = g(y) = f(y)$ so that $x = y$ since f is a bijection. If $x, y \notin F$, then $x = g(x) = g(y) = y$. So g is an injection.

It remains to show that $\mathrm{ran}(g) = B$. So fix $x \in B \subseteq C$. If $x \notin F$, then $g(x) = x$. If $x \in F$, then $x = f(y)$ for some $y \in F$, since otherwise $F \setminus \{x\}$ would be f-closed and extend $C \setminus B$, which is impossible by Lemma 7.8; now $g(y) = f(y) = x$. □

Finally, here is the proof of the main result. Recall that given a function h and set D, we define $h[D] = \{h(x) : x \in D\}$.

Proof of Schröder-Bernstein. Let $f : A \to B$ and $g : B \to A$ be injections. Since $f[A] \subseteq B$ we have that $g[f[A]] \subseteq g[B] \subseteq A$. Also, $g \circ f : A \to g[f[A]]$ is an injection since both g and f are; and indeed $g \circ f$ is a bijection, just by the way we defined its codomain. So $g[f[A]] \approx A$, and hence by Proposition 7.9 there is a bijection $h : A \to g[B]$. Moreover, g^{-1} is a bijection $g[B] \to B$. So $g^{-1} \circ h : A \to B$ is a bijection. □

PART III

The Iterative Conception

Introduction to Part III

Part II discussed sets in a naïve, informal way. It is now time to tighten this up, and provide a formal theory of sets. That is the aim of part III.

Our formal theory is a first-order theory with just one two-place predicate, $\in$. We will lay down several axioms that govern the behaviour of the membership relation. However, we will introduce these axioms only as we need them, and consider how we might justify them as we encounter them. As a result, we will introduce our axioms *throughout* the ensuing chapters.

It might, though, be helpful for the reader to have a list of all the axioms in one place. So, here are *all* the axioms that we will consider in part III. As in part II, the choice of lowercase and uppercase letters is guided only by readability:

Extensionality.

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

Union.

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

i.e., $\bigcup A$ exists for any set A.

Pairs.

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

i.e., $\{a, b\}$ exists for any a and b.

Powersets.

$$\forall A \exists P \forall x(x \in P \leftrightarrow (\forall z \in x)z \in A),$$

i.e., $\wp(A)$ exists for any set A.

Infinity.

$$\exists I((\exists o \in I)\forall x(x \notin o) \wedge (\forall x \in I)(\exists s \in I)\forall z(z \in s \leftrightarrow (z \in x \vee z = x))),$$

i.e., there is a set with $\emptyset$ as a member and which is closed under $x \mapsto x \cup \{x\}$.

Foundation.

$$\forall A(\forall x\, x \notin A \vee (\exists b \in A)(\forall x \in A)x \notin b),$$

i.e., $A = \emptyset$ or $(\exists b \in A)A \cap b = \emptyset$.

Well-Ordering. For every set A, there is a relation that well-orders A. (Writing this one out in first-order logic is too painful to bother with.)

Separation Scheme. For any formula $\varphi(x)$ which does not contain "S":

$$\forall A\exists S\forall x(x \in S \leftrightarrow (\varphi(x) \wedge x \in A)),$$

i.e., $\{x \in A : \varphi(x)\}$ exists for any set A.

Replacement Scheme. For any formula $\varphi(x,y)$ which does not contain "B":

$$\forall A((\forall x \in A)\exists! y\, \varphi(x,y) \rightarrow \exists B\forall y(y \in B \leftrightarrow (\exists x \in A)\varphi(x,y))),$$

i.e., $\{y : (\exists x \in A)\varphi(x,y)\}$ exists for any A, if φ is "functional."

In both schemes, the formulas may contain parameters. Indeed, throughout part III, we follow the convention that any formula can contain parameters. (See section 7.3 for a reminder of what it means to say that a formula may contain parameters.)

Z$^-$ is Extensionality, Union, Pairs, Powersets, Infinity, Separation.

Z is **Z$^-$** plus Foundation.

ZF$^-$ is **Z** plus Replacement.

ZF is **ZF$^-$** plus Foundation.

ZFC is **ZF** plus Well-Ordering.

CHAPTER 8

The Iterative Conception

8.1 Extensionality

The very first thing to say is that sets are individuated by their members. More precisely:

Axiom (Extensionality). If sets A and B have the same members, then A and B are the same set.

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

We assumed this throughout part II. But it bears repeating. The Axiom of Extensionality expresses the basic idea that a set is determined by its members. (So sets might be contrasted with *concepts*, where precisely the same objects might fall under many different concepts.)

Why embrace this principle? Well, it is plausible to say that any denial of Extensionality is a decision to abandon anything which might even be called *set theory*. Set theory is no more nor less than the theory of extensional collections.

The real challenge in part III, though, is to lay down principles which tell us *which sets exist*. And it turns out that the only

truly "obvious" answer to this question is provably wrong.

8.2 Russell's Paradox (again)

In part II, we worked with a naïve set theory. But according to a *very* naïve conception, sets are just the extensions of predicates. This naïve thought would mandate the following principle:

Naïve Comprehension. $\{x : \varphi(x)\}$ exists for any formula φ.

Tempting as this principle is, it is provably inconsistent. We saw this in section 2.6, but the result is so important, and so straightforward, that it's worth repeating. Verbatim.

Theorem 8.1 (Russell's Paradox). *There is no set* $R = \{x : x \notin x\}$

Proof. If $R = \{x : x \notin x\}$ exists, then $R \in R$ iff $R \notin R$, which is a contradiction. □

Russell discovered this result in June 1901. (He did not, though, put the paradox in quite the form we just presented it, since he was considering Frege's set theory, as outlined in *Grundgesetze*. We will return to this in section 8.6.) Russell wrote to Frege on June 16, 1902, explaining the inconsistency in Frege's system. For the correspondence, and a bit of background, see Heijenoort (1967, pp. 124–8).

It is worth emphasising that this two-line proof is a result of *pure logic*. Granted, we implicitly used a (non-logical?) axiom, Extensionality, in our notation $\{x : x \notin x\}$; for $\{x : \varphi(x)\}$ is to be *the unique* (by Extensionality) set of the φs, if one exists. But we can avoid even the hint of Extensionality, just by stating the result as follows: *there is no set whose members are exactly the non-self-membered sets.* And this has nothing much to do with sets. As Russell himself observed, exactly similar reasoning will lead you

to conclude: *no man shaves exactly the men who do not shave themselves.* Or: *no pug sniffs exactly the pugs which don't sniff themselves.* And so on. Schematically, the shape of the result is just:

$$\neg\exists x\forall z(Rzx \leftrightarrow \neg Rzz).$$

And that's just a theorem (scheme) of first-order logic. Consequently, we can't avoid Russell's Paradox just by tinkering with our set theory; it arises before we even *get* to set theory. If we're going to use (classical) first-order logic, we simply have to *accept* that there is no set $R = \{x : x \notin x\}$.

The upshot is this. If you want to accept Naïve Comprehension whilst *avoiding* inconsistency, you cannot just tinker with the *set theory.* Instead, you would have to overhaul your *logic.*

Of course, set theories with non-classical logics have been presented. But they are—to say the least—non-standard. The standard approach to Russell's Paradox is to treat it as a straightforward non-existence proof, and then to try to learn how to live with it. That is the approach we will follow.

8.3 Predicative and Impredicative

The Russell set, R, was defined via $\{x : x \notin x\}$. Spelled out more fully, R would be the set which contains all and only those sets which are not non-self-membered. So in defining R, we quantify over the domain which would contain R (if it existed).

This is an *impredicative* definition. More generally, we might say that a definition is impredicative iff it quantifies over a domain which contains the object that is being defined.

In the wake of the paradoxes, Whitehead, Russell, Poincaré and Weyl rejected such impredicative definitions as "viciously circular":

> An analysis of the paradoxes to be avoided shows that they all result from a kind of vicious circle. The vicious circles in question arise from supposing that

> a collection of objects may contain members which can only be defined by means of the collection as a whole[. . . . ¶]
>
> The principle which enables us to avoid illegitimate totalities may be stated as follows: 'Whatever involves *all* of a collection must not be one of the collection'; or, conversely: 'If, provided a certain collection had a total, it would have members only definable in terms of that total, then the said collection has no total.' We shall call this the 'vicious-circle principle,' because it enables us to avoid the vicious circles involved in the assumption of illegitimate totalities. (Whitehead and Russell, 1910, p. 37)

If we follow them in rejecting the *vicious-circle principle*, then we might attempt to replace the disastrous Naïve Comprehension Scheme (of section 8.2) with something like this:

Predicative Comprehension. For every formula φ quantifying only over sets: the set′ $\{x : \varphi(x)\}$ exists.

So long as sets′ are not sets, no contradiction will ensue.

Unfortunately, Predicative Comprehension is not very *comprehensive.* After all, it introduces us to new entities, sets′. So we will have to consider formulas which quantify over sets′. If they always yield a set′, then Russell's paradox will arise again, just by considering the set′ of all non-self-membered sets′. So, pursuing the same thought, we must say that a formula quantifying over sets′ yields a corresponding set″. And then we will need sets‴, sets⁗, etc. To prevent a rash of primes, it will be easier to think of these as sets$_0$, sets$_1$, sets$_2$, sets$_3$, sets$_4$,. . . . And this would give us a way into the (simple) theory of types.

There are a few obvious objections against such a theory (though it is not obvious that they are *overwhelming* objections). In brief: the resulting theory is cumbersome to use; it is profligate

in postulating different kinds of objects; and it is not clear, in the end, that impredicative definitions are even *all that bad.*

To bring out the last point, consider this remark from Ramsey:

> we may refer to a man as the tallest in a group, thus identifying him by means of a totality of which he is himself a member without there being any vicious circle. (Ramsey, 1925)

Ramsey's point is that "the tallest man in the group" *is* an impredicative definition; but it is obviously perfectly kosher.

One might respond that, in this case, we could pick out the tallest person by *predicative* means. For example, maybe we could just point at the man in question. The objection against impredicative definitions, then, would clearly need to be limited to entities which can *only* be picked out impredicatively. But even then, we would need to hear more, about why such "essential impredicativity" would be so bad.[1]

Admittedly, impredicative definitions are extremely bad news, if we want our definitions to provide us with something like a recipe for *creating* an object. For, given an impredicative definition, one would genuinely be caught in a vicious circle: to create the impredicatively specified object, one would *first* need to create all the objects (including the impredicatively specified object), since the impredicatively specified object is specified in terms of all the objects; so one would need to create the impredicatively specified object before one had created it itself. But again, this is only a serious objection against "essentially impredicatively" specified sets, if we think of sets as things that we *create.* And we (probably) don't.

As such—for better or worse—the approach which became common does not involve taking a hard line concerning (im)predicativity. Rather, it involves what is now regarded as the cumulative-iterative approach. In the end, this will allow us to

[1] For more, see Linnebo (2010).

stratify our sets into "stages"—a *bit* like the predicative approach stratifies entities into sets_0, sets_1, sets_2, ...—but we will not postulate any difference in kind between them.

8.4 The Cumulative-Iterative Approach

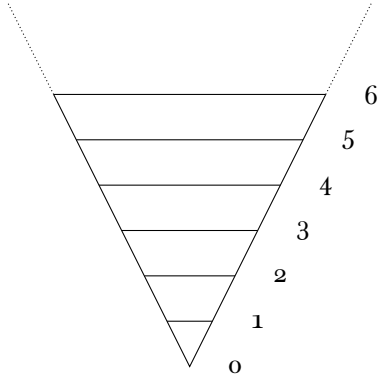
Here is a slightly fuller statement of how we will stratify sets into stages:

> Sets are formed in *stages*. For each stage S, there are certain stages which are *before* S. At stage S, each collection consisting of sets formed at stages before S is formed into a set. There are no sets other than the sets which are formed at stages. (Shoenfield, 1977, p. 323)

This is a sketch of the *cumulative-iterative conception of set*. It will underpin the formal set theory that we present in part III.

Let's explore this in a little more detail. As Shoenfield describes the process, at every stage, we form new sets from the sets which were available to us from earlier stages. So, on Shoenfield's picture, at the initial stage, stage 0, there are no *earlier* stages, and so *a fortiori* there are no sets available to us from earlier stages.[2] So we form only one set: the set with no members $\emptyset$. At stage 1, exactly one set is available to us from earlier stages, so only one new set is $\{\emptyset\}$. At stage 2, two sets are available to us from earlier stages, and we form two new sets $\{\{\emptyset\}\}$ and $\{\emptyset, \{\emptyset\}\}$. At stage 3, four sets are available to us from earlier stages, so we form twelve new sets.... As such, the cumulative-iterative picture of the sets will look a bit like this (with numbers indicating stages):

[2] Why should we assume that there *is* a first stage? See the footnote to *Stages-are-ordered* in section 9.1.



So: why should we embrace this story?

One reason is that it is a nice, tractable story. Given the demise of the most obvious story, i.e., Naïve Comprehension, we are in want of something nice.

But the story is not *just* nice. We have a good reason to believe that any set theory based on this story will be *consistent*. Here is why.

Given the cumulative-iterative conception of set, we form sets at stages; and their members must be objects which were available *already*. So, for any stage S, we can form the set

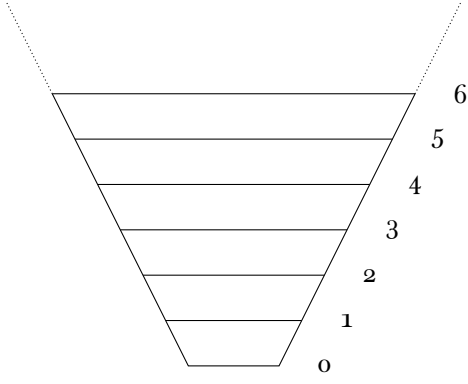
$$R_S = \{x : x \notin x \text{ and } x \text{ was available before } S\}$$

The reasoning involved in proving Russell's Paradox will now establish that R_S itself is not available before stage S. And that's not a contradiction. Moreover, if we embrace the cumulative-iterative conception of set, then we shouldn't even have *expected* to be able to form the Russell set itself. For that would be the set of all non-self-membered sets that "will ever be available". In short: the fact that we (provably) can't form the Russell set isn't *surprising*, given the cumulative-iterative story; it's what we would *predict*.

8.5 Urelements or Not?

In the next few chapters, we will try to extract axioms from the cumulative-iterative conception of set. But, before going any further, we need to say something more about *urelements*.

The picture of section 8.4 allowed us only to form new sets from old *sets*. However, we might want to allow that certain *non-sets*—cows, pigs, grains of sand, or whatever—can be members of sets. In that case, we would start with certain basic elements, *urelements*, and then say that at each stage S we would form "all possible" sets consisting of urelements or sets formed at stages before S (in any combination). The resulting picture would look more like this:



So now we have a decision to take: *Should we allow urelements?*

Philosophically, it makes sense to include urelements in our theorising. The main reason for this is to make our set theory *applicable*. To illustrate the point, recall from chapter 5 that we say that two sets A and B have the same size, i.e., $A \approx B$, iff there is a bijection between them. Now, if the cows in the field and the pigs in the sty both form sets, we can offer a set-theoretical treatment of the claim "there are as many cows as pigs". But if we ban urelements, so that the cows and the pigs do *not* form sets, then that set-theoretical treatment will be unavailable. Indeed, we will have no straightforward ability to apply set theory to any-

thing other than sets themselves. (For more reasons to include urelements, see Potter 2004, pp. vi, 24, 50–1.)

Mathematically, however, it is quite rare to allow urelements. In part, this is because it is *very slightly* easier to formulate set theory without urelements. But, occasionally, one finds more interesting justifications for excluding urelement from set theory:

> In accordance with the belief that set theory is the foundation of mathematics, we should be able to capture all of mathematics by just talking about sets, so our variable should not range over objects like cows and pigs. (Kunen, 1980, p. 8)

So: a focus on applicability would suggest *including* urelements; a focus on a reductive foundational goal (reducing mathematics to pure set theory) might suggest *excluding* them. Mild laziness, too, points in the direction of excluding urelements.

We will follow the laziest path. Partly, though, there is a pedagogical justification. Our aim is to introduce you to the elements of set theory that you would need in order to get started on the philosophy of set theory. And most of that philosophical literature discusses set theories formulated *without* urelements. So this book will, perhaps, be of more use, if it hews fairly closely to that literature.

8.6 Appendix: Frege's Basic Law V

In section 8.2, we explained that Russell's formulated his paradox as a problem for the system Frege outlined in his *Grundgesetze*. Frege's system did not include a direct formulation of Naïve Comprehension. So, in this appendix, we will very briefly explain what Frege's system *did* include, and how it relates to Naïve Comprehension and how it relates to Russell's Paradox.

Frege's system is *second-order*, and was designed to formulate the notion of an *extension of a concept*.[3] Using notation inspired

[3]Strictly speaking, Frege attempts to formalize a more general notion: the

by Frege, we will write $\epsilon x\, F(x)$ for *the extension of the concept* F. This is a device which takes a *predicate*, "F", and turns it into a (first-order) *term*, "$\epsilon x\, F(x)$". Using this device, Frege offered the following *definition* of membership:

$$a \in b =_{\text{df}} \exists G(b = \epsilon x\, G(x) \wedge Ga)$$

roughly: $a \in b$ iff a falls under a concept whose extension is b. (Note that the quantifier "$\exists G$" is second-order.) Frege also maintained the following principle, known as *Basic Law V*:

$$\epsilon x\, F(x) = \epsilon x\, G(x) \leftrightarrow \forall x(Fx \leftrightarrow Gx)$$

roughly: concepts have identical extensions iff they are coextensive. (Again, both "F" and "G" are in predicate position.) Now a simple principle connects membership with property-satisfaction:

Lemma 8.2 (in *Grundgesetze*). $\forall F \forall a(a \in \epsilon x\, F(x) \leftrightarrow Fa)$

Proof. Fix F and a. Now $a \in \epsilon x\, F(x)$ iff $\exists G(\epsilon x\, F(x) = \epsilon x\, G(x) \wedge Ga)$ (by the definition of membership) iff $\exists G(\forall x(Fx \leftrightarrow Gx) \wedge Ga)$ (by Basic Law V) iff Fa (by elementary second-order logic). □

And this yields Naïve Comprehension almost immediately:

Lemma 8.3 (in *Grundgesetze*.). $\forall F \exists s \forall a(a \in s \leftrightarrow Fa)$

Proof. Fix F; now Lemma 8.2 yields $\forall a(a \in \epsilon x\, F(x) \leftrightarrow Fa)$; so $\exists s \forall a(a \in s \leftrightarrow Fa)$ by existential generalisation. The result follows since F was arbitrary. □

Russell's Paradox follows by taking F as given by $\forall x(Fx \leftrightarrow x \notin x)$.

"value-range" of a function. Extensions of concepts are a special case of the more general notion. See Heck (2012, pp. 8–9) for the details.

CHAPTER 9

Steps towards Z

In the previous chapter, we considered the iterative conception of set. In this chapter, we will attempt to extract most of the axioms of Zermelo's set theory, i.e., **Z**. The approach is *entirely* inspired by Boolos (1971), Scott (1974), and Shoenfield (1977).

9.1 The Story in More Detail

In section 8.4, we quoted Schoenfield's description of the process of set-formation. We now want to write down a few more principles, to make this story a bit more precise. Here they are:

> *Stages-are-key.* Every set is formed at some stage.
>
> *Stages-are-ordered.* Stages are ordered: some come *before* others.[1]

[1] We will actually assume—tacitly—that the stages are *well-ordered.* What this amounts to is explained in chapter 10. This is a substantial assumption. In fact, using a very clever technique due to Scott (1974), this assumption can be *avoided* and then *derived.* (This will also explain why we should think that there is an initial stage.) We cannot go into that here; for more, see Button (2021).

> *Stages-accumulate.* For any stage S, and for any sets which were formed *before* stage S: a set is formed at stage S whose members are exactly those sets. Nothing else is formed at stage S.

These are informal principles, but we will be able to use them to vindicate several of the axioms of Zermelo's set theory.

(We should offer a word of caution. Although we will be presenting some completely standard axioms, with completely standard names, the italicized principles we have just presented have no particular names in the literature. We simply monikers which we hope are helpful.)

9.2 Separation

We start with a principle to replace Naïve Comprehension:

Axiom (Scheme of Separation). For every formula $\varphi(x)$, this is an axiom: for any A, the set $\{x \in A : \varphi(x)\}$ exists.

Note that this is not a single axiom. It is a *scheme* of axioms. There are *infinitely many* Separation axioms; one for every formula $\varphi(x)$. The scheme can equally well be (and normally is) written down as follows:

For any formula $\varphi(x)$ which does not contain "S", this is an axiom:

$$\forall A \exists S \forall x (x \in S \leftrightarrow (\varphi(x) \wedge x \in A)).$$

In keeping with the convention noted at the start of part III, the formulas φ in the Separation axioms may have parameters.[2]

Separation is immediately justified by our cumulative-iterative conception of sets we have been telling. To see why, let A be a set.

[2]For an explanation of what this means, see the discussion immediately after Corollary 7.7.

So A is formed by some stage S (by *Stages-are-key*). Since A was formed at stage S, all of A's members were formed before stage S (by *Stages-accumulate*). Now in particular, consider all the sets which are members of A and which also satisfy φ; clearly all of these sets, too, were formed before stage S. So they are formed into a set $\{x \in A : \varphi(x)\}$ at stage S too (by *Stages-accumulate*).

Unlike Naïve Comprehension, this avoid Russell's Paradox. For we cannot simply assert the existence of the set $\{x : x \notin x\}$. Rather, *given* some set A, we can assert the existence of the set $R_A = \{x \in A : x \notin x\}$. But all this proves is that $R_A \notin R_A$ and $R_A \notin A$, none of which is very worrying.

However, Separation has an immediate and striking consequence:

Theorem 9.1. *There is no* universal *set, i.e., $\{x : x = x\}$ does not exist.*

Proof. For reductio, suppose V is a universal set. Then by Separation, $R = \{x \in V : x \notin x\} = \{x : x \notin x\}$ exists, contradicting Russell's Paradox. □

The absence of a universal set—indeed, the open-endedness of the hierarchy of sets—is one of the most fundamental ideas behind the cumulative-iterative conception. So it is worth seeing that, intuitively, we could reach it via a different route. A universal set must be a member of itself. But, on our cumulative-iterative conception, every set appears (for the first time) in the hierarchy at the first stage immediately after all of its members. But this entails that *no* set is self-membered. For any self-membered set would have to first occur immediately after the stage at which it first occurred, which is absurd. (We will see in Definition 11.15 how to make this explanation more rigorous, by using the notion of the "rank" of a set. However, we will need to have a few more axioms in place to do this.)

Here are a few more consequences of Separation and Extensionality.

Proposition 9.2. *If any set exists, then $\emptyset$ exists.*

Proof. If A is a set, $\emptyset = \{x \in A : x \neq x\}$ exists by Separation. □

Proposition 9.3. *$A \setminus B$ exists for any sets A and B*

Proof. $A \setminus B = \{x \in A : x \notin B\}$ exists by Separation. □

It also turns out that (almost) arbitrary intersections exist:

Proposition 9.4. *If $A \neq \emptyset$, then $\bigcap A = \{x : (\forall y \in A) x \in y\}$ exists.*

Proof. Let $A \neq \emptyset$, so there is some $c \in A$. Then $\bigcap A = \{x : (\forall y \in A) x \in y\} = \{x \in c : (\forall y \in A) x \in y\}$, which exists by Separation. □

Note the condition that $A \neq \emptyset$, though; for $\bigcap \emptyset$ would be the universal set, vacuously, contradicting Theorem 9.1.

9.3 Union

Proposition 9.4 gave us intersections. But if we want arbitrary unions to exist, we need to lay down another axiom:

Axiom (Union). For any set A, the set $\bigcup A = \{x : (\exists b \in A) x \in b\}$ exists.

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

This axiom is also justified by the cumulative-iterative conception. Let A be a set, so A is formed at some stage S (by *Stages-are-key*). Every member of A was formed *before* S (by *Stages-accumulate*); so, reasoning similarly, every member of every member of A was formed before S. Thus all of *those* sets are available before S, to be formed into a set at S. And that set is just $\bigcup A$.

9.4 Pairs

The next axiom to consider is the following:

Axiom (Pairs). For any sets a, b, the set $\{a, b\}$ exists.

$$\forall a \forall b \exists P \forall x (x \in P \leftrightarrow (x = a \vee x = b))$$

Here is how to justify this axiom, using the iterative conception. Suppose a is available at stage S, and b is available at stage T. Let M be whichever of stages S and T comes later. Then since a and b are both available at stage M, the set $\{a, b\}$ is a possible collection available at any stage after M (whichever is the greater).

But hold on! Why assume that there *are* any stages after M? If there are none, then our justification will fail. So, to justify Pairs, we will have to add another principle to the story we told in section 9.1, namely:

> *Stages-keep-going.* There is no last stage.

Is this principle justified? Nothing in Shoenfield's story stated *explicitly* that there is no last stage. Still, even if it is (strictly speaking) an extra addition to our story, it fits well with the basic idea that sets are formed in stages. We will simply accept it in what follows. And so, we will accept the Axiom of Pairs too.

Armed with this new Axiom, we can prove the existence of plenty more sets. For example:

Proposition 9.5. *For any sets a and b, the following sets exist:*

1. $\{a\}$
2. $a \cup b$
3. $\langle a, b \rangle$

Proof. (1). By Pairs, $\{a, a\}$ exists, which is $\{a\}$ by Extensionality.

(2). By Pairs, $\{a,b\}$ exists. Now $a \cup b = \bigcup\{a,b\}$ exists by Union.

(3). By (1), $\{a\}$ exists. By Pairs, $\{a,b\}$ exists. Now $\{\{a\},\{a,b\}\} = \langle a,b\rangle$ exists, by Pairs again. □

9.5 Powersets

We will proceed with another axiom:

Axiom (Powersets). For any set A, the set $\wp(A) = \{x : x \subseteq A\}$ exists.

$$\forall A \exists P \forall x (x \in P \leftrightarrow (\forall z \in x) z \in A)$$

Our justification for this is pretty straightforward. Suppose A is formed at stage S. Then all of A's members were available before S (by *Stages-accumulate*). So, reasoning as in our justification for Separation, every subset of A is formed by stage S. So they are all available, to be formed into a single set, at any stage after S. And we know that there is some such stage, since S is not the last stage (by *Stages-keep-going*). So $\wp(A)$ exists.

Here is a nice consequence of Powersets:

Proposition 9.6. *Given any sets A, B, their Cartesian product $A \times B$ exists.*

Proof. The set $\wp(\wp(A \cup B))$ exists by Powersets and Proposition 9.5. So by Separation, this set exists:

$$C = \{z \in \wp(\wp(A \cup B)) : (\exists x \in A)(\exists y \in B) z = \langle x,y\rangle\}.$$

Now, for any $x \in A$ and $y \in B$, the set $\langle x,y\rangle$ exists by Proposition 9.5. Moreover, since $x,y \in A \cup B$, we have that $\{x\},\{x,y\} \in \wp(A \cup B)$, and $\langle x,y\rangle \in \wp(\wp(A \cup B))$. So $A \times B = C$. □

In this proof, Powerset interacts with Separation. And that is no surprise. Without Separation, Powersets wouldn't be a very *powerful* principle. After all, Separation tells us which subsets of a set exist, and hence determines just how "fat" each Powerset is.

9.6 Infinity

We already have enough axioms to ensure that there are infinitely many sets (if there are any). For suppose some set exists, and so $\emptyset$ exists (by Proposition 9.2). Now for any set x, the set $x \cup \{x\}$ exists by Proposition 9.5. So, applying this a few times, we will get sets as follows:

0. $\emptyset$
1. $\{\emptyset\}$
2. $\{\emptyset, \{\emptyset\}\}$
3. $\{\emptyset, \{\emptyset\}, \{\emptyset, \{\emptyset\}\}\}$
4. $\{\emptyset, \{\emptyset\}, \{\emptyset, \{\emptyset\}\}, \{\emptyset, \{\emptyset\}, \{\emptyset, \{\emptyset\}\}\}\}$

and we can check that each of these sets is distinct.

We have started the numbering from 0, for a few reasons. But one of them is this. It is not that hard to check that the set we have labelled "n" has exactly n members, and (intuitively) is formed at the nth stage.

But. This gives us *infinitely many* sets, but it does not guarantee that there is an *infinite set*, i.e., a set with infinitely many members. And this really matters: unless we can find a (Dedekind) infinite set, we cannot construct a Dedekind algebra. But we want a Dedekind algebra, so that we can treat it as the set of natural numbers. (Compare section 7.4.)

Importantly, the axioms we have laid down so far do *not* guarantee the existence of any infinite set. So we have to lay down a new axiom:

Axiom (Infinity). There is a set, I, such that $\emptyset \in I$ and $x \cup \{x\} \in$

I whenever $x \in I$.

$$\exists I((\exists o \in I)\forall x\ x \notin o \wedge (\forall x \in I)(\exists s \in I)\forall z(z \in s \leftrightarrow (z \in x \vee z = x)))$$

It is easy to see that the set I given to us by the Axiom of Infinity is Dedekind infinite. Its distinguished element is $\emptyset$, and the injection on I is given by $s(x) = x \cup \{x\}$. Now, Theorem 7.5 showed how to extract a Dedekind Algebra from a Dedekind infinite set; and we will treat this as our set of natural numbers. More precisely:

Definition 9.7. Let I be any set given to us by the Axiom of Infinity. Let s be the function $s(x) = x \cup \{x\}$. Let $\omega = \mathrm{clo}_s(\emptyset)$. We call the members of ω the *natural numbers*, and say that n is the result of n-many applications of s to $\emptyset$.

You can now look back and check that the set labelled "n", a few paragraphs earlier, will be treated *as* the number n.

We will discuss this significance of this stipulation in section 9.8. For now, it enables us to prove an intuitive result:

Proposition 9.8. *No natural number is Dedekind infinite.*

Proof. The proof is by induction, i.e., Theorem 7.6. Clearly $0 = \emptyset$ is not Dedekind infinite. For the induction step, we will establish the contrapositive: if (absurdly) $s(n)$ is Dedekind infinite, then n is Dedekind infinite.

So suppose that $s(n)$ is Dedekind infinite, i.e., there is some injection f with $\mathrm{ran}(f) \subsetneq \mathrm{dom}(f) = s(n) = n \cup \{n\}$. There are two cases to consider.

