Inquiry-based notes on Zermelo-Fraenkel set theory

Kenny Easwaran

October 28, 2025

Many students and instructors of mathematics find it useful to learn a subject through working exercises rather than being lectured. Students are given some definitions and axioms, but the instructor does not present any proofs of theorems. Instead, students are given a series of problems set by the instructor, and they are expected to discover the proofs of all results themselves. Students explain their answers to each other until the others are satisfied. More work is thus required on the students's part, and progress through the material often starts out slower, but the hope is that students will have deeper mastery of the material by the end.

Versions of this method have often been called the "Moore Method", after Robert Lee Moore, a topologist who championed it for several decades at the University of Texas. Moore was prominent in the history of American mathematics, with himself and five of his PhD students all becoming president of the Mathematics Association of America at various points in the middle of the 20th century. Moore himself insisted that students work individually on their answers before presenting them, while other instructors have found that it is beneficial for students to work collaboratively in coming up with their proofs. Moore was also an arch-segregationist, as one might expect for someone named after a Confederate general. Even though the University of Texas was required by court order to allow Black students during the last 20 years of his time there, he never allowed Black people in his classrooms. (However his first PhD student, John Kline, became the supervisor of the second and third Black people to earn PhDs in mathematics in the United States.) Moore's influence as a mathematician and educator was great enough that the University of Texas named its largest classroom building after him, but in July, 2020, the university decided that it was inappropriate to give such an honor to a racist like him. Similarly, although this influential style of math education was for many decades named after him, it seems appropriate to change the name, particularly if there is less emphasis on the individual and competitive aspects of it that he pursued.

These notes are a series of problems that will help students learn the basics of set theory, starting with the Zermelo-Fraenkel axioms (the most common foundation for mathematics as a whole), and leading to the basics of ordinal and cardinal arithmetic of the infinite.

1 Terminology and Logical Background

1.1 Formal logic and notation

For any sentence φ, $\neg\varphi$ will mean 'it is not the case that φ'.

'If φ then ψ' means the same thing as $\neg(\varphi$ and $\neg\psi)$, which is the same as $(\neg\varphi$ or $\psi)$. This is not the same as the ordinary usage of the word 'if'. 'Iff' is an abbreviation for 'if and only if', so something like 'φ iff ψ' will mean the same as $((\varphi$ and $\psi)$ or $(\neg\varphi$ and $\neg\psi))$, or $((\text{if } \varphi \text{ then } \psi)$ and $(\text{if } \psi \text{ then } \varphi))$ or $((\text{if } \varphi$ then $\psi)$ and $(\text{if } \neg\varphi \text{ then } \neg\psi))$. These notions will be abbreviated as $\varphi \to \psi$ and $\varphi \leftrightarrow \psi$ respectively.

I will also write 'if φ then ψ' as $\varphi \to \psi$ and 'φ iff ψ' as $\varphi \leftrightarrow \psi$.

$\forall x\varphi(x)$ means that $\varphi(t)$ is true for any t. For instance, if I were talking about natural numbers and asserted $\forall x(x \geq 0)$, then I would be able to assert $5 \geq 0$ or $0 \geq 0$. In addition, if you can prove $\varphi(x)$ without any assumptions about x, then you can conclude $\forall x\varphi(x)$.

Similarly, $\exists x\varphi(x)$ means that there exists an t for which $\varphi(t)$ is true. If at any point you can derive $\varphi(t)$ for some t, then you can derive $\exists x\varphi(x)$ as a consequence. It's useful to note that $\neg\exists x\neg\varphi$ is the same as $\forall x\varphi$ and similarly that $\neg\forall x\neg\varphi$ is the same as $\exists x\varphi$.

In addition, if you can prove $\forall x\varphi$ and $\forall x\psi$, then you can prove $\forall x(\varphi$ and $\psi)$ and vice versa. Note that this is not true for 'or', but 'or' does have this property with $\exists$.

'=' is a notation mean that the notations on either side of it actually stand for the same object. That means that if we know '$x = y$' then we can substitute 'x' for 'y' in any context where either term appears.

Ordinary parentheses like) and (are used only for grouping symbols in formulas; curly braces like } and { enclose the elements of a set (introduced at Axiom 2); $\rangle$ and $\langle$ enclose the elements of an ordered pair (introduced in Exercise 2.9).

If a problem is just a mathematical statement, then the exercise is to prove that statement. All proofs should be expressed clearly and in full sentences. Some exercises list a specific set of axioms to use — for these, one should not need to appeal to any other axioms, though one will need to use some general logical reasoning. For any exercise that doesn't specifically mention any axioms, one can use all the axioms that have been mentioned so far.

A few exercises are marked with a '*'. These are quite a bit more complicated, and can be skipped without any interruption of the flow, but are historically significant. When I last checked, the only such exercises were 2.29, 4.21and 5.65.

Tautologies (like '$\mathcal{P}(x) \in y$ or $\neg(\mathcal{P}(x) \in y)$') and logical truths (like '$\forall x\forall y((x \in y) \to (x \in y))$') can be introduced at any time without comment.

From here on, all variables and other terms will refer to sets, except 'φ' and 'ψ', which will refer to sentences or formulas of the formal language.

1.2 Membership and subsets

'$x \in y$' is read as saying that x is an *element* (or *member*) of y. The only actual definition of this notion comes from the axioms.

'$x \subseteq y$' means '$\forall z(z \in x \to z \in y)$'. In words, x is a *subset* of y iff every element of x is also an element of y.

Exercise 1.1 $x \subseteq x$.

Exercise 1.2 (Transitivity of $\subseteq$) If $x \subseteq y$ and $y \subseteq z$ then $x \subseteq z$.

2 Zermelo-Fraenkel Axioms

The standard Zermelo-Fraenkel axioms will be introduced over several subsections.

2.1 Basic set identity: the Axiom of Extensionality

First, let us make the definition $\forall x \forall y(x \approx y \leftrightarrow \forall z(z \in x \leftrightarrow z \in y))$. This definition says that $x \approx y$ iff they have the same elements.

Exercise 2.1 $x \approx y$ iff ($x \subseteq y$ and $y \subseteq x$).

Now, we will show that $\approx$ satisfies the usual properties that we're used to for equality, which make it an *equivalence relation*.

Exercise 2.2 $x \approx x$. (Reflexivity)

Exercise 2.3 If $x \approx y$ then $y \approx x$. (Symmetry)

Exercise 2.4 If $x \approx y$ and $y \approx z$ then $x \approx z$. (Transitivity)

Axiom 1 Extensionality

$$\forall x \forall y(x \approx y \leftrightarrow x = y)$$

That is, if x and y have the same elements, then they are in fact the same set, so they can be substituted for one another in any locations where they appear. Because we already know that $\approx$ is an equivalence relation, we see why this axiom might be plausible. This axiom gives set theory its 'extensional' quality, meaning that we don't care how you name a set or where you got it from, just what elements it has. This allows most of the future axioms, where we define a set by saying what its elements are — this axiom tells us that that is all you need to identify it uniquely.

Among other things, as we will see later, the Extensionality axiom also means that the only entities we talk about are sets, and all the members of sets are themselves sets. This means we are talking about a relatively "pure" set theory, rather than really talking about sets of numbers or sets of points or any of the other more familiar mathematical concepts. However, it will turn out that within this universe of pure sets, we are able to construct entities that behave like all of these mathematical concepts, so that this isn't really a limitation.

2.2 Construction of sets from other sets: the Axioms of Pairing, Union, and Power Set

The axioms we will introduce in this sub-section all play the role of constructing new sets out of sets that we are given. That is, they all take the form of saying that for all x there exists a y that bears some relation to the initial x.

Because of Axiom 1, Extensionality, we are able to define what y is by saying exactly which things are its elements. That is, the inner statement of the relation y bears to x will take the form $\forall z(z \in y \leftrightarrow \quad)$, where the blank is filled in with the condition any z needs to satisfy to be a member of y.

Axiom 2 Pairing

$$\forall x_1 \forall x_2 \exists y(\forall z(z \in y \leftrightarrow (z = x_1 \text{ or } z = x_2)))$$

In this first example, the condition to be a member of y just is being one of the two sets we started with. We can explain this axiom in words by saying that for any x_1 and x_2, there is a y that has both x_1 and x_2 as members, and nothing else. We will give this set the name '$\{x_1, x_2\}$'. Such a set is called an *unordered pair*.

Exercise 2.5 For any x, there is a set $\{x\}$ whose unique element is x. (Hint: in Axiom 2, there is no rule that says that x_1 and x_2 have to be different objects.)

Such a set is called a *singleton*. (We occasionally refer to an unordered pair as a 'doubleton', though this is less common.)

Exercise 2.6 $\{x\} = \{y\}$ iff $x = y$.

Exercise 2.7 $\{x, y\} = \{y, x\}$.

Exercise 2.8 $\{w, x\} = \{y, z\}$ iff ($x = y$ and $z = w$) or ($x = z$ and $y = w$).

These two results justify the name 'unordered pair', by saying that it doesn't matter in what order we list the two elements.

In many other branches of mathematics, we often want to talk about *ordered* pairs $\langle x_1, x_2 \rangle$[1], with the property that $\langle x_1, x_2 \rangle = \langle y_1, y_2 \rangle$ iff $x_1 = y_1$ and $x_2 = y_2$. For instance, in Descartes' algebraic representation of geometry, the pairs $\langle 2, 3 \rangle$ and $\langle 3, 2 \rangle$ represent two different points in the plane, the first one 2 units to the right and 3 units forward from the origin, and the second one 3 units to the right and 2 units forward.

Strictly speaking, set theory just talks about sets that have no ordering to their elements, but in the next exercises, we use a trick introduced by Kazimierz Kuratowski in 1921 to prove that there are entities (involving doubly-nested sets) that do have the properties we need out of an ordered pair.

[1] This ordered pair might more commonly be notated with parentheses as '(x_1, x_2)', but the angle brackets are also sometimes used. I will always use the angle brackets, because the parentheses have so many uses that they introduce a lot of ambiguity.

Exercise 2.9 If x_1, x_2, y_1, and y_2 are sets, then the sets $z_1 = \{\{x_1\}, \{x_1, x_2\}\}$ and $z_2 = \{\{y_1\}, \{y_1, y_2\}\}$ are equal iff $x_1 = y_1$ and $x_2 = y_2$. (Hint: either apply Exercise 2.8 directly, or start by considering the cases where $x_1 = x_2$ and $x_1 \neq x_2$ separately.)

From now on, we will use the notation '$\langle x_1, x_2 \rangle$' as an abbreviation for $\{\{x_1\}, \{x_1, x_2\}\}$.

Axiom 3 Union

$$\forall x \exists y (\forall z (z \in y \leftrightarrow \exists w (z \in w \text{ and } w \in x)))$$

While the Pairing axiom starts with two sets, and makes a set whose members are just those two, the Union axiom starts with a set x and 'cracks open' its members, to create a set y whose members are anything that appears in any of the members of x.

This set y, called the *union* of the members of x, is also written as '$\bigcup x$'. We could also write it as $\bigcup x = \{z \mid \exists w{\in}x (z \in w)\}$.[2] In the case where $x = \{a, b\}$ this can be written either as $\bigcup\{a, b\}$ or $a \cup b$, and we can think of it as the set of everything that is in either a or b. The relationship of the symbol '$\bigcup$' to the symbol '$\cup$' is a lot like the relationship of the symbol '$\sum$' to '$+$' in other areas of mathematics.

This first exercise is just an exercise in remembering definitions.

Exercise 2.10 If $w \in x$ then $w \subseteq \bigcup x$. (Every element of x is a subset of $\bigcup x$).

Exercise 2.11 Assume that a and b are distinct sets. (Note: we have not yet proven that two distinct sets exist, though we will in Exercise 2.17!) Let $c = \{a, b\}$, $d = \{\{a\}, c\}$, $e = \{c, d\}$, $f = \{c, e\}$. What are the members of $\bigcup d$? of $\bigcup e$? of $\bigcup f$?

Note that if all of these sets are distinct from each other, then some of these sets have three or four members. Just as we have names for singletons and doubletons that involve writing names of each of the members in between the curly brackets '}' and '{', we have such a name for any set with finitely many members. The next two exercises show that for any name we can write down this way, the Pairing and Union axioms allow us to prove that the relevant set exists.

Exercise 2.12 If x, y, and z are sets, then there is a unique set $\{x, y, z\}$ whose elements are exactly these three.

[2] This sort of notation $\{z|\quad\}$ is often called 'set-builder' notation, and it is used in naive set theory to name the set of all objects z that satisfy a condition that is put into the blank. We will discuss this notation more when we introduce Axiom 7, Separation. As we will show in Exercise 2.28, this sort of notation can lead to paradoxes if we aren't careful with it. The official axioms that we state can all be thought of as instances of this notation that have been deemed safe to use.

For the next exercise (and many later ones), we use names 'x_i' that consist of a letter subscripted with a natural number. This notation is just to provide us with an infinite set of names that can stand for (possibly distinct) objects, but there is no specific relationship between the number that is in the subscript and the object that it stands for, and no specific relationship that must hold between objects with names involving the same letter.

This exercise also asks for a proof by induction. For a statement involving an arbitrary natural number n, a proof by induction proceeds in two parts. First (the 'base case'), we show that the statement is true when $n = 0$. Second (the 'induction step'), we show that *if* it is true for n, *then* it is also true for $n+1$. Having proved both of these, we conclude that it is true for every natural number. At this point, we just use the intuitive idea of the natural numbers, but when we get to Exercise 2.73, we actually give a set theoretic proof that proof by induction shows that every natural number has the property in question.

Exercise 2.13 Show (by induction on n) that if $x_0, \ldots, x_n$ are sets, then there is a unique set $\{x_0, \ldots, x_n\}$ whose elements are exactly these sets. (Note: our official notation will never use '...' — we use it here as an intuitive abbreviation, but official notation will always spell out every member of a set individually.)

Axiom 4 Null Set, or Empty Set

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

Where the previous two axioms showed how to construct a new set y from existing sets x, this axiom is the first one we have that guarantees the unconditional existence of a particular object.[3] The object here has no members.

The following exercise shows that this object is unique, and we will denote it by the symbol '$\emptyset$', and call it the *empty set* (or sometimes the *null set*).

Exercise 2.14 Justify the use of the definite article by showing that this set is the only one with no elements. That is, if y_1 has no members and y_2 has no members, then $y_1 = y_2$.

Exercise 2.15 $\emptyset \subseteq x$ for any x.

Some versions of set theory allow for the existence of many objects with no elements, so that we can have (for instance) both sets and numbers, with some sets having numbers as elements in addition to sets. These other objects, traditionally called *Ur-elemente* (German for "original elements"), would be objects distinct from the empty set that also have no elements. So these versions of set theory have a slightly weakened form of the Axiom of Extensionality, saying

[3] Most presentations of logic make it a logical truth that something exists, but doesn't guarantee anything about that object — conceivably, it could have itself as an element, or there could be some other large number of objects that have elements, or who knows what. But this axiom guarantees the existence of a specific object, and we will soon show that, together with the previous axioms, this one actually entails the existence of an infinite collection of distinct objects.

only that two objects *with elements* that have the same elements are identical, and allowing that there might be non-identical objects with no elements.

But for most mathematical purposes, this is unnecessary. As we will soon show, Zermelo-Fraenkel set theory enables us to prove the existence of objects that behave like the natural numbers (0, 1, 2, ...). Textbooks of real analysis usually provide methods for constructing objects that behave like the integers and rational numbers from ordered pairs of natural numbers, and objects that behave like the real numbers from certain sets of rational numbers. Geometry textbooks show how to construct objects that behave like geometric points from ordered pairs or triples of real number coordinates, and then construct lines, shapes, and spaces as sets of points. In general, nearly all mathematical objects can be understood in terms of set-theoretic constructions using "pure sets" without urelements.

Much of the construction of mathematical objects out of set-theoretic components was done by Richard Dedekind, in his 1889 *Was Sind und Was Sollen die Zahlen?* ("What are and what should be the numbers?"), Gottlob Frege in his two-volume 1893 and 1903 *Grundgesetze der Arithmetik* ("Basic laws of arithmetic"), and Bertrand Russell and Alfred North Whitehead in their 1913 two-volume *Principia Mathematica*. This set-theoretic foundation for other mathematical disciplines became standardized in the sequence of textbooks by the pseudonymous collective of French mathematicians using the name "Nicolas Bourbaki".

Exercise 2.16 Let us use the name 's_0' to denote $\emptyset$, and for any natural number n, use the numeral 's_{n+1}' to denote $s_n \cup \{s_n\}$. Show (by induction on n) that for each n, s_n is the set containing all s_i where $i < n$, and nothing else.[4]

Exercise 2.17 Show that if $m < n$ then $s_m \neq s_n$. (See hints below.)

For the above exercise, we need to work by *strong* induction on m. As a base case, we show that if $0 < n$, then $s_0 \neq s_n$. Next, we want to show that if $m + 1 < n$, then $s_{m+1} \neq s_n$. It turns out that just assuming that if $m < n$ then $s_m \neq s_n$ is not enough for this. Instead, we need to proceed by *strong* induction, assuming that this is true for *all* numbers up to m, not just for m itself. Once we have made this assumption, we will show that there is a particular z with $z \in s_n$ but $z \notin s_{m+1}$, namely the specific case that $z = s_{m+1}$. In showing that $s_{m+1} \notin s_{m+1}$, we make use of the full strong induction assumption.

Thus, the axioms so far suffice to show that there are infinitely many distinct sets. Now that we have shown that these sets are distinct, we can even use the numeral 'n' itself to stand for this set, rather than s_n. We haven't yet shown that there is a set of all these numbers, or that standard arithmetical functions and relations on the natural numbers can be defined on these sets, but we will in Sections 2.6 and 3.

[4] We might want to write this as $s_n = \{s_0, \ldots, s_{n-1}\}$. However this unofficial notation with '...' is hard to interpret for small values of n. By convention, when a sequence involves natural numbers, and is said to start with 0 and end with $n-1$, then the sequence is taken to have n elements, so it has one element when $n = 1$ and no elements at all when $n = 0$. Thus, '$s_n = \{s_0, \ldots, s_{n-1}\}$' means, among other things, that $s_0 = \{\}$ (that is, $\emptyset$) and $s_1 = \{s_0\}$.

Axiom 5 Power Set

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

Just like Pairing and Union, this axiom starts with a given set x and states that there exists a set y whose membership criterion is given in terms of x. In this case, we call y the *power set* of x, and its elements are all and only the subsets of x. We name this set '$\mathcal{P}(x)$', and could also write it as $\{z | z \subseteq x\}$, with the set-builder notation mentioned in footnote 2.

Exercise 2.18 Find $\mathcal{P}(\emptyset)$, $\mathcal{P}(\mathcal{P}(\emptyset))$, and $\mathcal{P}(\mathcal{P}(\mathcal{P}(\emptyset)))$ and write them explicitly using curly braces.

Exercise 2.19 If a and b are two distinct objects, then find $\mathcal{P}(\{a, b\})$ and $\mathcal{P}(\mathcal{P}(\{a, b\}))$ and write them explicitly using curly braces.

Exercise 2.20 For any x, $\emptyset \in \mathcal{P}(x)$ and $x \in \mathcal{P}(x)$. (Does that mean that $\mathcal{P}(x)$ always has at least two distinct members?)

Exercise 2.21 If $x \subseteq y$ then $\mathcal{P}(x) \subseteq \mathcal{P}(y)$.

Exercise 2.22 Show that for any natural number n, if x has exactly n distinct elements, then $\mathcal{P}(x)$ has exactly 2^n distinct elements. (Hint: Use Exercise 2.13 to construct a subset of x with exactly $n - 1$ distinct elements, and consider how its power set relates to that of x.)

Exercise 2.23 Consider $\bigcup \mathcal{P}(x)$, $\mathcal{P}(\bigcup x)$, and x — which of these are guaranteed to be subsets of each other, or identical to each other? (Hint: you'll use Exercise 2.10 several times.) Can you come up with a set x for which these three sets are not all identical? (Hint: Either use Exercise 2.15 to find an object that is guaranteed to be in one of these sets and choose x to not contain that object, or use Exercise 2.22 to find the possible numbers of elements for one of these sets and choose x to have a different number of elements.)

2.3 Limitations on the universe: the Axiom of Foundation (or Regularity)

This axiom has a very different character from the others. While the others can be interpreted as showing us that certain sets *do* exist, this one is most naturally interpreted as saying that certain types of sets *don't* exist.

This axiom is not actually needed for most mathematical applications of set theory. After this sub-section, basically no Exercises will make use of it except for some minor ones where I mention it explicitly. However, it is part of the basic intuition of "pure sets" that most set theorists work with, and it is part of the standard package of Zermelo-Fraenkel axioms.

Axiom 6 Regularity, or Foundation[5]

$$\forall x (x \neq \emptyset \rightarrow \exists y \in x \; \forall z (z \in x \rightarrow z \notin y))$$

[5] '$\exists y \in x$' is an abbreviation for '$\exists y (y \in x$ and $\quad)$'.

That is, every nonempty set has an element with whom it shares no common elements. Two sets with no common element are said to be *disjoint*, so this axiom can be re-stated as the assumption that every set has an element that is disjoint from it.

Exercise 2.24 There is no set x such that $x \in x$. (Hint: We can't directly get a contradiction from the assumption that *this* set exists, but we can use Exercise 2.5 to construct a set that *would* lead to a contradiction of Axiom 6.)

Exercise 2.25 As a corollary, show there is no set of all sets. That is, there is no set V such that every set is a member of V.

Exercise 2.26 There is no finite collection of sets $x_0, \ldots, x_n$ such that for all $0 \leq i < n$, $x_i \in x_{i+1}$ and $x_n = x_0$. (Hint: Just like Exercise 2.24, but use Exercise 2.13 instead of Exercise 2.5.)

These are the first limitative results we have about the existence of sets. The fact that no set is a member of itself, or a 'grand-member' of itself, or 'great-grand-member' of itself, etc., is a helpful intuition for working with the universe of set theory, but it isn't actually essential for the mathematical or philosophical purposes to which set theory is often put. Peter Aczel has developed alternate versions of 'non-well-founded' set theory that *do* allow for these sorts of sets, and can be used for all the same purposes that standard set theory can be used for.[6]

The one significant limitative result here, Exercise 2.25, will actually be proved in a very different way in Exercise 2.28 (and again in Exercise 2.52). This result states that there is no V that contains all sets.[7] This is our first example of a collection of sets (namely, the collection of all of them) that do not themselves constitute the membership of any particular set. The word 'class' is used for an arbitrary collection of sets, and if there is no set that has all and only those sets as members, then the collection is said to be a 'proper class'.

We will eventually build up some results leading to the intuition that the 'real' reason proper classes usually aren't sets is because they have 'too many' members. Clearly, if there were a set V that contained everything, it would have 'too many' members if anything does! But at this point, the only reason we have ruled it out is because it would contain itself. Still, it is helpful to rule this out early, to help develop intuitions for what sorts of collections can and can't be sets. Note that in Exercise 2.17, we showed that a particular collection of sets aren't self-members, and don't form loops of membership, without using this axiom — this axiom generalizes that to all sets.

[6]The construction in the Exercises leading up to 6.30 can in fact be used in a non-well-founded set theory to construct a definable class of sets that form a model of standard well-founded set theory, and Aczel shows how to expand a model of standard well-founded set theory into models of his non-well-founded set theories, so this really isn't a significant distinction.

[7]The explanation of why we use the letter 'V' for the 'universe of set theory' rather than 'U' comes in Section 6.2. In that section we examine the 'shape' of the universe that we get by starting with the empty set, and building up by iterating the power set operation, and show that it 'looks like' a big V with a point at the bottom, and that every set eventually appears if we iterate this operation sufficiently infinitely many times.

2.4 Cutting a set down to size: the Axiom Schema of Separation

In Axiom 7 (as well as 8 and 9 in Section 2.5) the Greek letter φ will stand for a well-defined 'formula' defining a property or relation that sets x and y might or might not have. A *formula* φ is anything that can be written with just logical symbols (such as 'and', 'or', 'not', '$\rightarrow$', '$\leftrightarrow$', '$\forall$', '$\exists$') and the two relations '$=$' and '$\in$', as well as names for some particular sets. Some variables in the formula are bound by the quantifiers, while others are 'free'. A formula is usually interpreted as being 'about' its free variables, and it is standard when representing a formula in the abstract to write its free variables in parentheses afterwards. Thus '$\varphi(x)$' is a formula defining some property of an unspecified object x, while '$\varphi(x, y)$' is a formula defining some relation that unspecified objects x and y might bear.

Because $x \subseteq y$ can be written as $\forall z(z \in x \rightarrow z \in y)$, '$\subseteq$' is also a symbol that can be used in defining properties. Similarly, we can use the other notations introduced so far, like $\bigcup x$, $\{x, y\}$, $\langle x, y\rangle$, and so on.

Because the 'Axioms' 7, 8, and 9 include unspecified formulas, it is best to interpret each of them as an axiom *schema* (plural 'schemas' or 'schemata'), consisting of infinitely many axioms — one for each formula φ with the appropriate set of free variables. Thus, instead of three axioms, these are (strictly speaking) three separate sets of infinitely many axioms.

Axiom 7 Separation Schema

$$\forall x \exists y(\forall z(z \in y \leftrightarrow (z \in x \text{ and } \varphi(z))))$$

This set y is more commonly written as $\{z{\in}x \mid \varphi(z)\}$. (As mentioned in footnote 2, when the set of *all* sets satisfying φ exists, we often write it as $\{z \mid \varphi(z)\}$, but as you will show in problem 2.28, this notation sometimes refers to a proper class rather than a set.)

Exercise 2.27 Show that we can prove Axiom 4 (the Null Set Axiom) just using Axiom 7 and the assumption that at least one set exists. (Hint: Is there some condition that holds only of the members of the empty set?)

Since logic already guarantees that at least one object exists (and furthermore, Axiom 10 will explicitly assume the existence of a specific set), some presentations of Zermelo-Fraenkel set theory omit Axiom 4 as part of the standard set of axioms. Nevertheless, it is often included as a separate axiom, in order to investigate the strength of various subsystems of these axioms.

Exercise 2.28 (Russell's Paradox) Using only Axiom 7, show that there is no set of all sets. (Hint: let $\varphi(z)$ be $z \notin z$.)

Gottlob Frege gave one of the first formalized set theories in the 1893 first volume of his book *Grundgesetze der Arithmetik* ("The Basic Laws of Arithmetic"). In his set theory, every property $\phi(z)$ whatsoever picked out a set

$\{z|\phi(z)\}$ of all and only the objects satisfying that property. He wanted to use this set theory to give a full logical foundation for all of mathematics. As he was completing the second volume of this book, he received a note from Bertrand Russell, who showed him the result of Exercise 2.28, which showed that his set theory was inconsistent, and he included this paradox in the second volume, which led him to never finish the project.

Later formalized set theories have all been developed in ways that avoid this paradox. The method that has been followed by most set theorists is that of Ernst Zermelo, who included the specific restriction of the axiom schema of Separation to subsets of sets that already exist, in his 1908 paper "Untersuchungen über die Grundlagen der Mathematik". It is because of this restriction that we need Pairing, Union, and Power Set as separate axioms, because the sets they produce are not themselves subsets of sets that are already known to exist.

(Different methods of avoiding this paradox were also proposed by Bertrand Russell and Alfred North Whitehead in their 1913 *Principia Mathematica*, and by Willard Van Orman Quine in his 1937 "New Foundations for Mathematical Logic". Those systems are still studied by some logicians but are not nearly as standard in contemporary mathematics as Zermelo's system, with modifications due to Abraham Fraenkel and others.)

Exercise 2.29 *Use Axiom 7 to show that Axioms 2 (Pairing), 3 (Union), and 5 (Power Set) can all be proven from their weaker forms in which '$\leftrightarrow$' is replaced by '$\leftarrow$'.

Exercise 2.30 Let $x \setminus y$ be the collection of all elements of x that are not in y. Show that this collection is in fact a set. (Hint: for this Exercise, you need to use the fact that ϕ can include a name for a particular set, in this case y.)

Exercise 2.31 Show that for any non-empty set X there is a set $\bigcap X$ (the *intersection* of the elements of X) that contains exactly the sets that are in all elements of X. If $X = \{A, B\}$ then we also write $\bigcap X$ as $A \cap B$.

Exercise 2.32 Assuming that $\varphi(x)$ is a property that at least one set satisfies, show that there is a set $\bigcap \varphi$ such that

$$\forall z(z \in \bigcap \varphi \leftrightarrow \forall y(\varphi(y) \rightarrow z \in y)).$$

This shows that we can take the intersection of any non-empty collection of sets, including proper classes, not just of a *set* of sets. The previous problem is a special case when $\varphi(x)$ is $x \in X$.

Exercise 2.33 For any set X, show that $\{\{y\} \mid y \in X\}$ (ie, the set of all singletons of elements of X) exists. (You will have to find a set x that contains all the $\{y\}$, and find a property φ separating them out.)

Exercise 2.34 For any set X, show that the set $X \times X = \{\langle x_1, x_2\rangle \mid x_1, x_2 \in X\}$ of all ordered pairs from X exists. (Recall the definition of ordered pairs in Exercise 2.9.)

Exercise 2.35 For any sets X and Y, show that the set $X \times Y$ of all ordered pairs $\{\langle x, y\rangle \mid x \in X, y \in Y\}$ exists. (This set is called the *cartesian product* of X and Y.)

2.5 Relations and Functions: the Axiom of Replacement

When Ernst Zermelo gave his axioms for set theory, he didn't include the Axiom of Replacement. It was added to Zermelo's set theory by Abraham Fraenkel. This axiom is inessential for most of the classical results of set theory, so it is reasonable to skip this section and jump directly to Section 2.6.

Readers who are skipping this axiom may still want to work through Section 2.5.1 first, going over the idea of representing functions and relations as sets, and defining the mathematical concepts of totality, injection (one-to-one), surjection (onto), and bijection. But Section 2.5.2 is more complex and can be skipped without loss until Replacement actually gets used. The first time Replacement is used after this section is in Exercise 2.93 about the Axiom of Choice, and then not again until 3.52.