Case 1: $n \notin \mathrm{ran}(f)$. So $\mathrm{ran}(f) \subseteq n$, and $f(n) \in n$. Let $g = f\restriction_n$; now $\mathrm{ran}(g) = \mathrm{ran}(f) \setminus \{f(n)\} \subsetneq n = \mathrm{dom}(g)$. Hence n is Dedekind infinite.

Case 2: $n \in \text{ran}(f)$. Fix $m \in \text{dom}(f) \setminus \text{ran}(f)$, and define a function h with domain $s(n) = n \cup \{n\}$:

$$h(x) = \begin{cases} f(x) & \text{if } f(x) \neq n \\ m & \text{if } f(x) = n \end{cases}$$

So h and f agree everywhere, except that $h(f^{-1}(n)) = m \neq n = f(f^{-1}(n))$. Since f is an injection, $n \notin \text{ran}(h)$; and $\text{ran}(h) \subsetneq \text{dom}(h) = s(n)$. Now n is Dedekind infinite, using the argument of Case 1. □

The question remains, though, of how we might *justify* the Axiom of Infinity. The short answer is that we will need to add another principle to the story we have been telling. That principle is as follows:

> *Stages-hit-infinity.* There is an infinite stage. That is, there is a stage which (a) is not the first stage, and which (b) has some stages before it, but which (c) has no immediate predecessor.

The Axiom of Infinity follows straightforwardly from this principle. We know that natural number n is formed at stage n. So the set ω is formed at the first infinite stage. And ω itself witnesses the Axiom of Infinity.

This, however, simply pushes us back to the question of how we might justify *Stages-hit-infinity*. As with *Stages-keep-going*, it was not an explicit part of the story we told about the cumulative-iterative hierarchy. But more than that: nothing in the very idea of an iterative hierarchy, in which sets are formed stage by stage, forces us to think that the process involves an *infinite* stage. It seems perfectly coherent to think that the stages are ordered like the natural numbers.

This, however, gives rise to an obvious problem. In section 7.4, we considered Dedekind's "proof" that there is a Dedekind infinite set (of thoughts). This may not have struck you as very satisfying. But if *Stages-hit-infinity* is not "forced upon

us" by the iterative conception of set (or by "the laws of thought"), then we are still left without an intrinsic justification for the claim that there is a Dedekind infinite set.

There is much more to say here, of course. But hopefully you are now at a point to start thinking about what it might *take* to justify an axiom (or principle). In what follows we will simply take *Stages-hit-infinity* for granted.

9.7 $\mathbf{Z}^-$: a Milestone

We will revisit *Stages-hit-infinity* in the next section. However, with the Axiom of Infinity, we have reached an important milestone. We now have all the axioms required for the theory $\mathbf{Z}^-$. In detail:

Definition 9.9. The theory $\mathbf{Z}^-$ has these axioms: Extensionality, Union, Pairs, Powersets, Infinity, and all instances of the Separation scheme.

The name stands for *Zermelo* set theory (*minus* something which we will come to later). Zermelo deserves the honour, since he essentially formulated this theory in his 1908a.[3]

This theory is powerful enough to allow us to do an enormous amount of mathematics. In particular, you *should* look back through part II, and convince yourself that everything we did, naïvely, could be done more formally within $\mathbf{Z}^-$. (Once you have done that for a bit, you might want to skip ahead and read section 9.9.) So, henceforth, and without any further comment, we will take ourselves to be working in $\mathbf{Z}^-$ (at least).

9.8 Selecting our Natural Numbers

In Definition 9.7, we explicitly defined the expression "natural numbers". How should you understand this stipulation? It is not

[3]For interesting comments on the history and technicalities, see Potter (2004, Appendix A).

a metaphysical claim, but just a decision to *treat* certain sets as the natural numbers. We touched upon reasons for thinking this in section 3.2, section 6.5 and section 7.4. But we can make these reasons even more pointed.

Our Axiom of Infinity follows von Neumann (1925). But here is another axiom, which we could have adopted instead:

Zermelo's 1908a Axiom of Infinity. There is a set A such that $\emptyset \in A$ and $(\forall x \in A)\{x\} \in A$.

Had we used Zermelo's axiom, instead of our (von Neumann-inspired) Axiom of Infinity, we would equally well have been given a Dedekind infinite set, and so a Dedekind algebra. On Zermelo's approach, the distinguished element of our algebra would again have been $\emptyset$ (our surrogate for 0), but the injection would have been given by the map $x \mapsto \{x\}$, rather than $x \mapsto x \cup \{x\}$. The simplest upshot of this is that Zermelo treats 2 as $\{\{\emptyset\}\}$, whereas we (with von Neumann) treat 2 as $\{\emptyset, \{\emptyset\}\}$.

Why choose one axiom of Infinity rather than the other? The main practical reason is that von Neumann's approach "scales up" to handle transfinite numbers rather well. We will explore this from chapter 10 onwards. However, from the simple perspective of *doing arithmetic*, both approaches would do equally well. So if someone tells you that the natural numbers *are* sets, the obvious question is: *Which sets are they?*

This precise question was made famous by Benacerraf (1965). But it is worth emphasising that it is just the most famous example of a phenomenon that we have encountered many times already. The basic point is this. Set theory gives us a way to *simulate* a bunch of "intuitive" kinds of entities: the reals, rationals, integers, and naturals, yes; but also ordered pairs, functions, and relations. However, set theory never provides us with a *unique* choice of simulation. There are *always* alternatives which—straightforwardly—would have served us just as well.

9.9 Appendix: Closure, Comprehension, and Intersection

In section 9.7, we suggested that you should look back through the naïve work of part II and check that it can be carried out in $\mathbf{Z}^-$. If you followed that advice, *one* point might have tripped you up: the use of *intersection* in Dedekind's treatment of *closures*.

Recall from Definition 7.2 that

$$\mathrm{clo}_f(o) = \bigcap\{X : o \in X \text{ and } X \text{ is } f\text{-closed}\}.$$

The general shape of this is a definition of the form:

$$C = \bigcap\{X : \varphi(X)\}.$$

But this should ring alarm bells: since Naïve Comprehension fails, there is no guarantee that $\{X : \varphi(X)\}$ exists. It looks dangerously, then, like such definitions are *cheating*.

Fortunately, they are not cheating; or rather, if they *are* cheating as they stand, then we can engage in some honest toil to render them kosher. That honest toil was foreshadowed in Proposition 9.4, when we explained why $\bigcap A$ exists for any $A \neq \emptyset$. But we will spell it out explicitly.

Given Extensionality, if we attempt to define C as $\bigcap\{X : \varphi(X)\}$, all we are really asking is for an object C which obeys the following:

$$\forall x(x \in C \leftrightarrow \forall X(\varphi(X) \rightarrow x \in X)) \qquad (*)$$

Now, suppose there is *some* set, S, such that $\varphi(S)$. Then to deliver eq. (*), we can simply define C using *Separation*, as follows:

$$C = \{x \in S : \forall X(\varphi(X) \rightarrow x \in X)\}.$$

We leave it as an exercise to check that this definition yields eq. (*), as desired. And this general strategy will allow us to circumvent any apparent use of Naïve Comprehension in defining intersections. In the particular case which got us started on

this line of thought, namely that of $\mathrm{clo}_f(o)$, here is how that would work. We began the proof of Lemma 7.3 by noting that $o \in \mathrm{ran}(f) \cup \{o\}$ and that $\mathrm{ran}(f) \cup \{o\}$ is f-closed. So, we can define what we want thus:

$$\mathrm{clo}_f(o) = \{x \in \mathrm{ran}(f) \cup \{o\} : (\forall X \ni o)(X \text{ is } f\text{-closed} \to x \in X)\}.$$

Problems

Problem 9.1. Show that, for any sets a, b, c, the set $\{a, b, c\}$ exists.

Problem 9.2. Show that, for any sets $a_1, \ldots, a_n$, the set $\{a_1, \ldots, a_n\}$ exists.

Problem 9.3. Show that, for any sets A, B: (i) the set of all relations with domain A and range B exists; and (ii) the set of all functions from A to B exists.

Problem 9.4. Let A be a set, and let $\sim$ be an equivalence relation on A. Prove that the set of equivalence classes under $\sim$ on A, i.e., $A/_\sim$, exists.

CHAPTER 10

Ordinals

10.1 Introduction

In chapter 9, we postulated that there is an infinite-th stage of the hierarchy, in the form of *Stages-hit-infinity* (see also our axiom of Infinity). However, given *Stages-keep-going*, we can't stop at the infinite-th stage; we have to keep going. So: at the next stage after the first infinite stage, we form all possible collections of sets that were available at the first infinite stage; and repeat; and repeat; and repeat; ...

Implicitly what has happened here is that we have started to invoke an "intuitive" notion of number, according to which there can be numbers *after* all the natural numbers. In particular, the notion involved is that of a *transfinite ordinal*. The aim of this chapter is to make this idea more rigorous. We will explore the general notion of an ordinal, and then explicitly define certain sets to be our ordinals.

10.2 The General Idea of an Ordinal

Consider the natural numbers, in their usual order:

$$0 < 1 < 2 < 3 < 4 < 5 < \cdots$$

We call this, in the jargon, an ω-sequence. And indeed, this general ordering is mirrored in our initial construction of the stages of the set hierarchy. But, now suppose we move 0 to the end of this sequence, so that it comes after all the other numbers:

$$1 < 2 < 3 < 4 < 5 < \cdots < 0$$

We have the same entities here, but ordered in a fundamentally different way: our first ordering had no last element; our new ordering does. Indeed, our new ordering consists of an ω-sequence of entities $(1, 2, 3, 4, 5, \ldots)$, followed by another entity. It will be an $\omega + 1$-sequence.

We can generate even more types of ordering, using just these entities. For example, consider all the even numbers (in their natural order) followed by all the odd numbers (in their natural order):

$$0 < 2 < 4 < \cdots < 1 < 3 < \cdots$$

This is an ω-sequence followed by another ω-sequence; an $\omega + \omega$-sequence.

Well, we can keep going. But what we would like is a general way to understand this talk about *orderings*.

10.3 Well-Orderings

The fundamental notion is as follows:

Definition 10.1. The relation $<$ *well-orders* A iff it meets these two conditions:

1. $<$ is connected, i.e., for all $a, b \in A$, either $a < b$ or $a = b$ or $b < a$;
2. every non-empty subset of A has a $<$-minimal member, i.e., if $\emptyset \neq X \subseteq A$ then $(\exists m \in X)(\forall z \in X) z \not< m$

It is easy to see that three examples we just considered were indeed well-ordering relations.

Here are some elementary but extremely important observations concerning well-ordering.

Proposition 10.2. *If $<$ well-orders A, then every non-empty subset of A has a unique $<$-least member, and $<$ is irreflexive, asymmetric and transitive.*

Proof. If X is a non-empty subset of A, it has a $<$-minimal member m, i.e., $(\forall z \in X) z \not< m$. Since $<$ is connected, $(\forall z \in X) m \leq z$. So m is the $<$-least member of X.

For irreflexivity, fix $a \in A$; the $<$-least member of $\{a\}$ is a, so $a \not< a$. For transitivity, if $a < b < c$, then since $\{a, b, c\}$ has a $<$-least member, $a < c$. Asymmetry follows from irreflexivity and transitivity □

Proposition 10.3. *If $<$ well-orders A, then for any formula $\varphi(x)$:*

$$\textit{if } (\forall a \in A)((\forall b < a)\varphi(b) \to \varphi(a)), \textit{ then } (\forall a \in A)\varphi(a).$$

Proof. We will prove the contrapositive. Suppose $\neg(\forall a \in A)\varphi(a)$, i.e., that $X = \{x \in A : \neg\varphi(x)\} \neq \emptyset$. Then X has an $<$-minimal member, a. So $(\forall b < a)\varphi(b)$ but $\neg\varphi(a)$. □

This last property should remind you of the principle of strong induction on the naturals, i.e.: if $(\forall n \in \omega)((\forall m < n)\varphi(m) \to \varphi(n))$, then $(\forall n \in \omega)\varphi(n)$. And this property makes well-ordering into a very *robust* notion.[1]

10.4 Order-Isomorphisms

To explain *how* robust well-ordering is, we will start by introducing a method for comparing well-orderings.

[1] A reminder: all formulas can have parameters (unless explicitly stated otherwise).

Definition 10.4. A *well-ordering* is a pair $\langle A, < \rangle$, such that $<$ well-orders A. The well-orderings $\langle A, < \rangle$ and $\langle B, \lessdot \rangle$ are *order-isomorphic* iff there is a bijection $f : A \to B$ such that: $x < y$ iff $f(x) \lessdot f(y)$. In this case, we write $\langle A, < \rangle \cong \langle B, \lessdot \rangle$, and say that f is an *order-isomorphism.*

In what follows, for brevity, we will speak of "isomorphisms" rather than "order-isomorphisms". Intuitively, isomorphisms are structure-preserving bijections. Here are some simple facts about isomorphisms.

Lemma 10.5. *Compositions of isomorphisms are isomorphisms, i.e.: if $f : A \to B$ and $g : B \to C$ are isomorphisms, then $(g \circ f) : A \to C$ is an isomorphism.*

Proof. Left as an exercise. □

Corollary 10.6. *$X \cong Y$ is an equivalence relation.*

Proposition 10.7. *If $\langle A, < \rangle$ and $\langle B, \lessdot \rangle$ are isomorphic well-orderings, then the isomorphism between them is unique.*

Proof. Let f and g be isomorphisms $A \to B$. We will prove the result by induction, i.e. using Proposition 10.3. Fix $a \in A$, and suppose (for induction) that $(\forall b < a) f(b) = g(b)$. Fix $x \in B$.

If $x \lessdot f(a)$, then $f^{-1}(x) < a$, so $g(f^{-1}(x)) \lessdot g(a)$, invoking the fact that f and g are isomorphisms. But since $f^{-1}(x) < a$, by our supposition $x = f(f^{-1}(x)) = g(f^{-1}(x))$. So $x \lessdot g(a)$. Similarly, if $x \lessdot g(a)$ then $x \lessdot f(a)$.

Generalising, $(\forall x \in B)(x \lessdot f(a) \leftrightarrow x \lessdot g(a))$. It follows that $f(a) = g(a)$ by Proposition 3.26. So $(\forall a \in A) f(a) = g(a)$ by Proposition 10.3. □

This gives some sense that well-orderings are robust. But to continue explaining this, it will help to introduce some more notation.

Definition 10.8. When $\langle A, <\rangle$ is a well-ordering with $a \in A$, let $A_a = \{x \in A : x < a\}$. We say that A_a is a proper *initial segment* of A (and allow that A itself is an improper initial segment of A). Let $<_a$ be the restriction of $<$ to the initial segment, i.e., $< \restriction_{A_a^2}$.

Using this notation, we can state and prove that no well-ordering is isomorphic to any of its proper initial segments.

Lemma 10.9. *If $\langle A, <\rangle$ is a well-ordering with $a \in A$, then $\langle A, <\rangle \not\cong \langle A_a, <_a\rangle$*

Proof. For reductio, suppose $f : A \to A_a$ is an isomorphism. Since f is a bijection and $A_a \subsetneq A$, using Proposition 10.2 let $b \in A$ be the $<$-least member of A such that $b \neq f(b)$. We'll show that $(\forall x \in A)(x < b \leftrightarrow x < f(b))$, from which it will follow by Proposition 3.26 that $b = f(b)$, completing the reductio.

Suppose $x < b$. So $x = f(x)$, by the choice of b. And $f(x) < f(b)$, as f is an isomorphism. So $x < f(b)$.

Suppose $x < f(b)$. So $f^{-1}(x) < b$, since f is an isomorphism, and so $f^{-1}(x) = x$ by the choice of b. So $x < b$. □

Our next result shows, roughly put, that an "initial segment" of an isomorphism is an isomorphism:

Lemma 10.10. *Let $\langle A, <\rangle$ and $\langle B, \lessdot\rangle$ be well-orderings. If $f : A \to B$ is an isomorphism and $a \in A$, then $f \restriction_{A_a} : A_a \to B_{f(a)}$ is an isomorphism.*

Proof. Since f is an isomorphism:

$$\begin{aligned} f[A_a] &= f[\{x \in A : x < a\}] \\ &= f[\{f^{-1}(y) \in A : f^{-1}(y) < a\}] \\ &= \{y \in B : y \lessdot f(a)\} \\ &= B_{f(a)} \end{aligned}$$

And $f \restriction_{A_a}$ preserves order because f does. □

Our next two results establish that well-orderings are always comparable:

Lemma 10.11. *Let $\langle A, <\rangle$ and $\langle B, \lessdot\rangle$ be well-orderings. If $\langle A_{a_1}, <_{a_1}\rangle \cong \langle B_{b_1}, \lessdot_{b_1}\rangle$ and $\langle A_{a_2}, <_{a_2}\rangle \cong \langle B_{b_2}, \lessdot_{b_2}\rangle$, then $a_1 < a_2$ iff $b_1 \lessdot b_2$*

Proof. We will prove *left to right*; the other direction is similar. Suppose both $\langle A_{a_1}, <_{a_1}\rangle \cong \langle B_{b_1}, \lessdot_{b_1}\rangle$ and $\langle A_{a_2}, <_{a_2}\rangle \cong \langle B_{b_2}, \lessdot_{b_2}\rangle$, with $f : A_{a_2} \to B_{b_2}$ our isomorphism. Let $a_1 < a_2$; then $\langle A_{a_1}, <_{a_1}\rangle \cong \langle B_{f(a_1)}, \lessdot_{f(a_1)}\rangle$ by Lemma 10.10. So $\langle B_{b_1}, \lessdot_{b_1}\rangle \cong \langle B_{f(a_1)}, \lessdot_{f(a_1)}\rangle$, and so $b_1 = f(a_1)$ by Lemma 10.9. Now $b_1 \lessdot b_2$ as f's domain is B_{b_2}. □

Theorem 10.12. *Given any two well-orderings, one is isomorphic to an initial segment (not necessarily proper) of the other.*

Proof. Let $\langle A, <\rangle$ and $\langle B, \lessdot\rangle$ be well-orderings. Using Separation, let

$$f = \{\langle a, b\rangle \in A \times B : \langle A_a, <_a\rangle \cong \langle B_b, \lessdot_b\rangle\}.$$

By Lemma 10.11, $a_1 < a_2$ iff $b_1 \lessdot b_2$ for all $\langle a_1, b_1\rangle, \langle a_2, b_2\rangle \in f$. So $f : \mathrm{dom}(f) \to \mathrm{ran}(f)$ is an isomorphism.

If $a_2 \in \mathrm{dom}(f)$ and $a_1 < a_2$, then $a_1 \in \mathrm{dom}(f)$ by Lemma 10.10; so $\mathrm{dom}(f)$ is an initial segment of A. Similarly, $\mathrm{ran}(f)$ is an initial segment of B. For reductio, suppose both are *proper* initial segments. Then let a be the $<$-least member of $A \setminus \mathrm{dom}(f)$, so that $\mathrm{dom}(f) = A_a$, and let b be the $\lessdot$-least member of $B \setminus \mathrm{ran}(f)$, so that $\mathrm{ran}(f) = B_b$. So $f : A_a \to B_b$ is an isomorphism, and hence $\langle a, b\rangle \in f$, a contradiction. □

10.5 Von Neumann's Construction of the Ordinals

Theorem 10.12 gives rise to a thought. We could introduce certain objects, called *order types*, to go proxy for the well-orderings.

Writing $\text{ord}(A,<)$ for the order type of the well-ordering $\langle A,<\rangle$, we would hope to secure the following two principles:

$$\text{ord}(A,<) = \text{ord}(B,\lessdot) \text{ iff } \langle A,<\rangle \cong \langle B,\lessdot\rangle$$
$$\text{ord}(A,<) < \text{ord}(B,\lessdot) \text{ iff } \langle A,<\rangle \cong \langle B_b,\lessdot_b\rangle \text{ for some } b \in B$$

Moreover, we might hope to introduce order-types *as certain sets*, just as we can introduce the natural numbers as certain sets.

The most common way to do this—and the approach we will follow—is to define these order-types via certain *canonical* well-ordered sets. These canonical sets were first introduced by von Neumann:

Definition 10.13. The set A is *transitive* iff $(\forall x \in A)x \subseteq A$. Then A is an *ordinal* iff A is transitive and well-ordered by $\in$.

In what follows, we will use Greek letters for ordinals. It follows immediately from the definition that, if α is an ordinal, then $\langle \alpha, \in_\alpha \rangle$ is a well-ordering, where $\in_\alpha = \{\langle x,y\rangle \in \alpha^2 : x \in y\}$. So, abusing notation a little, we can just say that α *itself* is a well-ordering.

Here are our first few ordinals:

$$\emptyset, \{\emptyset\}, \{\emptyset, \{\emptyset\}\}, \{\emptyset, \{\emptyset\}, \{\emptyset, \{\emptyset\}\}\}, \ldots$$

You will note that these are the first few ordinals that we encountered in our Axiom of Infinity, i.e., in von Neumann's definition of ω (see section 9.6). This is no coincidence. Von Neumann's definition of the ordinals treats natural numbers as ordinals, but allows for transfinite ordinals too.

As always, we can now ask: *are* these the ordinals? Or has von Neumann simply given us some sets that we can *treat* as the ordinals? The kinds of discussions one might have about this question are similar to the discussions we had in section 3.2, section 6.5, section 7.4, and section 9.8, so we will not belabour the point. Instead, in what follows, we will simply use "the ordinals" to speak of "the von Neumann ordinals".

10.6 Basic Properties of the Ordinals

We observed that the first few ordinals are the natural numbers. The main reason for developing a theory of ordinals is to extend the principle of induction which holds on the natural numbers. We will build up to this via a sequence of elementary results.

Lemma 10.14. *Every member of an ordinal is an ordinal.*

Proof. Let α be an ordinal with $b \in \alpha$. Since α is transitive, $b \subseteq \alpha$. So $\in$ well-orders b as $\in$ well-orders α.

To see that b is transitive, suppose $x \in c \in b$. So $c \in \alpha$ as $b \subseteq \alpha$. Again, as α is transitive, $c \subseteq \alpha$, so that $x \in \alpha$. So $x, c, b \in \alpha$. But $\in$ well-orders α, so that $\in$ is a transitive relation on α by Proposition 10.2. So since $x \in c \in b$, we have $x \in b$. Generalising, $c \subseteq b$ □

Corollary 10.15. $\alpha = \{\beta \in \alpha : \beta \text{ is an ordinal}\}$, *for any ordinal* α

Proof. Immediate from Lemma 10.14. □

The rough gist of the next two main results, Theorem 10.16 and Theorem 10.17, is that the ordinals themselves are well-ordered by membership:

Theorem 10.16 (Transfinite Induction). *For any formula* $\varphi(x)$*:*

$$\text{if } \exists\alpha\varphi(\alpha), \text{ then } \exists\alpha(\varphi(\alpha) \wedge (\forall\beta \in \alpha)\neg\varphi(\beta))$$

where the displayed quantifiers are implicitly restricted to ordinals.

Proof. Suppose $\varphi(\alpha)$, for some ordinal α. If $(\forall\beta \in \alpha)\neg\varphi(\beta)$, then we are done. Otherwise, as α is an ordinal, it has some $\in$-least member which is φ, and this is an ordinal by Lemma 10.14. □

Note that we can equally express Theorem 10.16 as the scheme:

$$\text{if } \forall\alpha((\forall\beta \in \alpha)\varphi(\beta) \to \varphi(\alpha)), \text{ then } \forall\alpha\varphi(\alpha)$$

just by taking $\neg\varphi(\alpha)$ in Theorem 10.16, and then performing elementary logical manipulations.

Theorem 10.17 (Trichotomy). *$\alpha \in \beta \vee \alpha = \beta \vee \beta \in \alpha$, for any ordinals α and β.*

Proof. The proof is by double induction, i.e., using Theorem 10.16 twice. Say that x is *comparable* with y iff $x \in y \vee x = y \vee y \in x$.

For induction, suppose that every ordinal in α is comparable with *every* ordinal. For further induction, suppose that α is comparable with every ordinal in β. We will show that α is comparable with β. By induction on β, it will follow that α is comparable with every ordinal; and so by induction on α, *every* ordinal is comparable with *every* ordinal, as required. It suffices to assume that $\alpha \notin \beta$ and $\beta \notin \alpha$, and show that $\alpha = \beta$.

To show that $\alpha \subseteq \beta$, fix $\gamma \in \alpha$; this is an ordinal by Lemma 10.14. So by the first induction hypothesis, γ is comparable with β. But if either $\gamma = \beta$ or $\beta \in \gamma$ then $\beta \in \alpha$ (invoking the fact that α is transitive if necessary), contrary to our assumption; so $\gamma \in \beta$. Generalising, $\alpha \subseteq \beta$.

Exactly similar reasoning, using the second induction hypothesis, shows that $\beta \subseteq \alpha$. So $\alpha = \beta$. □

As such, we will sometimes write $\alpha < \beta$ rather than $\alpha \in \beta$, since $\in$ is behaving as an ordering relation. There are no deep reasons for this, beyond familiarity, and because it is easier to write $\alpha \leq \beta$ than $\alpha \in \beta \vee \alpha = \beta$.[2]

Here are two quick consequences of our last results, the first of which puts our new notation into action:

[2] We could write $\alpha \trianglelefteq \beta$; but that would be wholly non-standard.

Corollary 10.18. *If $\exists\alpha\varphi(\alpha)$, then $\exists\alpha(\varphi(\alpha)\wedge\forall\beta(\varphi(\beta)\rightarrow\alpha\leq\beta))$. Moreover, for any ordinals α,β,γ, both $\alpha\notin\alpha$ and $\alpha\in\beta\in\gamma\rightarrow\alpha\in\gamma$.*

Proof. Just like Proposition 10.2. □

Corollary 10.19. *A is an ordinal iff A is a transitive set of ordinals.*

Proof. Left-to-right. By Lemma 10.14. *Right-to-left.* If A is a transitive set of ordinals, then $\in$ well-orders A by Theorem 10.16 and Theorem 10.17. □

Now, we glossed Theorem 10.16 and Theorem 10.17 as telling us that $\in$ well-orders the ordinals. However, we have to be *very cautious* about this sort of claim, thanks to the following result:

Theorem 10.20 (Burali-Forti Paradox). *There is no set of all the ordinals*

Proof. For reductio, suppose O is the set of all ordinals. If $\alpha\in\beta\in O$, then α is an ordinal, by Lemma 10.14, so $\alpha\in O$. So O is transitive, and hence O is an ordinal by Corollary 10.19. Hence $O\in O$, contradicting Corollary 10.18. □

This result is named after Burali-Forti. But, it was Cantor in 1899—in a letter to Dedekind—who first saw clearly the *contradiction* in supposing that there is a set of all the ordinals. As van Heijenoort explains:

> Burali-Forti himself considered the contradiction as establishing, by *reductio ad absurdum*, the result that the natural ordering of the ordinals is just a partial ordering. (Heijenoort, 1967, p. 105)

Setting Burali-Forti's mistake to one side, we can summarize the foregoing as follows. Ordinals are sets which are individually well-ordered by membership, and collectively well-ordered by membership (without collectively constituting a set).

Rounding this off, here are some more basic properties about the ordinals which follow from Theorem 10.16 and Theorem 10.17.

Proposition 10.21. *Any strictly descending sequence of ordinals is finite.*

Proof. Any infinite strictly descending sequence of ordinals $\alpha_0 > \alpha_1 > \alpha_2 > \ldots$ has no $<$-minimal member, contradicting Theorem 10.16. □

Proposition 10.22. $\alpha \subseteq \beta \vee \beta \subseteq \alpha$, *for any ordinals* α, β.

Proof. If $\alpha \in \beta$, then $\alpha \subseteq \beta$ as β is transitive. Similarly, if $\beta \in \alpha$, then $\beta \subseteq \alpha$. And if $\alpha = \beta$, then $\alpha \subseteq \beta$ and $\beta \subseteq \alpha$. So by Theorem 10.17 we are done. □

Proposition 10.23. $\alpha = \beta$ *iff* $\alpha \cong \beta$, *for any ordinals* α, β.

Proof. The ordinals are well-orders; so this is immediate from Trichotomy (Theorem 10.17) and Lemma 10.9. □

10.7 Replacement

In section 10.5, we motivated the introduction of ordinals by suggesting that we could treat them as order-types, i.e., canonical proxies for well-orderings. In order for that to work, we would need to prove that *every well-ordering is isomorphic to some ordinal.* This would allow us to define $\mathrm{ord}(A, <)$ as the ordinal α such that $\langle A, < \rangle \cong \alpha$.

Unfortunately, we *cannot* prove the desired result only the Axioms we provided introduced so far. (We will see why in section 12.2, but for now the point is: we can't.) We need a new thought, and here it is:

Axiom (Scheme of Replacement). For any formula $\varphi(x,y)$, the following is an axiom:

> for any A, if $(\forall x \in A)\exists! y\, \varphi(x,y)$, then $\{y : (\exists x \in A)\varphi(x,y)\}$ exists.

As with Separation, this is a scheme: it yields infinitely many axioms, for each of the infinitely many different φ's. And it can equally well be (and normally is) written down thus:

For any formula $\varphi(x,y)$ which does not contain "B", the following is an axiom:

$$\forall A[(\forall x \in A)\exists! y\, \varphi(x,y) \rightarrow \exists B \forall y(y \in B \leftrightarrow (\exists x \in A)\varphi(x,y))]$$

On first encounter, however, this is quite a tangled formula. The following quick consequence of Replacement probably gives a *clearer* expression to the intuitive idea we are working with:

Corollary 10.24. *For any term $\tau(x)$, and any set A, this set exists:*

$$\{\tau(x) : x \in A\} = \{y : (\exists x \in A) y = \tau(x)\}.$$

Proof. Since τ is a *term*, $\forall x \exists! y\, \tau(x) = y$. A fortiori, $(\forall x \in A)\exists! y\, \tau(x) = y$. So $\{y : (\exists x \in A)\tau(x) = y\}$ exists by Replacement. □

This suggests that "Replacement" is a good name for the Axiom: given a set A, you can form a new set, $\{\tau(x) : x \in A\}$, by replacing every member of A with its image under τ. Indeed, following the notation for the image of a set under a function, we might write $\tau[A]$ for $\{\tau(x) : x \in A\}$.

Crucially, however, τ is a *term*. It need not be (a name for) a *function*, in the sense of section 4.3, i.e., a certain set of ordered pairs. After all, if f is a function (in that sense), then the set $f[A] = \{f(x) : x \in A\}$ is just a particular subset of $\mathrm{ran}(f)$, and

that is already guaranteed to exist, just using the axioms of $\mathbf{Z}^-$.[3] Replacement, by contrast, is a *powerful* addition to our axioms, as we will see in chapter 12.

10.8 $\mathbf{ZF}^-$: a milestone

The question of how to justify Replacement (if at all) is not straightforward. As such, we will reserve that for chapter 12. However, with the addition of Replacement, we have reached another important milestone. We now have all the axioms required for the theory $\mathbf{ZF}^-$. In detail:

Definition 10.25. The theory $\mathbf{ZF}^-$ has these axioms: Extensionality, Union, Pairs, Powersets, Infinity, and all instances of the Separation and Replacement schemes. Otherwise put, $\mathbf{ZF}^-$ adds Replacement to $\mathbf{Z}^-$.

This stands for *Zermelo–Fraenkel* set theory (*minus* something which we will come to later). Fraenkel gets the honour, since he is credited with the formulation of Replacement in 1922, although the first precise formulation was due to Skolem (1922).

10.9 Ordinals as Order-Types

Armed with Replacement, and so now working in $\mathbf{ZF}^-$, we can finally prove the result we have been aiming for:

Theorem 10.26. *Every well-ordering is isomorphic to a unique ordinal.*

Proof. Let $\langle A, < \rangle$ be a well-order. By Proposition 10.23, it is isomorphic to at most one ordinal. So, for reductio, suppose $\langle A, < \rangle$ is not isomorphic to *any* ordinal. We will first "make $\langle A, < \rangle$ as small as possible". In detail: if some proper initial segment

[3] Just consider $\{y \in \bigcup\bigcup f : (\exists x \in A) y = f(x)\}$.

$\langle A_a, <_a\rangle$ is not isomorphic to any ordinal, there is a least $a \in A$ with that property; then let $B = A_a$ and $\lessdot = <_a$. Otherwise, let $B = A$ and $\lessdot = <$.

By definition, every proper initial segment of B is isomorphic to some ordinal, which is unique as above. So by Replacement, the following set exists, and is a function:

$$f = \{\langle \beta, b\rangle : b \in B \text{ and } \beta \cong \langle B_b, \lessdot_b\rangle\}$$

To complete the reductio, we'll show that f is an isomorphism $\alpha \to B$, for some ordinal α.

It is obvious that $\mathrm{ran}(f) = B$. And by Lemma 10.11, f preserves ordering, i.e., $\gamma \in \beta$ iff $f(\gamma) \lessdot f(\beta)$. To show that $\mathrm{dom}(f)$ is an ordinal, by Corollary 10.19 it suffices to show that $\mathrm{dom}(f)$ is transitive. So fix $\beta \in \mathrm{dom}(f)$, i.e., $\beta \cong \langle B_b, \lessdot_b\rangle$ for some b. If $\gamma \in \beta$, then $\gamma \in \mathrm{dom}(f)$ by Lemma 10.10; generalising, $\beta \subseteq \mathrm{dom}(f)$. □

This result licenses the following definition, which we have wanted to offer since section 10.5:

Definition 10.27. If $\langle A, <\rangle$ is a well-ordering, then its order type, $\mathrm{ord}(A, <)$, is the unique ordinal α such that $\langle A, <\rangle \cong \alpha$.

Moreover, this definition licenses two nice principles:

Corollary 10.28. *Where $\langle A, <\rangle$ and $\langle B, \lessdot\rangle$ are well-orderings:*

$\mathrm{ord}(A, <) = \mathrm{ord}(B, \lessdot)$ *iff* $\langle A, <\rangle \cong \langle B, \lessdot\rangle$
$\mathrm{ord}(A, <) \in \mathrm{ord}(B, \lessdot)$ *iff* $\langle A, <\rangle \cong \langle B_b, \lessdot_b\rangle$ *for some* $b \in B$

Proof. The identity holds by Proposition 10.23. To prove the second claim, let $\mathrm{ord}(A, <) = \alpha$ and $\mathrm{ord}(B, \lessdot) = \beta$, and let $f : \beta \to \langle B, \lessdot\rangle$ be our isomorphism. Then:

$$\begin{aligned} \alpha \in \beta &\text{ iff } f\restriction_\alpha : \alpha \to B_{f(\alpha)} \text{ is an isomorphism} \\ &\text{iff } \langle A, <\rangle \cong \langle B_{f(\alpha)}, \lessdot_{f(\alpha)}\rangle \end{aligned}$$

$$\text{iff } \langle A, < \rangle \cong \langle B_b, <_b \rangle \text{ for some } b \in B$$

by Proposition 10.7, Lemma 10.10, and Corollary 10.15. □

10.10 Successor and Limit Ordinals

In the next few chapters, we will use ordinals a great deal. So it will help if we introduce some simple notions.

Definition 10.29. For any ordinal α, its *successor* is $\alpha^+ = \alpha \cup \{\alpha\}$. We say that α is a *successor* ordinal if $\beta^+ = \alpha$ for some ordinal β. We say that α is a *limit* ordinal iff α is neither empty nor a successor ordinal.

The following result shows that this is the right notion of *successor*:

Proposition 10.30. *For any ordinal α:*

1. $\alpha \in \alpha^+$;
2. α^+ *is an ordinal;*
3. *there is no ordinal β such that $\alpha \in \beta \in \alpha^+$.*

Proof. Trivially, $\alpha \in \alpha \cup \{\alpha\} = \alpha^+$. Equally, α^+ is a transitive set of ordinals, and hence an ordinal by Corollary 10.19. And it is impossible that $\alpha \in \beta \in \alpha^+$, since then either $\beta \in \alpha$ or $\beta = \alpha$, contradicting Corollary 10.18. □

This also licenses a variant of proof by transfinite induction:

Theorem 10.31 (Simple Transfinite Induction). *Let $\varphi(x)$ be a formula such that:*

1. $\varphi(\emptyset)$*; and*
2. *for any ordinal α, if $\varphi(\alpha)$ then $\varphi(\alpha^+)$; and*
3. *if α is a limit ordinal and $(\forall \beta \in \alpha)\varphi(\beta)$, then $\varphi(\alpha)$.*

Then $\forall\alpha\varphi(\alpha)$.

Proof. We prove the contrapositive. So, suppose there is some ordinal which is $\neg\varphi$; let γ be the least such ordinal. Then either $\gamma = \emptyset$, or $\gamma = \alpha^+$ for some α such that $\varphi(\alpha)$; or γ is a limit ordinal and $(\forall\beta \in \gamma)\varphi(\beta)$. □

A final bit of notation will prove helpful later on:

Definition 10.32. If X is a set of ordinals, then $\mathrm{lsub}(X) = \bigcup_{\alpha\in X} \alpha^+$.

Here, "lsub" stands for "least strict upper bound".[4] The following result explains this:

Proposition 10.33. *If X is a set of ordinals,* $\mathrm{lsub}(X)$ *is the least ordinal greater than every ordinal in X.*

Proof. Let $Y = \{\alpha^+ : \alpha \in X\}$, so that $\mathrm{lsub}(X) = \bigcup Y$. Since ordinals are transitive and every member of an ordinal is an ordinal, $\mathrm{lsub}(X)$ is a transitive set of ordinals, and so is an ordinal by Corollary 10.19.

If $\alpha \in X$, then $\alpha^+ \in Y$, so $\alpha^+ \subseteq \bigcup Y = \mathrm{lsub}(X)$, and hence $\alpha \in \mathrm{lsub}(X)$. So $\mathrm{lsub}(X)$ is strictly greater than every ordinal in X.

Conversely, if $\alpha \in \mathrm{lsub}(X)$, then $\alpha \in \beta^+ \in Y$ for some $\beta \in X$, so that $\alpha \leq \beta \in X$. So $\mathrm{lsub}(X)$ is the *least* strict upper bound on X. □

[4] Some books use "$\sup(X)$" for this. But other books use "$\sup(X)$" for the least *non-strict* upper bound, i.e., simply $\bigcup X$. If X has a greatest element, α, these notions come apart: the least *strict* upper bound is α^+, whereas the least *non-strict* upper bound is just α.

Problems

Problem 10.1. Section 10.2 presented three example orderings on the natural numbers. Check that each is a well-ordering.

Problem 10.2. Prove Lemma 10.5.

Problem 10.3. Complete the "exactly similar reasoning" in the proof of Theorem 10.17.

Problem 10.4. Prove that, if every member of X is an ordinal, then $\bigcup X$ is an ordinal.

CHAPTER 11

Stages and Ranks

11.1 Defining the Stages as the V_αs

In chapter 10, we defined well-orderings and the (von Neumann) ordinals. In this chapter, we will use these to characterise the hierarchy of sets *itself*. To do this, recall that in section 10.10, we defined the idea of successor and limit ordinals. We use these ideas in following definition:

Definition 11.1.

$$\begin{aligned} V_\emptyset &:= \emptyset \\ V_{\alpha^+} &:= \wp(V_\alpha) && \text{for any ordinal } \alpha \\ V_\alpha &:= \bigcup_{\gamma<\alpha} V_\gamma && \text{when } \alpha \text{ is a limit ordinal} \end{aligned}$$

This will be a definition by *transfinite recursion* on the ordinals. In this regard, we should compare this with recursive definitions of functions on the natural numbers.[1] As when dealing with natural

[1] Cf. the definitions of addition, multiplication, and exponentiation in section 7.2.

numbers, one defines a base case and successor cases; but when dealing with ordinals, we also need to describe the behaviour of *limit* cases.

This definition of the V_αs will be an important milestone. We have informally motivated our hierarchy of sets as forming sets by *stages*. The V_αs are, in effect, just those stages. Importantly, though, this is an *internal* characterisation of the stages. Rather than suggesting a possible *model* of the theory, we will have defined the stages *within* our set theory.

11.2 The Transfinite Recursion Theorem(s)

The first thing we must do, though, is confirm that Definition 11.1 is a successful definition. More generally, we need to prove that any attempt to offer a transfinite by (transfinite) recursion will succeed. That is the aim of this section.

Warning: this is tricky material. The overarching moral, though, is quite simple: Transfinite Induction plus Replacement guarantee the legitimacy of (several versions of) transfinite recursion.[2]

Definition 11.2. Let $\tau(x)$ be a term; let f be a function; let α be an ordinal. We say that f is an *α-approximation* for τ iff both $\mathrm{dom}(f) = \alpha$ and $(\forall \beta \in \alpha) f(\beta) = \tau(f \restriction_\beta)$.

Lemma 11.3 (Bounded Recursion). *For any term $\tau(x)$ and any ordinal α, there is a unique α-approximation for τ.*

Proof. We will show that, for any $\gamma \leq \alpha$, there is a unique γ-approximation.

We first establish uniqueness. Let g and h (respectively) be γ- and δ-approximations. A transfinite induction on their arguments

[2] A reminder: all formulas and terms can have parameters (unless explicitly stated otherwise).

shows that $g(\beta) = h(\beta)$ for any $\beta \in \mathrm{dom}(g) \cap \mathrm{dom}(h) = \gamma \cap \delta = \min(\gamma, \delta)$. So our approximations are unique (if they exist), and agree on all values.

To establish existence, we now use a simple transfinite induction (Theorem 10.31) on ordinals $\delta \leq \alpha$.

The empty function is trivially an $\emptyset$-approximation.

If g is a γ-approximation, then $g \cup \{\langle \gamma^+, \tau(g)\rangle\}$ is a γ^+-approximation.

If γ is a limit ordinal and g_δ is a δ-approximation for all $\delta < \gamma$, let $g = \bigcup_{\delta\in\gamma} g_\delta$. This is a function, since our various g_δs agree on all values. And if $\delta \in \gamma$ then $g(\delta) = g_{\delta^+}(\delta) = \tau(g_{\delta^+}{\restriction_\delta}) = \tau(g{\restriction_\delta})$.

This completes the proof by transfinite induction. □

If we allow ourselves to define a *term* rather than a function, then we can remove the bound α from the previous result. In the statement and proof of the following result, when σ is a term, we let $\sigma{\restriction_\alpha} = \{\langle \beta, \sigma(\beta)\rangle : \beta \in \alpha\}$.

Theorem 11.4 (General Recursion). *For any term $\tau(x)$, we can explicitly define a term $\sigma(x)$, such that $\sigma(\alpha) = \tau(\sigma{\restriction_\alpha})$ for any ordinal α.*

Proof. For each α, by Lemma 11.3 there is a unique α-approximation, f_α, for τ. Define $\sigma(\alpha)$ as $f_{\alpha^+}(\alpha)$. Now:

$$\begin{aligned}\sigma(\alpha) &= f_{\alpha^+}(\alpha)\\ &= \tau(f_{\alpha^+}{\restriction_\alpha})\\ &= \tau(\{\langle \beta, f_{\alpha^+}(\beta)\rangle : \beta \in \alpha\})\\ &= \tau(\{\langle \beta, f_{\beta^+}(\beta)\rangle : \beta \in \alpha\})\\ &= \tau(\sigma{\restriction_\alpha})\end{aligned}$$

noting that $f_{\beta^+}(\beta) = f_{\alpha^+}(\beta)$ for all $\beta < \alpha$, as in Lemma 11.3. □

Note that Theorem 11.4 is a *schema*. Crucially, we cannot expect σ to define a function, i.e., a certain kind of *set*, since then $\mathrm{dom}(\sigma)$ would be the set of all ordinals, contradicting the Burali-Forti Paradox (Theorem 10.20).

It still remains to show, though, that Theorem 11.4 vindicates our definition of the V_αs. This may not be immediately obvious; but it will become apparent with a last, simple, version of transfinite recursion.

Theorem 11.5 (Simple Recursion). *For any terms $\tau(x)$ and $\theta(x)$ and any set A, we can explicitly define a term $\sigma(x)$ such that:*

$$\begin{aligned} \sigma(\emptyset) &= A \\ \sigma(\alpha^+) &= \tau(\sigma(\alpha)) && \text{for any ordinal } \alpha \\ \sigma(\alpha) &= \theta(\mathrm{ran}(\sigma\restriction_\alpha)) && \text{when } \alpha \text{ is a limit ordinal} \end{aligned}$$

Proof. We start by defining a term, $\xi(x)$, as follows:

$$\xi(x) = \begin{cases} A & \text{if } x \text{ is not a function whose} \\ & \quad \text{domain is an ordinal; otherwise:} \\ \tau(x(\alpha)) & \text{if } \mathrm{dom}(x) = \alpha^+ \\ \theta(\mathrm{ran}(x)) & \text{if } \mathrm{dom}(x) \text{ is a limit ordinal} \end{cases}$$

By Theorem 11.4, there is a term $\sigma(x)$ such that $\sigma(\alpha) = \xi(\sigma\restriction_\alpha)$ for every ordinal α; moreover, $\sigma\restriction_\alpha$ is a function with domain α. We show that σ has the required properties, by simple transfinite induction (Theorem 10.31).

First, $\sigma(\emptyset) = \xi(\emptyset) = A$.

Next, $\sigma(\alpha^+) = \xi(\sigma\restriction_{\alpha^+}) = \tau(\sigma\restriction_{\alpha^+}(\alpha)) = \tau(\sigma(\alpha))$.

Last, $\sigma(\alpha) = \xi(\sigma\restriction_\alpha) = \theta(\mathrm{ran}(\sigma\restriction_\alpha))$, when α is a limit. □

Now, to vindicate Definition 11.1, just take $A = \emptyset$ and $\tau(x) = \wp(x)$ and $\theta(x) = \bigcup x$. At long last, this vindicates the definition of the V_αs!

11.3 Basic Properties of Stages

To bring out the foundational importance of the definition of the V_αs, we will present a few basic results about them. We start with

a definition:[3]

Definition 11.6. The set A is *potent* iff $\forall x((\exists y \in A)x \subseteq y \rightarrow x \in A)$.

Lemma 11.7. *For each ordinal α:*

1. *Each V_α is transitive.*
2. *Each V_α is potent.*
3. *If $\gamma \in \alpha$, then $V_\gamma \in V_\alpha$ (and hence also $V_\gamma \subseteq V_\alpha$ by (1))*

Proof. We prove this by a (simultaneous) transfinite induction. For induction, suppose that (1)–(3) holds for each ordinal $\beta < \alpha$.

The case of $\alpha = \emptyset$ is trivial.

Suppose $\alpha = \beta^+$. To show (3), if $\gamma \in \alpha$ then $V_\gamma \subseteq V_\beta$ by hypothesis, so $V_\gamma \in \wp(V_\beta) = V_\alpha$. To show (2), suppose $A \subseteq B \in V_\alpha$ i.e., $A \subseteq B \subseteq V_\beta$; then $A \subseteq V_\beta$ so $A \in V_\alpha$. To show (1), note that if $x \in A \in V_\alpha$ we have $A \subseteq V_\beta$, so $x \in V_\beta$, so $x \subseteq V_\beta$ as V_β is transitive by hypothesis, and so $x \in V_\alpha$.

Suppose α is a limit ordinal. To show (3), if $\gamma \in \alpha$ then $\gamma \in \gamma^+ \in \alpha$, so that $V_\gamma \in V_{\gamma^+}$ by assumption, hence $V_\gamma \in \bigcup_{\beta \in \alpha} V_\beta = V_\alpha$. To show (1) and (2), just observe that a union of transitive (respectively, potent) sets is transitive (respectively, potent). □

Lemma 11.8. *For each ordinal α, $V_\alpha \notin V_\alpha$.*

Proof. By transfinite induction. Evidently $V_\emptyset \notin V_\emptyset$.

If $V_{\alpha^+} \in V_{\alpha^+} = \wp(V_\alpha)$, then $V_{\alpha^+} \subseteq V_\alpha$; and since $V_\alpha \in V_{\alpha^+}$ by Lemma 11.7, we have $V_\alpha \in V_\alpha$. Conversely: if $V_\alpha \notin V_\alpha$ then $V_{\alpha^+} \notin V_{\alpha^+}$

[3] There's no standard terminology for "potent"; this is the name used by Button (2021).

If α is a limit and $V_\alpha \in V_\alpha = \bigcup_{\beta\in\alpha} V_\beta$, then $V_\alpha \in V_\beta$ for some $\beta \in \alpha$; but then also $V_\beta \in V_\alpha$ so that $V_\beta \in V_\beta$ by Lemma 11.7 (twice). Conversely, if $V_\beta \notin V_\beta$ for all $\beta \in \alpha$, then $V_\alpha \notin V_\alpha$. □

Corollary 11.9. *For any ordinals α, β: $\alpha \in \beta$ iff $V_\alpha \in V_\beta$*

Proof. Lemma 11.7 gives one direction. Conversely, suppose $V_\alpha \in V_\beta$. Then $\alpha \neq \beta$ by Lemma 11.8; and $\beta \notin \alpha$, for otherwise we would have $V_\beta \in V_\alpha$ and hence $V_\beta \in V_\beta$ by Lemma 11.7 (twice), contradicting Lemma 11.8. So $\alpha \in \beta$ by Trichotomy. □

All of this allows us to think of each V_α as the αth stage of the hierarchy. Here is why.

Certainly our V_αs can be thought of as being formed in an *iterative* process, for our use of ordinals tracks the notion of iteration. Moreover, if one stage is formed before the other, i.e., $V_\beta \in V_\alpha$, i.e., $\beta \in \alpha$, then our process of formation is *cumulative*, since $V_\beta \subseteq V_\alpha$. Finally, we are indeed forming *all* possible collections of sets that were available at any earlier stage, since any successor stage V_{α^+} is the power-set of its predecessor V_α.

In short: with $\mathbf{ZF}^-$, we are *almost* done, in articulating our vision of the cumulative-iterative hierarchy of sets. (Though, of course, we still need to justify Replacement.)

11.4 Foundation

We are only *almost* done—and not *quite* finished—because nothing in $\mathbf{ZF}^-$ guarantees that *every* set is in some V_α, i.e., that every set is formed at some stage.

Now, there is a fairly straightforward (mathematical) sense in which we don't *care* whether there are sets outside the hierarchy. (If there are any there, we can simply ignore them.) But we have motivated our *concept* of set with the thought that every set is formed at some stage (see *Stages-are-key* in section 9.1). So we will want to preclude the possibility of sets which fall outside of

the hierarchy. Accordingly, we must add a new axiom, which ensures that every set occurs somewhere in the hierarchy.

Since the V_αs are our stages, we might simply consider adding the following as an axiom:

Regularity. $\forall A \exists \alpha\ A \subseteq V_\alpha$

This would be a perfectly reasonable approach. However, for reasons that will be explained in the next section, we will instead adopt an alternative axiom:

Axiom (Foundation). $(\forall A \neq \emptyset)(\exists B \in A) A \cap B = \emptyset$.

With some effort, we can show (in $\mathbf{ZF}^-$) that Foundation entails Regularity:

Definition 11.10. For each set A, let:

$$\begin{aligned} \mathrm{cl}_0(A) &= A, \\ \mathrm{cl}_{n+1}(A) &= \bigcup \mathrm{cl}_n(A), \\ \mathrm{trcl}(A) &= \bigcup_{n<\omega} \mathrm{cl}_n(A). \end{aligned}$$

We call $\mathrm{trcl}(A)$ the *transitive closure* of A.

The name "transitive closure" is apt:

Proposition 11.11. *$A \subseteq \mathrm{trcl}(A)$ and $\mathrm{trcl}(A)$ is a transitive set.*

Proof. Evidently $A = \mathrm{cl}_0(A) \subseteq \mathrm{trcl}(A)$. And if $x \in b \in \mathrm{trcl}(A)$, then $b \in \mathrm{cl}_n(A)$ for some n, so $x \in \mathrm{cl}_{n+1}(A) \subseteq \mathrm{trcl}(A)$. □

Lemma 11.12. *If A is a transitive set, then there is some α such that $A \subseteq V_\alpha$.*

Proof. Recalling the definition of "lsub(X)" from Definition 10.32, define two sets:

$$D = \{x \in A : \forall \delta\ x \not\subseteq V_\delta\}$$
$$\alpha = \text{lsub}\{\delta : (\exists x \in A)(x \subseteq V_\delta \wedge (\forall \gamma \in \delta) x \not\subseteq V_\gamma)\}$$

Suppose $D = \emptyset$. So if $x \in A$, then there is some δ such that $x \subseteq V_\delta$ and, by the well-ordering of the ordinals, $(\forall \gamma \in \delta) x \not\subseteq V_\gamma$; hence $\delta \in \alpha$ and so $x \in V_\alpha$ by Lemma 11.7. Hence $A \subseteq V_\alpha$, as required.

So it suffices to show that $D = \emptyset$. For reductio, suppose otherwise. By Foundation, there is some $B \in D \subseteq A$ such that $D \cap B = \emptyset$. If $x \in B$ then $x \in A$, since A is transitive, and since $x \notin D$, it follows that $\exists \delta\ x \subseteq V_\delta$. So now let

$$\beta = \text{lsub}\{\delta : (\exists x \in b)(x \subseteq V_\delta \wedge (\forall \gamma < \delta) x \not\subseteq V_\gamma)\}.$$

As before, $B \subseteq V_\beta$, contradicting the claim that $B \in D$. □

Theorem 11.13. *Regularity holds.*

Proof. Fix A; now $A \subseteq \text{trcl}(A)$ by Proposition 11.11, which is transitive. So there is some α such that $A \subseteq \text{trcl}(A) \subseteq V_\alpha$ by Lemma 11.12 □

These results show that $\mathbf{ZF}^-$ proves the conditional *Foundation* $\Rightarrow$ *Regularity*. In Proposition 11.22, we will show that $\mathbf{ZF}^-$ proves *Regularity* $\Rightarrow$ *Foundation*. As such, Foundation and Regularity are *equivalent* (modulo $\mathbf{ZF}^-$). But this means that, given $\mathbf{ZF}^-$, we can justify Foundation by noting that it is equivalent to Regularity. And we can justify Regularity immediately on the basis of *Stages-are-key*.

11.5 Z and ZF: A Milestone

With Foundation, we reach another important milestone. We have considered theories $\mathbf{Z}^-$ and $\mathbf{ZF}^-$, which we said were certain theories "minus" a certain something. That certain something is Foundation. So:

Definition 11.14. The theory **Z** adds Foundation to $\mathbf{Z}^-$. So its axioms are Extensionality, Union, Pairs, Powersets, Infinity, Foundation, and all instances of the Separation scheme.

The theory **ZF** adds Foundation to $\mathbf{ZF}^-$. Otherwise put, **ZF** adds all instances of Replacement to **Z**.

Still, one question might have occurred to you. If Regularity is equivalent over $\mathbf{ZF}^-$ to Foundation, and Regularity's justification is clear, why bother to go around the houses, and take Foundation as our basic axiom, rather than Regularity?

Setting aside historical reasons (to do with who formulated what and when), the basic reason is that Foundation can be presented without employing the definition of the V_αs. That definition relied upon all of the work of section 11.2: we needed to prove Transfinite Recursion, to show that it was justified. But our proof of Transfinite Recursion employed *Replacement*. So, whilst Foundation and Regularity are equivalent modulo $\mathbf{ZF}^-$, they are not equivalent modulo $\mathbf{Z}^-$.

Indeed, the matter is more drastic than this simple remark suggests. Though it goes well beyond this book's remit, it turns out that both $\mathbf{Z}^-$ and **Z** are too weak to define the V_αs. So, if you are working only in **Z**, then Regularity (as we have formulated it) does not even make *sense*. This is why our official axiom is Foundation, rather than Regularity.

From now on, we will work in **ZF** (unless otherwise stated), without any further comment.

11.6 Rank

Now that we have defined the stages as the V_α's, and we know that every set is a subset of some stage, we can define the *rank* of a set. Intuitively, the rank of A is the first moment at which A is formed. More precisely:

Definition 11.15. For each set A, $\mathrm{rank}(A)$ is the least ordinal α such that $A \subseteq V_\alpha$.

Proposition 11.16. $\mathrm{rank}(A)$ *exists, for any* A.

Proof. Left as an exercise. □

The well-ordering of ranks allows us to prove some important results:

Proposition 11.17. *For any ordinal* α, $V_\alpha = \{x : \mathrm{rank}(x) \in \alpha\}$.

Proof. If $\mathrm{rank}(x) \in \alpha$ then $x \subseteq V_{\mathrm{rank}(x)} \in V_\alpha$, so $x \in V_\alpha$ as V_α is potent (invoking Lemma 11.7 multiple times). Conversely, if $x \in V_\alpha$ then $x \subseteq V_\alpha$, so $\mathrm{rank}(x) \leq \alpha$; now a simple transfinite induction shows that $x \notin V_\alpha$. □

Proposition 11.18. *If* $B \in A$, *then* $\mathrm{rank}(B) \in \mathrm{rank}(A)$.

Proof. $A \subseteq V_{\mathrm{rank}(A)} = \{x : \mathrm{rank}(x) \in \mathrm{rank}(A)\}$ by Proposition 11.17. □

Using this fact, we can establish a result which allows us to prove things about *all sets* by a form of induction:

Theorem 11.19 (∈-Induction Scheme). *For any formula φ:*

$$\forall A((\forall x \in A)\varphi(x) \to \varphi(A)) \to \forall A\varphi(A).$$

Proof. We will prove the contrapositive. So, suppose $\neg\forall A\varphi(A)$. By Transfinite Induction (Theorem 10.16), there is some non-φ of least possible rank; i.e. some A such that $\neg\varphi(A)$ and $\forall x(\text{rank}(x) \in \text{rank}(A) \to \varphi(x))$. Now if $x \in A$ then $\text{rank}(x) \in \text{rank}(A)$, by Proposition 11.18, so that $\varphi(x)$; i.e. $(\forall x \in A)\varphi(x) \wedge \neg\varphi(A)$. □

Here is an informal way to gloss this powerful result. Say that φ is *hereditary* iff whenever every member of a set is φ, the set itself is φ. Then ∈-Induction tells you the following: if φ is hereditary, every set is φ.

To wrap up the discussion of ranks (for now), we'll prove a few claims which we have foreshadowed a few times.

Proposition 11.20. $\text{rank}(A) = \text{lsub}_{x\in A}\,\text{rank}(x)$.

Proof. Let $\alpha = \text{lsub}_{x\in A}\,\text{rank}(x)$. By Proposition 11.18, $\alpha \leq \text{rank}(A)$. But if $x \in A$ then $\text{rank}(x) \in \alpha$, so that $x \in V_\alpha$ by Proposition 11.17, and hence $A \subseteq V_\alpha$, i.e., $\text{rank}(A) \leq \alpha$. Hence $\text{rank}(A) = \alpha$. □

Corollary 11.21. *For any ordinal α,* $\text{rank}(\alpha) = \alpha$.

Proof. Suppose for transfinite induction that $\text{rank}(\beta) = \beta$ for all $\beta \in \alpha$. Now $\text{rank}(\alpha) = \text{lsub}_{\beta\in\alpha}\,\text{rank}(\beta) = \text{lsub}_{\beta\in\alpha}\,\beta = \alpha$ by Proposition 11.20. □

Finally, here is a quick proof of the result promised at the end of section 11.4, that $\mathbf{ZF}^-$ proves the conditional *Regularity* ⇒ *Foundation*. (Note that the notion of "rank" and Proposition 11.18 are available for use in this proof since—as mentioned at the start of this section—they can be presented using $\mathbf{ZF}^-$ + Regularity.)

Proposition 11.22 (working in ZF$^-$ + Regularity). *Foundation holds.*

Proof. Fix $A \neq \emptyset$, and some $B \in A$ of least possible rank. If $c \in B$ then $\mathrm{rank}(c) \in \mathrm{rank}(B)$ by Proposition 11.18, so that $c \notin A$ by choice of B. □

Problems

Problem 11.1. Prove Proposition 11.16.

Problem 11.2. Complete the simple transfinite induction mentioned in Proposition 11.17.

CHAPTER 12

Replacement

12.1 Introduction

Replacement is the axiom scheme which makes the difference between **ZF** and **Z**. We helped ourselves to it throughout chapters 10 to 11. In this chapter, we will finally consider the question: is Replacement justified?

To make the question sharp, it is worth observing that Replacement is really rather *strong*. We will get a sense of just how strong it is, during this chapter (and again in section 15.5). But this will suggest that justification really is required.