2.5.1 Set relations and functions

Any subset of $X \times Y$ is called a *relation from X to Y*. (We will sometimes call it a *set relation* to distinguish it from the *class relations* defined in Section 2.5.2.) If $Y = X$, then we call it a relation *on* X. We can think of $\mathcal{P}(X \times Y)$ as the set of all relations from X to Y.

If R is a relation, then we often write "xRy" as an abbreviation for "$\langle x, y\rangle \in R$", following the notation we use for the class relations $=$ and $\in$.

Exercise 2.36 For any set X, show that the set $\{\langle x, x\rangle \mid x \in X\}$ exists. (This set is called the *identity* relation on X.)

Exercise 2.37 Show that if R is a set of ordered pairs, then the set $X = \{x \mid \exists y(xRy)\}$ exists, as does the set $Y = \{y \mid \exists x(xRy)\}$. X is called the *domain* of R and Y is called the *range* of R.

Exercise 2.38 Every set of ordered pairs is a relation from a set to a set.

Exercise 2.39 If R is a relation then the relation R^{-1} (the *inverse* relation) given by $\{\langle y, x\rangle \mid xRy\}$ exists. If R is a relation from X to Y, then R^{-1} is a relation from Y to X.

Exercise 2.40 If R and S are relations then there is a relation $S \circ R = \{\langle x, z\rangle \mid \exists y(\langle x, y\rangle \in R \text{ and } \langle y, z\rangle \in S)\}$, called the *composition* of the two relations.

Exercise 2.41 If R is a relation from X to Y and S is a relation from Y to Z then $(S \circ R)$ is a relation from X to Z.

Exercise 2.42 If R and S are relations, then $(R \circ S)^{-1} = (S^{-1} \circ R^{-1})$.

Exercise 2.43 If R, S, T are relations, then $(R \circ S) \circ T = R \circ (S \circ T)$.

We use the symbol '$\exists!x$' to mean 'there exists a *unique* x such that...'. Thus, $\exists!x\varphi(x)$ is an abbreviation for

$$\exists x(\varphi(x) \text{ and } \forall y(\varphi(y) \to y = x)).$$

A relation $R \subseteq X \times Y$ such that $\forall x{\in}X\ \exists!y{\in}Y\ xRy$ is called a *total set function from X to Y*. If $X' \subsetneq X$ we sometimes say that a total set function from X' to Y is a *partial set function from X to Y*. We sometimes omit the words "partial" or "total" or "set" if it is clear from context which is intended.

We tend to use f, g for relations that are known to be functions, and R, S for relations that aren't known to be functions.

Exercise 2.44 There is a set ${}^{X}Y$ of all total functions from X to Y.

Exercise 2.45 There is a set of all total or partial functions from X to Y.

We often write $f\colon X \longrightarrow Y$ to mean that f is a function from X to Y. (We sometimes mean this to imply that it is total, and sometimes allow it to be partial — usually it is clear from context.) If f is a function and $\langle x, y\rangle \in f$ then we often write $f(x)$ for y.

Exercise 2.46 If f and g are functions and the domain of f contains the range of g then $(f \circ g)(x) = f(g(x))$. What goes wrong if the domain of f does not contain the range of g?

If f and f^{-1} are both functions, then both are said to be *injective* or *invertible* functions.

Exercise 2.47 f is injective iff $\forall x\forall y(f(x) = f(y) \to x = y)$.

If $f\colon X \longrightarrow Y$ is a function and the range of f is Y, then f is said to be *surjective* on Y, and it is also called a *surjection* to Y, or even just said to be *onto* Y. If f is injective, it is said to be *bijective* from its domain to its range, and it is also called a *bijection* between its domain and range. (The noun "injection" and preposition "into" are also occasionally used for injective functions, but less commonly than "surjection" and "onto", or "bijection".)

Exercise 2.48 If $f \subseteq X \times Y$ is surjective on Y and $g \subseteq Y \times Z$ is surjective on Z, then $g \circ f \subseteq X \times Z$ is surjective on Z.

Exercise 2.49 If $f \subseteq X \times Y$ and $g \subseteq Y \times Z$ are both injective functions, then so is $g \circ f \subseteq X \times Z$.

Exercise 2.50 Show that for any X, the identity relation on X (see Exercise 2.36) is a bijective function between X and X.

Exercise 2.51 Show that for any set x there is no surjective function f with domain x and range $\mathcal{P}(x)$. (Hint: consider the set $\{z{\in}x \mid z \notin f(z)\}$. Because this is a subset of x, it is a member of $\mathcal{P}(x)$. Is it in the range of f?)

In a sense that will be explained in more detail in Exercise 5.18, developed in the 1870's, by Georg Cantor, this last result shows that a power set is 'bigger than' the set itself. You already showed this for finite sets in problem 2.22. It is this result of Cantor's that led Russell to his paradox for Frege's system:

Exercise 2.52 Using only Exercises 2.50 and 2.51, show that there is no set of all sets. (Hint: If V were the set of all sets, what would be its power set? Compare Exercise 2.28.)

2.5.2 Relational and functional formulas

Although set theorists most often talk about relations and functions as particular sets of ordered pairs, it is more natural in many areas of mathematics to think about relations and functions as defined linguistically. That is, we can think of any formula $\varphi(x, y)$ with two free variables as implicitly defining a binary relation. If there is a *set* consisting of all and only the ordered pairs satisfying this formula, then we call this set the *set relation*, and if there is no such set, then we call the formula a *class relation*.

If a formula $\varphi(x, y)$ satisfies $\forall x \exists! y\ \varphi(x, y)$, then we sometimes say it is a *functional formula* or that it defines a *class function*. We sometimes say it defines a *total class function*, to contrast it with a formula $\varphi(x, y)$ that merely satisfies $\forall x\ (\exists y\ \varphi(x, y) \rightarrow \exists! y\ \varphi(x, y))$, which is said to define a *partial class function*. (Slightly confusingly, every total function is a partial function, even though we might have expected that "partial function" only applies to functional formulas that are not total.) If there is a set consisting of all the ordered pairs $\langle x, y\rangle$ satisfying $\varphi(x, y)$, then we call this set function the *graph* of φ. We sometimes omit the word "class" when it is clear from context that we are not talking about the set function.

Exercise 2.53 No total function is a set function. (Hint: Use Exercise 2.37.)

Exercise 2.54 Consider the following eight formulas that have only x and y free. Which of them define total functions, partial functions, or relations?

'$y = x$'	'$\forall w(w \in y \leftrightarrow w = x)$'
'$y = x$ and $\exists w(w \in x)$'	'$\forall w(w \in x \leftrightarrow w = y)$'
'$x \in y$'	'$\forall w(w \in y \leftrightarrow w = x)$ or $\forall w(w \in x \leftrightarrow w = y)$'
'$y \in x$'	'$\forall w(w \in x \leftrightarrow w = y)$ or $(\neg\exists! w(w \in x)$ and $y = \emptyset)$'

For φ defining a (total or partial) function, if $\forall x{\in}X\ \exists y\ \varphi(x, y)$[8], then we say the function is *defined on* X. For φ defining a relation, if $\forall x{\in}X\ (\exists y\ \varphi(x, y) \rightarrow \exists! y\ \varphi(x, y)$, then we say that φ is *functional on* X.

Exercise 2.55 In Exercise 2.54, two of the formulas defined partial functions. Let S be a set with multiple elements that each have two or more members. Are either of these functions defined on S? What about $\mathcal{P}(S)$? What about the set of all singletons in $\mathcal{P}(S)$?

[8] '$\forall x{\in}X$' is an abbreviation for '$\forall x(x \in X \rightarrow \quad)$'.

Exercise 2.56 In Exercise 2.54, three of the formulas defined relations that were not partial functions. Let S be a set with multiple elements that each have two or more members. Are any of these relations functional on S? What about $\mathcal{P}(S)$? What about the set of all singletons in $\mathcal{P}(S)$?

Exercise 2.57 If $\varphi(z, w)$ is a total function, and $\psi(z)$ is any formula, then '$\varphi(z, w)$ and $\psi(z)$' is a partial function.

Exercise 2.58 If $\varphi(z, w)$ is a relation that is functional on X, then $\varphi'(z, w)$ defined by '$z \in X$ and $\varphi(z, w)$' is a partial function.

Exercise 2.59 If $\varphi(z, w)$ is a partial function, then $\varphi'(z, w)$ defined by '$\varphi(z, w)$ or $(\neg\exists w'\varphi(z, w')$ and $w = \emptyset)$' is a total function.

Exercise 2.60 Show that if R is a set of ordered pairs that is a set function, as defined in Section 2.5.1, then $\varphi(z, w)$ given by '$\langle z, w\rangle \in R$' defines a partial function defined on the domain of R. (That is, every set function naturally defines a class function.)

We use total and partial class functions to introduce the next axiom schema. It formalizes the practice that occurs commonly in mathematics, whereby one has a set of objects, and some function, and applies the function to each of those objects to get a new set. We consider two different versions of it and investigate how they are related.

Axiom 8 Replacement Schema for Total Functions

$$(\forall z\exists! w\varphi(z, w)) \rightarrow \forall x\exists y(\forall w(w \in y \leftrightarrow \exists z{\in}x\ \varphi(z, w)))$$

Axiom 9 Replacement Schema for Partial Functions

$$(\forall z\ (\exists w\ \varphi(z, w) \rightarrow \exists! w\ \varphi(z, w))) \rightarrow \forall x\exists y(\forall w(w \in y \leftrightarrow \exists z{\in}x\ \varphi(z, w)))$$

Like several of the earlier axioms, these two axiom schemas show a way to start with a set x and construct a set y whose members are defined somehow in terms of x. In these cases, the way y is constructed is by taking each element of x as an input to the function defined by ϕ, and collecting together all the resulting outputs of ϕ (though in the case of Axiom 9, in some cases there is no output of ϕ).

Exercise 2.61 In Exercise 2.54, some of the formulas defined partial or total functions. What is the result of applying the relevant Replacement axiom, with these functions, to the set ω?

Note that Axiom 9 trivially entails Axiom 8, because any formula satisfying the antecedent of Axiom 8 also satisfies the antecedent of Axiom 9.

Exercise 2.62 Show that Axiom 9 implies Axiom 7 (Separation). (Hint: Use Exercise 2.57.)

Exercise 2.63 Show that Axioms 7 and 8 together imply Axiom 9. (Hint: Use Exercise 2.59.)

Some presentations of ZF set theory thus include only Axiom 9, though most include Axioms 7 and 8 separately. It might be an interesting bonus exercise to show that Axioms 4 (Null Set) and 8 suffice to prove Axiom 9, and thus 7, which demonstrates one more reason why Axiom 4 is sometimes listed separately, even though we saw in Exercise 2.27 that it is a simple implication of Axiom 7.

Exercise 2.64 Show (using just Axiom 2 (Pairing) and some form of Replacement, but *not* Axiom 5 (Power Set)) that if φ is a partial function defined on x then there is a set function $f = \{\langle z, w\rangle \mid z \in x \text{ and } \varphi(z, w)\}$. (Hint: Define some formula $\psi(z, y)$ that is satisfied iff $y = \langle z, w\rangle$ and show that it is a function on x, and use Replacement.) (f is often called the *graph* of φ on x.)

The full generality of this result about the existence of graphs of functions requires the Axiom of Replacement. However, in most mathematical contexts, the graphs of the functions one works with can be constructed just by using Separation from the cartesian product of the domain and range of the function. Thus, Zermelo did not actually include any form of Replacement in his set theory, and it was only explicitly stated in 1922 by Abraham Fraenkel. We will start to see some of its essential power in proving things that one can't prove without it later in Section 3 (notably with Exercise 3.52). But for now, we will note that Replacement allows us to eliminate several uses of other earlier axioms:

Exercise 2.65 Using just Axiom 2 (Pairing) and some form of Replacement, show that for any sets X and y there is a set of ordered pairs $X \otimes y = \{\langle x, y\rangle \mid x \in X\}$. (Hint: Write down a formula $\varphi(z, w)$ that is satisfied iff $w = \langle x, y\rangle$ in the right sort of way.)

Exercise 2.66 Use the previous result and Replacement again to show that for any X and Y, the set $\{X \otimes y \mid y \in Y\}$ exists.

Exercise 2.67 Use the previous result and Axiom 3 (Union) to show that the cartesian product $X \times Y$ exists without having to use Axiom 5 (Power Set) at all.

Exercise 2.68 Using only Extensionality, Power Set, Empty Set, and Replacement, prove the Pairing axiom. (Hint: prove what the elements are in the set $\mathcal{P}(\mathcal{P}(\emptyset))$, and define an appropriate function with it as the domain.)

2.6 The Axiom of Infinity

The Axiom of Infinity is the only axiom other than Axiom 4 (Null Set) that unconditionally states the existence of a specific set, and as we noted, Axiom 4 is often left out. All other axioms have a much more general character, stating that *if* certain sets exist, *then* certain other related sets exist.

Axiom 10 Infinity

$$\exists y(\emptyset \in y \text{ and } \forall z(z \in y \to (z \cup \{z\}) \in y))$$

Since $\emptyset$ is an element of y, we see that $\emptyset \cup \{\emptyset\} = \{\emptyset\}$ is also. Similarly $\{\emptyset\} \cup \{\{\emptyset\}\} = \{\emptyset, \{\emptyset\}\} \in y$, and the same holds for $\{\emptyset, \{\emptyset\}, \{\emptyset, \{\emptyset\}\}\}$ and $\{\emptyset, \{\emptyset\}, \{\emptyset, \{\emptyset\}\}, \{\emptyset, \{\emptyset\}, \{\emptyset, \{\emptyset\}\}\}\}$ and so on. There may be additional random stuff in there too, but these intuitively guarantee there are infinitely many elements of y.

Exercise 2.69 Let $\phi(x)$ be the property $\emptyset \in x$ and $\forall z(z \in x \to (z \cup \{z\}) \in x))$, i.e., the property of being the type of object asserted to exist by Axiom 10 (Infinity). Use Exercise 2.32 to define

$$\omega = \bigcap \phi.$$

Show that $\phi(\omega)$. (That is, show that $\emptyset \in \omega$ and $\forall z(z \in \omega \to (z \cup \{z\}) \in \omega)$.)

Exercise 2.70 Recall the definitions we gave in Exercise 2.17, that $0 = \emptyset$, and $n + 1 = n \cup \{n\}$. Show (by induction on n) that each of these sets is in ω.

It would be nice to be able to show that these are the *only* sets that are in ω. However, to even *state* this fact, we would need an independent characterization of the natural numbers, which we don't have. It is common in set theory to use the fact that ω has a precise definition to say that its elements just *are* the natural numbers. Thus, we will typically henceforth use variables m, n to refer to members of ω.

An early attempt to precisely define the natural numbers was made by Richard Dedekind in his 1888 book, "Was sind und was sollen die Zahlen?" ("What are numbers and what should they be?") He presented five basic postulates that he claimed defined everything we know about the natural numbers. These postulates were discussed at length by Giuseppe Peano in his 1889 book, *Arithmetices principia, nova methodo exposita* (*The principles of arithmetic, presented by a new method*), so they are sometimes called the 'Peano postulates' or 'Peano axioms' or 'Dedekind-Peano axioms'.