We will discuss two kinds of justification. Roughly: an *extrinsic* justification is an attempt to justify an axiom by its fruits; an *intrinsic* justification is an attempt to justify an axiom by suggesting that it is vindicated by the mathematical concepts in question. We will get a greater sense of what this means during this chapter, but it is just the tip of an iceberg. For more, see in particular Maddy (1988a and 1988b).

12.2 The Strength of Replacement

We begin with a simple observation about the strength of Replacement: unless we go beyond **Z**, we cannot prove the existence of any von Neumann ordinal greater than or equal to $\omega + \omega$.

Here is a sketch of why. Working in **ZF**, consider the set $V_{\omega+\omega}$. This set acts as the domain for a *model* for **Z**. To see this, we introduce some notation for the *relativization* of a formula:

Definition 12.1. For any set M, and any formula φ, let φ^M be the formula which results by restricting all of φ's quantifiers to M. That is, replace "$\exists x$" with "$(\exists x \in M)$", and replace "$\forall x$" with "$(\forall x \in M)$".

It can be shown that, for every axiom φ of **Z**, we have that $\mathbf{ZF} \vdash \varphi^{V_{\omega+\omega}}$. But $\omega + \omega$ is not *in* $V_{\omega+\omega}$, by Corollary 11.21. So **Z** is consistent with the non-existence of $\omega + \omega$.

This is why we said, in section 10.7, that Theorem 10.26 cannot be proved without Replacement. For it is easy, within **Z**, to define an explicit well-ordering which intuitively *should* have order-type $\omega + \omega$. Indeed, we gave an informal example of this in section 10.2, when we presented the ordering on the natural numbers given by:

$$n \lessdot m \text{ iff either } n < m \text{ and } m - n \text{ is even,}$$
$$\text{or } n \text{ is even and } m \text{ is odd.}$$

But if $\omega + \omega$ does not exist, this well-ordering is not isomorphic to any ordinal. So **Z** does *not* prove Theorem 10.26.

Flipping things around: Replacement allows us to prove the existence of $\omega+\omega$, and hence must allow us to prove the existence of $V_{\omega+\omega}$. And not just that. For *any* well-ordering we can define, Theorem 10.26 tells us that there is some α isomorphic with that well-ordering, and hence that V_α exists. In a straightforward way, then, Replacement guarantees that the hierarchy of sets must be *very tall.*

Over the next few sections, and then again in section 15.5, we'll get a better sense of better just *how* tall Replacement forces the hierarchy to be. The simple point, for now, is that Replacement really *does* stand in need of justification!

12.3 Extrinsic Considerations about Replacement

We start by considering an *extrinsic* attempt to justify Replacement. Boolos suggests one, as follows.

> [...] the reason for adopting the axioms of replacement is quite simple: they have many desirable consequences and (apparently) no undesirable ones. In addition to theorems about the iterative conception, the consequences include a satisfactory if not ideal theory of infinite numbers, and a highly desirable result that justifies inductive definitions on well-founded relations. (Boolos, 1971, 229)

The gist of Boolos's idea is that we should justify Replacement by its fruits. And the specific fruits he mentions are the things we have discussed in the past few chapters. Replacement allowed us to prove that the von Neumann ordinals were excellent surrogates for the idea of a well-ordering type (this is our "satisfactory if not ideal theory of infinite numbers"). Replacement also allowed us to define the V_αs, establish the notion of rank, and prove $\in$-Induction (this amounts to our "theorems about the iterative conception"). Finally, Replacement allows us to prove the Transfinite Recursion Theorem (this is the "inductive definitions on well-founded relations").

These are, indeed, desirable consequences. But do these desirable consequences suffice to *justify* Replacement? *No.* Or at least, not straightforwardly.

Here is a simple problem. Whilst we have stated some desirable consequences of Replacement, we could have obtained many of them via other means. This is not as well known as it ought to be, though, so we should pause to explain the situation.

There is a simple theory of sets, Level Theory, or **LT** for short.[1] **LT**'s axioms are just Extensionality, Separation, and the

[1] The first versions of **LT** are offered by Montague (1965) and Scott (1974);

claim that every set is a subset of some *level*, where "level" is cunningly defined so that the levels behave like our friends, the V_αs. So **ZF** proves **LT**; but **LT** is *much* weaker than **ZF**. In fact, **LT** does not give you Pairs, Powersets, Infinity, or Replacement. Let **Zr** be the result of adding Infinity and Powersets to **LT**; this delivers Pairs too, so, **Zr** is at least as strong as **Z**. But, in fact, **Zr** is strictly stronger than **Z**, since it adds the claim that every set has a rank (hence my suggestion that we call it **Zr**). Indeed, **Zr** delivers: a perfectly satisfactory theory of ordinals; results which stratify the hierarchy into well-ordered stages; a proof of $\in$-Induction; and a *version* of Transfinite Recursion.

In short: although Boolos didn't know this, all of the desirable consequences which he mentions could have been arrived at *without* Replacement; he simply needed to use **Zr** rather than **Z**.

(Given all of this, why did we follow the conventional route, of teaching you **ZF**, rather than **LT** and **Zr**? There are two reasons. First: for purely historical reasons, starting with **LT** is rather nonstandard; we wanted to equip you to be able to read more standard discussions of set theory. Second: when you are ready to appreciate **LT** and **Zr**, you can simply read Potter 2004 and Button 2021.)

Of course, since **Zr** is strictly weaker than **ZF**, there are results which **ZF** proves which **Zr** leaves open. So one could try to justify Replacement on extrinsic grounds by pointing to one of these results. But, once you know how to use **Zr**, it is quite hard to find many examples of things that are (a) settled by Replacement but not otherwise, and (b) are intuitively true. (For more on this, see Potter 2004, §13.2.)

The bottom line is this. To provide a compelling extrinsic justification for Replacement, one would need to find a result which *cannot* be achieved without Replacement. And that's not an easy enterprise.

Let's consider a further problem which arises for any attempt

this was simplified, and given a book-length treatment, by Potter (2004); and Button (2021) has recently simplified **LT** further.

to offer a purely extrinsic justification for Replacement. (This problem is perhaps more fundamental than the first.) Boolos does not just point out that Replacement has many desirable consequences. He also states that Replacement has "(apparently) no undesirable" consequences. But this parenthetical caveat, "apparently," is surely absolutely crucial.

Recall how we ended up here: Naïve Comprehension ran into inconsistency, and we responded to this inconsistency by embracing the cumulative-iterative conception of set. This conception comes equipped with a story which, we hope, assures us of its consistency. But if we cannot justify Replacement from within that story, then we have (as yet) no reason to believe that **ZF** is consistent. Or rather: we have no reason to believe that **ZF** is consistent, apart from the (perhaps merely contingent) fact that no one has discovered a contradiction *yet*. In exactly that sense, Boolos's comment seems to come down to this: "(apparently) **ZF** is consistent". We should demand greater reassurance of consistency than this.

This issue will affect any *purely* extrinsic attempt to justify Replacement, i.e., any justification which is couched solely in terms of the (known) consequences of **ZF**. As such, we will want to look for an *intrinsic* justification of Replacement, i.e., a justification which suggests that the story which we told about sets somehow "already" commits us to Replacement.

12.4 Limitation-of-size

Perhaps the most common attempt to offer an "intrinsic" justification of Replacement comes via the following notion:

> *Limitation-of-size.* Any things form a set, provided that there are not too many of them.

This principle will immediately vindicate Replacement. After all, any set formed by Replacement cannot be any larger than any set from which it was formed. Stated precisely: suppose you form a

set $\tau[A] = \{\tau(x) : x \in A\}$ using Replacement; then $\tau[A] \preceq A$; so if the members of A were not too numerous to form a set, their images are not too numerous to form $\tau[A]$.

The obvious difficulty with invoking *Limitation-of-size* to justify Replacement is that we have *not* yet laid down any principle like *Limitation-of-size*. Moreover, when we told our story about the cumulative-iterative conception of set in chapters 8 to 9, nothing ever *hinted* in the direction of *Limitation-of-size*. This, indeed, is precisely why Boolos at one point wrote: "Perhaps one may conclude that there are at least two thoughts 'behind' set theory" (1989, p. 19). On the one hand, the ideas surrounding the cumulative-iterative conception of set are meant to vindicate **Z**. On the other hand, *Limitation-of-size* is meant to vindicate Replacement.

But the issue it is not just that we have thus far been *silent* about *Limitation-of-size*. Rather, the issue is that *Limitation-of-size* (as just formulated) seems to sit quite badly with the cumulative-iterative notion of set. After all, it mentions nothing about the idea of sets as formed in *stages*.

This is really not much of a surprise, given the history of these "two thoughts" (i.e., the cumulative-iterative conception of set, and *Limitation-of-size*). These "two thoughts" ultimately amount to two rather different projects for blocking the set-theoretic paradoxes. The cumulative-iterative notion of set blocks Russell's paradox by saying, roughly: *we should never have expected a Russell set to exist, because it would not be "formed" at any stage.* By contrast, *Limitation-of-size* is meant to rule out the Russell set, by saying, roughly: *we should never have expected a Russell set to exist, because it would have been too big.*

Put like this, then, let's be blunt: considered as a reply to the paradoxes, *Limitation-of-size* stands in need of *much* more justification. Consider, for example, this version of Russell's Paradox: *no pug sniffs exactly the pugs which don't sniff themselves* (see section 8.2). If you ask "why is there no such pug?", it is not a good answer to be told that such a pug would have to sniff too many pugs. So why would it be a good intuitive explanation, of the

non-existence of a Russell set, that it would have to be "too big" to exist?

In short, it's forgivable if you are a bit mystified concerning the "intuitive" motivation for *Limitation-of-size*.

12.5 Replacement and "Absolute Infinity"

We will now put *Limitation-of-size* behind us, and explore a different family of (intrinsic) attempts to justify Replacement, which do take seriously the idea of the sets as formed in stages.

When we first outlined the iterative process, we offered some principles which explained what happens at each stage. These were *Stages-are-key*, *Stages-are-ordered*, and *Stages-accumulate*. Later, we added some principles which told us something about the number of stages: *Stages-keep-going* told us that the process of set-formation never ends, and *Stages-hit-infinity* told us that the process goes through an infinite-th stage.

It is reasonable to suggest that these two latter principles fall out of some a broader principle, like:

> *Stages-are-inexhaustible.* There are absolutely infinitely many stages; the hierarchy is as tall as it could possibly be.

Obviously this is an informal principle. But even if it is not immediately *entailed* by the cumulative-iterative conception of set, it certainly seems *consonant* with it. At the very least, and unlike *Limitation-of-size*, it retains the idea that sets are formed stage-by-stage.

The hope, now, is to leverage *Stages-are-inexhaustible* into a justification of Replacement. So let us see how this might be done.

In section 10.2, we saw that it is easy to construct a well-ordering which (morally) should be isomorphic to $\omega + \omega$. Otherwise put, we can easily imagine a stage-by-stage iterative process, whose order-type (morally) is $\omega + \omega$. As such, if we have accepted *Stages-are-inexhaustible*, then we should surely accept that there is

at least an $\omega + \omega$-th stage of the hierarchy, i.e., $V_{\omega+\omega}$, for the hierarchy surely *could* continue thus far.

This thought generalizes as follows: for any well-ordering, the process of building the iterative hierarchy should run at least as far as that well-ordering. And we could guarantee this, just by treating Theorem 10.26 as an *axiom*. This would tell us that any well-ordering is isomorphic to a von Neumann ordinal. Since each von Neumann ordinal will be equal to its own rank, Theorem 10.26 will then tell us that, whenever we can describe a well-ordering in our set theory, the iterative process of set building must outrun that well-ordering.

This idea certainly seems like a corollary of *Stages-are-inexhaustible*. Unfortunately, if our aim is to extract Replacement from this idea, then we face a simple, technical, barrier: Replacement is strictly stronger than Theorem 10.26. (This observation is made by Potter (2004, §13.2); we will prove it in section 12.8.)

The upshot is that, if we are going to understand *Stages-are-inexhaustible* in such a way as to yield Replacement, then it cannot *merely* say that the hierarchy outruns any well-ordering. It must make a stronger claim than that. To this end, Shoenfield (1977) proposed a very natural strengthening of the idea, as follows: the hierarchy is not *cofinal* with any set.[2] In slightly more detail: if τ is a mapping which sends sets to stages of the hierarchy, the image of any set A under τ does not exhaust the hierarchy. Otherwise put (schematically):

> *Stages-are-super-cofinal.* If A is a set and $\tau(x)$ is a stage for every $x \in A$, then there is a stage which comes after each $\tau(x)$ for $x \in A$.

It is obvious that **ZF** proves a suitably formalised version of *Stages-are-super-cofinal*. Conversely, we can informally argue that

[2] Gödel seems to have proposed a similar thought; see Potter (2004, p. 223). For discussion of Gödel and Shoenfield, see Incurvati (2020, 90–5).

Stages-are-super-cofinal justifies Replacement.[3] For suppose $(\forall x \in A)\exists! y\, \varphi(x,y)$. Then for each $x \in A$, let $\sigma(x)$ be the y such that $\varphi(x,y)$, and let $\tau(x)$ be the stage at which $\sigma(x)$ is first formed. By *Stages-are-super-cofinal*, there is a stage V such that $(\forall x \in A)\tau(x) \in V$. Now since each $\tau(x) \in V$ and $\sigma(x) \subseteq \tau(x)$, by Separation we can obtain $\{y \in V : (\exists x \in A)\sigma(x) = y\} = \{y : (\exists x \in A)\varphi(x,y)\}$.

So *Stages-are-super-cofinal* vindicates Replacement. And it is at least plausible that *Stages-are-inexhaustible* vindicates *Stages-are-super-cofinal*. For suppose *Stages-are-super-cofinal* fails. So the hierarchy is cofinal with some set A, i.e., we have a map τ such that for any stage S there is some $x \in A$ such that $S \in \tau(x)$. In that case, we do have a way to get a handle on the supposed "absolute infinity" of the hierarchy: it is *exhausted* by the range of τ applied to A. And that compromises the thought that the hierarchy is "absolutely infinite". Contraposing: *Stages-are-inexhaustible* entails *Stages-are-super-cofinal*, which in turn justifies Replacement.

This represents a genuinely promising attempt to provide an intrinsic justification for Replacement. But whether it ultimately works, or not, we will have to leave to you to decide.

12.6 Replacement and Reflection

Our last attempt to justify Replacement, via *Stages-are-inexhaustible*, begins with a deep and lovely result:[4]

Theorem 12.2 (Reflection Schema). *For any formula φ:*

$$\forall\alpha\exists\beta > \alpha(\forall x_1 \ldots, x_n \in V_\beta)(\varphi(x_1,\ldots,x_n) \leftrightarrow \varphi^{V_\beta}(x_1,\ldots,x_n))$$

[3] It would be harder to prove Replacement using some formalisation of *Stages-are-super-cofinal*, since **Z** on its own is not strong enough to define the stages, so it is not clear how one would formalise *Stages-are-super-cofinal*. One option, though, is to work in some extension of **LT**, as discussed in section 12.3.

[4] A reminder: all formulas can have parameters (unless explicitly stated otherwise).

As in Definition 12.1, φ^{V_β} is the result of restricting every quantifier in φ to the set V_β. So, intuitively, Reflection says this: if φ is true in the entire hierarchy, then φ is true in arbitrarily many *initial segments* of the hierarchy.

Montague (1961) and Lévy (1960) showed that (suitable formulations of) Replacement and Reflection are equivalent, modulo **Z**, so that adding either gives you **ZF**. (We prove these results in section 12.7.) Given this equivalence, one might hope to justify Reflection and Replacement via *Stages-are-inexhaustible* as follows: given *Stages-are-inexhaustible*, the hierarchy should be very, very tall; so tall, in fact, that nothing we can say about it is sufficient to bound its height. And we can understand this as the thought that, if any sentence φ is true in the entire hierarchy, then it is true in arbitrarily many initial segments of the hierarchy. And that is just Reflection.

Again, this seems like a genuinely promising attempt to provide an intrinsic justification for Replacement. But there is much too much to say about it here. You must now decide for yourself whether it succeeds.[5]

12.7 Appendix: Results surrounding Replacement

In this section, we will prove Reflection within **ZF**. We will also prove a sense in which Reflection is equivalent to Replacement. And we will prove an interesting consequence of all this, concerning the strength of Reflection/Replacement. *Warning: this is easily the most advanced bit of mathematics in this textbook.*

We'll start with a lemma which, for brevity, employs the notational device of *overlining* to deal with sequences of variables or objects. So: "$\overline{a}_k$" abbreviates "$a_{k_1}, \ldots, a_{k_n}$", where n is determined by context.

[5] Though you might like to continue by reading Incurvati (2020, 95–100).

Lemma 12.3. *For each $1 \leq i \leq k$, let $\varphi_i(\overline{v}_i, x)$ be a formula. Then for each α there is some $\beta > \alpha$ such that, for any $\overline{a}_1, \ldots, \overline{a}_k \in V_\beta$ and each $1 \leq i \leq k$:*

$$\exists x \varphi_i(\overline{a}_i, x) \rightarrow (\exists x \in V_\beta)\varphi_i(\overline{a}_i, x)$$

Proof. We define a term μ as follows: $\mu(\overline{a}_1, \ldots, \overline{a}_k)$ is the least stage, V, which satisfies all of the following conditionals, for $1 \leq i \leq k$:

$$\exists x \varphi_i(\overline{a}_i, x) \rightarrow (\exists x \in V)\varphi_i(\overline{a}_i, x))$$

It is easy to confirm that $\mu(\overline{a}_1, \ldots, \overline{a}_k)$ exists for all $\overline{a}_1, \ldots, \overline{a}_k$. Now, using Replacement and our recursion theorem, define:

$$\begin{aligned} S_0 &= V_{\alpha+1} \\ S_{n+1} &= S_n \cup \bigcup \{\mu(\overline{a}_1, \ldots, \overline{a}_k) : \overline{a}_1, \ldots, \overline{a}_k \in S_n\} \\ S &= \bigcup_{m<\omega} S_n. \end{aligned}$$

Each S_n, and hence S itself, is a stage after V_α. Now fix $\overline{a}_1$, …, $\overline{a}_k \in S$; so there is some $n < \omega$ such that $\overline{a}_1$, …, $\overline{a}_k \in S_n$. Fix some $1 \leq i \leq k$, and suppose that $\exists x \varphi_i(\overline{a}_i, x)$. So $(\exists x \in \mu(\overline{a}_1, \ldots, \overline{a}_k))\varphi_i(\overline{a}_i, x)$ by construction, so $(\exists x \in S_{n+1})\varphi_i(\overline{a}_i, x)$ and hence $(\exists x \in S)\varphi_i(\overline{a}_i, x)$. So S is our V_β. □

We can now prove Theorem 12.2 quite straightforwardly:

Proof. Fix α. Without loss of generality, we can assume φ's only connectives are $\exists$, $\neg$ and $\wedge$ (since these are expressively adequate). Let $\psi_1, \ldots, \psi_k$ enumerate each of φ's subformulas according to complexity, so that $\psi_k = \varphi$. By Lemma 12.3, there is a $\beta > \alpha$ such that, for any $\overline{a}_i \in V_\beta$ and each $1 \leq i \leq k$:

$$\exists x \psi_i(\overline{a}_i, x) \rightarrow (\exists x \in V_\beta)\psi_i(\overline{a}_i, x) \qquad (*)$$

By induction on complexity of ψ_i, we will show that $\psi_i(\overline{a}_i) \leftrightarrow \psi_i^{V_\beta}(\overline{a}_i)$, for any $\overline{a}_i \in V_\beta$. If ψ_i is atomic, this is trivial. The

biconditional also establishes that, when ψ_i is a negation or conjunction of subformulas satisfying this property, ψ_i itself satisfies this property. So the only interesting case concerns quantification. Fix $\overline{a}_i \in V_\beta$; then:

$$\begin{aligned}
(\exists x \psi_i(\overline{a}_i, x))^{V_\beta} \text{ iff } & (\exists x \in V_\beta)\psi_i^{V_\beta}(\overline{a}_i, x) && \text{by definition} \\
\text{iff } & (\exists x \in V_\beta)\psi_i(\overline{a}_i, x) && \text{by hypothesis} \\
\text{iff } & \exists x \psi_i(\overline{a}_i, x) && \text{by (*)}
\end{aligned}$$

This completes the induction; the result follows as $\psi_k = \varphi$. □

We have proved Reflection in **ZF**. Our proof essentially followed Montague (1961). We now want to prove in **Z** that Reflection entails Replacement. The proof follows Lévy (1960), but with a simplification.

Since we are working in **Z**, we cannot present Reflection in exactly the form given above. After all, we formulated Reflection using the "V_α" notation, and that cannot be defined in **Z** (see section 11.5). So instead we will offer an apparently weaker formulation of Replacement, as follows:

Weak-Reflection. For any formula φ, there is a transitive set S such that 0, 1, and any parameters to φ are members of S, and $(\forall \overline{x} \in S)(\varphi \leftrightarrow \varphi^S)$.

To use this to prove Replacement, we will first follow Lévy (1960, first part of Theorem 2) and show that we can "reflect" two formulas at once:

Lemma 12.4 (in Z + Weak-Reflection.). *For any formulas ψ, χ, there is a transitive set S such that 0 and 1 (and any parameters to the formulas) are members of S, and $(\forall \overline{x} \in S)((\psi \leftrightarrow \psi^S) \wedge (\chi \leftrightarrow \chi^S))$.*

Proof. Let φ be the formula $(z = 0 \wedge \psi) \vee (z = 1 \wedge \chi)$.

Here we use an abbreviation; we should spell out "$z = 0$" as "$\forall t\ t \notin z$" and "$z = 1$" as "$\forall s(s \in z \leftrightarrow \forall t\ t \notin s)$". But since $0, 1 \in S$

and S is transitive, these formulas are *absolute* for S; that is, they will apply to the same object whether we restrict their quantifiers to S.[6]

By Weak-Reflection, we have some appropriate S such that:

$$\begin{aligned}
&(\forall z, \overline{x} \in S)(\varphi \leftrightarrow \varphi^S)\\
\text{i.e. } &(\forall z, \overline{x} \in S)(((z = 0 \wedge \psi) \vee (z = 1 \wedge \chi)) \leftrightarrow\\
&\qquad ((z = 0 \wedge \psi) \vee (z = 1 \wedge \chi))^S)\\
\text{i.e. } &(\forall z, \overline{x} \in S)(((z = 0 \wedge \psi) \vee (z = 1 \wedge \chi)) \leftrightarrow\\
&\qquad ((z = 0 \wedge \psi^S) \vee (z = 1 \wedge \chi^S)))\\
\text{i.e. } &(\forall \overline{x} \in S)((\psi \leftrightarrow \psi^S) \wedge (\chi \leftrightarrow \chi^S))
\end{aligned}$$

The second claim entails the third because "$z = 0$" and "$z = 1$" are absolute for S; the fourth claim follows since $0 \neq 1$. □

We can now obtain Replacement, just by following and simplifying Lévy (1960, Theorem 6):

Theorem 12.5 (in Z + Weak-Reflection). *For any formula $\varphi(v, w)$, and any A, if $(\forall x \in A)\exists! y\, \varphi(x, y)$, then $\{y : (\exists x \in A)\varphi(x, y\}$ exists.*

Proof. Fix A such that $(\forall x \in A)\exists! y\, \varphi(x, y)$, and define formulas:

$$\begin{aligned}
\psi &\text{ is } (\varphi(x, z) \wedge A = A)\\
\chi &\text{ is } \exists y\, \varphi(x, y)
\end{aligned}$$

Using Lemma 12.4, since A is a parameter to ψ, there is a transitive S such that $0, 1, A \in S$ (along with any other parameters), and such that:

$$(\forall x, z \in S)((\psi \leftrightarrow \psi^S) \wedge (\chi \leftrightarrow \chi^S))$$

So in particular:

$$(\forall x, z \in S)(\varphi(x, z) \leftrightarrow \varphi^S(x, z))$$

[6] More formally, letting ξ be either of these formulas, $\xi(z) \leftrightarrow \xi^S(z)$.

$$(\forall x \in S)(\exists y \varphi(x,y) \leftrightarrow (\exists y \in S)\varphi^S(x,y))$$

Combining these, and observing that $A \subseteq S$ since $A \in S$ and S is transitive:

$$(\forall x \in A)(\exists y \varphi(x,y) \leftrightarrow (\exists y \in S)\varphi(x,y))$$

Now $(\forall x \in A)(\exists! y \in S)\varphi(x,y)$, because $(\forall x \in A)\exists! y\, \varphi(x,y)$. Now Separation yields $\{y \in S : (\exists x \in A)\varphi(x,y)\} = \{y : (\exists x \in A)\varphi(x,y)\}$. □

12.8 Appendix: Finite axiomatizability

We close this chapter by extracting some results from Replacement. The first result is due to Montague (1961); note that it is not a proof *within* **ZF**, but a proof *about* **ZF**:

Theorem 12.6. **ZF** *is not finitely axiomatizable. More generally: if* **T** *is finite and* $\mathbf{T} \vdash \mathbf{ZF}$*, then* **T** *is inconsistent.*

(Here, we tacitly restrict ourselves to first-order sentences whose only non-logical primitive is $\in$*, and we write* $\mathbf{T} \vdash \mathbf{ZF}$ *to indicate that* $\mathbf{T} \vdash \varphi$ *for all* $\varphi \in \mathbf{ZF}$*.)*

Proof. Fix finite **T** such that $\mathbf{T} \vdash \mathbf{ZF}$. So, **T** proves Reflection, i.e. Theorem 12.2. Since **T** is finite, we can rewrite it as a single conjunction, θ. Reflecting with this formula, $\mathbf{T} \vdash \exists \beta(\theta \leftrightarrow \theta^{V_\beta})$. Since trivially $\mathbf{T} \vdash \theta$, we find that $\mathbf{T} \vdash \exists \beta\ \theta^{V_\beta}$.

Now, let $\psi(X)$ abbreviate:

$$\theta^X \wedge X \text{ is transitive} \wedge (\forall Y \in X)(Y \text{ is transitive} \rightarrow \neg\theta^Y)$$

roughly this says: X is a transitive model of θ, and $\in$-minimal in this regard. Now, recalling that $\mathbf{T} \vdash \exists \beta\ \theta^{V_\beta}$, by basic facts about ranks within **ZF** and hence within **T**, we have:

$$\mathbf{T} \vdash \exists M \psi(M). \qquad (*)$$

Using the first conjunct of $\psi(X)$, whenever $\mathbf{T} \vdash \sigma$, we have that $\mathbf{T} \vdash \forall X(\psi(X) \to \sigma^X)$. So, by (*):

$$\mathbf{T} \vdash \forall X(\psi(X) \to (\exists N \psi(N))^X)$$

Using this, and (*) again:

$$\mathbf{T} \vdash \exists M(\psi(M) \wedge (\exists N \psi(N))^M)$$

In particular, then:

$$\mathbf{T} \vdash \exists M(\psi(M) \wedge (\exists N \in M)((N \text{ is transitive})^N \wedge (\theta^N)^M))$$

So, by elementary reasoning concerning transitivity:

$$\mathbf{T} \vdash \exists M(\psi(M) \wedge (\exists N \in M)(N \text{ is transitive} \wedge \theta^N))$$

So that $\mathbf{T}$ is inconsistent.[7] □

Here is a similar result, noted by Potter (2004, 223):

Proposition 12.7. *Let* $\mathbf{T}$ *extend* $\mathbf{Z}$ *with finitely many new axioms. If* $\mathbf{T} \vdash \mathbf{ZF}$*, then* $\mathbf{T}$ *is inconsistent. (Here we use the same tacit restrictions as for Theorem 12.6.)*

Proof. Use θ for the conjunction of all of $\mathbf{T}$'s axioms *except* for the (infinitely many) instances of Separation. Defining ψ from θ as in Theorem 12.6, we can show that $\mathbf{T} \vdash \exists M \psi(M)$.

As in Theorem 12.6, we can establish the schema that, whenever $\mathbf{T} \vdash \sigma$, we have that $\mathbf{T} \vdash \forall X(\psi(X) \to \sigma^X)$. We then finish our proof, exactly as in Theorem 12.6.

However, establishing the schema involves a little more work than in Theorem 12.6. After all, the Separation-instances are in $\mathbf{T}$, but they are not conjuncts of θ. However, we can overcome this obstacle by proving that $\mathbf{T} \vdash \forall X(X \text{ is transitive} \to \sigma^X)$, for every Separation-instance σ. We leave this to the reader. □

[7] This "elementary reasoning" involves proving certain "absoluteness facts" for transitive sets.

As remarked in section 12.5, this shows that Replacement is strictly stronger than Theorem 10.26. Or, slightly more strictly: if **Z** + "every well-ordering is isomorphic to a unique ordinal" is consistent, then it fails to prove some Replacement-instance.

Problems

Problem 12.1. Formalize *Stages-are-super-cofinal* within **ZF**.

Problem 12.2. Show that, for every Separation-instance σ, we have: $\mathbf{Z} \vdash \forall X(X \text{ is transitive} \rightarrow \sigma^X)$. (We used this schema in Proposition 12.7.)

Problem 12.3. Show that, for every $\varphi \in \mathbf{Z}$, we have $\mathbf{ZF} \vdash \varphi^{V_{\omega+\omega}}$.

Problem 12.4. Confirm the remaining schematic results invoked in the proofs of Theorem 12.6 and Proposition 12.7.

CHAPTER 13

Ordinal Arithmetic

13.1 Introduction

In chapter 10, we developed a theory of ordinal numbers. We saw in chapter 11 that we can think of the ordinals as a spine around which the remainder of the hierarchy is constructed. But that is not the only role for the ordinals. There is also the task of performing ordinal arithmetic.

We already gestured at this, back in section 10.2, when we spoke of ω, $\omega + 1$ and $\omega + \omega$. At the time, we spoke informally; the time has come to spell it out properly. However, we should mention that there is not much philosophy in this chapter; just technical developments, coupled with a (mildly) interesting observation that we can do the same thing in two different ways.

13.2 Ordinal Addition

Suppose we want to add α and β. We can simply put a copy of β immediately after a copy of α. (We need to take *copies*, since we know from Proposition 10.22 that either $\alpha \subseteq \beta$ or $\beta \subseteq \alpha$.) The intuitive effect of this is to run through an α-sequence of steps,

and then to run through a β-sequence. The resulting sequence will be well-ordered; so by Theorem 10.26 it is isomorphic to a (unique) ordinal. That ordinal can be regarded as the *sum* of α and β.