We can now prove that, if we interpret ω as the set of natural numbers, 0 as $\emptyset$, and the successor of a natural number n as $n \cup \{n\}$, the five Dedekind-Peano axioms are all true.

The first two axioms, "0 is a natural number", and "the successor of any natural number is a natural number", were already shown in Exercise 2.69. The other three axioms will each be an Exercise.

The third axiom states, "No two different natural numbers have the same successor."

Exercise 2.71 $\forall n{\in}\omega \forall m{\in}\omega(((n \cup \{n\}) = (m \cup \{m\})) \to m = n))$. (Hint: Note that n is an element of $n \cup \{n\}$, and any element of $m \cup \{m\}$ is either an element of m, or m itself.)

The fourth axiom states, "0 is not the successor of any natural number."

Exercise 2.72 $\forall n(0 \neq n \cup \{n\})$.

The fifth axiom is the principle of induction.

Exercise 2.73 Assume that $X \subseteq \omega$, that $\emptyset \in X$, and $\forall n(n \in X \rightarrow (n \cup \{n\}) \in X)$. Show that $X = \omega$. (Hint: Look back at the definition of ω in Exercise 2.69.)

The idea that the natural numbers really *are* these sets, rather than just behaving like them, has been strongly disputed by philosophers like Paul Benacerraf, in his 1965 paper "What Numbers Could Not Be". (He argues that there is no better case that the numbers *are* the sets given by $n + 1 = n \cup \{n\}$ than that they *are* the sets given by $n + 1 = \{n\}$.) Nevertheless, these sets can be used to do all the things that we do with natural numbers, so they are a good enough replacement, even if they aren't the actual numbers themselves.

Whether or not ω is taken to actually *be* the natural numbers, the proof of the existence of a model still has several other types of significance. For instance, all the logical rules are supported by any model — if every axiom of Peano arithmetic is true in the model, then anything *provable* from these axioms will also be true in the model. But no contradiction will be true in the model. So if there is a model that makes all the axioms true, then the axioms can not lead to a contradiction. So the construction of a model shows that the axioms are consistent.

Gödel's important incompleteness results show that no theory can prove itself to be consistent. Thus, because ZF proves the existence of a model of Peano arithmetic, ZF must in fact be a stronger theory than PA. In Section 6, we will eventually investigate models of fragments of ZF itself, and show what a model of ZF as a whole would have to look like. This will enable us to investigate the consistency of these axioms, and understand how strong an axiom will need to be to extend this theory of sets.

Theoretically, someone who didn't have the concept of a natural number, and had skipped the Exercises that mention them, could use ω as a definition of the set of natural numbers, and Exercise 2.73 as a proof that the proof method of induction works.

Exercise 2.74 For every $n \in \omega$ either $n = \emptyset$ or there is an $m \in \omega$ such that $m \cup \{m\} = n$. (Hint: this is a simple proof by induction, letting X be the set $\{n \in \omega \mid n = \emptyset \text{ or } \exists m{\in}\omega(m \cup \{m\} = n)\}$.)

We call this m the *predecessor* of n.

Exercise 2.75 For any $n \in \omega$ with $n \neq \emptyset$, $\bigcup n$ is the predecessor of n.

Exercise 2.76 $\bigcup \omega = \omega$.

2.6.1 Recursive Definition

Just as the ordinary natural numbers admit proof by induction, they also allow us to define functions by recursion. That is, let X be any set, and let g be a total function from X to X, and let $x \in X$. We want to show that there is a

total function $f\colon \omega \longrightarrow X$ with $f(0) = x$, $f(1) = g(x)$, $f(2) = g(g(x))$, and in general, $f(S(n)) = g(f(n))$. The following exercises show that such a set exists.

Exercise 2.77 There is a set of all partial functions from ω to X.

Exercise 2.78 There is a set of all partial functions from ω to X whose domain is some member of ω.

Exercise 2.79 There is a set of all partial functions from ω to X whose domain is some member of ω and with $f(0) = x$.

Exercise 2.80 There is a set of all partial functions from ω to X whose domain is some member of ω, with $f(0) = x$, and for every $n \neq 0$ in the domain of f, if m is the predecessor of n, then $f(n) = g(f(m))$.

Exercise 2.81 Every member of ω is the domain of some function in the set from the previous exercise. (Hint: Work by induction. Show that the set of domains of functions in that domain contains 0 and is closed under successor.)

Exercise 2.82 The union of the previous set of partial functions is in fact the desired total function from ω to X. (Hint: The main thing to prove here is that this union is in fact a *function*, i.e., that two different outputs don't both get associated with the same input by different members of this set.)

Exercise 2.83 The previous set of exercises can be repeated for two-place functions, three-place functions, and so on. That is, if $h\colon \omega^k \longrightarrow X$ and $g\colon (\omega^k \times X) \longrightarrow X$, then we can use a parallel procedure to define $f\colon \omega^{k+1} \to X$ with $f(n_1, \ldots, n_k, 0) = h(n_1, \ldots, n_k)$ and $f(n_1, \ldots, n_k, S(n)) = g(n_1, \ldots, n_k, f(n_1, \ldots, n_k, n))$.

This procedure allows us to define many familiar functions on the natural numbers. If we let $h\colon \omega \longrightarrow \omega$ be the identity function, and $g\colon \omega^2 \longrightarrow \omega$ be the successor of the second input, then the resulting $f\colon \omega^2 \longrightarrow \omega$ is addition. If we let $h\colon \omega \longrightarrow \omega$ be the function that is everywhere 0, and $g\colon \omega^2 \longrightarrow \omega$ be addition, then the resulting $f\colon \omega^2 \longrightarrow \omega$ is multiplication. If we let $h\colon \omega \longrightarrow \omega$ be the function that is everywhere 1, and $g\colon \omega^2 \longrightarrow \omega$ be multiplication, then the resulting $f\colon \omega^2 \longrightarrow \omega$ is exponentiation.

2.6.2 Dedekind's Definition of Finiteness

We say that a set X is *Dedekind-infinite* iff there is a function $g\colon X \longrightarrow X$ that is injective but not surjective, and it is *Dedekind-finite* iff there is no such function.

Exercise 2.84 If there is a bijection between two sets, then one of them is Dedekind-infinite iff the other is.

Exercise 2.85 ω is Dedekind-infinite. (Hint: think about the successor function.)

Exercise 2.86 If some subset of X has a bijection to ω, then X is Dedekind-infinite.

Exercise 2.87 $\emptyset$ is not Dedekind-infinite. (You have to consider some tricky questions about what it means for something to be a function from $\emptyset$ to $\emptyset$ — as a hint, there is only one such function, and it is called the "empty function".)

Exercise 2.88 If x is not Dedekind-infinite, then $x \cup \{x\}$ is not Dedekind-infinite. (Hint: The contrapositive might be easier — if $x \cup \{x\}$ *is* Dedekind-infinite, then you can consider where the relevant function sends x, and what it sends to x, and use this to construct a function showing that x itself is Dedekind-infinite.)

Exercise 2.89 No member of ω is Dedekind-infinite. (Hint: this follows trivially from the previous two.)

Modern set theorists say that X is *finite* iff there is a bijection between X and some member of ω, and *infinite* otherwise. What we have shown here is that if a set is finite, then it is Dedekind-finite, and if a set is Dedekind-infinite then it is infinite. However, we haven't shown the converse.

We know that a set is finite iff it has a bijection to some member of ω, and the next exercise shows us that it is Dedekind-infinite iff some subset has a bijection to ω.

Exercise 2.90 If X is Dedekind-infinite, then there is a bijection between ω and some subset of X. (Hint: Apply Exercise 2.82. Let x be some member of X that is not in the range of g, and define $f(0) = x$ and $f(S(n)) = g(f(n))$, and show that f is injective.)

The next axiom will allow us to show that there is no set in between — every set is either finite or Dedekind-infinite; every set either has a bijection to some natural number, or has a subset with a bijection to ω.

2.7 The Axiom of Choice

The axioms mentioned so far are essential for Section 3, but the Axiom of Choice will not come back until Section 5. Many applications of the Axiom of Choice are only possible once we have studied the ordinals (i.e., once we have completed Section 3). So readers (particularly those that skipped the section on functions and Replacement) can skip from here to Section 3 and come back to the Axiom of Choice later if they want. But I include it here because it is the last of the standard axioms of set theory.

Axiom 11 Axiom of Choice

$$\forall x(\emptyset \notin x \to \exists F(\exists z(F\colon x \longrightarrow z) \text{ and } \forall y{\in}x(F(y) \in y)))$$

This F is called a *choice function* on x — for each set in x (which are all assumed to be non-empty), F "chooses" an element of that set. Bertrand Russell characterized the content of this axiom with a memorable image. If x is an infinite set of pairs of shoes, then we can take the union of these sets, then use Separation to pick out all the left shoes, and then use Pairing and

Replacement to construct a choice function picking out the left shoe from each pair, without having to appeal to this axiom. But if x were an infinite set of pairs of *socks* (where there is no systematic way of distinguishing the two socks in each pair) then it turns out that we need the Axiom of Choice to prove that a choice function exists.

This image highlights one of the distinguishing features of the Axiom of Choice. With all the other axioms, for every choice of input sets and formulas, there is a *specific* set whose existence the axiom entails. But with the Axiom of Choice, there are generally many, though we need the Axiom of Choice to prove that there are any.

Choice was introduced by Zermelo in the same paper where he included all the axioms above, other than Replacement, but because of various issues around it, Choice is not considered one of "the Zermelo Axioms" or "the Zermelo-Fraenkel Axioms". Without Choice, the axioms are collectively known as ZF, and with it they are called ZFC. The Axiom of Choice (often called "AC") contradicts some desirable axioms in higher set theory and leads to odd results, like the Banach-Tarski paradox. However, without it, we can't prove certain standard results in other areas of mathematics in full generality (results like "every vector space has a basis", "every ring has a prime ideal", which involve showing that an object of a certain type exists, even though there is no convenient way to pick out a *particular* object of that type).

Exercise 2.91 Show that

$$\forall x \exists F(\exists z(F\colon x \longrightarrow z) \text{ and } \forall y{\in}x(F(y) \in y))$$

is false. (Hint: Look specifically at how this statement differs from Axiom 11.)

Exercise 2.92 Show that

$$\forall x \exists F(\exists z(F\colon (x \setminus \{\emptyset\}) \longrightarrow z) \text{ and } \forall y{\in}(x \setminus \{\emptyset\})(F(y) \in y))$$

is equivalent to the Axiom of Choice.

As we will show in the next two exercises, this is an equivalent way to formulate the Axiom of Choice, which may be easier to understand conceptually, but harder to write down symbolically (and hence I write it primarily in English):

Axiom 12 Axiom of Disjoint Choice

> If X is a collection of non-empty sets that are pairwise disjoint (that is, $\forall x{\in}X\ \forall y{\in}X(x \neq y \rightarrow x \cap y = \emptyset)$) then there is a Z containing exactly one member of each member of X.

Exercise 2.93 Show that the Axiom of Choice implies the Axiom of Disjoint Choice. (Hint: Let X be a collection of non-empty, pairwise disjoint sets, let F be the choice function, and consider the range of F.)

Exercise 2.94 Show that the Axiom of Disjoint Choice implies the Axiom of Choice. (Hint: Let x be a collection of non-empty sets, and for each $y \in x$ show that there is a set $Y = \{\langle y, w\rangle \mid w \in y\}$ (this is $y \otimes y$ as defined in Exercise 2.65 — the notation is annoyingly asymmetrical in this application). Let X be the set you get by Replacing each y by Y, and show that X is a collection of pairwise disjoint non-empty sets. Apply the Axiom of Disjoint Choice to X, and show that the resulting Z is a choice function on x.)

Now let us use this axiom to prove that every infinite set is Dedekind-infinite.

Exercise 2.95 (AC) Let X be an infinite set. Let g be a choice function on the set of non-empty subsets of X. Define $f\colon \omega \longrightarrow X$ recursively (as in Exercise 2.82), by letting $f(0) = g(X)$ and $f(n \cup \{n\}) = g(X \setminus \{f(m) \mid m \in (n \cup \{n\})\})$. Show that f is a total, injective function from ω into X. (Hint: Once you decipher all the notation, injectivity should be obvious. Totality is where you need to use the fact that X is infinite — we observe that the only way this could fail to be total is if f ended up being a bijection with a natural number.)

3 Ordinals

This section introduces Georg Cantor's 'ordinals', which are an extremely important class of sets for investigating the structure of orderings, and will eventually give us a bigger taste of infinity. The presentation here is not Cantor's own, but instead follows more modern notation, due to John von Neumann.

3.1 Transitive sets, Partial Orderings, and Linear Orderings

We say that a set x is *transitive* iff it satisfies $\forall a \forall b((b \in a \text{ and } a \in x) \to b \in x)$, so that all elements of its elements are already its elements, or equivalently $\forall a(a \in x \to a \subseteq x)$.

Exercise 3.1 Which of the following sets are transitive?

$A = \emptyset$ $\quad B = \{\emptyset\}$ $\quad C = \{\emptyset, \{\emptyset\}\}$ $\quad D = \{\emptyset, \{\emptyset\}, \{\{\emptyset\}\}\}$

$\{\emptyset, \{\{\emptyset\}\}\}$ $\quad C \cup \{C\}$ $\quad \mathcal{P}(C)$ $\quad D \cup \{D\}$ $\quad \mathcal{P}(D)$

Exercise 3.2 If x is transitive, then $x \cup \{x\}$ and $\mathcal{P}(x)$ are transitive.

Exercise 3.3 Come up with a set x where $x \cup \{x\}$ is not transitive, and one where $\mathcal{P}(x)$ is not transitive.

Exercise 3.4 Every member of ω is transitive. (Hint: This should be a straightforward proof by induction.)

Exercise 3.5 ω is transitive.

Exercise 3.6 If x is transitive, then either $x = \emptyset$ or $\emptyset \in x$. (This is a rare Exercise that essentially relies on Axiom 6 (Foundation). Several others in the definition of ordinals will as well.)

Exercise 3.7 If x and y are transitive, then $x \cap y$ is transitive.

3.2 Definition of Ordinals

We officially define a set to be an *ordinal* iff it is transitive, and every one of its members is itself transitive. We generally use Greek letters from the beginning of the alphabet, $\alpha, \beta, \gamma, \ldots$ to represent ordinals.

Exercise 3.8 Which of the sets named in exercise 3.1 are ordinals?

Exercise 3.9 ω is an ordinal.

Exercise 3.10 Every member of an ordinal is an ordinal.

Exercise 3.11 If α is an ordinal, then $\alpha \cup \alpha$ is an ordinal.

Exercise 3.12 A set is an ordinal iff it is a member of an ordinal.

Note that this means we could have given a recursive definition stating "An ordinal is a transitive set of ordinals."

Exercise 3.13 If x is a set of ordinals, then $\bigcup x$ is an ordinal.

Exercise 3.14 For every two ordinals, there is an ordinal containing both of them.