That is the intuitive idea behind ordinal addition. To define it rigorously, we start with the idea of taking *copies* of sets. The idea here is to use arbitrary tags, 0 and 1, to keep track of which object came from where:

Definition 13.1. The *disjoint sum* of A and B is $A \sqcup B = (A \times \{0\}) \cup (B \times \{1\})$.

We next define an ordering on pairs of ordinals:

Definition 13.2. For any ordinals $\alpha_1, \alpha_2, \beta_1, \beta_2$, say that:

$$\langle \alpha_1, \alpha_2 \rangle \lessdot \langle \beta_1, \beta_2 \rangle \text{ iff either } \alpha_2 \in \beta_2 \text{ or both } \alpha_2 = \beta_2 \text{ and } \alpha_1 \in \beta_1$$

This is a *reverse lexicographic* ordering, since you order by the second element, then by the first. Now recall that we wanted to define $\alpha + \beta$ as the order type of a copy of α followed by a copy of β. To achieve that, we say:

Definition 13.3. For any ordinals α, β, their sum is $\alpha + \beta = \mathrm{ord}(\alpha \sqcup \beta, \lessdot)$.

Note that we slightly abused notation here; strictly we should write "$\{\langle x, y \rangle \in \alpha \sqcup \beta : x \lessdot y\}$" in place of "$\lessdot$". For brevity, though, we will continue to abuse notation in this way in what follows.

The following result, together with Theorem 10.26, confirms that our definition is well-formed:

Lemma 13.4. *$\langle \alpha \sqcup \beta, \lessdot \rangle$ is a well-order, for any ordinals α and β.*

Proof. Obviously $\lessdot$ is connected on $\alpha \sqcup \beta$. To show it is well-founded, fix a non-empty $X \subseteq \alpha \sqcup \beta$. Let Y be the subset of X whose second coordinate is as small as possible, i.e. $Y = \{\langle \gamma, i \rangle \in X : (\forall \langle \delta, j \rangle \in X) i \leq j\}$. Now choose the element of Y with smallest first coordinate. □

So we have a nice, explicit definition of ordinal addition. Here is an unsurprising fact (recall that $1 = \{0\}$, by Definition 9.7):

Proposition 13.5. *$\alpha + 1 = \alpha^+$, for any ordinal α.*

Proof. Consider the isomorphism f from $\alpha^+ = \alpha \cup \{\alpha\}$ to $\alpha \sqcup 1 = (\alpha \times \{0\}) \sqcup (\{0\} \times \{1\})$ given by $f(\gamma) = \langle \gamma, 0 \rangle$ for $\gamma \in \alpha$, and $f(\alpha) = \langle 0, 1 \rangle$. □

Moreover, it is easy to show that addition obeys certain recursive conditions:

Lemma 13.6. *For any ordinals α, β, we have:*

$$\begin{aligned} \alpha + 0 &= \alpha \\ \alpha + (\beta + 1) &= (\alpha + \beta) + 1 \\ \alpha + \beta &= \operatorname*{lsub}_{\delta < \beta} (\alpha + \delta) \qquad \text{if } \beta \text{ is a limit ordinal} \end{aligned}$$

Proof. We check case-by-case; first:

$$\begin{aligned} \alpha + 0 &= \operatorname{ord}((\alpha \times \{0\}) \cup (0 \times \{1\}), \lessdot) \\ &= \operatorname{ord}((\alpha \times \{0\}) \cup \{0\}, \lessdot) \\ &= \alpha \\ \alpha + (\beta + 1) &= \operatorname{ord}((\alpha \times \{0\}) \cup (\beta^+ \times \{1\}), \lessdot) \\ &= \operatorname{ord}((\alpha \times \{0\}) \cup (\beta \times \{1\}), \lessdot) + 1 \\ &= (\alpha + \beta) + 1 \end{aligned}$$

Now let $\beta \neq \emptyset$ be a limit. If $\delta < \beta$ then also $\delta + 1 < \beta$, so $\alpha + \delta$ is a proper initial segment of $\alpha + \beta$. So $\alpha + \beta$ is a strict upper bound on $X = \{\alpha + \delta : \delta < \beta\}$. Moreover, if $\alpha \leq \gamma < \alpha + \beta$, then clearly $\gamma = \alpha + \delta$ for some $\delta < \beta$. So $\alpha + \beta = \text{lsub}_{\delta<\beta}(\alpha + \delta)$. □

But here is a striking fact. To define ordinal addition, we could *instead* have simply used the Transfinite Recursion Theorem, and laid down the recursion equations, exactly as given in Lemma 13.6 (though using "β^+" rather than "$\beta + 1$").

There are, then, two different ways to define operations on the ordinals. We can define them *synthetically*, by explicitly constructing a well-ordered set and considering its order type. Or we can define them *recursively*, just by laying down the recursion equations. Done correctly, though, the outcome is identical. For Theorem 10.26 guarantees that these recursion equations pin down *unique* ordinals.

In many ways, ordinal arithmetic behaves just like addition of the natural numbers. For example, we can prove the following:

Lemma 13.7. *If α, β, γ are ordinals, then:*

1. *if $\beta < \gamma$, then $\alpha + \beta < \alpha + \gamma$*
2. *if $\alpha + \beta = \alpha + \gamma$, then $\beta = \gamma$*
3. *$\alpha + (\beta + \gamma) = (\alpha + \beta) + \gamma$, i.e., addition is associative*
4. *If $\alpha \leq \beta$, then $\alpha + \gamma \leq \beta + \gamma$*

Proof. We prove (3), leaving the rest as an exercise. The proof is by Simple Transfinite Induction on γ, using Lemma 13.6. When $\gamma = 0$:

$$(\alpha + \beta) + 0 = \alpha + \beta = \alpha + (\beta + 0)$$

When $\gamma = \delta + 1$, suppose for induction that $(\alpha + \beta) + \delta = \alpha + (\beta + \delta)$; now using Lemma 13.6 three times:

$$(\alpha + \beta) + (\delta + 1) = ((\alpha + \beta) + \delta) + 1$$

$$= (\alpha + (\beta + \delta)) + 1$$
$$= \alpha + ((\beta + \delta) + 1)$$
$$= \alpha + (\beta + (\delta + 1))$$

When γ is a limit ordinal, suppose for induction that if $\delta \in \gamma$ then $(\alpha + \beta) + \delta = \alpha + (\beta + \delta)$; now:

$$\begin{aligned}(\alpha + \beta) + \gamma &= \operatorname*{lsub}_{\delta<\gamma}((\alpha + \beta) + \delta) \\ &= \operatorname*{lsub}_{\delta<\gamma}(\alpha + (\beta + \delta)) \\ &= \alpha + \operatorname*{lsub}_{\delta<\gamma}(\beta + \delta) \\ &= \alpha + (\beta + \gamma)\end{aligned}$$

□

In these ways, ordinal addition should be very familiar. But, there is a crucial way in which ordinal addition is *not* like addition on the natural numbers.

Proposition 13.8. *Ordinal addition is not commutative;* $1 + \omega = \omega < \omega + 1$.

Proof. Note that $1 + \omega = \operatorname{lsub}_{n<\omega}(1 + n) = \omega \in \omega \cup \{\omega\} = \omega^+ = \omega + 1$. □

Whilst this may initially come as a surprise, *it shouldn't.* On the one hand, when you consider $1+\omega$, you are thinking about the order type you get by putting an extra element *before* all the natural numbers. Reasoning as we did with Hilbert's Hotel in section 7.1, intuitively, this extra first element shouldn't make any difference to the overall order type. On the other hand, when you consider $\omega+1$, you are thinking about the order type you get by putting an extra element *after* all the natural numbers. And that's a radically different beast!

13.3 Using Ordinal Addition

Using addition on the ordinals, we can explicitly calculate the ranks of various sets, in the sense of Definition 11.15:

Lemma 13.9. *If* $\text{rank}(A) = \alpha$ *and* $\text{rank}(B) = \beta$*, then:*

1. $\text{rank}(\wp(A)) = \alpha + 1$
2. $\text{rank}(\{A, B\}) = \max(\alpha, \beta) + 1$
3. $\text{rank}(A \cup B) = \max(\alpha, \beta)$
4. $\text{rank}(\langle A, B\rangle) = \max(\alpha, \beta) + 2$
5. $\text{rank}(A \times B) \leq \max(\alpha, \beta) + 2$
6. $\text{rank}(\bigcup A) = \alpha$ *when* α *is empty or a limit;* $\text{rank}(\bigcup A) = \gamma$ *when* $\alpha = \gamma + 1$

Proof. Throughout, we invoke Proposition 11.20 repeatedly.

(1). If $x \subseteq A$ then $\text{rank}(x) \leq \text{rank}(A)$. So $\text{rank}(\wp(A)) \leq \alpha + 1$. Since $A \in \wp(A)$ in particular, $\text{rank}(\wp(A)) = \alpha + 1$.

(2). By Proposition 11.20

(3). By Proposition 11.20.

(4). By (2), twice.

(5). Note that $A \times B \subseteq \wp(\wp(A \cup B))$, and invoke (4).

(6). If $\alpha = \gamma + 1$, there is some $c \in A$ with $\text{rank}(c) = \gamma$, and no member of A has higher rank; so $\text{rank}(\bigcup A) = \gamma$. If α is a limit ordinal, then A has members with rank arbitrarily close to (but strictly less than) α, so that $\bigcup A$ also has members with rank arbitrarily close to (but strictly less than) α, so that $\text{rank}(\bigcup A) = \alpha$. □

We leave it as an exercise to show why (5) involves an *in*equality.

We are also now in a position to show that several reasonable notions of what it might mean to describe an ordinal as "finite" or "infinite" coincide:

Lemma 13.10. *For any ordinal* α*, the following are equivalent:*

1. $\alpha \notin \omega$*, i.e.,* α *is not a natural number*

2. $\omega \leq \alpha$
3. $1 + \alpha = \alpha$
4. $\alpha \approx \alpha + 1$, *i.e.,* α *and* $\alpha + 1$ *are equinumerous*
5. α *is Dedekind infinite*

So we have five provably equivalent ways to understand what it takes for an ordinal to be (in)finite.

Proof. *(1)* $\Rightarrow$ *(2).* By Trichotomy.

(2) $\Rightarrow$ *(3).* Fix $\alpha \geq \omega$. By Transfinite Induction, there is some least ordinal γ (possibly 0) such that there is a limit ordinal β with $\alpha = \beta + \gamma$. Now:

$$1 + \alpha = 1 + (\beta + \gamma) = (1 + \beta) + \gamma = \operatorname*{lsub}_{\delta<\beta}(1 + \delta) + \gamma = \beta + \gamma = \alpha.$$

(3) $\Rightarrow$ *(4).* There is clearly a bijection $f : (\alpha \sqcup 1) \to (1 \sqcup \alpha)$. If $1+\alpha = \alpha$, there is an isomorphism $g : (1 \sqcup \alpha) \to \alpha$. Now consider $g \circ f$.

(4) $\Rightarrow$ *(5).* If $\alpha \approx \alpha + 1$, there is a bijection $f : (\alpha \sqcup 1) \to \alpha$. Define $g(\gamma) = f(\gamma, 0)$ for each $\gamma < \alpha$; this injection witnesses that α is Dedekind infinite, since $f(0,1) \in \alpha \setminus \mathrm{ran}(g)$.

(5) $\Rightarrow$ *(1).* This is Proposition 9.8. □

13.4 Ordinal Multiplication

We now turn to ordinal multiplication, and we approach this much like ordinal addition. So, suppose we want to multiply α by β. To do this, you might imagine a rectangular grid, with width α and height β; the product of α and β is now the result of moving along each row, then moving through the next row... until you have moved through the entire grid. Otherwise put, the product of α and β arises by replacing *each* element in β with a copy of α.

To make this formal, we simply use the reverse lexicographic ordering on the Cartesian product of α and β:

Definition 13.11. For any ordinals α, β, their product $\alpha \cdot \beta = \mathrm{ord}(\alpha \times \beta, \lhd)$.

We must again confirm that this is a well-formed definition:

Lemma 13.12. *$\langle \alpha \times \beta, \lhd \rangle$ is a well-order, for any ordinals α and β.*

Proof. Exactly as for Lemma 13.4. □

And it is not hard to prove that multiplication behaves thus:

Lemma 13.13. *For any ordinals α, β:*

$$\begin{aligned} \alpha \cdot 0 &= 0 \\ \alpha \cdot (\beta + 1) &= (\alpha \cdot \beta) + \alpha \\ \alpha \cdot \beta &= \operatorname*{lsub}_{\delta < \beta}(\alpha \cdot \delta) \qquad \text{when } \beta \text{ is a limit ordinal.} \end{aligned}$$

Proof. Left as an exercise. □

Indeed, just as in the case of addition, we could have defined ordinal multiplication via these recursion equations, rather than offering a direct definition. Equally, as with addition, certain behaviour is familiar:

Lemma 13.14. *If α, β, γ are ordinals, then:*

1. *if $\alpha \neq 0$ and $\beta < \gamma$, then $\alpha \cdot \beta < \alpha \cdot \gamma$;*
2. *if $\alpha \neq 0$ and $\alpha \cdot \beta = \alpha \cdot \gamma$, then $\beta = \gamma$;*
3. *$\alpha \cdot (\beta \cdot \gamma) = (\alpha \cdot \beta) \cdot \gamma$;*
4. *If $\alpha \leq \beta$, then $\alpha \cdot \gamma \leq \beta \cdot \gamma$;*

5. $\alpha \cdot (\beta + \gamma) = (\alpha \cdot \beta) + (\alpha \cdot \gamma)$.

Proof. Left as an exercise. □

You can prove (or look up) other results, to your heart's content. But, given Proposition 13.8, the following should not come as a surprise:

Proposition 13.15. *Ordinal multiplication is not commutative:* $2 \cdot \omega = \omega < \omega \cdot 2$

Proof. $2 \cdot \omega = \text{lsub}_{n<\omega}(2 \cdot n) = \omega \in \text{lsub}_{n<\omega}(\omega + n) = \omega + \omega = \omega \cdot 2$. □

Again, the intuitive rationale is quite straightforward. To compute $2 \cdot \omega$, you replace each natural number with two entities. You would get the same order type if you simply inserted all the "half" numbers into the natural numbers, i.e., you considered the natural ordering on $\{n/2 : n \in \omega\}$. And, put like that, the order type is plainly the same as that of ω itself. But, to compute $\omega \cdot 2$, you place down two copies of ω, one after the other.

13.5 Ordinal Exponentiation

We now move to ordinal exponentiation. Sadly, there is no *nice* synthetic definition for ordinal exponentiation.

Sure, there *are* explicit synthetic definitions. Here is one. Let $\text{finfun}(\alpha, \beta)$ be the set of all functions $f : \alpha \to \beta$ such that $\{\gamma \in \alpha : f(\gamma) \neq 0\}$ is equinumerous with some natural number. Define a well-ordering on $\text{finfun}(\alpha, \beta)$ by $f \sqsubset g$ iff $f \neq g$ and $f(\gamma_0) < g(\gamma_0)$, where $\gamma_0 = \max\{\gamma \in \alpha : f(\gamma) \neq g(\gamma)\}$. Then we can define $\alpha^{(\beta)}$ as $\text{ord}(\text{finfun}(\alpha, \beta), \sqsubset)$. Potter employs this explicit definition, and then immediately explains:

> The choice of this ordering is determined purely by our desire to obtain a definition of ordinal exponentiation which obeys the appropriate recursive condition..., and it is much harder to picture than either

> the ordered sum or the ordered product. (Potter, 2004, p. 199)

Quite. We explained addition as "a copy of α followed by a copy of β", and multiplication as "a β-sequence of copies of α". But we have nothing pithy to say about finfun(α, γ). So instead, we'll offer the definition of ordinal exponentiation just *by* transfinite recursion, i.e.:

Definition 13.16.

$$\begin{aligned} \alpha^{(0)} &= 1 \\ \alpha^{(\beta+1)} &= \alpha^{(\beta)} \cdot \alpha \\ \alpha^{(\beta)} &= \bigcup_{\delta<\beta} \alpha^{(\delta)} \qquad \text{when } \beta \text{ is a limit ordinal} \end{aligned}$$

If we were working *as* set theorists, we might want to explore some of the properties of ordinal exponentiation. But we have nothing much more to add, except to note the unsurprising fact that ordinal exponentiation does not commute. Thus $2^{(\omega)} = \bigcup_{\delta<\omega} 2^{(\delta)} = \omega$, whereas $\omega^{(2)} = \omega \cdot \omega$. But then, we should not *expect* exponentiation to commute, since it does not commute with natural numbers: $2^{(3)} = 8 < 9 = 3^{(2)}$.

Problems

Problem 13.1. Prove the remainder of Lemma 13.7.

Problem 13.2. Produce sets A and B such that $\text{rank}(A \times B) = \max(\text{rank}(A), \text{rank}(B))$. Produce sets A and B such that $\text{rank}(A \times B)$ $\max(\text{rank}(A), \text{rank}(B)) + 2$. Are any other ranks possible?

Problem 13.3. Prove Lemma 13.12, Lemma 13.13, and Lemma 13.14

Problem 13.4. Using Transfinite Induction, prove that, if we define $\alpha^{(\beta)} = \mathrm{ord}(\mathrm{finfun}(\alpha,\beta),\sqsubset)$, we obtain the recursion equations of Definition 13.16.

CHAPTER 14

Cardinals

14.1 Cantor's Principle

Cast your mind back to section 10.5. We were discussing well-ordered sets, and suggested that it would be nice to have objects which go proxy for well-orders. With this is mind, we introduced ordinals, and then showed in Corollary 10.28 that these behave as we would want them to, i.e.:

$$\mathrm{ord}(A, <) = \mathrm{ord}(B, \lessdot) \text{ iff } \langle A, < \rangle \cong \langle B, \lessdot \rangle.$$

Cast your mind back even further, to section 5.7. There, working naïvely, we introduced the notion of the "size" of a set. Specifically, we said that two sets are equinumerous, $A \approx B$, just in case there is a bijection $f : A \to B$. This is an intrinsically simpler notion than that of a well-ordering: we are only interested in bijections, and not (as with order-isomorphisms) whether the bijections "preserve any structure".

This all gives rise to an obvious thought. Just as we introduced certain objects, *ordinals*, to calibrate well-orders, we can introduce certain objects, *cardinals*, to calibrate size. That is the aim of this chapter.

Before we say what these cardinals will be, we should lay down a principle which they ought to satisfy. Writing $|X|$ for the cardinality of the set X, we would want them to obey:

$$|A| = |B| \text{ iff } A \approx B.$$

We'll call this *Cantor's* Principle, since Cantor was probably the first to have it very clearly in mind. (We'll say more about its relationship to *Hume's* Principle in section 14.5.) So our aim is to define $|X|$, for each X, in such a way that it delivers Cantor's Principle.

14.2 Cardinals as Ordinals

In fact, our theory of cardinals will just make (shameless) use of our theory of ordinals. That is: we will just define cardinals as certain specific ordinals. In particular, we will offer the following:

Definition 14.1. If A can be well-ordered, then $|A|$ is the least ordinal γ such that $A \approx \gamma$. For any ordinal γ, we say that γ is a *cardinal* iff $\gamma = |\gamma|$.

We just used the phrase "A can be well-ordered". As is almost always the case in mathematics, the modal locution here is just a hand-waving gloss on an existential claim: to say "A can be well-ordered" is just to say "there is a relation which well-orders A".

But there is a snag with Definition 14.1. We would like it to be the case that *every* set has a size, i.e., that $|A|$ exists for every A. The definition we just gave, though, begins with a conditional: "*If* A can be well-ordered…". If there is some set A which cannot be well-ordered, then our definition will simply fail to define an object $|A|$.

So, to use Definition 14.1, we need a guarantee that every set can be well-ordered. Sadly, though, this guarantee is unavailable in **ZF**. So, if we want to use Definition 14.1, there is no alternative but to add a new axiom, such as:

Axiom (Well-Ordering). Every set can be well-ordered.

We will discuss whether the Well-Ordering Axiom is acceptable in chapter 16. From now on, though, we will simply help ourselves to

it. And, using it, it is quite straightforward to prove that cardinals (as defined in Definition 14.1) exist and behave nicely:

Lemma 14.2. *For every set A:*

1. *$|A|$ exists and is unique;*
2. *$|A| \approx A$;*
3. *$|A|$ is a cardinal, i.e., $|A| = ||A||$;*

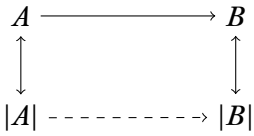
Proof. Fix A. By Well-Ordering, there is a well-ordering $\langle A, R\rangle$. By Theorem 10.26, $\langle A, R\rangle$ is isomorphic to a unique ordinal, β. So $A \approx \beta$. By Transfinite Induction, there is a uniquely least ordinal, γ, such that $A \approx \gamma$. So $|A| = \gamma$, establishing (1) and (2). To establish (3), note that if $\delta \in \gamma$ then $\delta \prec A$, by our choice of γ, so that also $\delta \prec \gamma$ since equinumerosity is an equivalence relation (Proposition 5.12). So $\gamma = |\gamma|$. □

The next result guarantees Cantor's Principle, and more besides. (Note that cardinals inherit their ordering from the ordinals, i.e., $\mathfrak{a} < \mathfrak{b}$ iff $\mathfrak{a} \in \mathfrak{b}$. In formulating this, we will use Fraktur letters for objects we know to be cardinals. This is fairly standard. A common alternative is to use Greek letters, since cardinals are ordinals, but to choose them from the middle of the alphabet, e.g.: κ, λ.):

Lemma 14.3. *For any sets A and B:*

$$A \approx B \text{ iff } |A| = |B|$$
$$A \preceq B \text{ iff } |A| \leq |B|$$
$$A \prec B \text{ iff } |A| < |B|$$

Proof. We will prove the left-to-right direction of the second claim (the other cases are similar, and left as an exercise). So, consider the following diagram:



The double-headed arrows indicate bijections, whose existence is guaranteed by Lemma 14.2. In assuming that $A \preceq B$, there is an injection $A \to B$. Now, chasing the arrows around from $|A|$ to A to B to $|B|$, we obtain an injection $|A| \to |B|$ (the dashed arrow). □

We can also use Lemma 14.3 to re-prove Schröder–Bernstein. This is the claim that if $A \preceq B$ and $B \preceq A$ then $A \approx B$. We stated this as Theorem 5.17, but first proved it—with some effort—in section 7.5. Now consider:

Re-proof of Schröder-Bernstein. If $A \preceq B$ and $B \preceq A$, then $|A| \leq |B|$ and $|B| \leq |A|$ by Lemma 14.3. So $|A| = |B|$ and $A \approx B$ by Trichotomy and Lemma 14.3. □

Whilst this is a very simple proof, it implicitly relies on both Replacement (to secure Theorem 10.26) and on Well-Ordering (to guarantee Lemma 14.3). By contrast, the proof of section 7.5 was much more self-standing (indeed, it can be carried out in $\mathbf{Z}^-$).

14.3 ZFC: A Milestone

With the addition of Well-Ordering, we have reached the final theoretical milestone. We now have all the axioms required for **ZFC**. In detail:

Definition 14.4. The theory **ZFC** has these axioms: Extensionality, Union, Pairs, Powersets, Infinity, Foundation, Well-Ordering and all instances of the Separation and Replacement schemes. Otherwise put, **ZFC** adds Well-Ordering to **ZF**.

ZFC stands for *Zermelo–Fraenkel* set theory with *Choice*. Now this might seem slightly odd, since the axiom we added was called "Well-Ordering", not "Choice". But, when we later formulate Choice, it will turn out that Well-Ordering is equivalent (modulo **ZF**) to Choice (see Theorem 16.6). So which to take as our "basic" axiom is a matter of indifference. And the name "**ZFC**" is entirely standard in the literature.

14.4 Finite, Countable, Uncountable

Now that we have been introduced to cardinals, it is worth spending a little time talking about different varieties of cardinals; specifically, finite, countable, and uncountable cardinals.

Our first two results entail that the finite cardinals will be exactly the finite ordinals, which we defined as our *natural numbers* back in Definition 9.7:

Proposition 14.5. *Let $n, m \in \omega$. Then $n = m$ iff $n \approx m$.*

Proof. *Left-to-right* is trivial. To prove *right-to-left*, suppose $n \approx m$ although $n \neq m$. By Trichotomy, either $n \in m$ or $m \in n$; suppose $n \in m$ without loss of generality. Then $n \subsetneq m$ and there is a bijection $f : m \to n$, so that m is Dedekind infinite, contradicting Proposition 9.8. □

Corollary 14.6. *If $n \in \omega$, then n is a cardinal.*

Proof. Immediate. □

It also follows that several reasonable notions of what it might mean to describe a cardinal as "finite" or "infinite" coincide:

Theorem 14.7. *For any set A, the following are equivalent:*

1. *$|A| \notin \omega$, i.e., A is not a natural number;*

2. $\omega \leq |A|$;

3. *A is Dedekind infinite.*

Proof. From Lemma 13.10, Lemma 14.3, and Corollary 14.6. □

This licenses the following *definition* of some notions which we used rather informally in part II:

Definition 14.8. We say that A is *finite* iff $|A|$ is a natural number, i.e., $|A| \in \omega$. Otherwise, we say that A is *infinite*.

But note that this definition is presented against the background of **ZFC**. After all, we needed Well-Ordering to guarantee that every set has a cardinality. And indeed, without Well-Ordering, there can be a set which is neither finite nor Dedekind infinite. We will return to this sort of issue in chapter 16. For now, we continue to rely upon Well-Ordering.

Let us now turn from the finite cardinals to the infinite cardinals. Here are two elementary points:

Corollary 14.9. *ω is the least infinite cardinal.*

Proof. ω is a cardinal, since ω is Dedekind infinite and if $\omega \approx n$ for any $n \in \omega$ then n would be Dedekind infinite, contradicting Proposition 9.8. Now ω is the least infinite cardinal by definition. □

Corollary 14.10. *Every infinite cardinal is a limit ordinal.*

Proof. Let α be an infinite successor ordinal, so $\alpha = \beta + 1$ for some β. By Proposition 14.5, β is also infinite, so $\beta \approx \beta + 1$ by Lemma 13.10. Now $|\beta| = |\beta + 1| = |\alpha|$ by Lemma 14.3, so that $\alpha \neq |\alpha|$. □

Now, as early as Definition 5.2, we flagged we can distinguish between countable and uncountable infinite sets. That definition naturally leads to the following:

Proposition 14.11. *A is countable iff $|A| \leq \omega$, and A is uncountable iff $\omega < |A|$.*

Proof. By Trichotomy, the two claims are equivalent, so it suffices to prove that A is countable iff $|A| \leq \omega$. For *right-to-left*: if $|A| \leq \omega$, then $A \preceq \omega$ by Lemma 14.3 and Corollary 14.9. For *left-to-right*: suppose A is countable; then by Definition 5.2 there are three possible cases:

1. if $A = \emptyset$, then $|A| = 0 \in \omega$, by Corollary 14.6 and Lemma 14.3.
2. if $n \approx A$, then $|A| = n \in \omega$, by Corollary 14.6 and Lemma 14.3.
3. if $\omega \approx A$, then $|A| = \omega$, by Corollary 14.9.

So in all cases, $|A| \leq \omega$. □

Indeed, ω has a special place. Whilst there are many countable ordinals:

Corollary 14.12. *ω is the only countable infinite cardinal.*

Proof. Let $\mathfrak{a}$ be a countable infinite cardinal. Since $\mathfrak{a}$ is infinite, $\omega \leq \mathfrak{a}$. Since $\mathfrak{a}$ is a countable cardinal, $\mathfrak{a} = |\mathfrak{a}| \leq \omega$. So $\mathfrak{a} = \omega$ by Trichotomy. □

Of course, there are infinitely many cardinals. So we might ask: *How many cardinals are there?* The following results show that we might want to reconsider that question.

Proposition 14.13. *If every member of X is a cardinal, then $\bigcup X$ is a cardinal.*

Proof. It is easy to check that $\bigcup X$ is an ordinal. Let $\alpha \in \bigcup X$ be an ordinal; then $\alpha \in \mathfrak{b} \in X$ for some cardinal $\mathfrak{b}$. Since $\mathfrak{b}$ is a cardinal, $\alpha \prec \mathfrak{b}$. Since $\mathfrak{b} \subseteq \bigcup X$, we have $\mathfrak{b} \preceq \bigcup X$, and so $\alpha \not\approx \bigcup X$. Generalising, $\bigcup X$ is a cardinal. □

Theorem 14.14. *There is no largest cardinal.*

Proof. For any cardinal $\mathfrak{a}$, Cantor's Theorem (Theorem 5.16) and Lemma 14.2 entail that $\mathfrak{a} < |\wp(\mathfrak{a})|$. □

Theorem 14.15. *The set of all cardinals does not exist.*

Proof. For reductio, suppose $C = \{\mathfrak{a} : \mathfrak{a} \text{ is a cardinal}\}$. Now $\bigcup C$ is a cardinal by Proposition 14.13, so by Theorem 14.14 there is a cardinal $\mathfrak{b} > \bigcup C$. By definition $\mathfrak{b} \in C$, so $\mathfrak{b} \subseteq \bigcup C$, so that $\mathfrak{b} \leq \bigcup C$, a contradiction. □

You should compare this with both Russell's Paradox and Burali-Forti.

14.5 Appendix: Hume's Principle

In section 14.1, we described Cantor's Principle. This was:

$$|A| = |B| \text{ iff } A \approx B.$$

This is very similar to what is now called *Hume's Principle*, which says:

$$\#x\, F(x) = \#x\, G(x) \text{ iff } F \sim G$$

where '$F \sim G$' abbreviates that there are exactly as many Fs as Gs, i.e., the Fs can be put into a bijection with the Gs, i.e.:

$$\begin{aligned}&\exists R(\forall v \forall y(Rvy \rightarrow (Fv \wedge Gy)) \wedge \\ &\quad \forall v(Fv \rightarrow \exists! y\, Rvy) \wedge \\ &\quad \forall y(Gy \rightarrow \exists! v\, Rvy))\end{aligned}$$

But there is a type-difference between Hume's Principle and Cantor's Principle. In the statement of Cantor's Principle, the variables "A" and "B" are first-order terms which stand for *sets*. In the statement of Hume's Principle, "F", "G" and "R" are *not* first-order terms; rather, they are in *predicate position*. (Maybe they stand for *properties*.) So we might gloss Hume's Principle in English as: the number of Fs is the number of Gs iff the Fs are bijective with the Gs. This is called *Hume's Principle*, because Hume once wrote this:

> When two numbers are so combined as that the one has always an unit answering to every unit of the other, we pronounce them equal. (Hume, 1740, Pt.III Bk.1 §1)

And Hume's Principle was brought to contemporary mathematico-logical prominence by Frege (1884, §63), who quoted this passage from Hume, before (in effect) sketching (what we have called) Hume's Principle.