Say that two ordinals α, β are *comparable* iff $\alpha \in \beta$, or $\alpha = \beta$, or $\beta \in \alpha$. Say that they are *incomparable* otherwise.

Exercise 3.15 If α and β are incomparable ordinals, then there is a $\beta' \subseteq \alpha$ that is incomparable to α. (Hint: Find an ordinal γ that contains α and β. Consider the set of all members of γ that are incomparable to α. Apply Axiom 6 (Foundation) to this set, and show that every element of the resulting set must be a member of α.)

Exercise 3.16 If γ is an ordinal, and x is the set of members of γ that are incomparable to some other member of γ, then show that x must be empty. (Hint: If x is non-empty, then we can apply Axiom 6 (Foundation) to it to find some minimal member α of x. Apply the above exercise to find β' incomparable to α that is a subset of it. Use minimality of α to show that $\alpha \subseteq \beta'$, so that $\alpha = \beta'$, contradicting their incomparability.)

Exercise 3.17 Any two ordinals are comparable.

Exercise 3.18 For any ordinal α, x is a transitive subset of α iff $x \in \alpha$ or $x = \alpha$. (Hint: Recall that an ordinal is a transitive set of ordinals.)

Exercise 3.19 For every set of ordinals, there is an ordinal containing all of them. (Hint: Apply Exercise 3.17 to see how the union of this set of ordinals relates to each of them. If they're not all in this union, figure out what to do.)

3.3 Partial Orderings and Linear Orderings

A relation $R \subseteq X \times X$ is called a *strict partial ordering* of X iff it is irreflexive

$$\forall x{\in}X\ (\langle x, x\rangle \notin R)$$

transitive

$$\forall x{\in}X\ \forall y{\in}X\ \forall z{\in}X\ ((\langle x, y\rangle \in R \text{ and } \langle y, z\rangle \in R) \to \langle x, z\rangle \in R)$$

and antisymmetric

$$\forall x{\in}X\ \forall y{\in}X\ (\langle x, y\rangle \in R \to \langle y, x\rangle \notin R).$$

[9]

Exercise 3.20 Show that for any X, the strict subset relation on X (that is, the set $R = \{\langle x, y\rangle \in X \times X \mid x \subseteq y \text{ and } x \neq y\}$) is a strict partial ordering.

Exercise 3.21 Consider the set $\mathbb{R}^2$ of all ordered pairs of real numbers. Say that $(x, y)R(x', y')$ iff $x < x'$ and $y < y'$. Show that this is a strict partial ordering.

If a partial ordering satisfies *trichotomy*

$$\forall x{\in}X\ \forall y{\in}X\ (\langle x, y\rangle \in R \text{ or } \langle y, x\rangle \in R \text{ or } x = y)$$

then it is called a *linear ordering* or *total ordering*. For any relation R, we often write xRy instead of $\langle x, y\rangle \in R$.

In that notation, irreflexivity is $\forall x{\in}X\ (\neg xRx)$, transitivity is $\forall x{\in}X\ \forall y{\in}X\ \forall z{\in}X\ ((xRy) \text{ and } (yRz) \to (xRz))$, antisymmetry is $\forall x{\in}X\ \forall y{\in}X\ (xRy \to \neg yRx)$, and trichotomy is $\forall x{\in}X\ \forall y{\in}X\ (xRy \text{ or } yRx \text{ or } x = y)$.

If R is a (partial or total) ordering of X then we also say that X is (partially or totally) ordered by R.

Exercise 3.22 Let a, b, c be distinct sets. Show that $\mathcal{P}(\{a, b, c\})$ is *not* linearly ordered by the strict subset relation.

Exercise 3.23 Let a, b, c be distinct sets. Show that $\{\emptyset, \{a\}, \{a, b\}, \{a, b, c\}\}$ *is* linearly ordered by the strict subset relation.

Exercise 3.24 The partial ordering in Exercise 3.21 is *not* a linear ordering.

[9] It is also common to define a *weak partial ordering* as a relation that is reflexive

$$\forall x{\in}X\ (\langle x, x\rangle \in R)$$

transitive

$$\forall x{\in}X\ \forall y{\in}X\ \forall z{\in}X\ ((\langle x, y\rangle \in R \text{ and } \langle y, z\rangle \in R) \to \langle x, z\rangle \in R)$$

and antisymmetric in the slightly different sense that

$$\forall x{\in}X\ \forall y{\in}X\ ((\langle x, y\rangle \in R \text{ and } \langle y, x\rangle \in R) \to y = x).$$

Exercise 3.25 The orderings standardly represented by "$<$" on $\mathbb{N}$ (the set of natural numbers), $\mathbb{Z}$ (the set of integers), $\mathbb{Q}$ (the set of rational numbers), and $\mathbb{R}$ (the set of real numbers) are linear orderings.

Exercise 3.26 If α is an ordinal, then $\in$ is a linear ordering on α. (Hint: Transitivity follows from the definition of an ordinal, irreflexivity and antisymmetry follow from Axiom 6, and trichotomy follows from Exercise 3.17.)

If X is ordered by R, then an *initial segment* of X is a set $Y \subseteq X$ such that $\forall x{\in}X(\exists y(y \in Y \text{ and } xRy) \rightarrow x \in Y)$. Equivalently $\forall y{\in}Y(\forall x{\in}X(xRy \rightarrow x \in Y)$. The definition gives meaning to the word 'initial'—an initial segment is a subset that doesn't leave out anything earlier than anything it contains.

Exercise 3.27 If $Y \subseteq X$ and Y is *not* an initial segment of X then there are $x, y \in X$ with xRy, $y \in Y$, and $x \notin Y$.

Exercise 3.28 For any real number x, the set of rational numbers less than x is an initial segment of $\mathbb{Q}$, and these sets are distinct.

In fact, Richard Dedekind *defined* $\mathbb{R}$ to be the set of non-trivial initial segments of $\mathbb{Q}$ — though again, many philosophers object to the idea that these numbers can be identified with sets.

Exercise 3.29 $\emptyset$ is an initial segment of any ordering.

Exercise 3.30 For any two ordinals, one is an initial segment of the other under the $\in$ ordering.

3.4 Well-Orderings

Cantor's initial approach to ordinals came through studying well-orderings.

A linear ordering R of X is called a *well-ordering* iff for any non-empty $Y \subseteq X$ there is an $x \in Y$ such that $\neg\exists z{\in}Y(zRx)$. That is, every non-empty subset of the ordering has a least element.

Exercise 3.31 Which of these sets are well-ordered by the $<$ relation?

- $\mathbb{N}$ (the set of natural numbers, $\{0, 1, 2, \dots\}$)
- $\mathbb{Z}$ (the set of *all* integers, $\{\dots, -1, 0, 1, \dots\}$)
- $\mathbb{Q}^{\geq 0}$ (the set of non-negative rational numbers)
- A set consisting of 17 distinct rational numbers.

Exercise 3.32 Let $E \subseteq \mathbb{N} \times \mathbb{N}$ be the set of all pairs $\langle m, n\rangle$, where m, n are both even and $m < n$. Let $O \subseteq \mathbb{N} \times \mathbb{N}$ be the set of all pairs $\langle m, n\rangle$, where m, n are both odd and $m < n$. Let $X \subseteq \mathbb{N} \times \mathbb{N}$ be the set of all pairs $\langle m, n\rangle$, where m is even and n is odd. Let $R = E \cup O \cup X$. Show that R is a well-ordering of $\mathbb{N}$. What happens if we use $E \cup O^{-1} \cup X$ instead?

Exercise 3.33 Let $S \subseteq \mathbb{N} \times \mathbb{N}$ be the set of all pairs $\langle m, n \rangle$, where there exists a natural number k with m not divisible by 2^k but n is. Let T be the set of all pairs $\langle m, n \rangle$ where m and n are divisible by the same powers of 2 and $m < n$. Intuitively, m comes before n if n is "more even" than m, or they are "equally even" and $m < n$. Show that R is a well-ordering of $\mathbb{N} \times \mathbb{N}$. What happens if we use $S^{-1} \cup T$ instead?

Exercise 3.34 Any ordinal is well-ordered by the $\in$ relation. (Hint: Use several earlier Exercises and Axiom 6 (Foundation).)

Some sources actually use this as a definition of ordinals — that an ordinal is a transitive set whose elements are well-ordered by $\in$. It is straightforward to prove that if a set is transitive, and its elements are well-ordered by $\in$, then its elements are transitive as well, so that the two definitions are equivalent. This alternate definition also allows us to eliminate the many uses of the Axiom of Foundation that we have had so far

Exercise 3.35 Every non-empty set of ordinals has a least element under the $\in$ relation; that is, the collection of all ordinals is in fact well-ordered. (Hint: Use Exercise 3.19.)

Exercise 3.36 (Burali-Forti Paradox) No set contains all the ordinals. (Hint: use Exercise 2.24.)

Exercise 3.37 If α and β are ordinals, then $\alpha \subseteq \beta$ iff $\alpha \in \beta$ or $\alpha = \beta$.

From now on, we will use '$\in$' and '$<$' interchangeably for ordinals and their elements (which are also ordinals), and use words like 'less' and 'smaller' and 'earlier' for it. Using Exercise 3.37, we can also see that '$\subseteq$' and '$\leq$' are interchangeable for ordinals as well.

Exercise 3.38 If X is an initial segment of the ordinals then X is an ordinal. (Hint: show that there is some ordinal $\alpha \notin X$, that $X \subseteq \alpha$, and that X is an initial segment of α.)

Exercise 3.39 If α is an ordinal, then $S(\alpha)$ (that is, $\alpha \cup \{\alpha\}$) is the least ordinal greater than α, justifying the use of the term 'successor'.

Exercise 3.40 If X is a non-empty set of ordinals, then $\bigcap X$ is the least ordinal in X, and $\bigcup X$ is the least ordinal at least as large as every element of X (the least upper bound of X). (Hint: For the second part, show that $\bigcup X$ is an ordinal, is at least as large as every element of X, and that every ordinal strictly less than it is strictly less than some element of X.)

3.5 The Structure of Well-Orderings

Now we will show some more general results about well-orderings. If X is well-ordered by R and Y is well-ordered by S, we will say that $\langle X, R \rangle$ is *isomorphic* to $\langle Y, S \rangle$ (written $\langle X, R \rangle \cong \langle Y, S \rangle$) iff there is a bijection F between X and Y, such that for any $x, x' \in X$, xRx' iff $F(x)SF(x')$. This F is called an *isomorphism* between $\langle X, R \rangle$ and $\langle Y, S \rangle$.

Exercise 3.41 $\langle X, R\rangle \cong \langle X, R\rangle$

Exercise 3.42 If $\langle X, R\rangle \cong \langle Y, S\rangle$ then $\langle Y, S\rangle \cong \langle X, R\rangle$.

Exercise 3.43 If $\langle X, R\rangle \cong \langle Y, S\rangle$ and $\langle Y, S\rangle \cong \langle Z, T\rangle$ then $\langle X, R\rangle \cong \langle Z, T\rangle$.

Exercise 3.44 If $\langle X, R\rangle$ is a well-ordering and $a \in X$ then the set X_a given by $\{x \in X \mid xRa\}$ is an initial segment of X.

Exercise 3.45 If $\langle X, R\rangle$ is a well-ordering and Y is an initial segment of X then $Y = X$ or $Y = X_a$ for some $a \in X$. (Hint: Consider the least element *not* included in Y, if there is one.)

That is, the initial segments of a well-ordered set (other than the set as a whole) are in one-to-one correspondence with the members of the set.

Exercise 3.46 If $\langle X, R\rangle$ is a well-ordering and $a \in X$ then $\langle X, R\rangle$ is not isomorphic to $\langle X_a, R_a\rangle$, where R_a is just R restricted to the elements before a. (Hint: work by contradiction, assuming there *is* an isomorphism F, show that there is an x with $F(x) \neq x$, and consider the least such x.)

Exercise 3.47 If α and β are ordinals, and $\langle \alpha, <_\alpha\rangle \cong \langle \beta, <_\beta\rangle$, then $\alpha = \beta$.

Exercise 3.48 If F is an isomorphism from $\langle X, R\rangle$ to itself, then F is the identity. (Hint: Again, consider the least element moved by the isomorphism.)

Exercise 3.49 If $\langle X, R\rangle$ and $\langle Y, S\rangle$ are isomorphic well-orderings, then the isomorphism between them is unique. (Hint: If F and G are two distinct isomorphisms, then consider $F \circ G^{-1}$.)

Exercise 3.50 If $F\colon X \longrightarrow Y$ is an isomorphism of $\langle X, R\rangle$ to $\langle Y, S\rangle$, then for any $a \in X$, the restriction of F to X_a is an isomorphism of $\langle X_a, R_a\rangle$ to $\langle Y_{F(a)}, S_{F(a)}\rangle$.

Exercise 3.51 If $\langle X, R\rangle$ and $\langle Y, S\rangle$ are well-orderings, and $\langle X_x, R_x\rangle \cong \langle Y_y, S_y\rangle$ and $\langle X_{x'}, R_{x'}\rangle \cong \langle Y_{y'}, S_{y'}\rangle$, then xRx' iff ySy'.

Exercise 3.52 Every initial segment of a well-ordering $\langle X, R\rangle$ is isomorphic to a unique ordinal, and therefore $\langle X, R\rangle$ is isomorphic to a unique ordinal. (Hint: let a be the least element such that X_a is *not* isomorphic to a unique ordinal, and consider the set of ordinals that are isomorphic to the earlier initial segments. Constructing this *set* of ordinals is our first essential use of the Axiom of Replacement. Then apply the previous two exercises to show that this set is an initial segment of the ordinals, and that the association is an isomorphism.)

This was Cantor's initial characterization of ordinals—as things that every well-ordered set is isomorphic to. Von Neumann's contribution was to notice that the transitive sets of transitive sets provide a unique set isomorphic to each well-ordering.

Exercise 3.53 If $\langle X, R\rangle$ and $\langle Y, S\rangle$ are well-orderings, then either they are isomorphic, or one is isomorphic to an initial segment of the other.

3.6 Introduction to Ordinal Arithmetic

We now begin to investigate the ways that ordinals behave like numbers, and generalize arithmetic operations to apply to them. Several problems from this and later subsections of Section 3 are taken from Chapter 1 of Kenneth Kunen's *Set Theory* (1980, Amsterdam: Elsevier).

Exercise 3.54 ω is the least ordinal other than $\emptyset$ that is not the successor of any ordinal. (Hint: This is basically the same as Exercise 2.74.)

Exercise 3.55 If β is an ordinal that is the successor of an ordinal α then $\bigcup \beta = \alpha$. (That is, $\bigcup S(\alpha) = \alpha$.)

Exercise 3.56 If λ is an ordinal that is not the successor of any ordinal, then $\bigcup \lambda = \lambda$. (Hint: Consider $\alpha \in \lambda$ and figure out whether $S(\alpha)$ is less than, greater than, or equal to λ.)

Exercise 3.57 $\bigcup \omega = \bigcup \{\omega\} = \bigcup S(\omega) = \omega$.