You should note the structural similarity between Hume's Principle and Basic Law V. We formulated this in section 8.6 as follows:

$$\epsilon x\, F(x) = \epsilon x\, G(x) \text{iff } \forall x\, (F(x) \leftrightarrow G(x)).$$

And, at this point, some commentary and comparison might help.

There are two ways to take a principle like Hume's Principle or Basic Law V: *predicatively* or *impredicatively* (recall section 8.3). On the impredicative reading of Basic Law V, for each F, the object $\epsilon x\, F(x)$ falls within the domain of quantification that we used

in formulating Basic Law V itself. Similarly, on the impredicative reading of Hume's Principle, for each F, the object $\#x\,F(x)$ falls within the domain of quantification that we used in formulating Hume's Principle. By contrast, on the *predicative* understanding, the objects $\epsilon x\,F(x)$ and $\#x\,F(x)$ would be entities from some *different* domain.

Now, if we read Basic Law V impredicatively, it leads to inconsistency, via Naïve Comprehension (for the details, see section 8.6). Much like Naïve Comprehension, it can be rendered consistent by reading it *predicatively*. But it probably will not do everything that we wanted it to.

Hume's Principle, however, *can* consistently be read impredicatively. And, read thus, it is quite powerful.

To illustrate: consider the predicate "$x \neq x$", which obviously nothing satisfies. Hume's Principle now yields an object $\#x(x \neq x)$. We might treat this as the number 0. Now, on the *impredicative* understanding—but *only* on the impredicative understanding—this entity 0 falls within our original domain of quantification. So we can sensibly apply Hume's Principle with the predicate "$x = 0$" to obtain an object $\#x(x = 0)$. We might treat this as the number 1. Moreover, Hume's Principle entails that $0 \neq 1$, since there cannot be a bijection from the non-self-identical objects to the objects identical with 0 (there are none of the former, but one of the latter). Now, working impredicatively again, 1 falls within our original domain of quantification. So we can sensibly apply Hume's Principle with the predicate "$(x = 0 \vee x = 1)$" to obtain an object $\#x(x = 0 \vee x = 1)$. We might treat this as the number 2, and we can show that $0 \neq 2$ and $1 \neq 2$ and so on.

In short, taken impredicatively, Hume's Principle entails that there are *infinitely many objects*. And this has encouraged *neo-Fregean logicists* to take Hume's Principle as the foundation for arithmetic.

Frege *himself*, though, did not take Hume's Principle as his foundation for arithmetic. Instead, Frege proved Hume's Principle from an explicit definition: $\#x\,F(x)$ is defined as the extension of the concept $F \sim \Phi$. In modern terms, we might attempt to ren-

der this as $\#x\, F(x) = \{G : F \sim G\}$; but this will pull us back into the problems of Naïve Comprehension.

CHAPTER 15

Cardinal Arithmetic

In chapter 14, we developed a theory of cardinals. Our next step is to outline a theory of cardinal arithmetic. This chapter briefly summarises some of the elementary facts, and then points to some of the difficulties, which turn out to be fascinating and philosophically rich.

15.1 Defining the Basic Operations

Since we do not need to keep track of order, cardinal arithmetic is rather easier to define than ordinal arithmetic. We will define addition, multiplication, and exponentiation simultaneously.

Definition 15.1. When $\mathfrak{a}$ and $\mathfrak{b}$ are cardinals:

$$\mathfrak{a} \oplus \mathfrak{b} := |\mathfrak{a} \sqcup \mathfrak{b}|$$
$$\mathfrak{a} \otimes \mathfrak{b} := |\mathfrak{a} \times \mathfrak{b}|$$
$$\mathfrak{a}^{\mathfrak{b}} := \left|{}^{\mathfrak{b}}\mathfrak{a}\right|$$

where ${}^{X}Y = \{f : f \text{ is a function } X \to Y\}$. (It is easy to show

that ${}^{X}Y$ exists for any sets X and Y; we leave this as an exercise.)

It might help to explain this definition. Concerning addition: this uses the notion of disjoint sum, $\sqcup$, as defined in Definition 13.1; and it is easy to see that this definition gives the right verdict for finite cases. Concerning multiplication: Proposition 2.27 tells us that if A has n members and B has m members then $A \times B$ has $n \cdot m$ members, so our definition simply generalises the idea to transfinite multiplication. Exponentiation is similar: we are simply generalising the thought from the finite to the transfinite. Indeed, in certain ways, transfinite cardinal arithmetic looks much more like "ordinary" arithmetic than does transfinite ordinal arithmetic:

Proposition 15.2. *$\oplus$ and $\otimes$ are commutative and associative.*

Proof. For commutativity, by Lemma 14.3 it suffices to observe that $(\mathfrak{a} \sqcup \mathfrak{b}) \approx (\mathfrak{b} \sqcup \mathfrak{a})$ and $(\mathfrak{a} \times \mathfrak{b}) \approx (\mathfrak{b} \times \mathfrak{a})$. We leave associativity as an exercise. □

Proposition 15.3. *A is infinite iff $|A| \oplus 1 = 1 \oplus |A| = |A|$.*

Proof. As in Theorem 14.7, from Lemma 13.10 and Lemma 14.3. □

This explains why we need to use different symbols for ordinal versus cardinal addition/multiplication: these are genuinely *different* operations. This next pair of results shows that ordinal versus cardinal exponentiation are also different operations. (Recall that Definition 9.7 entails that $2 = \{0,1\}$):

Lemma 15.4. *$|\wp(A)| = 2^{|A|}$, for any A.*

Proof. For each subset $B \subseteq A$, let $\chi_B \in {}^{A}2$ be given by:

$$\chi_B(x) := \begin{cases} 1 & \text{if } x \in B \\ 0 & \text{otherwise.} \end{cases}$$

Now let $f(B) = \chi_B$; this defines a bijection $f : \wp(A) \to {}^A2$. So $\wp(A) \approx {}^A2$. Hence $\wp(A) \approx {}^{|A|}2$, so that $|\wp(A)| = \left|{}^{|A|}2\right| = 2^{|A|}$. □

This snappy proof essentially subsumes the discussion of section 5.6. There, we showed how to "reduce" the uncountability of $\wp(\omega)$ to the uncountability of the set of infinite binary strings, $\mathbb{B}^\omega$. In effect, $\mathbb{B}^\omega$ is just ${}^\omega 2$; and the preceding proof showed that the reasoning we went through in section 5.6 will go through using any set A in place of ω. The result also yields a quick fact about cardinal exponentiation:

Corollary 15.5. $\mathfrak{a} < 2^\mathfrak{a}$ *for any cardinal* $\mathfrak{a}$.

Proof. From Cantor's Theorem (Theorem 5.16) and Lemma 15.4. □

So $\omega < 2^\omega$. But note: this is a result about *cardinal* exponentiation. It should be contrasted with *ordinal* exponentiation, since in the latter case $\omega = 2^{(\omega)}$ (see section 13.5).

Whilst we are on the topic of cardinal exponentiation, we can also be a bit more precise about the "way" in which $\mathbb{R}$ is uncountable.

Theorem 15.6. $|\mathbb{R}| = 2^\omega$

Proof skeleton. There are plenty of ways to prove this. The most straightforward is to argue that $\wp(\omega) \preceq \mathbb{R}$ and $\mathbb{R} \preceq \wp(\omega)$, and then use Schröder-Bernstein to infer that $\mathbb{R} \approx \wp(\omega)$, and Lemma 15.4 to infer that $|\mathbb{R}| = 2^\omega$. We leave it as an (illuminating) exercise to define injections $f : \wp(\omega) \to \mathbb{R}$ and $g : \mathbb{R} \to \wp(\omega)$. □

15.2 Simplifying Addition and Multiplication

It turns out that transfinite cardinal addition and multiplication is *extremely* easy. This follows from the fact that cardinals are (certain) ordinals, and so well-ordered, and so can be manipulated

in a certain way. Showing this, though, is *not* so easy. To start, we need a tricksy definition:

Definition 15.7. We define a *canonical ordering*, $\lhd$, on pairs of ordinals, by stipulating that $\langle \alpha_1, \alpha_2 \rangle \lhd \langle \beta_1, \beta_2 \rangle$ iff either:

1. $\max(\alpha_1, \alpha_2) < \max(\beta_1, \beta_2)$; or
2. $\max(\alpha_1, \alpha_2) = \max(\beta_1, \beta_2)$ and $\alpha_1 < \beta_1$; or
3. $\max(\alpha_1, \alpha_2) = \max(\beta_1, \beta_2)$ and $\alpha_1 = \beta_1$ and $\alpha_2 < \beta_2$

Lemma 15.8. *$\langle \alpha \times \alpha, \lhd \rangle$ is a well-order, for any ordinal α.*

Proof. Evidently $\lhd$ is connected on $\alpha \times \alpha$. For suppose that neither $\langle \alpha_1, \alpha_2 \rangle$ nor $\langle \beta_1, \beta_2 \rangle$ is $\lhd$-less than the other. Then $\max(\alpha_1, \alpha_2) = \max(\beta_1, \beta_2)$ and $\alpha_1 = \beta_1$ and $\alpha_2 = \beta_2$, so that $\langle \alpha_1, \alpha_2 \rangle = \langle \beta_1, \beta_2 \rangle$.

To show well-ordering, let $X \subseteq \alpha \times \alpha$ be non-empty. Since α is an ordinal, some δ is the least member of $\{\max(\gamma_1, \gamma_2) : \langle \gamma_1, \gamma_2 \rangle \in X\}$. Now discard all pairs from $\{\langle \gamma_1, \gamma_2 \rangle \in X : \max(\gamma_1, \gamma_2) = \delta\}$ except those with least first coordinate; from among these, the pair with least second coordinate is the $\lhd$-least element of X. □

Now for a teensy, simple observation:

Proposition 15.9. *If $\alpha \approx \beta$, then $\alpha \times \alpha \approx \beta \times \beta$.*

Proof. Just let $f : \alpha \to \beta$ induce $\langle \gamma_1, \gamma_2 \rangle \mapsto \langle f(\gamma_1), f(\gamma_2) \rangle$. □

And now we will put all this to work, in proving a crucial lemma:

Lemma 15.10. *$\alpha \approx \alpha \times \alpha$, for any infinite ordinal α*

Proof. For reductio, let α be the least infinite ordinal for which this is false. Proposition 5.6 shows that $\omega \approx \omega \times \omega$, so $\omega \in \alpha$. Moreover, α is a cardinal: suppose otherwise, for reductio; then

$|\alpha| \in \alpha$, so that $|\alpha| \approx |\alpha| \times |\alpha|$, by hypothesis; and $|\alpha| \approx \alpha$ by definition; so that $\alpha \approx \alpha \times \alpha$ by Proposition 15.9.

Now, for each $\langle \gamma_1, \gamma_2 \rangle \in \alpha \times \alpha$, consider the segment:

$$\mathrm{Seg}(\gamma_1, \gamma_2) = \{\langle \delta_1, \delta_2 \rangle \in \alpha \times \alpha : \langle \delta_1, \delta_2 \rangle \lhd \langle \gamma_1, \gamma_2 \rangle\}$$

Letting $\gamma = \max(\gamma_1, \gamma_2)$, note that $\langle \gamma_1, \gamma_2 \rangle \lhd \langle \gamma + 1, \gamma + 1 \rangle$. So, when γ is infinite, observe:

$$\begin{aligned} \mathrm{Seg}(\gamma_1, \gamma_2) &\preccurlyeq ((\gamma + 1) \cdot (\gamma + 1)) \\ &\approx (\gamma \cdot \gamma), \text{ by Lemma 13.10 and Proposition 15.9} \\ &\approx \gamma, \text{ by the induction hypothesis} \\ &\prec \alpha, \text{ since } \alpha \text{ is a cardinal} \end{aligned}$$

So $\mathrm{ord}(\alpha \times \alpha, \lhd) \leq \alpha$, and hence $\alpha \times \alpha \preceq \alpha$. Since of course $\alpha \preceq \alpha \times \alpha$, the result follows by Schröder-Bernstein. □

Finally, we get to our simplifying result:

Theorem 15.11. *If* $\mathfrak{a}, \mathfrak{b}$ *are infinite cardinals, then:*

$$\mathfrak{a} \otimes \mathfrak{b} = \mathfrak{a} \oplus \mathfrak{b} = \mathit{max}(\mathfrak{a}, \mathfrak{b}).$$

Proof. Without loss of generality, suppose $\mathfrak{a} = \max(\mathfrak{a}, \mathfrak{b})$. Then invoking Lemma 15.10, $\mathfrak{a} \otimes \mathfrak{a} = \mathfrak{a} \leq \mathfrak{a} \oplus \mathfrak{b} \leq \mathfrak{a} \oplus \mathfrak{a} \leq \mathfrak{a} \otimes \mathfrak{a}$. □

Similarly, if $\mathfrak{a}$ is infinite, an $\mathfrak{a}$-sized union of $\leq \mathfrak{a}$-sized sets has size $\leq \mathfrak{a}$:

Proposition 15.12. *Let* $\mathfrak{a}$ *be an infinite cardinal. For each ordinal* $\beta \in \mathfrak{a}$*, let* X_β *be a set with* $\left|X_\beta\right| \leq \mathfrak{a}$*. Then* $\left|\bigcup_{\beta \in \mathfrak{a}} X_\beta\right| \leq \mathfrak{a}$.

Proof. For each $\beta \in \mathfrak{a}$, fix an injection $f_\beta \colon X_\beta \to \mathfrak{a}$.[1] Define an injection $g \colon \bigcup_{\beta \in \mathfrak{a}} X_\beta \to \mathfrak{a} \times \mathfrak{a}$ by $g(v) = \langle \beta, f_\beta(v) \rangle$, where $v \in X_\beta$ and $v \notin X_\gamma$ for any $\gamma \in \beta$. Now $\bigcup_{\beta \in \mathfrak{a}} X_\beta \preceq \mathfrak{a} \times \mathfrak{a} \approx \mathfrak{a}$ by Theorem 15.11. □

[1] How are these "fixed"? See section 16.5.

15.3 Some Simplification with Cardinal Exponentiation

Whilst defining $\lhd$ was a little involved, the upshot is a useful result concerning cardinal addition and multiplication, Theorem 15.11. Transfinite exponentiation, however, cannot be simplified so straightforwardly. To explain why, we start with a result which extends a familiar pattern from the finitary case (though its proof is at a high level of abstraction):

Proposition 15.13. $\mathfrak{a}^{\mathfrak{b}\oplus\mathfrak{c}} = \mathfrak{a}^{\mathfrak{b}} \otimes \mathfrak{a}^{\mathfrak{c}}$ *and* $(\mathfrak{a}^{\mathfrak{b}})^{\mathfrak{c}} = \mathfrak{a}^{\mathfrak{b}\otimes\mathfrak{c}}$*, for any cardinals* $\mathfrak{a},\mathfrak{b},\mathfrak{c}$.

Proof. For the first claim, consider a function $f\colon (\mathfrak{b} \sqcup \mathfrak{c}) \to \mathfrak{a}$. Now "split this", by defining $f_{\mathfrak{b}}(\beta) = f(\beta,0)$ for each $\beta \in \mathfrak{b}$, and $f_{\mathfrak{c}}(\gamma) = f(\gamma,1)$ for each $\gamma \in \mathfrak{c}$. The map $f \mapsto (f_{\mathfrak{b}} \times f_{\mathfrak{c}})$ is a bijection ${}^{\mathfrak{b}\sqcup\mathfrak{c}}\mathfrak{a} \to ({}^{\mathfrak{b}}\mathfrak{a} \times {}^{\mathfrak{c}}\mathfrak{a})$.

For the second claim, consider a function $f\colon \mathfrak{c} \to ({}^{\mathfrak{b}}\mathfrak{a})$; so for each $\gamma \in \mathfrak{c}$ we have some function $f(\gamma)\colon \mathfrak{b} \to \mathfrak{a}$. Now define $f^*(\beta,\gamma) = (f(\gamma))(\beta)$ for each $\langle\beta,\gamma\rangle \in \mathfrak{b} \times \mathfrak{c}$. The map $f \mapsto f^*$ is a bijection ${}^{\mathfrak{c}}({}^{\mathfrak{b}}\mathfrak{a}) \to {}^{\mathfrak{b}\otimes\mathfrak{c}}\mathfrak{a}$. □

Now, what we would *like* is an easy way to compute $\mathfrak{a}^{\mathfrak{b}}$ when we are dealing with infinite cardinals. Here is a nice step in this direction:

Proposition 15.14. *If* $2 \leq \mathfrak{a} \leq \mathfrak{b}$ *and* $\mathfrak{b}$ *is infinite, then* $\mathfrak{a}^{\mathfrak{b}} = 2^{\mathfrak{b}}$

Proof.

$$\begin{aligned}
2^{\mathfrak{b}} &\leq \mathfrak{a}^{\mathfrak{b}}, \text{ as } 2 \leq \mathfrak{a} \\
&\leq (2^{\mathfrak{a}})^{\mathfrak{b}}, \text{ by Lemma 15.4} \\
&= 2^{\mathfrak{a}\otimes\mathfrak{b}}, \text{ by Proposition 15.13} \\
&= 2^{\mathfrak{b}}, \text{ by Theorem 15.11}
\end{aligned}$$

□

We should not really expect to be able to simplify this any further, since $\mathfrak{b} < 2^{\mathfrak{b}}$ by Lemma 15.4. However, this does not tell us what to say about $\mathfrak{a}^{\mathfrak{b}}$ when $\mathfrak{b} < \mathfrak{a}$. Of course, if $\mathfrak{b}$ is *finite*, we know what to do.

Proposition 15.15. *If $\mathfrak{a}$ is infinite and $n \in \omega$ then $\mathfrak{a}^n = \mathfrak{a}$*

Proof. $\mathfrak{a}^n = \mathfrak{a} \otimes \mathfrak{a} \otimes \ldots \otimes \mathfrak{a} = \mathfrak{a}$, by Theorem 15.11. □

Additionally, in some other cases, we can control the size of $\mathfrak{a}^{\mathfrak{b}}$:

Proposition 15.16. *If $2 \leq \mathfrak{b} < \mathfrak{a} \leq 2^{\mathfrak{b}}$ and $\mathfrak{b}$ is infinite, then $\mathfrak{a}^{\mathfrak{b}} = 2^{\mathfrak{b}}$*

Proof. $2^{\mathfrak{b}} \leq \mathfrak{a}^{\mathfrak{b}} \leq (2^{\mathfrak{b}})^{\mathfrak{b}} = 2^{\mathfrak{b} \otimes \mathfrak{b}} = 2^{\mathfrak{b}}$, reasoning as in Proposition 15.14. □

But, beyond this point, things become rather more subtle.

15.4 The Continuum Hypothesis

The previous result hints (correctly) that cardinal exponentiation would be quite *easy*, if infinite cardinals are guaranteed to "play straightforwardly" with powers of 2, i.e., (by Lemma 15.4) with taking powersets. But we cannot assume that infinite cardinals *do* play straightforwardly powersets.

To start unpacking this, we introduce some nice notation.

Definition 15.17. Where $\mathfrak{a}^{\oplus}$ is the least cardinal strictly greater than $\mathfrak{a}$, we define two infinite sequences:

$$\begin{aligned} \aleph_0 &:= \omega & \beth_0 &:= \omega \\ \aleph_{\alpha+1} &:= (\aleph_\alpha)^{\oplus} & \beth_{\alpha+1} &:= 2^{\beth_\alpha} \\ \aleph_\alpha &:= \bigcup_{\beta<\alpha} \aleph_\beta & \beth_\alpha &:= \bigcup_{\beta<\alpha} \beth_\beta \quad \text{when } \alpha \text{ is a limit ordinal.} \end{aligned}$$

The definition of $\mathfrak{a}^{\oplus}$ is in order, since Theorem 14.14 tells us that, for each cardinal $\mathfrak{a}$, there is some cardinal greater than $\mathfrak{a}$, and Transfinite Induction guarantees that there is a *least* cardinal greater than $\mathfrak{a}$. The rest of the definition of $\mathfrak{a}$ is provided by transfinite recursion.

Cantor introduced this "$\aleph$" notation; this is *aleph*, the first letter in the Hebrew alphabet and the first letter in the Hebrew word for "infinite". Peirce introduced the "$\beth$" notation; this is *beth*, which is the second letter in the Hebrew alphabet.[2] Now, these notations provide us with infinite cardinals.

Proposition 15.18. *$\aleph_\alpha$ and $\beth_\alpha$ are cardinals, for every ordinal α.*

Proof. Both results hold by a simple transfinite induction. $\aleph_0 = \beth_0 = \omega$ is a cardinal by Corollary 14.9. Assuming $\aleph_\alpha$ and $\beth_\alpha$ are both cardinals, $\aleph_{\alpha+1}$ and $\beth_{\alpha+1}$ are explicitly defined as cardinals. And the union of a set of cardinals is a cardinal, by Proposition 14.13. □

Moreover, every infinite cardinal is an $\aleph$:

Proposition 15.19. *If $\mathfrak{a}$ is an infinite cardinal, then $\mathfrak{a} = \aleph_\gamma$ for some unique γ.*

Proof. By transfinite induction on cardinals. For induction, suppose that if $\mathfrak{b} < \mathfrak{a}$ then $\mathfrak{b} = \aleph_{\gamma_\mathfrak{b}}$. If $\mathfrak{a} = \mathfrak{b}^{\oplus}$ for some $\mathfrak{b}$, then $\mathfrak{a} = (\aleph_{\gamma_\mathfrak{b}})^{\oplus} = \aleph_{\gamma_\mathfrak{b}+1}$. If $\mathfrak{a}$ is not the successor of any cardinal, then since cardinals are ordinals $\mathfrak{a} = \bigcup_{\mathfrak{b}<\mathfrak{a}} \mathfrak{b} = \bigcup_{\mathfrak{b}<\mathfrak{a}} \aleph_{\gamma_\mathfrak{b}}$, so $\mathfrak{a} = \aleph_\gamma$ where $\gamma = \bigcup_{\mathfrak{b}<\mathfrak{a}} \gamma_\mathfrak{b}$. □

Since every infinite cardinal is an $\aleph$, this prompts us to ask: is every infinite cardinal a $\beth$? Certainly if that *were* the case, then the infinite cardinals would "play straightforwardly" with

[2] Peirce used this notation in a letter to Cantor of December 1900. Unfortunately, Peirce also gave a bad argument there that $\beth_\alpha$ does not exist for $\alpha \geq \omega$.

the operation of taking powersets. Indeed, we would have the following:

Generalized Continuum Hypothesis (GCH). $\aleph_\alpha = \beth_\alpha$, for all α.

Moreover, if GCH held, then we could make some considerable simplifications with cardinal exponentiation. In particular, we could show that when $\mathfrak{b} < \mathfrak{a}$, the value of $\mathfrak{a}^\mathfrak{b}$ is trapped by $\mathfrak{a} \leq \mathfrak{a}^\mathfrak{b} \leq \mathfrak{a}^\oplus$. We could then go on to give precise conditions which determine which of the two possibilities obtains (i.e., whether $\mathfrak{a} = \mathfrak{a}^\mathfrak{b}$ or $\mathfrak{a}^\mathfrak{b} = \mathfrak{a}^\oplus$).[3]

But GCH is a *hypothesis*, not a *theorem*. In fact, Gödel (1938) proved that if **ZFC** is consistent, then so is **ZFC** + GCH. But it later turned out that we can equally add ¬GCH to **ZFC**. Indeed, consider the simplest non-trivial *instance* of GCH, namely:

Continuum Hypothesis (CH). $\aleph_1 = \beth_1$.

Cohen (1963) proved that if **ZFC** is consistent then so is **ZFC** + ¬CH. So the Continuum Hypothesis is independent from **ZFC**.

The Continuum Hypothesis is so-called, since "the continuum" is another name for the real line, $\mathbb{R}$. Theorem 15.6 tells us that $|\mathbb{R}| = \beth_1$. So the Continuum Hypothesis states that there is no cardinal between the cardinality of the natural numbers, $\aleph_0 = \beth_0$, and the cardinality of the continuum, $\beth_1$.

Given the *independence* of (G)CH from **ZFC**, what should say about their *truth*? Well, there is *much* to say. Indeed, and much fertile recent work in set theory has been directed at investigating these issues. But two very quick points are certainly worth emphasising.

First: it does not *immediately* follow from these formal independence results that either GCH or CH is *indeterminate* in truth value. After all, maybe we just need to add more axioms, which

[3] The condition is dictated by *cofinality*.

strike us as natural, and which will settle the question one way or another. Gödel himself suggested that this was the right response.

Second: the independence of CH from **ZFC** is certainly *striking*, but it is certainly not *incredible* (in the literal sense). The point is simply that, for all **ZFC** tells us, moving from cardinals to their successors may involve a less blunt tool than simply taking powersets.

With those two observations made, if you want to know more, you will now have to turn to the various philosophers and mathematicians with horses in the race.[4]

15.5 ℵ-Fixed Points

In chapter 11, we suggested that Replacement stands in need of justification, because it forces the hierarchy to be rather tall. Having done some cardinal arithmetic, we can give a little illustration of the height of the hierarchy.

Evidently $0 < \aleph_0$, and $1 < \aleph_1$, and $2 < \aleph_2 \ldots$ and, indeed, the difference in size only gets *bigger* with every step. So it is tempting to conjecture that $\kappa < \aleph_\kappa$ for every ordinal κ.

But this conjecture is *false*, given **ZFC**. In fact, we can prove that there are ℵ*-fixed-points*, i.e., cardinals κ such that $\kappa = \aleph_\kappa$.

Proposition 15.20. *There is an ℵ-fixed-point.*

Proof. Using recursion, define:

$$\begin{aligned} \kappa_0 &= 0 \\ \kappa_{n+1} &= \aleph_{\kappa_n} \\ \kappa &= \bigcup_{n<\omega} \kappa_n \end{aligned}$$

Now κ is a cardinal by Proposition 14.13. But now:

$$\kappa = \bigcup_{n<\omega} \kappa_{n+1} = \bigcup_{n<\omega} \aleph_{\kappa_n} = \bigcup_{\alpha<\kappa} \aleph_\alpha = \aleph_\kappa$$

□

[4] Though you might want to start by reading Potter (2004, §15.6).

Boolos once wrote an article about exactly the $\aleph$-fixed-point we just constructed. After noting the existence of κ, at the start of his article, he said:

> [κ is] a *pretty big* number, by the lights of those with no previous exposure to set theory, so big, it seems to me, that it calls into question the truth of any theory, one of whose assertions is the claim that there are at least κ objects. (Boolos, 2000, p. 257)

And he ultimately concluded his paper by asking:

> [do] we suspect that, however it may have been at the beginning of the story, by the time we have come thus far the wheels are spinning and we are no longer listening to a description of anything that is the case? (Boolos, 2000, p. 268)

If we have, indeed, outrun "anything that is the case", then we must point the finger of blame directly at Replacement. For it is this axiom which allows our proof to work. In which case, one assumes, Boolos would need to revisit the claim he made, a few decades earlier, that Replacement has "no undesirable" consequences (see section 12.3).

But is the existence of κ so bad? It might help, here, to consider Russell's *Tristram Shandy paradox*. Tristram Shandy documents his life in his diary, but it takes him a year to record a single day. With every passing year, Tristram falls further and further behind: after one year, he has recorded only one day, and has lived 364 days unrecorded days; after two years, he has only recorded two days, and has lived 728 unrecorded days; after three years, he has only recorded three days, and lived 1092 unrecorded days ...[5] Still, if Tristram is *immortal*, Tristram will manage to record every day, for he will record the nth day on the nth year of his life. And so, "at the end of time", Tristram will have a complete diary.

[5] Forgetting about leap years.

Now: why is this so different from the thought that α is smaller than $\aleph_\alpha$—and indeed, increasingly, desperately smaller—up until κ, at which point, we catch up, and $\kappa = \aleph_\kappa$?

Setting that aside, and assuming we accept **ZFC**, let's close with a little more fun concerning fixed-point constructions. The next three results establish, intuitively, that there is a (non-trivial) point at which the hierarchy is as wide as it is tall:

Proposition 15.21. *There is a $\beth$-fixed-point, i.e., a κ such that $\kappa = \beth_\kappa$.*

Proof. As in Proposition 15.20, using "$\beth$" in place of "$\aleph$". □

Proposition 15.22. *$|V_{\omega+\alpha}| = \beth_\alpha$. If $\omega \cdot \omega \leq \alpha$, then $|V_\alpha| = \beth_\alpha$.*

Proof. The first claim holds by a simple transfinite induction. The second claim follows, since if $\omega \cdot \omega \leq \alpha$ then $\omega + \alpha = \alpha$. To establish this, we use facts about ordinal arithmetic from chapter 13. First note that $\omega \cdot \omega = \omega \cdot (1 + \omega) = (\omega \cdot 1) + (\omega \cdot \omega) = \omega + (\omega \cdot \omega)$. Now if $\omega \cdot \omega \leq \alpha$, i.e., $\alpha = (\omega \cdot \omega) + \beta$ for some β, then $\omega + \alpha = \omega + ((\omega \cdot \omega) + \beta) = (\omega + (\omega \cdot \omega)) + \beta = (\omega \cdot \omega) + \beta = \alpha$.□

Corollary 15.23. *There is a κ such that $|V_\kappa| = \kappa$.*

Proof. Let κ be a $\beth$-fixed point, as given by Proposition 15.21. Clearly $\omega \cdot \omega < \kappa$. So $|V_\kappa| = \beth_\kappa = \kappa$ by Proposition 15.22. □

There are as many stages beneath V_κ as there are members of V_κ. Intuitively, then, V_κ is as wide as it is tall. This is very Tristram-Shandy-esque: we move from one stage to the next by taking *powersets*, thereby making our hierarchy *much* bigger with each step. But, "in the end", i.e., at stage κ, the hierarchy's width catches up with its height.