Exercise 3.58 If X and Y are sets of ordinals, and for every $\alpha \in X$ there is a $\beta \in Y$ with $\alpha \leq \beta$, then $\bigcup X \leq \bigcup Y$. (You might consider whether Exercise 3.40 is helpful.)

Exercise 3.59 In the previous exercise equality is possible, even if every element of Y is strictly greater than every element of X. (Hint: You can pick X and Y from the exercise before it.)

Because of these results, we classify ordinals into three types: $\emptyset$ (also known as 0), successor ordinals, and all others, which are called *limit ordinals*. We will often show that some statement is true of all ordinals by considering the least ordinal for which it is false, and showing that this ordinal can't be 0, a successor, or a limit.

ω is the first limit ordinal. But we will prove the existence of others shortly.

To simplify notation, we refer to $S(0)$ as 1, $S(1)$ as 2, and so on. To make the convention more explicit, we use variables like m and n to range over members of ω, rather than α and β, which range over all ordinals.

3.7 Ordinal Addition

Recall from Exercise 2.65 that $X \otimes y = \{\langle x, y\rangle \mid x \in X\}$. If $\langle X, R\rangle$ and $\langle Y, S\rangle$ are well-orderings, we define $X \oplus Y = X \otimes 0 \cup Y \otimes 1$ with the ordering given by $\langle a, b\rangle < \langle a', b'\rangle$ iff $((b = 0$ and $b' = 1)$ or $(b = b' = 0$ and $aRa')$ or $(b = b' = 1$ and $aSa'))$. That is, it is the ordering created by making disjoint copies of X and Y, and putting the copy of X before the copy of Y.

Exercise 3.60 If $\langle X, R\rangle$ and $\langle Y, S\rangle$ are well-orderings, then $\langle X \oplus Y, <\rangle$, defined as above, is as well.

Using Exercise 3.52, we see that any well-ordered set is isomorphic to a unique ordinal. Thus, for ordinals α and β, we can define $\alpha + \beta$ to be the unique ordinal isomorphic to $\alpha \oplus \beta$ under this ordering. You should be able to convince yourself at this point that this definition, applied to finite ordinals, agrees with the ordinary concept of addition you are already familiar with. The next three results show that even beyond the finite, ordinal addition behaves a lot like normal addition of natural numbers, but differences start to show up beyond there.

Exercise 3.61 The ordering in Exercise 3.32 is isomorphic to $\omega + \omega$.

Note that the use of Exercise 3.52 means that this result depends on Axiom 8 (Replacement). In Exercise 6.40, we show that the existence of the ordinal $\omega + \omega$ can't be proven in any way using just the other axioms, so that this dependence is essenntial.

Exercise 3.62 $\alpha + (\beta + \gamma) = (\alpha + \beta) + \gamma$. (Hint: Find an explicit isomorphism between $\alpha \oplus (\beta \oplus \gamma)$ and $(\alpha \oplus \beta) \oplus \gamma$.)

Exercise 3.63 $\alpha + 0 = 0 + \alpha = \alpha$.

Exercise 3.64 $\alpha + S(\beta) = S(\alpha + \beta)$ (and in particular, $\alpha + 1 = S(\alpha)$).

Exercise 3.65 If β is a successor ordinal, then $\alpha + \beta$ is too.

Exercise 3.66 If λ is a limit ordinal, then $\alpha + \lambda = \bigcup\{\alpha + \beta \mid \beta < \lambda\}$. (Hint: recall Exercise 3.56, so that $\alpha \oplus \lambda = \alpha \oplus \bigcup\{\beta \mid \beta < \lambda\}$.)

Exercise 3.67 If λ is a limit ordinal, then $\alpha + \lambda$ is too. (Hint: Show that for anything less than $\alpha + \lambda$, its successor is also less than $\alpha + \lambda$.)

Exercise 3.68 If α is a successor ordinal and β is a limit ordinal, then $\alpha + \beta \neq \beta + \alpha$.

Exercise 3.69 If $\alpha < \beta$ then $\gamma + \alpha < \gamma + \beta$. (Hint: Show that $\gamma \oplus \alpha$ is a proper initial segment of $\gamma \oplus \beta$ and use Exercise 3.46.)

Exercise 3.70 If $\alpha < \beta$ then $\alpha + \gamma \leq \beta + \gamma$. (Hint: For any fixed α and β consider the least ordinal γ for which this is false, and show that it can't be 0, a successor, or a limit. To show that it's not a limit, you will use Exercise 3.58.)

Exercise 3.71 Equality is possible in the previous problem. (Hint: Consider what happens with Exercise 3.59.)

Exercise 3.72 If $\alpha \leq \beta$ then there is a unique δ such that $\alpha + \delta = \beta$. (Hint: Show that there is some γ such that $\alpha + \gamma \geq \beta$, and then consider the least such ordinal.)

Pause for a moment to consider the beauty of ordinals like $\omega + \omega$, $\omega + \omega + 3$, $\omega + \omega + \omega$, and so on. (Why don't I mention $\omega + 3 + \omega$ or $3 + \omega + \omega$? Why don't I bother with parentheses for this binary operation of addition?)

3.8 Ordinal Multiplication

If $\langle X, R\rangle$ and $\langle Y, S\rangle$ are well-orderings, we can define a relation on $X \times Y$ by saying that $\langle a, b\rangle < \langle a', b'\rangle$ iff bSb' or $b = b'$ and aRa'. That is, it is the lexicographic (alphabetical) ordering where the second component of the pair is more significant than the first.

Exercise 3.73 Show that the relation so defined is a well-ordering of $X \times Y$.

For ordinals α and β, we define $\alpha \cdot \beta$ as the unique ordinal isomorphic to $\alpha \times \beta$ under this ordering. This ordering corresponds to one copy of α for each element of β. Again, you should be able to convince yourself that this operation applied to finite ordinals corresponds exactly to the concept of multiplication that you are familiar with for natural numbers.

Exercise 3.74 $\alpha \cdot 2 = \alpha + \alpha$.

Exercise 3.75 The ordering in Exercise 3.33 is isomorphic to $\omega \cdot \omega$.

The next four problems show that ordinal multiplication even beyond the finite behaves in some ways like normal multiplication of natural numbers, but differences start to show up beyond there.

Exercise 3.76 $\alpha \cdot (\beta \cdot \gamma) = (\alpha \cdot \beta) \cdot \gamma$.

Exercise 3.77 $\alpha \cdot 0 = 0 \cdot \alpha = 0$.

Exercise 3.78 $\alpha \cdot S(\beta) = (\alpha \cdot \beta) + \alpha$.

Exercise 3.79 If α and β are both successor ordinals, then so is $\alpha \cdot \beta$.

Exercise 3.80 $\alpha \cdot (\beta + \gamma) = (\alpha \cdot \beta) + (\alpha \cdot \gamma)$.

Exercise 3.81 If λ is a limit ordinal, then $\alpha \cdot \lambda = \bigcup\{\alpha \cdot \beta \mid \beta < \lambda\}$. (Again, recall that $\lambda = \bigcup\{\beta \mid \beta < \lambda\}$, so that it suffices to prove that $\alpha \cdot \bigcup\{\beta \mid \beta < \lambda\} = \bigcup\{\alpha \cdot \beta \mid \beta < \lambda\}$.)

Exercise 3.82 If λ is a limit ordinal, and $\alpha \neq 0$, then $\alpha \cdot \lambda$ is a limit ordinal. (Hint: Use the previous exercise to show that if $\beta < \alpha \cdot \lambda$, then so is $S(\beta)$.)

Exercise 3.83 If λ is a limit ordinal, and $\beta \neq 0$, then $\lambda \cdot \beta$ is a limit ordinal. (Hint: Consider the case where β is a successor and β is a limit separately.)

Exercise 3.84 There are α and β such that $\alpha \cdot \beta \neq \beta \cdot \alpha$. (Hint: This violation clearly isn't going to show up when α and β are both finite, or when they are equal, so think about what happens when one of them is finite and the other is the simplest limit ordinal.)

Exercise 3.85 There are α, β, and γ such that $(\alpha + \beta) \cdot \gamma \neq (\alpha \cdot \gamma) + (\beta \cdot \gamma)$.

Exercise 3.86 If $\gamma \neq 0$ and $\alpha < \beta$ then $\gamma \cdot \alpha < \gamma \cdot \beta$. (Hint: Note that $\gamma \times \alpha \subseteq \gamma \times \beta$ and is a proper initial segment of it under the ordering, and then use Exercise 3.46.)

Exercise 3.87 If $\alpha < \beta$ then $\alpha \cdot \gamma \leq \beta \cdot \gamma$. (Hint: Fix α, β and show that the least γ for which this is false is not 0, a successor, or a limit.)

Exercise 3.88 Equality is possible in the previous problem (even with $\gamma > 0$).

Exercise 3.89 If $\alpha > 0$ and β is any ordinal, then there are unique δ and ξ such that $\xi < \alpha$ and $\alpha \cdot \delta + \xi = \beta$. (Hint: show that there is a γ such that $\alpha \cdot \gamma + \alpha > \beta$ and let δ be the least such ordinal, and then apply Exercise 3.72.)

Exercise 3.90 If $\alpha < \omega \cdot \omega \cdot \omega$, then α can be written uniquely as $\omega \cdot \omega \cdot a + \omega \cdot b + c$ for some $a, b, c \in \omega$.

Exercise 3.91 Write the ordinals $(\omega + 2) \cdot (\omega + 3)$ and $(\omega + 3) \cdot (\omega + 2)$ in this form.

Pause again to consider the even greater beauty of ordinals like $\omega \cdot 3$, $\omega \cdot \omega$, $\omega \cdot \omega \cdot 4 + \omega \cdot 5 + 17$, $\omega \cdot \omega \cdot \omega$, and so on.

4 Transfinite Induction and Recursion

Several proofs that have been given already are proofs by what is called *transfinite induction*. Such a proof works to show that all elements of some well-ordering have a property by showing there can be no least element in the ordering without the property. Thus, since every non-empty subset of the well-ordering has a least element, we see that the subset consisting of the elements without the property must be empty.

Exercise 4.1 (Transfinite induction on ordinals.) Show that if there is no smallest ordinal α such that $\varphi(\alpha)$, then there is no ordinal α such that $\varphi(\alpha)$.

Exercise 4.2 Assume that $\varphi(0)$, that whenever $\varphi(\alpha)$ then $\varphi(S(\alpha))$, and that if $\varphi(\beta)$ for every $\beta \in X$ then $\varphi(\bigcup X)$. Prove that φ holds of every ordinal. (Hint: Consider whether the smallest ordinal for which φ *doesn't* hold is 0, a successor, or a limit.)

For the natural numbers, we showed in Exercise 2.82 that in addition to doing proof by induction, we can define functions by recursion — if there is a way to calculate the value of a function at 0, and to calculate the value at the next input from the value at the previous input, then there is a set that encodes all values of the function. The next several exercises will show an analogous thing for ordinals more generally.

Definition by recursion on the natural numbers used two separate clauses for 0 and successor steps. For ordinals, we will first do a version where we add a third clause for limit steps. We then show it is also possible to use a process that defines the value of the function at a given ordinal provided the values at *all* previous ordinals are given.

Let α be any ordinal, and X be any set. Let x be any member of X and g be a total function from X to X.

Exercise 4.3 There is a set of all partial functions from α to X.

Exercise 4.4 There is a set of all partial functions from α to X whose domain is some member of α.

Exercise 4.5 There is a set of all partial functions from α to X whose domain is some member of α and with $f(0) = x$.

Exercise 4.6 There is a set of all partial functions from α to X whose domain is some member of α, with $f(0) = x$, and for every successor ordinal β in the domain of f, if γ is the predecessor of β, then $f(\gamma) = g(f(\beta))$.

Exercise 4.7 There is a set of all partial functions from α to X whose domain is some limit ordinal member of α, with $f(0) = x$, and for every successor ordinal β in the domain of f, if γ is the predecessor of β, then $f(\gamma) = g(f(\beta))$.

Let ℓ be any function from this set to X. We can think of ℓ as saying how the value of the function at a limit ordinal will depend on the value of the function at all ordinals up to that limit ordinal. For f any function defined on an ordinal β, and $\gamma < \beta$, define $f \restriction \gamma$ to be the set $\{\langle x, y\rangle \in f \mid x \in \gamma\}$ — that is, it is the restriction of f to the domain γ.

Exercise 4.8 There is a set of all partial functions from α to X whose domain is some member of α, with $f(0) = x$, and for every successor ordinal β in the domain of f, if γ is the predecessor of β, then $f(\gamma) = g(f(\beta))$, and additionally, for every limit ordinal λ in the domain of f, $f(\lambda) = \ell(f \restriction \lambda)$.

Exercise 4.9 Every member of α is the domain of some function in the set from the previous exercise. (Hint: Work by induction. Show that the least member of α that is not the domain of any such function can't be 0, or a successor, or a limit ordinal.)

Exercise 4.10 The union of the previous set of partial functions is in fact the desired total function from α to X. (Hint: The main thing to prove here is that this union is in fact a *function*, i.e., that two different outputs don't both get associated with the same input by different members of this set.)

At this point, we have shown how to create a set function from any ordinal to X that encodes the function we want up to that ordinal. Now we show that this can be extended to a class function defined on all ordinals.

Exercise 4.11 There is a formula $\varphi(\beta, x)$ that holds iff there is some $\alpha > \beta$, and some f defined in the above way on α, with $f(\beta) = x$, and this defines the desired class function for all ordinals.

Just as in the finite case, we can generalize this to define functions of multiple ordinals, and we can see that Exercises 3.63, 3.64, 3.66 can be used to define addition recursively, and that Exercises 3.77, 3.78, 3.81 can be used to define multiplication recursively.

As mentioned above, we can also do this more generally. Instead of using x for the 0 step, g for the successor step, and ℓ for the limit step, we can just use ℓ for all steps, if we ask that it be defined not just for partial functions whose domain is a limit ordinal, but for partial functions whose domain is *any* member of α. (Note that any function with domain 0 is in fact the empty function!)

Exercise 4.12 With ℓ as just generalized, there is a set of all partial functions from α to X whose domain is some member of α, with $f(\beta) = \ell(f \restriction \beta)$ for all β in its domain. Every member of α is the domain of some such function, and the union of these functions is in fact the desired total function from α to X.

Exercise 4.13 Use the above construction to define a function on ordinals such that $\alpha^0 = 1$, $\alpha^{S(\beta)} = \alpha^\beta \cdot \alpha$, and $\alpha^\lambda = \bigcup\{\alpha^\beta \mid \beta < \lambda\}$ if λ is a limit ordinal. That is, write down the equivalent of ψ and φ from above, and use them to give the definition of $y = \alpha^x$.

Clearly this operation agrees with ordinary exponentiation when α and β are natural numbers (apart from some ambiguity about how to think of '0^0'). However, at infinite levels, it behaves differently. (It also has important differences from Cantor's cardinal exponentiation, to be defined in Section 5).[10]

Exercise 4.14 $\alpha^\beta \cdot \alpha^\gamma = \alpha^{(\beta+\gamma)}$. (Hint: Prove this by induction on γ, as in Exercise 4.2. You will use Exercises 3.63, 3.64, and 3.66, as well as 3.76 and 3.81.)