One might ask: *How often does the hierarchy's width match its height?* The answer is: *As often as there are ordinals.* But this needs a little explanation.

We define a term τ as follows. For any A, let:

$$\begin{aligned}\tau_0(A) &\coloneqq |A| \\ \tau_{n+1}(A) &\coloneqq \beth_{\tau_n(A)} \\ \tau(A) &\coloneqq \bigcup_{n<\omega} \tau_n(A)\end{aligned}$$

As in Proposition 15.21, $\tau(A)$ is a $\beth$-fixed point for any A, and trivially $|A| < \tau(A)$. So now consider this recursive definition:

$$\begin{aligned}W_0 &\coloneqq 0 \\ W_{\alpha+1} &\coloneqq \tau(W_\alpha) \\ W_\alpha &\coloneqq \bigcup_{\beta<\alpha} W_\beta, \text{ when } \alpha \text{ is a limit}\end{aligned}$$

The construction is defined for all ordinals. Intuitively, then, W is "an injection" from the ordinals to $\beth$-fixed points. And, exactly as before, V_{W_α} is as wide as it is tall, for any α.

Problems

Problem 15.1. Prove in $\mathbf{Z}^-$ that ${}^X Y$ exists for any sets X and Y. Working in **ZF**, compute $\mathrm{rank}({}^X Y)$ from $\mathrm{rank}(X)$ and $\mathrm{rank}(Y)$, in the manner of Lemma 13.9.

Problem 15.2. Prove that $\oplus$ and $\otimes$ are associative.

Problem 15.3. Complete the proof of Theorem 15.6, by showing that $\wp(\omega) \preceq \mathbb{R}$ and $\mathbb{R} \preceq \wp(\omega)$.

CHAPTER 16

Choice

16.1 Introduction

In chapters 14 to 15, we developed a theory of cardinals by treating cardinals as ordinals. That approach depends upon the Axiom of Well-Ordering. It turns out that Well-Ordering is equivalent to another principle—the Axiom of Choice—and there has been serious philosophical discussion of its acceptability. Our question for this chapter are: How is the Axiom used, and can it be justified?

16.2 The Tarski–Scott Trick

In Definition 14.1, we defined cardinals as ordinals. To do this, we assumed the Axiom of Well-Ordering. We did this, for no other reason than that it is the "industry standard".

Before we discuss any of the philosophical issues surrounding Well-Ordering, then, it is important to be clear that we *can* depart from the industry standard, and develop a theory of cardinals *without* assuming Well-Ordering. We can still employ the definitions of $A \approx B$, $A \preceq B$ and $A \prec B$, as they appeared in chapter 5. We will just need a new notion of *cardinal*.

A naïve thought would be to attempt to define A's cardinality

thus:

$$\{x : A \approx x\}.$$

You might want to compare this with Frege's definition of $\#xFx$, sketched at the very end of section 14.5. And, for reasons we gestured at there, this definition fails. Any singleton set is equinumerous with $\{\emptyset\}$. But new singleton sets are formed at every successor stage of the hierarchy (just consider the singleton of the previous stage). So $\{x : A \approx x\}$ does not exist, since it cannot have a rank.

To get around this problem, we use a trick due to Tarski and Scott:[1]

Definition 16.1 (Tarski–Scott). For any formula $\varphi(x)$, let $[x : \varphi(x)]$ be the set of all x, of least possible rank, such that $\varphi(x)$ (or $\emptyset$, if there are no φs).

We should check that this definition is legitimate. Working in **ZF**, Theorem 11.13 guarantees that $\mathrm{rank}(x)$ exists for every x. Now, if there are any entities satisfying φ, then we can let α be the least rank such that $(\exists x \subseteq V_\alpha)\varphi(x)$, i.e., $(\forall \beta \in \alpha)(\forall x \subseteq V_\beta)\neg\varphi(x)$. We can then define $[x : \varphi(x)]$ by Separation as $\{x \in V_{\alpha+1} : \varphi(x)\}$.

Having justified the Tarski–Scott trick, we can now use it to define a notion of cardinality:

Definition 16.2. The TS-cardinality of A is $\mathrm{tsc}(A) = [x : A \approx x]$.

The definition of a TS-cardinal does not use Well-Ordering. But, even without that Axiom, we can show that *TS-cardinals* behave rather like *cardinals* as defined in Definition 14.1. For example, if we restate Lemma 14.3 and Lemma 15.4 in terms of TS-cardinals, the proofs go through just fine in **ZF**, without assuming Well-Ordering.

[1] A reminder: all formulas may have parameters (unless explicitly stated otherwise).

Whilst we are on the topic, it is worth noting that we can also develop a theory of ordinals using the Tarski–Scott trick. Where $\langle A, <\rangle$ is a well-ordering, let $\mathrm{tso}(A, <) = [\langle X, R\rangle : \langle A, <\rangle \cong \langle X, R\rangle]$. For more on this treatment of cardinals and ordinals, see Potter (2004, chs. 9–12).

16.3 Comparability and Hartogs' Lemma

That's the plus side. Here's the minus side. Without Choice, things get *messy*. To see why, here is a nice result due to Hartogs (1915):

Lemma 16.3 (*in* ZF). *For any set A, there is an ordinal α such that $\alpha \not\preceq A$*

Proof. If $B \subseteq A$ and $R \subseteq B^2$, then $\langle B, R\rangle \subseteq V_{\mathrm{rank}(A)+4}$ by Lemma 13.9. So, using Separation, consider:

$$C = \{\langle B, R\rangle \in V_{\mathrm{rank}(A)+5} : B \subseteq A \text{ and } \langle B, R\rangle \text{ is a well-ordering}\}$$

Using Replacement and Theorem 10.26, form the set:

$$\alpha = \{\mathrm{ord}(B, R) : \langle B, R\rangle \in C\}.$$

By Corollary 10.19, α is an ordinal, since it is a transitive set of ordinals. After all, if $\gamma \in \beta \in \alpha$, then $\beta = \mathrm{ord}(B, R)$ for some $B \subseteq R$, whereupon $\gamma = \mathrm{ord}(B_b, R_b)$ for some $b \in B$ by Lemma 10.10, so that $\gamma \in \alpha$.

For reductio, suppose there is an injection $f : \alpha \to A$. Then, where:

$$B = \mathrm{ran}(f)$$
$$R = \{\langle f(\alpha), f(\beta)\rangle \in A \times A : \alpha \in \beta\}.$$

Clearly $\alpha = \mathrm{ord}(B, R)$ and $\langle B, R\rangle \in C$. So $\alpha \in \alpha$, which is a contradiction. □

This entails a deep result:

Theorem 16.4 (*in* ZF). *The following claims are equivalent:*

1. *The Axiom of Well-Ordering*
2. *Either $A \preceq B$ or $B \preceq A$, for any sets A and B*

Proof. (1) $\Rightarrow$ (2). Fix A and B. Invoking (1), there are well-orderings $\langle A, R\rangle$ and $\langle B, S\rangle$. Invoking Theorem 10.26, let $f : \alpha \to \langle A, R\rangle$ and $g : \beta \to \langle B, S\rangle$ be isomorphisms. By Proposition 10.22, either $\alpha \subseteq \beta$ or $\beta \subseteq \alpha$. If $\alpha \subseteq \beta$, then $g \circ f^{-1} : A \to B$ is an injection, and hence $A \preceq B$; similarly, if $\beta \subseteq \alpha$ then $B \preceq A$.

(2) $\Rightarrow$ (1). Fix A; by Lemma 16.3 there is some ordinal β such that $\beta \not\preceq A$. Invoking (2), we have $A \preceq \beta$. So there is some injection $f : A \to \beta$, and we can use this injection to well-order the elements of A, by defining an order $\{\langle a, b\rangle \in A \times A : f(a) \in f(b)\}$. □

As an immediate consequence: if Well-Ordering fails, then some sets are *literally incomparable* with regard to their size. So, if Well-Ordering fails, then transfinite cardinal arithmetic will be messy. For example, we will have to abandon the idea that if A and B are infinite then $A \sqcup B \approx A \times B \approx M$, where M is the larger of A and B (see Theorem 15.11). The problem is simple: if we cannot *compare* the size of A and B, then it is nonsensical to ask which is larger.

16.4 The Well-Ordering Problem

Evidently rather a lot hangs on whether we accept Well-Ordering. But the discussion of this principle has tended to focus on an equivalent principle, the Axiom of Choice. So we will now turn our attention to that (and prove the equivalence).

In 1883, Cantor expressed his support for the Axiom of Well-Ordering, calling it "a law of thought which appears to me to be fundamental, rich in its consequences, and particularly remarkable for its general validity" (cited in Potter 2004, p. 243). But

Cantor ultimately became convinced that the "Axiom" was in need of proof. So did the mathematical community.

The problem was "solved" by Zermelo in 1904. To explain his solution, we need some definitions.

Definition 16.5. A function f is a *choice function* iff $f(x) \in x$ for all $x \in \mathrm{dom}(f)$. We say that f is a *choice function for* A iff f is a choice function with $\mathrm{dom}(f) = A \setminus \{\emptyset\}$.

Intuitively, for every (non-empty) set $x \in A$, a choice function for A *chooses* a particular element, $f(x)$, from x. The Axiom of Choice is then:

Axiom (Choice). Every set has a choice function.

Zermelo showed that Choice entails well-ordering, and vice versa:

Theorem 16.6 (in ZF). *Well-Ordering and Choice are equivalent.*

Proof. Left-to-right. Let A be a set of sets. Then $\bigcup A$ exists by the Axiom of Union, and so by Well-Ordering there is some $<$ which well-orders $\bigcup A$. Now let $f(x) =$ the $<$-least member of x. This is a choice function for A.

Right-to-left. Fix A. By Choice, there is a choice function, f, for $\wp(A) \setminus \{\emptyset\}$. Using Transfinite Recursion, define a function:

$$g(0) = f(A)$$
$$g(\alpha) = \begin{cases} \text{stop!} & \text{if } A = g[\alpha] \\ f(A \setminus g[\alpha]) & \text{otherwise} \end{cases}$$

The indication to "stop!" is just a shorthand for what would otherwise be a more long-winded definition. That is, when $A = g[\alpha]$ for the first time, let $g(\delta) = A$ for all $\delta \leq \alpha$. Now, in the first instance, we can only be sure that this defines a *term* (see the remarks after Theorem 11.4); but we will show that we indeed have a function.

Since f is a choice function, for each α (when defined) we have $g(\alpha) = f(A \setminus g[\alpha]) \in A \setminus g[\alpha]$; i.e., $g(\alpha) \notin g[\alpha]$. So if $g(\alpha) = g(\beta)$ then $g(\beta) \notin g[\alpha]$, i.e., $\beta \notin \alpha$, and similarly $\alpha \notin \beta$. So $\alpha = \beta$, by Trichotomy. So g is injective.

Next, observe that we do stop!, i.e. that there is some (least) ordinal α such that $A = g[\alpha]$. For suppose otherwise; then as g is injective we would have $\alpha \prec \wp(A) \setminus \{\emptyset\}$ for every ordinal α, contradicting Lemma 16.3. Hence also $\mathrm{ran}(g) = A$.

Assembling these facts, g is a bijection from some ordinal to A. Now g can be used to well-order A. □

So Well-Ordering and Choice stand or fall together. But the question remains: do they stand or fall?

16.5 Countable Choice

It is easy to prove, without any use of Choice/Well-Ordering, that:

Lemma 16.7 (in $\mathbf{Z}^-$). *Every finite set has a choice function.*

Proof. Let $a = \{b_1, \ldots, b_n\}$. Suppose for simplicity that each $b_i \neq \emptyset$. So there are objects $c_1, \ldots, c_n$ such that $c_1 \in b_1, \ldots, c_n \in b_n$. Now by Proposition 9.5, the set $\{\langle b_1, c_1\rangle, \ldots, \langle b_n, c_n\rangle\}$ exists; and this is a choice function for a. □

But matters get murkier as soon as we consider infinite sets. For example, consider this "minimal" extension to the above:

Countable Choice. Every *countable* set has a choice function.

This is a special case of Choice. And it transpires that this principle was invoked fairly frequently, without an obvious awareness of its use. Here are two nice examples.[2]

[2] Due to Potter (2004, §9.4) and Luca Incurvati.

Example 16.8. Here is a natural thought: for any set A, either $\omega \preceq A$, or $A \approx n$ for some $n \in \omega$. This is one way to state the intuitive idea, that every set is either finite or infinite. Cantor, and many other mathematicians, made this claim without proving it. Cautious as we are, we proved this in Theorem 14.7. But in that proof we were working in **ZFC**, since we were assuming that any set A can be well-ordered, and hence that $|A|$ is guaranteed to exist. That is: we explicitly assumed Choice.

In fact, Dedekind (1888) offered his own proof of this claim, as follows:

Theorem 16.9 (in $\mathbf{Z}^-$ + Countable Choice). *For any A, either $\omega \preceq A$ or $A \approx n$ for some $n \in \omega$.*

Proof. Suppose $A \not\approx n$ for all $n \in \omega$. Then in particular for each $n < \omega$ there is subset $A_n \subseteq A$ with exactly 2^n elements. Using this sequence $A_0, A_1, A_2, \ldots$, we define for each n:

$$B_n = A_n \setminus \bigcup_{i<n} A_i.$$

Now note the following

$$\begin{aligned}\left|\bigcup_{i<n} A_n\right| &\leq |A_0| + |A_1| + \ldots + |A_{n-1}| \\ &= 1 + 2 + \ldots + 2^{n-1} \\ &= 2^n - 1 \\ &< 2^n = |A_n|\end{aligned}$$

Hence each B_n has at least one member, c_n. Moreover, the B_ns are pairwise disjoint; so if $c_n = c_m$ then $n = m$. But every $c_n \in A$. So the function $f(n) = c_n$ is an injection $\omega \to A$. □

Dedekind did not flag that he had used Countable Choice. But, did *you* spot its use? Look again. (Really: *look again.*)

The proof used Countable Choice twice. We used it once, to obtain our sequence of sets A_0, A_1, A_2, ... We then used it

again to select our elements c_n from each B_n. Moreover, this use of Choice is ineliminable. Cohen (1966, p. 138) proved that the result fails if we have no version of Choice. That is: it is consistent with **ZF** that there are sets which are *incomparable* with ω.

Example 16.10. In 1878, Cantor stated that a countable union of countable sets is countable. He did not present a proof, perhaps indicating that he took the proof to be obvious. Now, cautious as we are, we proved a more general version of this result in Proposition 15.12. But our proof explicitly assumed Choice. And even the proof of the less general result requires Countable Choice.

Theorem 16.11 (in $\mathbf{Z}^-$ + Countable Choice). *If A_n is countable for each $n \in \omega$, then $\bigcup_{n<\omega} A_n$ is countable.*

Proof. Without loss of generality, suppose that each $A_n \neq \emptyset$. So for each $n \in \omega$ there is a surjection $f_n : \omega \to A_n$. Define $f : \omega \times \omega \to \bigcup_{n<\omega} A_n$ by $f(m,n) = f_n(m)$. The result follows because $\omega \times \omega$ is countable (Proposition 5.6) and f is a surjection. □

Did you spot the use of the Countable Choice? It is used to choose our sequence of functions f_0, f_1, f_2, ...[3] And again, the result fails in the absence of any Choice principle. Specifically, Feferman and Levy (1963) proved that it is consistent with **ZF** that a countable union of countable sets has cardinality $\beth_1$. But here is a much funnier statement of the point, from Russell:

> This is illustrated by the millionaire who bought a pair of socks whenever he bought a pair of boots, and never at any other time, and who had such a passion for buying both that at last he had $\aleph_0$ pairs of boots and $\aleph_0$ pairs of socks... Among boots we can distinguish right and left, and therefore we can

[3] A similar use of Choice occurred in Proposition 15.12, when we gave the instruction "For each $\beta \in \mathfrak{a}$, fix an injection f_β".

> make a selection of one out of each pair, namely, we can choose all the right boots or all the left boots; but with socks no such principle of selection suggests itself, and we cannot be sure, unless we assume the multiplicative axiom [i.e., in effect Choice], that there is any class consisting of one sock out of each pair. (Russell, 1919, p. 126)

In short, some form of Choice is needed to prove the following: If you have countably many pairs of socks, then you have (only) countably many socks. And in fact, without Countable Choice (or something equivalent), a countable union of countable sets can fail to be countable.

The moral is that Countable Choice was used repeatedly, without much awareness of its users. The philosophical question is: How could we *justify* Countable Choice?

An attempt at an intuitive justification might invoke an appeal to a supertask. Suppose we make the first choice in $1/2$ a minute, our second choice in $1/4$ a minute, ..., our n-th choice in $1/2^n$ a minute, ... Then within 1 minute, we will have made an ω-sequence of choices, and defined a choice function.

But what, really, could such a thought-experiment tell us? For a start, it relies upon taking this idea of "choosing" rather literally. For another, it seems to bind up mathematics in metaphysical possibility.

More important: it is not going to give us any justification for Choice *tout court*, rather than *mere* Countable Choice. For if we need *every* set to have a choice function, then we'll need to be able to perform a "supertask of arbitrary ordinal length." Bluntly, that idea is laughable.

16.6 Intrinsic Considerations about Choice

The broader question, then, is whether Well-Ordering, or Choice, or indeed the comparability of all sets as regards their size—it

doesn't matter which—can be justified.

Here is an attempted *intrinsic* justification. Back in section 9.1, we introduced several principles about the hierarchy. One of these is worth restating:

> *Stages-accumulate.* For any stage S, and for any sets which were formed *before* stage S: a set is formed at stage S whose members are exactly those sets. Nothing else is formed at stage S.

In fact, many authors have suggested that the Axiom of Choice can be justified via (something like) this principle. We will briefly provide a gloss on that approach.

We will start with a simple little result, which offers *yet another* equivalent for Choice:

Theorem 16.12 (in ZF). *Choice is equivalent to the following principle. If the members of A are disjoint and non-empty, then there is some C such that $C \cap x$ is a singleton for every $x \in A$. (We call such a C a choice set for A.)*

The proof of this result is straightforward, and we leave it as an exercise for the reader.

The essential point is that a choice set for A is just the range of a choice function for A. So, to justify Choice, we can simply try to justify its equivalent formulation, in terms of the existence of choice sets. And we will now try to do exactly that.

Let A's members be disjoint and non-empty. By *Stages-are-key* (see section 9.1), A is formed at some stage S. Note that all the members of $\bigcup A$ are available before stage S. Now, by *Stages-accumulate*, for *any* sets which were formed before S, a set is formed whose members are exactly those sets. Otherwise put: every *possible* collections of earlier-available sets will exist at S. But it is certainly *possible* to select objects which could be formed into a choice set for A; that is just some very specific subset of $\bigcup A$. So: some such choice set exists, as required.

Well, that's a *very* quick attempt to offer a justification of Choice on intrinsic grounds. But, to pursue this idea further, you should read Potter's (2004, §14.8) neat development of it.

16.7 The Banach–Tarski Paradox

We might also attempt to justify Choice, as Boolos attempted to justify Replacement, by appealing to *extrinsic* considerations (see section 12.3). After all, adopting Choice has many desirable consequences: the ability to compare every cardinal; the ability to well-order every set; the ability to treat cardinals as a particular kind of ordinal; etc.

Sometimes, however, it is claimed that Choice has *undesirable* consequences. Mostly, this is due to a result by Banach and Tarski (1924).

Theorem 16.13 (Banach–Tarski Paradox (in ZFC)). *Any ball can be decomposed into finitely many pieces, which can be reassembled (by rotation and transportation) to form two copies of that ball.*

At first glance, this is a bit amazing. Clearly the two balls have *twice* the volume of the original ball. But rigid motions—rotation and transportation—do not change volume. So it looks as if Banach–Tarski allows us to magick new matter into existence.

It gets worse.[4] Similar reasoning shows that a pea can be cut into finitely many pieces, which can then be reassembled (by rotation and transportation) to form an entity the shape and size of Big Ben.

None of this, however, holds in **ZF** on its own.[5] So we face a decision: reject Choice, or learn to live with the "paradox".

We're going to suggest that we should learn to live with the "paradox". Indeed, we don't think it's much of a paradox at all.

[4] See Tomkowicz and Wagon (2016, Theorem 3.12).

[5] Though Banach–Tarski can be proved with principles which are strictly weaker than Choice; see Tomkowicz and Wagon (2016, 303).

In particular, we don't see why it is any more or less paradoxical than any of the following results:[6]

1. There are as many points in the interval $(0,1)$ as in $\mathbb{R}$.
   Proof: consider $\tan(\pi(r - 1/2))$.

2. There are as many points in a line as in a square.
   See section 1.3 and section 5.10.

3. There are space-filling curves.
   See section 1.3 and section 5.11.

None of these three results require Choice. Indeed, we now just regard them as surprising, lovely, bits of mathematics. Maybe we should adopt the same attitude to the Banach–Tarski Paradox.

To be sure, a technical observation is required here; but it only requires keeping a level head. Rigid motions preserve volume. Consequently, the five[7] pieces into which the ball is decomposed cannot all be *measurable*. Roughly put, then, it makes no sense to assign a volume to these individual pieces. You should think of these as unpicturable, "infinite scatterings" of points. Now, maybe it is "weird" to conceive of such "infinitely scattered" sets. But their existence seems to fall out from the injunction, embodied in *Stages-accumulate*, that you should form *all possible* collections of earlier-available sets.

If none of that convinces, here is a final (extrinsic) argument in favour of embracing the Banach–Tarski Paradox. It immediately entails the best math joke of all time:

> *Question.* What's an anagram of "Banach–Tarski"?
>
> *Answer.* "Banach–Tarski Banach–Tarski".

[6] Potter (2004, 276–7), Weston (2003, 16), Tomkowicz and Wagon (2016, 31, 308–9), make similar points, using other examples.

[7] We stated the Paradox in terms of "finitely many pieces". In fact, Robinson (1947) proved that the decomposition can be achieved with *five* pieces (but no fewer). For a proof, see Tomkowicz and Wagon (2016, pp. 66–7).

16.8 Appendix: Vitali's Paradox

To get a real sense of whether the Banach-Tarski construction is acceptable or not, we should examine its *proof*. Unfortunately, that would require much more algebra than we can present here. However, we can offer some quick remarks which might shed some insight on the proof of Banach-Tarski,[8] by focussing on the following result:

Theorem 16.14 (Vitali's Paradox (in ZFC)). *Any circle can be decomposed into countably many pieces, which can be reassembled (by rotation and transportation) to form two copies of that circle.*

Vitali's Paradox is much easier to prove than the Banach–Tarski Paradox. We have called it "Vitali's Paradox", since it follows from Vitali's 1905 construction of an unmeasurable set. But the set-theoretic aspects of the proof of Vitali's Paradox and the Banach-Tarski Paradox are very similar. The essential difference between the results is just that Banach-Tarski considers a *finite* decomposition, whereas Vitali's Paradox considers a *countably infinite* decomposition. As Weston (2003) puts it, Vitali's Paradox "is certainly not nearly as striking as the Banach–Tarski paradox, but it does illustrate that geometric paradoxes can happen even in 'simple' situations."

Vitali's Paradox concerns a two-dimensional figure, a circle. So we will work on the plane, $\mathbb{R}^2$. Let R be the set of (clockwise) rotations of points around the origin by *rational* radian values between $[0, 2\pi)$. Here are some algebraic facts about R (if you don't understand the statement of the result, the proof will make its meaning clear):

Lemma 16.15. *R forms an abelian group under composition of functions.*

[8] For a much fuller treatment, see Weston (2003) or Tomkowicz and Wagon (2016).

Proof. Writing 0_R for the rotation by 0 radians, this is an identity element for R, since $\rho \circ 0_R = 0_R \circ \rho = \rho$ for any $\rho \in R$.

Every element has an inverse. Where $\rho \in R$ rotates by r radians, $\rho^{-1} \in R$ rotates by $2\pi - r$ radians, so that $\rho \circ \rho^{-1} = 0_R$.

Composition is associative: $(\tau \circ \sigma) \circ \rho = \tau \circ (\sigma \circ \rho)$ for any $\rho, \sigma, \tau \in R$

Composition is commutative: $\sigma \circ \rho = \rho \circ \sigma$ for any $\rho, \sigma \in R$.□

In fact, we can split our group R in half, and then use either half to recover the whole group:

Lemma 16.16. *There is a partition of R into two disjoint sets, R_1 and R_2, both of which are a basis for R.*

Proof. Let R_1 consist of the rotations by rational radian values in $[0, \pi)$; let $R_2 = R \setminus R_1$. By elementary algebra, $\{\rho \circ \rho : \rho \in R_1\} = R$. A similar result can be obtained for R_2. □

We will use this fact about groups to establish Theorem 16.14. Let **S** be the unit circle, i.e., the set of points exactly 1 unit away from the origin of the plane, i.e., $\{\langle r, s\rangle \in \mathbb{R}^2 : \sqrt{r^2 + s^2} = 1\}$. We will split **S** into parts by considering the following relation on **S**:

$$r \sim s \textit{ iff } (\exists \rho \in R)\rho(r) = s.$$

That is, the points of **S** are linked by this relation iff you can get from one to the other by a rational-valued rotation about the origin. Unsurprisingly:

Lemma 16.17. *$\sim$ is an equivalence relation.*

Proof. Trivial, using Lemma 16.15. □

We now invoke Choice to obtain a set, C, containing exactly one member from each equivalence class of **S** under $\sim$. That is,

we consider a choice function f on the set of equivalence classes,[9]

$$E = \{[r]_\sim : r \in \mathbf{S}\},$$

and let $C = \mathrm{ran}(f)$. For each rotation $\rho \in R$, the set $\rho[C]$ consists of the points obtained by applying the rotation ρ to each point in C. These next two results show that these sets cover the circle completely and without overlap:

Lemma 16.18. $\mathbf{S} = \bigcup_{\rho \in R} \rho[C]$.

Proof. Fix $s \in \mathbf{S}$; there is some $r \in C$ such that $r \in [s]_\sim$, i.e., $r \sim s$, i.e., $\rho(r) = s$ for some $\rho \in R$. □

Lemma 16.19. *If $\rho_1 \neq \rho_2$ then $\rho_1[C] \cap \rho_2[C] = \emptyset$.*

Proof. Suppose $s \in \rho_1[C] \cap \rho_2[C]$. So $s = \rho_1(r_1) = \rho_2(r_2)$ for some $r_1, r_2 \in C$. Hence $\rho_2^{-1}(\rho_1(r_1)) = r_2$, and $\rho_2^{-1} \circ \rho_1 \in R$, so $r_1 \sim r_2$. So $r_1 = r_2$, as C selects exactly one member from each equivalence class under $\sim$. So $s = \rho_1(r_1) = \rho_2(r_1)$, and hence $\rho_1 = \rho_2$. □

We now apply our earlier algebraic facts to our circle:

Lemma 16.20. *There is a partition of $\mathbf{S}$ into two disjoint sets, D_1 and D_2, such that D_1 can be partitioned into countably many sets which can be rotated to form a copy of $\mathbf{S}$ (and similarly for D_2).*

Proof. Using R_1 and R_2 from Lemma 16.16, let:

$$D_1 = \bigcup_{\rho \in R_1} \rho[C] \qquad D_2 = \bigcup_{\rho \in R_2} \rho[C]$$

[9] Since R is countable, each member of E is countable. Since $\mathbf{S}$ is uncountable, it follows from Lemma 16.18 and Proposition 15.12 that E is uncountable. So this is a use of *un*countable Choice.

This is a partition of **S**, by Lemma 16.18, and D_1 and D_2 are disjoint by Lemma 16.19. By construction, D_1 can be partitioned into countably many sets, $\rho[C]$ for each $\rho \in R_1$. And these can be rotated to form a copy of **S**, since $\mathbf{S} = \bigcup_{\rho \in R} \rho[C] = \bigcup_{\rho \in R_1} (\rho \circ \rho)[C]$ by Lemma 16.16 and Lemma 16.18. The same reasoning applies to D_2. □

This immediately entails Vitali's Paradox. For we can generate *two* copies of **S** from **S**, just by splitting it up into countably many pieces (the various $\rho[C]$'s) and then rigidly moving them (simply rotate each piece of D_1, and first transport and then rotate each piece of D_2).

Let's recap the proof-strategy. We started with some algebraic facts about the group of rotations on the plane. We used this group to partition **S** into equivalence classes. We then arrived at a "paradox", by using Choice to select elements from each class.

We use exactly the same strategy to prove Banach–Tarski. The main difference is that the algebraic facts used to prove Banach–Tarski are significantly more complicated than those used to prove Vitali's Paradox. But those algebraic facts have nothing to do with Choice. We will summarise them quickly.

To prove Banach–Tarski, we start by establishing an analogue of Lemma 16.16: any *free group* can be split into four pieces, which intuitively we can "move around" to recover two copies of the whole group.[10] We then show that we can use two particular rotations around the origin of $\mathbb{R}^3$ to generate a free group of rotations, F.[11] (No Choice yet.) We now regard points on the surface of the sphere as "similar" iff one can be obtained from the other by a rotation in F. We then *use Choice* to select exactly one point from each equivalence class of "similar" points. Applying our division of F to the surface of the sphere, as in Lemma 16.20, we split that surface into four pieces, which we can "move around"

[10] The fact that we can use *four* pieces is due to Robinson (1947). For a recent proof, see Tomkowicz and Wagon (2016, Theorem 5.2). We follow Weston (2003, p. 3) in describing this as "moving" the pieces of the group.

[11] See Tomkowicz and Wagon (2016, Theorem 2.1).

to obtain two copies of the surface of the sphere. And this establishes (Hausdorff, 1914):

Theorem 16.21 (Hausdorff's Paradox (in ZFC)). *The surface of any sphere can be decomposed into finitely many pieces, which can be reassembled (by rotation and transportation) to form two disjoint copies of that sphere.*

A couple of further algebraic tricks are needed to obtain the full Banach-Tarski Theorem (which concerns not just the sphere's surface, but its interior too). Frankly, however, this is just icing on the algebraic cake. Hence Weston writes:

> [...] the result on free groups is the *key step* in the proof of the Banach-Tarski paradox. From this point of view, the Banach-Tarski paradox is not a statement about $\mathbb{R}^3$ so much as it is a statement about the complexity of the group [of translations and rotations in $\mathbb{R}^3$]. (Weston, 2003, p. 16)

That is: whether we can offer a *finite* decomposition (as in Banach–Tarski) or a *countably infinite* decomposition (as in Vitali's Paradox) comes down to certain group-theoretic facts about working in two-dimension or three-dimensions.

Admittedly, this last observation slightly spoils the joke at the end of section 16.7. Since it is two dimensional, "Banach-Tarski" must be divided into a countable *infinity* of pieces, if one wants to rearrange those pieces to form "Banach-Tarski Banach-Tarski". To repair the joke, one must write in three dimensions. We leave this as an exercise for the reader.

One final comment. In section 16.7, we mentioned that the "pieces" of the sphere one obtains cannot be *measurable*, but must be unpicturable "infinite scatterings". The same is true of our use of Choice in obtaining Lemma 16.20. And this is all worth explaining.

Again, we must sketch some background (but this is *just* a sketch; you may want to consult a textbook entry on *measure*).

To define a measure for a set X is to assign a value $\mu(E) \in \mathbb{R}$ for each E in some "σ-algebra" on X. Details here are not essential, except that the function μ must obey the principle of countable additivity: the measure of a countable union of disjoint sets is the sum of their individual measures, i.e., $\mu(\bigcup_{n<\omega} X_n) = \sum_{n<\omega} \mu(X_n)$ whenever the X_ns are disjoint. To say that a set is "unmeasurable" is to say that no measure can be suitably assigned. Now, using our R from before:

Corollary 16.22 (Vitali). *Let μ be a measure such that $\mu(\mathbf{S}) = 1$, and such that $\mu(X) = \mu(Y)$ if X and Y are congruent. Then $\rho[C]$ is unmeasurable for all $\rho \in R$.*

Proof. For reductio, suppose otherwise. So let $\mu(\sigma[C]) = r$ for some $\sigma \in R$ and some $r \in \mathbb{R}$. For any $\rho \in C$, $\rho[C]$ and $\sigma[C]$ are congruent, and hence $\mu(\rho[C]) = r$ for any $\rho \in C$. By Lemma 16.18 and Lemma 16.19, $\mathbf{S} = \bigcup_{\rho \in R} \rho[C]$ is a countable union of pairwise disjoint sets. So countable additivity dictates that $\mu(\mathbf{S}) = 1$ is the sum of the measures of each $\rho[C]$, i.e.,

$$1 = \mu(\mathbf{S}) = \sum_{\rho \in R} \mu(\rho[C]) = \sum_{\rho \in R} r$$

But if $r = 0$ then $\sum_{\rho \in R} r = 0$, and if $r > 0$ then $\sum_{\rho \in R} r = \infty$. □

Problems

Problem 16.1. Prove Theorem 16.12. If you struggle, you can find a proof in (Potter, 2004, pp. 242–3).

APPENDIX A

Biographies

A.1 Georg Cantor

Fig. A.1: Georg Cantor

An early biography of Georg Cantor (GAY-org KAHN-tor) claimed that he was born and found on a ship that was sailing for Saint Petersburg, Russia, and that his parents were unknown. This, however, is not true; although he was born in Saint Petersburg in 1845.

Cantor received his doctorate in mathematics at the University of Berlin in 1867. He is known for his work in set theory, and is credited with founding set theory as a distinctive research discipline. He was the first to prove that there are infinite sets of different sizes. His theories, and especially his theory of infinities, caused much debate among mathematicians at the time, and his work was controversial.

Cantor's religious beliefs and his mathematical work were in-

extricably tied; he even claimed that the theory of transfinite numbers had been communicated to him directly by God. In later life, Cantor suffered from mental illness. Beginning in 1894, and more frequently towards his later years, Cantor was hospitalized. The heavy criticism of his work, including a falling out with the mathematician Leopold Kronecker, led to depression and a lack of interest in mathematics. During depressive episodes, Cantor would turn to philosophy and literature, and even published a theory that Francis Bacon was the author of Shakespeare's plays.

Cantor died on January 6, 1918, in a sanatorium in Halle.

Further Reading For full biographies of Cantor, see Dauben (1990) and Grattan-Guinness (1971). Cantor's radical views are also described in the BBC Radio 4 program *A Brief History of Mathematics* (du Sautoy, 2014). If you'd like to hear about Cantor's theories in rap form, see Rose (2012).

A.2 Kurt Gödel

Kurt Gödel (GER-dle) was born on April 28, 1906 in Brünn in the Austro-Hungarian empire (now Brno in the Czech Republic). Due to his inquisitive and bright nature, young Kurtele was often called "Der kleine Herr Warum" (Little Mr. Why) by his family. He excelled in academics from primary school onward, where he got less than the highest grade only in mathematics. Gödel was often absent from school due to poor health and was exempt from physical education. He was diagnosed with rheumatic fever during his childhood. Throughout his life, he believed this permanently affected his heart despite medical assessment saying otherwise.

Gödel began studying at the University of Vienna in 1924 and completed his doctoral studies in 1929. He first intended to study physics, but his interests soon moved to mathematics and especially logic, in part due to the influence of the philosopher Rudolf Carnap. His dissertation, written under the supervision

of Hans Hahn, proved the completeness theorem of first-order predicate logic with identity (Gödel, 1929). Only a year later, he obtained his most famous results—the first and second incompleteness theorems (published in Gödel 1931). During his time in Vienna, Gödel was heavily involved with the Vienna Circle, a group of scientifically-minded philosophers that included Carnap, whose work was especially influenced by Gödel's results.

In 1938, Gödel married Adele Nimbursky. His parents were not pleased: not only was she six years older than him and already divorced, but she worked as a dancer in a nightclub. Social pressures did not affect Gödel, however, and they remained happily married until his death.

Fig. A.2: Kurt Gödel

After Nazi Germany annexed Austria in 1938, Gödel and Adele emigrated to the United States, where he took up a position at the Institute for Advanced Study in Princeton, New Jersey. Despite his introversion and eccentric nature, Gödel's time at Princeton was collaborative and fruitful. He published essays in set theory, philosophy and physics. Notably, he struck up a particularly strong friendship with his colleague at the IAS, Albert Einstein.

In his later years, Gödel's mental health deteriorated. His wife's hospitalization in 1977 meant she was no longer able to cook his meals for him. Having suffered from mental health issues throughout his life, he succumbed to paranoia. Deathly afraid of being poisoned, Gödel refused to eat. He died of starvation on January 14, 1978, in Princeton.

Further Reading For a complete biography of Gödel's life is available, see John Dawson (1997). For further biographical pieces, as well as essays about Gödel's contributions to logic and philosophy, see Wang (1990), Baaz et al. (2011), Takeuti et al. (2003), and Sigmund et al. (2007).

Gödel's PhD thesis is available in the original German (Gödel, 1929). The original text of the incompleteness theorems is (Gödel, 1931). All of Gödel's published and unpublished writings, as well as a selection of correspondence, are available in English in his *Collected Papers* Feferman et al. (1986, 1990).

For a detailed treatment of Gödel's incompleteness theorems, see Smith (2013). For an informal, philosophical discussion of Gödel's theorems, see Mark Linsenmayer's podcast (Linsenmayer, 2014).

A.3 Bertrand Russell

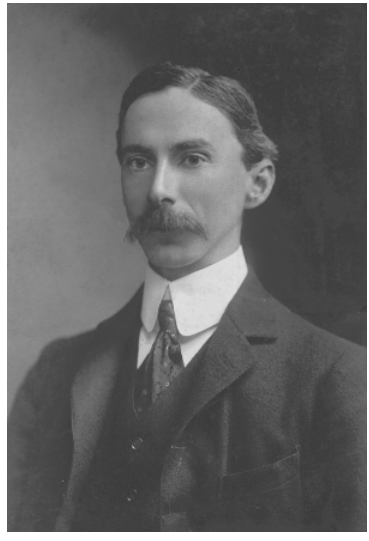


Fig. A.3: Bertrand Russell

Bertrand Russell is hailed as one of the founders of modern analytic philosophy. Born May 18, 1872, Russell was not only known for his work in philosophy and logic, but wrote many popular books in various subject areas. He was also an ardent political activist throughout his life.

Russell was born in Trellech, Monmouthshire, Wales. His parents were members of the British nobility. They were free-thinkers, and even made friends with the radicals in Boston at the time. Unfortunately, Russell's parents died when he was young, and Russell

was sent to live with his grandparents. There, he was given a religious upbringing (something his parents had wanted to avoid at all costs). His grandmother was very strict in all matters of morality. During adolescence he was mostly homeschooled by private tutors.

Russell's influence in analytic philosophy, and especially logic, is tremendous. He studied mathematics and philosophy at Trinity College, Cambridge, where he was influenced by the mathematician and philosopher Alfred North Whitehead. In 1910, Russell and Whitehead published the first volume of *Principia Mathematica*, where they championed the view that mathematics is reducible to logic. He went on to publish hundreds of books, essays and political pamphlets. In 1950, he won the Nobel Prize for literature.

Russell's was deeply entrenched in politics and social activism. During World War I he was arrested and sent to prison for six months due to pacifist activities and protest. While in prison, he was able to write and read, and claims to have found the experience "quite agreeable." He remained a pacifist throughout his life, and was again incarcerated for attending a nuclear disarmament rally in 1961. He also survived a plane crash in 1948, where the only survivors were those sitting in the smoking section. As such, Russell claimed that he owed his life to smoking. Russell was married four times, but had a reputation for carrying on extra-marital affairs. He died on February 2, 1970 at the age of 97 in Penrhyndeudraeth, Wales.

Further Reading Russell wrote an autobiography in three parts, spanning his life from 1872–1967 (Russell, 1967, 1968, 1969). The Bertrand Russell Research Centre at McMaster University is home of the Bertrand Russell archives. See their website at Duncan (2015), for information on the volumes of his collected works (including searchable indexes), and archival projects. Russell's paper *On Denoting* (Russell, 1905) is a classic of 20th century analytic philosophy.

The Stanford Encyclopedia of Philosophy entry on Russell (Irvine, 2015) has sound clips of Russell speaking on Desire and Political theory. Many video interviews with Russell are available online. To see him talk about smoking and being involved in a plane crash, e.g., see Russell (n.d.). Some of Russell's works, including his *Introduction to Mathematical Philosophy* are available as free audiobooks on LibriVox (n.d.).

A.4 Alfred Tarski

Fig. A.4: Alfred Tarski

Alfred Tarski was born on January 14, 1901 in Warsaw, Poland (then part of the Russian Empire). Described as "Napoleonic," Tarski was boisterous, talkative, and intense. His energy was often reflected in his lectures—he once set fire to a wastebasket while disposing of a cigarette during a lecture, and was forbidden from lecturing in that building again.

Tarski had a thirst for knowledge from a young age. Although later in life he would tell students that he studied logic because it was the only class in which he got a B, his high school records show that he got A's across the board—even in logic. He studied at the University of Warsaw from 1918 to 1924. Tarski first intended to study biology, but became interested in mathematics, philosophy, and logic, as the university was the center of the Warsaw School of Logic and Philosophy. Tarski earned his doctorate in 1924 under the supervision of Stanisław Leśniewski.

Before emigrating to the United States in 1939, Tarski completed some of his most important work while working as a secondary school teacher in Warsaw. His work on logical consequence and logical truth were written during this time. In 1939, Tarski was visiting the United States for a lecture tour. During his visit, Germany invaded Poland, and because of his Jewish heritage, Tarski could not return. His wife and children remained in Poland until the end of the war, but were then able to emigrate to the United States as well. Tarski taught at Harvard, the College of the City of New York, and the Institute for Advanced Study at Princeton, and finally the University of California, Berkeley. There he founded the multidisciplinary program in Logic and the Methodology of Science. Tarski died on October 26, 1983 at the age of 82.

Further Reading For more on Tarski's life, see the biography *Alfred Tarski: Life and Logic* (Feferman and Feferman, 2004). Tarski's seminal works on logical consequence and truth are available in English in (Corcoran, 1983). All of Tarski's original works have been collected into a four volume series, (Tarski, 1981).

A.5 Ernst Zermelo

Ernst Zermelo was born on July 27, 1871 in Berlin, Germany. He had five sisters, though his family suffered from poor health and only three survived to adulthood. His parents also passed away when he was young, leaving him and his siblings orphans when he was seventeen. Zermelo had a deep interest in the arts, and especially in poetry. He was known for being sharp, witty, and critical. His most celebrated mathematical achievements include the introduction of the axiom of choice (in 1904), and his axiomatization of set theory (in 1908).

Zermelo's interests at university were varied. He took courses in physics, mathematics, and philosophy. Under the supervision of Hermann Schwarz, Zermelo completed his dissertation *Inves-*

tigations in the Calculus of Variations in 1894 at the University of Berlin. In 1897, he decided to pursue more studies at the University of Göttigen, where he was heavily influenced by the foundational work of David Hilbert. In 1899 he became eligible for professorship, but did not get one until eleven years later—possibly due to his strange demeanour and "nervous haste."

Zermelo finally received a paid professorship at the University of Zurich in 1910, but was forced to retire in 1916 due to tuberculosis. After his recovery, he was given an honourary professorship at the University of Freiburg in 1921. During this time he worked on foundational mathematics. He became irritated with the works of Thoralf Skolem and Kurt Gödel, and publicly criticized their approaches in his papers. He was dismissed from his position at Freiburg in 1935, due to his unpopularity and his opposition to Hitler's rise to power in Germany.

Fig. A.5: Ernst Zermelo

The later years of Zermelo's life were marked by isolation. After his dismissal in 1935, he abandoned mathematics. He moved to the country where he lived modestly. He married in 1944, and became completely dependent on his wife as he was going blind. Zermelo lost his sight completely by 1951. He passed away in Günterstal, Germany, on May 21, 1953.

Further Reading For a full biography of Zermelo, see Ebbinghaus (2015). Zermelo's seminal 1904 and 1908 papers are available to read in the original German (Zermelo, 1904, 1908b). Zermelo's collected works, including his writing on physics, are avail-

able in English translation in (Ebbinghaus et al., 2010; Ebbinghaus and Kanamori, 2013).

Photo Credits

Georg Cantor, p. 230: Portrait of Georg Cantor by Otto Zeth courtesy of the Universitätsarchiv, Martin-Luther Universität Halle–Wittenberg. UAHW Rep. 40-VI, Nr. 3 Bild 102.

Kurt Gödel, p. 232: Portrait of Kurt Gödel, ca. 1925, photographer unknown. From the Shelby White and Leon Levy Archives Center, Institute for Advanced Study, Princeton, NJ, USA, on deposit at Princeton University Library, Manuscript Division, Department of Rare Books and Special Collections, Kurt Gödel Papers, (C0282), Box 14b, #110000. The Open Logic Project has obtained permission from the Institute's Archives Center to use this image for inclusion in non-commercial OLP-derived materials. Permission from the Archives Center is required for any other use.

Bertrand Russell, p. 233: Portrait of Bertrand Russell, ca. 1907, courtesy of the William Ready Division of Archives and Research Collections, McMaster University Library. Bertrand Russell Archives, Box 2, f. 4.

Alfred Tarski, p. 235: Passport photo of Alfred Tarski, 1939. Cropped and restored from a scan of Tarski's passport by Joel Fuller. Original courtesy of Bancroft Library, University of California, Berkeley. Alfred Tarski Papers, Banc MSS 84/49. The Open Logic Project has obtained permission to use this image for inclusion in non-commercial OLP-derived materials. Permission from Bancroft Library is required for any other use.

Ernst Zermelo, p. 237: Portrait of Ernst Zermelo, ca. 1922,

courtesy of the Abteilung für Handschriften und Seltene Drucke, Niedersächsische Staats- und Universitätsbibliothek Göttingen, Cod. Ms. D. Hilbert 754, Bl. 6 Nr. 25.

Bibliography

Baaz, Matthias, Christos H. Papadimitriou, Hilary W. Putnam, Dana S. Scott, and Charles L. Harper Jr. 2011. *Kurt Gödel and the Foundations of Mathematics: Horizons of Truth.* Cambridge: Cambridge University Press.

Banach, Stefan and Alfred Tarski. 1924. Sur la décomposition des ensembles de points en parties respectivement congruentes. *Fundamenta Mathematicae* 6: 244–77.

Benacerraf, Paul. 1965. What numbers could not be. *The Philosophical Review* 74(1): 47–73.

Berkeley, George. 1734. *The Analyst; or, a Discourse Adressed to an Infidel Mathematician.*

Boolos, George. 1971. The iterative conception of set. *The Journal of Philosophy* 68(8): 215–31.

Boolos, George. 1989. Iteration again. *Philosophical Topics* 17(2): 5–21.

Boolos, George. 2000. Must we believe in set theory? In *Between Logic and Intuition: Essays in Honor of Charles Parsons*, eds. Gila Sher and Richard Tieszen, 257–68. Cambridge: Cambridge University Press.

Burali-Forti, Cesare. 1897. Una questione sui numeri transfiniti. *Rendiconti del Circolo Matematico di Palermo* 11: 154–64.

Button, Tim. 2021. Level theory, part 1: Axiomatizing the bare idea of a cumulative hierarchy of sets. *The Bulletin of Symbolic Logic* 27(4): 436–460.

Cantor, Georg. 1878. Ein Beitrag zur Mannigfaltigkeitslehre. *Journal für die reine und angewandte Mathematik* 84: 242–58.

Cantor, Georg. 1883. *Grundlagen einer allgemeinen Mannigfaltigkeitslehre. Ein mathematisch-philosophischer Versuch in der Lehre des Unendlichen.* Leipzig: Teubner.

Cantor, Georg. 1892. Über eine elementare Frage der Mannigfaltigkeitslehre. *Jahresbericht der deutschen Mathematiker-Vereinigung* 1: 75–8.

Cohen, Paul J. 1963. The independence of the continuum hypothesis. *Proceedings of the National Academy of Sciences of the United States of America* 24: 556–557.

Cohen, Paul J. 1966. *Set Theory and the Continuum Hypothesis.* Reading, MA: Benjamin.

Conway, John. 2006. The power of mathematics. In *Power*, eds. Alan Blackwell and David MacKay, Darwin College Lectures. Cambridge: Cambridge University Press. URL http://www.cs.toronto.edu/~mackay/conway.pdf.

Corcoran, John. 1983. *Logic, Semantics, Metamathematics.* Indianapolis: Hackett, 2nd ed.

Dauben, Joseph. 1990. *Georg Cantor: His Mathematics and Philosophy of the Infinite.* Princeton: Princeton University Press.

Dedekind, Richard. 1888. *Was sind und was sollen die Zahlen?* Braunschweig: Vieweg.

du Sautoy, Marcus. 2014. A brief history of mathematics: Georg Cantor. URL http://www.bbc.co.uk/programmes/b00ss1j0. Audio Recording.

Duncan, Arlene. 2015. The Bertrand Russell Research Centre. URL http://russell.mcmaster.ca/.

Ebbinghaus, Heinz-Dieter. 2015. *Ernst Zermelo: An Approach to his Life and Work*. Berlin: Springer-Verlag.

Ebbinghaus, Heinz-Dieter, Craig G. Fraser, and Akihiro Kanamori. 2010. *Ernst Zermelo. Collected Works*, vol. 1. Berlin: Springer-Verlag.

Ebbinghaus, Heinz-Dieter and Akihiro Kanamori. 2013. *Ernst Zermelo: Collected Works*, vol. 2. Berlin: Springer-Verlag.

Feferman, Anita and Solomon Feferman. 2004. *Alfred Tarski: Life and Logic*. Cambridge: Cambridge University Press.

Feferman, Solomon, John W. Dawson Jr., Stephen C. Kleene, Gregory H. Moore, Robert M. Solovay, and Jean van Heijenoort. 1986. *Kurt Gödel: Collected Works. Vol. 1: Publications 1929–1936*. Oxford: Oxford University Press.

Feferman, Solomon, John W. Dawson Jr., Stephen C. Kleene, Gregory H. Moore, Robert M. Solovay, and Jean van Heijenoort. 1990. *Kurt Gödel: Collected Works. Vol. 2: Publications 1938–1974*. Oxford: Oxford University Press.

Feferman, Solomon and Azriel Levy. 1963. Independence results in set theory by Cohen's method II. *Notices of the American Mathematical Society* 10: 593.

Fraenkel, Abraham. 1922. Über den Begriff 'definit' und die Unabhängigkeit des Auswahlaxioms. *Sitzungsberichte der Preussischen Akadademie der Wissenschaften, Physikalisch-mathematische Klasse* 253–257.

Frege, Gottlob. 1884. *Die Grundlagen der Arithmetik: Eine logisch mathematische Untersuchung über den Begriff der Zahl*. Breslau: Wilhelm Koebner. Translation in Frege (1953).

Frege, Gottlob. 1953. *Foundations of Arithmetic*, ed. J. L. Austin. Oxford: Basil Blackwell & Mott, 2nd ed.

Giaquinto, Marcus. 2007. *Visual Thinking in Mathematics*. Oxford: Oxford University Press.

Gödel, Kurt. 1929. Über die Vollständigkeit des Logikkalküls [On the completeness of the calculus of logic]. Dissertation, Universität Wien. Reprinted and translated in Feferman et al. (1986), pp. 60–101.

Gödel, Kurt. 1931. über formal unentscheidbare Sätze der *Principia Mathematica* und verwandter Systeme I [On formally undecidable propositions of *Principia Mathematica* and related systems I]. *Monatshefte für Mathematik und Physik* 38: 173–198. Reprinted and translated in Feferman et al. (1986), pp. 144–195.

Gödel, Kurt. 1938. The consistency of the axiom of choice and the generalized continuum hypothesis. *Proceedings of the National Academy of Sciences of the United States of America* 50: 1143–8.

Gouvêa, Fernando Q. 2011. Was Cantor surprised? *American Mathematical Monthly* 118(3): 198–209.

Grattan-Guinness, Ivor. 1971. Towards a biography of Georg Cantor. *Annals of Science* 27(4): 345–391.

Hartogs, Friedrich. 1915. Über das Problem der Wohlordnung. *Mathematische Annalen* 76: 438–43.

Hausdorff, Felix. 1914. Bemerkung über den Inhalt von Punktmengen. *Mathematische Annalen* 75: 428–34.

Heck, Richard Kimberly. 2012. *Reading Frege's Grundgesetze*. Oxford: Oxford University Press.

Heijenoort, Jean van. 1967. *From Frege to Gödel: A Source Book in Mathematical Logic, 1879–1931.* Cambridge, MA: Harvard University Press.

Hilbert, David. 1891. Über die stetige Abbildung einer Linie auf ein Flächenstück. *Mathematische Annalen* 38(3): 459–460.

Hilbert, David. 2013. *David Hilbert's Lectures on the Foundations of Arithmetic and Logic 1917–1933*, eds. William Bragg Ewald and Wilfried Sieg. Heidelberg: Springer.

Hume, David. 1740. *A Treatise of Human Nature.* London.

Incurvati, Luca. 2020. *Conceptions of Set and the Foundations of Mathematics.* Cambridge: Cambridge University Press.

Irvine, Andrew David. 2015. Sound clips of Bertrand Russell speaking. URL http://plato.stanford.edu/entries/russell/russell-soundclips.html.

John Dawson, Jr. 1997. *Logical Dilemmas: The Life and Work of Kurt Gödel.* Boca Raton: CRC Press.

Katz, Karin Usadi and Mikhail G. Katz. 2012. Stevin numbers and reality. *Foundations of Science* 17(2): 109–23.

Kunen, Kenneth. 1980. *Set Theory: An Introduction to Independence Proofs.* New York: North Holland.

Lévy, Azriel. 1960. Axiom schemata of strong infinity in axiomatic set theory. *Pacific Journal of Mathematics* 10(1): 223–38.

LibriVox. n.d. Bertrand Russell. URL https://librivox.org/author/1508?primary_key=1508&search_category=author&search_page=1&search_form=get_results. Collection of public domain audiobooks.

Linnebo, Øystein. 2010. Predicative and impredicative definitions. *Internet Encyclopedia of Philosophy* URL http://www.iep.utm.edu/predicat/.

Linsenmayer, Mark. 2014. The partially examined life: Gödel on math. URL http://www.partiallyexaminedlife.com/2014/06/16/ep95-godel/. Podcast audio.

Maddy, Penelope. 1988a. Believing the axioms I. *The Journal of Symbolic Logic* 53(2): 481–511.

Maddy, Penelope. 1988b. Believing the axioms II. *The Journal of Symbolic Logic* 53(3): 736–64.

Montague, Richard. 1961. Semantic closure and non-finite axiomatizability I. In *Infinitistic Methods: Proceedings of the Symposium on Foundations of Mathematics (Warsaw 1959)*, 45–69. New York: Pergamon.

Montague, Richard. 1965. Set theory and higher-order logic. In *Formal systems and recursive functions*, eds. John Crossley and Michael Dummett, 131–48. Amsterdam: North-Holland. Proceedings of the Eight Logic Colloquium, July 1963.

O'Connor, John J. and Edmund F. Robertson. 2005. The real numbers: Stevin to Hilbert URL http://www-history.mcs.st-and.ac.uk/HistTopics/Real_numbers_2.html.

Peano, Giuseppe. 1890. Sur une courbe, qui remplit toute une aire plane. *Mathematische Annalen* 36(1): 157–60.

Potter, Michael. 2004. *Set Theory and its Philosophy*. Oxford: Oxford University Press.

Ramsey, Frank Plumpton. 1925. The foundations of mathematics. *Proceedings of the London Mathematical Society* 25: 338–384.

Robinson, Raphael. 1947. On the decomposition of spheres. *Fundamenta Mathematicae* 34(1): 246–60.

Rose, Daniel. 2012. A song about Georg Cantor. URL https://www.youtube.com/watch?v=QUP5Z4Fb5k4. Audio Recording.

Rose, Nicholas J. 2010. Hilbert-type space-filling curves. URL https://web.archive.org/web/20151010184939/http://www4.ncsu.edu/~njrose/pdfFiles/HilbertCurve.pdf.

Russell, Bertrand. 1905. On denoting. *Mind* 14: 479–493.

Russell, Bertrand. 1919. *Introduction to Mathematical Philosophy*. London: Allen & Unwin.

Russell, Bertrand. 1967. *The Autobiography of Bertrand Russell*, vol. 1. London: Allen and Unwin.

Russell, Bertrand. 1968. *The Autobiography of Bertrand Russell*, vol. 2. London: Allen and Unwin.

Russell, Bertrand. 1969. *The Autobiography of Bertrand Russell*, vol. 3. London: Allen and Unwin.

Russell, Bertrand. n.d. Bertrand Russell on smoking. URL https://www.youtube.com/watch?v=80oLTiVW_lc. Video Interview.

Scott, Dana. 1974. Axiomatizing set theory. In *Axiomatic Set Theory II*, ed. Thomas Jech, 207–14. American Mathematical Society. Proceedings of the Symposium in Pure Mathematics of the American Mathematical Society, July–August 1967.

Shoenfield, Joseph R. 1977. Axioms of set theory. In *Handbook of Mathematical Logic*, ed. Jon Barwise, 321–44. London: North-Holland.

Sigmund, Karl, John Dawson, Kurt Mühlberger, Hans Magnus Enzensberger, and Juliette Kennedy. 2007. Kurt Gödel: Das Album–The Album. *The Mathematical Intelligencer* 29(3): 73–76.

Skolem, Thoralf. 1922. Einige Bemerkungen zur axiomatischen Begründung der Mengenlehre. In *Wissenschaftliche Vorträge*

gehalten auf dem fünften Kongress der skandanivschen Mathematiker in Helsingfors vom 4. bis zum 7. Juli 1922, 137–52. Akademiska Bokhandeln.

Smith, Peter. 2013. *An Introduction to Gödel's Theorems.* Cambridge: Cambridge University Press.

Takeuti, Gaisi, Nicholas Passell, and Mariko Yasugi. 2003. *Memoirs of a Proof Theorist: Gödel and Other Logicians.* Singapore: World Scientific.

Tarski, Alfred. 1981. *The Collected Works of Alfred Tarski*, vol. I–IV. Basel: Birkhäuser.

Tomkowicz, Grzegorz and Stan Wagon. 2016. *The Banach-Tarski Paradox.* Cambridge: Cambridge University Press.

Vitali, Giuseppe. 1905. *Sul problema della misura dei gruppi di punti di una retta.* Bologna: Gamberini e Parmeggiani.

von Neumann, John. 1925. Eine Axiomatisierung der Mengenlehre. *Journal für die reine und angewandte Mathematik* 154: 219–40.

Wang, Hao. 1990. *Reflections on Kurt Gödel.* Cambridge: MIT Press.

Weston, Tom. 2003. The Banach-Tarski paradox URL `http://people.math.umass.edu/~weston/oldpapers/banach.pdf`.

Whitehead, Alfred North and Bertrand Russell. 1910. *Principia Mathematica*, vol. 1. Cambridge: Cambridge University Press.

Zermelo, Ernst. 1904. Beweis, daß jede Menge wohlgeordnet werden kann. *Mathematische Annalen* 59: 514–516. English translation in (Ebbinghaus et al., 2010, pp. 115–119).

Zermelo, Ernst. 1908a. Untersuchungen über die Grundlagen der Mengenlehre I. *Mathematische Annalen* 65: 261–81.

Zermelo, Ernst. 1908b. Untersuchungen über die Grundlagen der Mengenlehre I. *Mathematische Annalen* 65(2): 261–281. English translation in (Ebbinghaus et al., 2010, pp. 189-229).

About the Open Logic Project

The *Open Logic Text* is an open-source, collaborative textbook of formal meta-logic and formal methods, starting at an intermediate level (i.e., after an introductory formal logic course). Though aimed at a non-mathematical audience (in particular, students of philosophy and computer science), it is rigorous.

Coverage of some topics currently included may not yet be complete, and many sections still require substantial revision. We plan to expand the text to cover more topics in the future. We also plan to add features to the text, such as a glossary, a list of further reading, historical notes, pictures, better explanations, sections explaining the relevance of results to philosophy, computer science, and mathematics, and more problems and examples. If you find an error, or have a suggestion, please let the project team know.

The project operates in the spirit of open source. Not only is the text freely available, we provide the LaTeX source under the Creative Commons Attribution license, which gives anyone the right to download, use, modify, re-arrange, convert, and re-distribute our work, as long as they give appropriate credit. Please see the Open Logic Project website at openlogicproject.org for additional information.