Exercise 4.15 $(\alpha^\beta)^\gamma = \alpha^{(\beta\cdot\gamma)}$.

Exercise 4.16 If $\alpha > \beta$ and $\gamma > 1$ then $\gamma^\alpha > \gamma^\beta$. (Hint: Write $\alpha = \beta + \delta$.)

Exercise 4.17 If $\alpha > \beta$ then $\alpha^\gamma \geq \beta^\gamma$. (Hint: Fix α, β and prove this by induction on γ.)

Exercise 4.18 Equality is possible in the previous problem, even with $\gamma > 0$.

Exercise 4.19 If $\alpha, \beta > 1$, then $\alpha^\beta > \alpha$ and $\alpha^\beta \geq \beta$.

Exercise 4.20 There are $\alpha, \beta > 1$ such that $\alpha^\beta = \beta$.

[10] There is a combinatorial way to characterize α^β as a well-ordering, but it is slightly awkward to characterize, so it is easier to just explicitly define it by transfinite recursion. But here is the characterization, if you are interested.

Let $F = \{f\colon \beta \longrightarrow \alpha \mid$ For all but finitely many γ, $f(\gamma) = 0\}$. For two such functions f_1 and f_2, let γ be the greatest such that $f_1(\gamma) \neq f_2(\gamma)$ and define $f_1 < f_2$ iff $f_1(\gamma) < f_2(\gamma)$. We need to use the fact that both functions are zero at all but finitely many points in order to prove that there *is* a greatest such γ. It is moderately annoying to prove that this is a well-ordering.

It is straightforward to prove that if $\beta = 0$ then there is only one such function (namely, the empty function), and that taking the successor of β is equivalent to taking α many copies of this well-ordering in sequence. Proving that this function is continuous on limits is another point where it becomes essential that the function is non-zero on only finitely many inputs.

Exercise 4.21 *(Cantor Normal Form) Use transfinite induction to show that every ordinal can be written uniquely as a finite sum:

$$\omega^{\alpha_0} \cdot m_0 + \cdots + \omega^{\alpha_n} \cdot m_n$$

where $\alpha_0 > \cdots > \alpha_n$ and $m_i < \omega$ for each i. (Hint: To show that a limit of ordinals representable in this form can also be written in this form, consider the set of ordinals that appear as α_0 in these expansions and take the least upper bound α of this set. If none of the individual ordinals has $\alpha_0 = \alpha$ then show that the limit must equal ω^α, and otherwise restrict attention to the ones with $\alpha_0 = \alpha$ and show that they must have the same limit. Then in this case, consider the set of ordinals that appear as m_0 in these expansions and take its least upper bound m. If none of the individual ordinals has $m_0 = m$ then show that $m = \omega$ and the limit is $\omega^{S(\alpha)}$, and otherwise restrict attention to the ones with $m_0 = m$ and show that they must have the same limit. Then show that a minimal counterexample can't be a limit of ordinals all written with the same first term, because dropping that term would provide a smaller counterexample.)

Exercise 4.22 Write $(\omega^2 \cdot 2 + \omega \cdot 3 + 4) \cdot (\omega^\omega + \omega^3 + 5)$ and $(\omega^\omega + \omega^3 + 5) \cdot (\omega^2 \cdot 2 + \omega \cdot 3 + 4)$ in Cantor Normal Form.

This sounds like it will give us a nice notation scheme for any ordinal — we write it in "base ω", and do the same for the exponents on ω. This works as long as the exponent on ω is less than the ordinal we are trying to write in this notation, but as the next exercise shows, there are ordinals for which this fails — the exponent needed on ω is sometimes equal to the ordinal itself that we are trying to notate.

Exercise 4.23 Show that for every α there is an $\beta \geq \alpha$ such that $\omega^\beta = \beta$. (Hint: consider the limit of the sequence $\alpha, \omega^\alpha, \omega^{(\omega^\alpha)}, \ldots$.)

Such ordinals are called *epsilon numbers*. Because of well-ordering we can use transfinite recursion to define ε_β to be the first epsilon number strictly greater than ε_α for all $\alpha < \beta$. ε_0 is the first number of this sort, namely $\omega^{\omega^{\omega^{\omega^{\cdots}}}}$, but there are more.

Exercise 4.24 If λ is a limit ordinal, then $\epsilon_\lambda = \bigcup\{\epsilon_\beta \mid \beta < \lambda\}$. (Hint: The important point is to prove that this union is in fact an epsilon number.)

We call a function f on ordinals *increasing* iff $\forall\alpha\forall\beta(\alpha \geq \beta \to f(\alpha) \geq f(\beta))$. We call an increasing function *continuous* if $f(\lambda) = \bigcup\{f(\beta) \mid \beta < \lambda\}$ whenever λ is a limit ordinal. This concept of continuity is not exactly the same as the one that might be familiar from real analysis, but it is in fact an application of the same general concept from topology. To say a function is continuous is to say that if $x_0, x_1, x_2, \ldots$ converges to a limit x, then $f(x_0), f(x_1), f(x_2), \ldots$ must also converge to a limit, which must in fact equal $f(x)$. If you are familiar with the 'order topology', then this is exactly the concept of continuity that is being used.

Exercise 4.25 Addition, multiplication, and exponentiation of ordinals are increasing in both arguments, but only continuous in the second.

Exercise 4.26 Every continuous, increasing function has arbitrarily large fixed points. That is, for any α there is $\beta \geq \alpha$ such that $f(\beta) = \beta$. (Note that Exercises 3.71, 3.88, 4.18, and particularly 4.23 are special instances of this.)

Exercise 4.27 The function $f(\alpha) = \varepsilon_\alpha$ is a continuous, increasing function, so there is a β such that $\varepsilon_\beta = \beta$.

Pause to try to imagine what ϵ_0, ϵ_1, ϵ_ω, ϵ_{ϵ_0} and other epsilon numbers are like, and then let the unimaginably sublime beauty of this fixed point of the epsilon function wash over you.

5 Cardinals

Cantor's 'cardinals' play the more familiar role of saying when there are the 'same number' of elements in two sets, rather than tracking the ordinal structure of the sets. But some of the structure of cardinals depends on the Axiom of Choice, and we will start to apply this axiom here. Any exercise that requires either direct application of the Axiom of Choice, or of an earlier exercise that uses it, will henceforth be marked '(AC)'. Exercises that aren't marked this way should be proved without using any form of it.

Recall, from Section 2.5, that a (partial) (set) function $f\colon X \longrightarrow Y$ is a set $f \subseteq X \times Y$, where left elements of the pairs are interpreted as 'inputs' and right elements are interpreted as 'outputs', and each input has at most one output. A function is total on set X iff there is exactly one output associated with each member of X. The function is injective iff no two inputs are associated with the same output, it is surjective on Y iff every element of Y is the output associated to at least one input, and it is bijective (a 'bijection between X and Y') iff it is total on X, injective, and surjective on Y.

$f^{-1}\colon Y \longrightarrow X$, the inverse of the function, is the set of ordered pairs in the reverse order. f^{-1} is a function iff f is injective, f^{-1} is surjective on X iff f is total on X, and f^{-1} is total on Y iff f is surjective on Y.

We say that sets X and Y are *equinumerous* or *have the same cardinality*, and write $|X| = |Y|$ iff there is a bijection between X and Y.[11]

Exercise 5.1 $|X| = |X|$.

Exercise 5.2 If $|X| = |Y|$ then $|Y| = |X|$.

[11] Note that a 'bijection' here is considered as a set of ordered pairs in $X \times Y$. Thus, the existence or non-existence of a bijection is something that can be stated in the language, by means of quantifying over the sets. However, there is no need for this bijection to be able to be expressed in any simple way within the language. In Section 6, when we begin to talk about models of set theory, there will be some models that have different sets of ordered pairs in them than others, and this could change whether or not certain pairs of sets are equinumerous.

Exercise 5.3 If $|X| = |Y|$ and $|Y| = |Z|$ then $|X| = |Z|$.

Exercise 5.4 If $|X| = |Y|$ and $|X'| = |Y'|$, and $X \cap X' = Y \cap Y' = \emptyset$ then $|X \cup X'| = |Y \cup Y'|$. (Hint: Think about the bijections $f\colon X \longrightarrow Y$ and $g\colon X' \longrightarrow Y'$ and how to use them to construct a bijection $h\colon (X \cup X') \longrightarrow (Y \cup Y')$.)

We define $|X| \leq |Y|$ iff there is a total injection from X to Y (equivalently, iff there is a bijection between X and a subset of Y). The Cantor-Schröder-Bernstein theorem states that if $|X| \leq |Y|$ and $|Y| \leq |X|$ then $|X| = |Y|$, and is an important result for justifying this notation, but it is a bit tricky to prove.[12]

Exercise 5.5 $|\mathbb{N}| = |\mathbb{Z}| = |\mathbb{Q}| = |\omega| = |\omega + \omega| = |\omega \cdot \omega|$.

5.1 Cardinals and the Axiom of Choice

This last exercise shows that some sets are equinumerous with proper subsets of themselves. Richard Dedekind observed that this is impossible for finite sets (which we will show in a moment), and tried to use this to define what it is to be infinite. However, modern set theorists have noted that there is a slight problem here.

We say that a set is *finite* if it has a bijection to some member of ω, and *infinite* otherwise. We say that a set X is *Dedekind-infinite* if there is a total function $f\colon X \longrightarrow X$ that is injective but not surjective (equivalently, there is a bijection between X and some proper subset of X), and *Dedekind-finite* otherwise.

Exercise 5.6 If $|X| = |Y|$ then X is finite iff Y is finite.

Exercise 5.7 If $|X| = |Y|$ then X is Dedekind-finite iff Y is Dedekind-finite.

Exercise 5.8 If X is finite, then it is Dedekind-finite. (Hint: Prove, by induction, that each member of ω is Dedekind-finite and conclude that every finite set is.)

We would like to prove the converse, but the next few exercises end up needing the Axiom of Choice to show this. It turns out that this is essential, though showing that it is essential is beyond the scope of these notes.

Exercise 5.9 ω is infinite and Dedekind-infinite.

As a reminder, exercises marked (AC) require the use of the Axiom of Choice (either directly, or indirectly through the use of some other exercise marked (AC)) while exercises not marked (AC) should be done without appealing to it directly or indirectly.

[12] Let $f\colon X \longrightarrow Y$ and $g\colon Y \longrightarrow X$ be injections. Let $X_0 = \{x \in X | \neg\exists y \in Y (g(y) = x)\}$ and $Y_0 = \{y \in Y \mid \neg\exists x \in X (f(x) = y)\}$. For each n, let X_{n+1} be the result of applying Replacement to g on Y_n, and let Y_{n+1} be the result of applying Replacement to f on X_n. We can check that the X_n are pairwise disjoint, and so are the Y_n. Let $X_E = X_0 \cup X_2 \cup X_4 \cup \ldots$, $X_O = X_1 \cup X_3 \cup X_5 \cup \ldots$, $Y_E = Y_0 \cup Y_2 \cup Y_4 \cup \ldots$, and $Y_O = Y_1 \cup Y_3 \cup Y_5 \cup \ldots$. Then f shows that $|X_E| = |Y_O|$ and g shows that $|Y_E| = |X_O|$, and then the previous exercise shows that $|X| = |Y|$.

Exercise 5.10 X is Dedekind-infinite iff $|\omega| \leq |X|$. (Hint: For one direction, let f be a function on X that is injective but not surjective, and try to find a sequence of elements of this set to put in bijection with ω. For the other direction, use the successor function on ω to create f.)

Exercise 5.11 (AC) Let X be an infinite set. Let f be a choice function on the set of non-empty subsets of X. Define $g\colon \omega \longrightarrow X$ recursively, by letting $g(0) = f(X)$ and $g(S(n)) = f(X \setminus \{g(m) \mid m < S(n)\})$. Show that g is a total, injective function from ω into X.

Putting these last two exercises together we see that the Axiom of Choice implies that any infinite set is in fact Dedekind-infinite, so the two definitions coincide.

Exercise 5.12 (AC) For any set X, use transfinite recursion to define a bijective function from some initial segment of the ordinals to X. (Hint: Continue the above process to limit ordinals. Show that there must be some α such that $\{g(\beta) \mid \beta < \alpha\} = X$.))

Exercise 5.13 (AC) (Well-Ordering Principle) Every set can be well-ordered.

This result, that every set can be well-ordered, was first "proved" by Cantor in the 1870's, working intuitively with his conception of set. This result was so counterintuitive that many mathematicians were convinced that Cantor's intuitive conception of set had some hidden contradiction, and this belief became stronger as Russell's paradox for Frege's axioms became well-known. However, Ernst Zermelo was convinced Cantor was right, and formalized the proof. In this formalization, he listed all the axioms we have mentioned so far, except the various forms of Replacement (which were added later by Abraham Fraenkel, and aren't essential for the proof of the Well-Ordering Principle). Because all the other axioms seemed so clearly innocuous, this result set off several decades of argument about the Axiom of Choice.

Exercise 5.14 The Well-Ordering Principle implies Axiom 11 (the Axiom of Choice). (Hint: Well-order the union of a set of non-empty sets, and use this well-ordering to define a choice function.)

Thus, the Well-Ordering Principle is an equivalent to the Axiom of Choice, and can be used in place of it.

Exercise 5.15 (Injection Principle) If X is non-empty, and f is a total injection from X to Y, then there is a total surjection g from Y to X such that f is a right-inverse of g (i.e., $g(f(x)) = x$, for all $x \in X$). (Hint: Consider the inverse of the injection, and figure out how to make it total.)

Exercise 5.16 (AC) (Surjection Principle) If f a total surjection from X to Y then there is a total injection g from Y to X such that f is a left-inverse of g (i.e., $f(g(y)) = y$, for all $y \in Y$). (Hint: If there is a surjection $f\colon X \longrightarrow Y$ then the sets $X_y = \{x \in X \mid f(x) = y\}$ are non-empty sets.)

Exercise 5.17 Show that the Surjection Principle implies Axiom 12 (the Axiom of Disjoint Choices), and therefore the Axiom of Choice. (Hint: Start with a set of non-empty sets that are pairwise disjoint, and show that there is a surjection from the union of these sets to this set.)

Recall that we defined $|X| \leq |Y|$ iff there is an injection from X to Y. We might have naturally thought that we could also define this by stating there is a surjection from Y to X, but these results show that although it is true that if there is a surjection from Y to X, then $|X| \leq |Y|$, the converse is equivalent to the Axiom of Choice.

5.2 Increasing cardinality

We now define $|X| < |Y|$ to mean $|X| \leq |Y|$ (that is, there is an injection from X to Y, or equivalently, a bijection between X and a subset of Y) but $|Y| \not\leq |X|$ (that is, there is no injection from Y to X, or equivalently, no bijection between Y and a subset of X).

Exercise 5.18 $|X| < |\mathcal{P}(X)|$. (Hint: Use Exercise 2.51.)

Note that this notation strictly speaking is only about the existence of functions between sets, and doesn't define what a 'cardinality' is, or give an object for '$|X|$' to stand for. However, if X can be put in bijection with an ordinal, then we traditionally identify $|X|$ with the least such ordinal.

We define: α is a *well-ordered cardinal* iff α is an ordinal, and there is no bijection between α and any ordinal less than α.

The Well-Ordering Principle states that every set is equinumerous with some ordinal, and therefore all cardinalities can be represented by well-ordered cardinals. If we don't assume the Axiom of Choice, then we need to distinguish these 'well-ordered cardinals' from the more general class of cardinals. (There is a clever trick due to Dana Scott that allows us to find a set to represent each cardinality even *without* the Axiom of Choice, but describing this trick requires some concepts we won't discuss until Section 6.2.)

Just as we use lowercase Greek letters from the beginning of the alphabet as variables to represent infinite ordinals, it is customary to use Greek letters starting with κ as variables to denote infinite cardinals (well-ordered or otherwise). (λ is sometimes used for a general limit ordinal, and sometimes specifically for a cardinal, but hopefully context disambiguates things.)

Exercise 5.19 If κ and λ are well-ordered cardinals, then $\kappa < \lambda$ as ordinals iff $\kappa < \lambda$ as cardinals.

Exercise 5.20 If $\alpha < \beta$ as ordinals then $|\alpha| \leq |\beta|$ as cardinals, but equality is possible.

The next several exercises show that no set is larger than all the well-ordered cardinals. At each step, note that showing something is a subset of a power set essentially proves that the thing is itself a set that exists, by Axiom 7 (Separation).

Exercise 5.21 For any set X, any binary relation on X (including a well-ordering of any subset of X) is an element of $\mathcal{P}(\mathcal{P}(\mathcal{P}(X)))$.

If S is a set of well-orderings, then an *isomorphism class* of S is a set containing all and only the well-orderings in S that are isomorphic to a given well-ordering.

Exercise 5.22 For any set S of well-orderings, the set consisting of all the isomorphism classes of S is a subset of $\mathcal{P}(S)$.

Exercise 5.23 For any set X, the set $\aleph(X)$ ($\aleph$ is aleph, the first letter of the Hebrew alphabet) consisting of all isomorphism classes of well-orderings of subsets of X is an element of $\mathcal{P}(\mathcal{P}(\mathcal{P}(\mathcal{P}(X))))$. (It's possible I have the wrong number of iterations of the power set operation here, but the important thing is just showing that there *is* a number of iterations that suffices.)

Exercise 5.24 $\aleph(X)$ is well-orderable. (Hint: Consider the function that assigns to each isomorphism class the unique ordinal that is isomorphic to the members of the class.)

Exercise 5.25 $|\aleph(X)|$ is the least ordinal that is not equinumerous with any subset of X. (Hint: Show that the function in the previous exercise is an isomorphism to an ordinal, and that every ordinal equinumerous with a subset of X is less than this ordinal, and that no other ordinal is, and conclude that this ordinal is in fact a well-ordered cardinal.)

We showed in Exercise 3.36 that there is no set containing all the ordinals. With the Axiom of Choice, since every set can be well-ordered, this means that no set is larger than every ordinal, and also, for every set, there is an ordinal larger than it. This previous exercise now gives us Hartogs's Theorem — even without the Axiom of Choice, there is no set larger than every ordinal. But if a set is not well-orderable, then there is no ordinal larger than it, but at least there is an ordinal that is too large to be a subset of it.

With the Axiom of Choice, the power set of any ordinal gives us a set that has higher cardinality, which can be well-ordered, giving us a well-ordered cardinality strictly greater than the ordinal we started with. But even without the Axiom of Choice, we showed that iterating the power set operation enough times gives us a set that contains a well-orderable cardinality strictly greater than the ordinal we started with.

We might wonder how the power set cardinalities compare to these greater well-orderable cardinalities. One final implication of Hartogs's Theorem is that trichotomy for cardinals (that is, for every two sets X and Y, either $|X| < |Y|$ or $|X| = |Y|$ or $|X| > |Y|$) implies the Axiom of Choice (because no set is larger than every ordinal, and so if trichotomy holds, then every set must be smaller than some ordinal, and thus well-orderable). Thus, it could well be that the power set of an ordinal, and the first well-ordered cardinality greater than it, are incomparable with each other!

Using transfinite induction, define $\aleph_\alpha$ as follows: let $\aleph_0 = \omega$, let $\aleph_{S(\alpha)} = |\aleph(\aleph_\alpha)|$, and let $\aleph_\lambda = \bigcup\{\aleph_\beta \mid \beta < \lambda\}$ if λ is a limit ordinal.

Exercise 5.26 Every infinite well-ordered cardinal is $\aleph_\alpha$ for some ordinal α, and every $\aleph_\alpha$ is an infinite well-ordered cardinal. (Hint: For the first part, let κ be an infinite well-ordered cardinal, and use Replacement to create the set of all β such that $\aleph_\beta < \kappa$. Consider the least ordinal α not in this set. For the second part, only the limit ordinal stage needs any work.)

Exercise 5.27 There is a cardinal κ such that $\kappa = \aleph_\kappa$. (Hint: Show that the function that takes α to $\aleph_\alpha$ is a continuous, increasing function on the ordinals and apply Exercise 4.26.)

5.3 Cardinal Addition and Multiplication

We define addition and multiplication of cardinals similarly to how we defined them for ordinals. We let $\kappa + \lambda = |\kappa \otimes 0 \cup \lambda \otimes 1|$ and $\kappa \cdot \lambda = |\kappa \times \lambda|$. For well-ordered cardinals, these are the cardinalities of the corresponding ordinals, but these definitions also work for non-well-ordered cardinals (which we don't explicitly discuss).

Exercise 5.28 Cardinal addition and multiplication are in fact commutative, even though the ordinal operations are not.

Exercise 5.29 $\aleph_0 + \aleph_0 = \aleph_0$ as cardinals, but $\omega + \omega \neq \omega$ as ordinals, even though the cardinal $\aleph_0$ *is* the ordinal ω.

These results show that we have to be careful about whether addition and multiplication are applied to sets considered as ordinals or considered as cardinals. Thus, strictly speaking, we shouldn't have used the symbols '+' and '$\cdot$' for both ordinal and cardinal operations. We should have used one symbol for ordinal addition and a different symbol for cardinal addition. But it is standard to abuse notation this way, and write the same symbol for both, and hope that the reader understands which operation is intended by noticing whether we are writing things with early alphabet Greek letters, or with κ or Hebrew letters.

Exercise 5.30 $|\alpha| + |\beta| = |(\alpha + \beta)|$ and $|\alpha| \cdot |\beta| = |(\alpha \cdot \beta)|$, where we use cardinal operations on the left side and ordinal operations on the right side of each equation. (Hint: Show that if $|\alpha_1| = |\alpha_2|$ and $|\beta_1| = |\beta_2|$, then $|\alpha_1 + \beta_1| = |\alpha_2 + \beta_2|$ and $|\alpha_1 \cdot \beta_1| = |\alpha_2 \cdot \beta_2|$, where "+" and "$\cdot$" are read as the ordinal operations.)

Thus, although the ordinal and cardinal operations differ, they at least give results of the same cardinality.

Exercise 5.31 For any cardinals $\kappa \geq \lambda > 1$, the following inequalities hold: $\kappa \leq \kappa + \lambda \leq \kappa \cdot 2 \leq \kappa \cdot \lambda \leq \kappa \cdot \kappa$.

Exercise 5.32 For κ a well-ordered cardinal, show that the ordering on $\kappa \times \kappa$ given by $\langle \alpha, \alpha' \rangle < \langle \beta, \beta' \rangle$ iff $(\alpha \cup \alpha') < (\beta \cup \beta')$ or $((\alpha \cup \alpha') = (\beta \cup \beta')$ and $\alpha < \beta)$ or $((\alpha \cup \alpha') = (\beta \cup \beta')$, $\alpha = \beta$ and $\alpha' < \beta')$ is a well-ordering. (Hint: Consider finite κ and draw a picture of this ordering of the $\kappa \times \kappa$ square, to get a better

understanding of this ordering. The actual proof that this is a well-ordering is similar to the proof that the other ordering we have considered on $\kappa \times \kappa$ is a well-ordering.)

Exercise 5.33 Show by transfinite induction, using this well-ordering, that $\kappa \cdot \kappa = \kappa$ for all infinite well-ordered cardinals. (Hint: Consider the above ordering of $\kappa \times \kappa$, and show that it is isomorphic to the standard well-ordering on κ. To do this show that for any $\alpha, \beta < \kappa$, the pair $\langle \alpha, \beta \rangle$ has fewer than κ many predecessors.)

Using Exercise 5.31, we can see that $\max(\kappa, \lambda) \leq \kappa \cdot \lambda \leq \max(\kappa, \lambda) \cdot \max(\kappa, \lambda)$. Since Exercise 5.33 shows that $\max(\kappa, \lambda) \cdot \max(\kappa, \lambda) = \max(\kappa, \lambda)$, we see that all infinitary cardinal addition and multiplication of well-ordered cardinals (and thus, assuming AC, of all cardinals) is trivial — we just get the larger of the two cardinals involved.

Exercise 5.34 (AC) For infinite κ, if $|X| \leq \kappa$ and for each $x \in X$, $|x| \leq \kappa$, then $|\bigcup X| \leq \kappa$. (Hint: Consider the function f defined on X such that $f(x)$ is the set of well-orderings of x. Apply Replacement to this function, then AC to this set of sets, and use these to find a bijection between the union and some subset of $\kappa \times \kappa$.)

Exercise 5.35 (AC(?)) $|\alpha^\beta| \leq \max(|\alpha|, |\beta|)$ (with ordinal exponentiation — we haven't defined cardinal exponentiation yet). (Hint: Prove this by induction on β. I *think* AC is used at limit ordinal steps, but it may not be necessary. The other exercises in this sequence only depend on AC if this one does.)

Exercise 5.36 (AC(?)) For every infinite ordinal α, there is an epsilon number greater than α whose cardinality is the same as that of α. (Hint: Recall the construction of the next epsilon number in Exercise 4.23.)

Exercise 5.37 (AC(?)) If α is an infinite ordinal, then $|\varepsilon_\alpha| = |\alpha|$.

Exercise 5.38 (AC(?)) If $\kappa > \aleph_0$ is a well-ordered cardinal, and $\alpha < \kappa$ is an ordinal, then $\varepsilon_\alpha < \kappa$.

Exercise 5.39 (AC(?)) There is an ordinal β with $\beta = \varepsilon_\beta$, and $\beta < \aleph_1$.

Recall the sublime beauty of this ordinal β and shudder in horror at what this means about uncountable sets like $\aleph_1$, let alone cardinals like $\aleph_\omega$, $\aleph_{\aleph_1}$, and any κ with $\kappa = \aleph_\kappa$. (We will consider where $|\mathbb{R}|$ is with respect to these cardinals in a moment.)

5.4 Cardinal Exponentiation

We define cardinal exponentiation in a new way that doesn't agree with the ordinal exponentiation we've already defined. Since there is a formula that says when a set of ordered pairs is a function, we can define ${}^X Y$ to be the set of all total functions from X to Y.

Exercise 5.40 For natural numbers m, n, $|^n m| = m^n$. (Hint: work by induction on n.)

We then define cardinal exponentiation by letting $\kappa^\lambda = |^\lambda \kappa|$ for all cardinals κ, λ. Just as with addition and multiplication, cardinal exponentiation differs quite a bit from ordinal exponentiation once either input is infinite, but it turns out to differ even more strongly.

Exercise 5.41 If κ is an infinite well-ordered cardinal, then $\kappa^2 = \kappa$. (Hint: Find a bijection between $^2\kappa$ and $\kappa \times \kappa$ and recall Exercise 5.33.)

Exercise 5.42 $|\mathcal{P}(\kappa)| = 2^\kappa$. (Hint: Find a bijection between $^\kappa 2$ and $\mathcal{P}(\kappa)$.)

Exercise 5.43 $2^\kappa > \kappa$. (Hint: Use Exercise 2.51 or 5.18.)

Note that this contrasts with ordinal exponentiation, on which $2^\omega = \omega$, while in cardinal exponentiation, $2^{\aleph_0} > \aleph_0$. As mentioned in footnote 10, this is because ordinal exponentiation only considers the functions that are 0 at all but finitely many inputs, while cardinal exponentiation considers the set of *all* functions.

Using transfinite induction, define $\beth_0 = \omega$ ($\beth$ is beth, the second letter of the Hebrew alphabet), and $\beth_{S(\alpha)} = |\mathcal{P}(\beth_\alpha)|$, and let $\beth_\lambda = \bigcup\{\beth_\beta \mid \beta < \lambda\}$ if λ is a limit ordinal.

Exercise 5.44 (AC) $\aleph_\alpha \leq \beth_\alpha$ for all ordinals α. (At what step in the proof do we need AC?)

Exercise 5.45 (AC) There is a cardinal κ with $\kappa = \beth_\kappa$, and for any such cardinal, $\aleph_\kappa = \beth_\kappa$.

The *Continuum Hypothesis* (CH) is the statement that $\aleph_1 = \beth_1$. The *Generalized Continuum Hypothesis* (GCH) is the statement that $\aleph_\alpha = \beth_\alpha$ for all ordinals α. As we will show later, GCH entails AC, but this isn't obvious yet. Kurt Gödel's proved in the 1930's that an axiom $V = L$ entails GCH, and that if ZF is consistent, then so is ZF+$V = L$. Thus, he proved that AC and GCH were consistent with ZF, while it took the work of Paul Cohen in the 1960's on *forcing* to show that their negations are also consistent with ZF and one another.

Exercise 5.46 If $|X| = |X'|$ and $|Y| = |Y'|$, then $|^X Y| = |^{X'} Y'|$. (Hint: Fix a bijection between X and X' and a bijection between Y and Y', and use them to construct a bijection between $^X Y$ and $^{X'} Y'$.)

Exercise 5.47 Cardinal exponentiation, considered as a function extended to all ordinals as $|^\beta \alpha|$, is not continuous in β. (Hint: the smallest non-trivial test will show that continuity fails. Proving it is not continuous in α is much harder, but we can easily find a counterexample if we assume CH, and let $\alpha = 2$, $\beta = \aleph_1 = \beth_1$.)

Exercise 5.48 $\kappa^\lambda \cdot \kappa^\mu = \kappa^{(\lambda+\mu)}$, with all operations interpreted as cardinal operations. (Hint: Find a bijection.)

Exercise 5.49 $(\kappa^\lambda)^\mu = \kappa^{(\lambda \cdot \mu)}$, with all operations interpreted as cardinal operations. (Hint: Find a bijection.)

Exercise 5.50 If $\kappa < \kappa'$ then $\kappa^\lambda \leq \kappa'^\lambda$ and $\lambda^\kappa \leq \lambda^{\kappa'}$.

Exercise 5.51 If $2 \leq \lambda \leq 2^\kappa$ and κ is infinite then $\lambda^\kappa = 2^\kappa$. Thus, equality is possible in the previous exercise. (Hint: Compare 2^κ, λ^κ, and $(2^\kappa)^\kappa$.)

5.5 Cofinality

(This section can safely be skipped until you reach Section 6.3.)

If Y is any ordered set, and X is a subset of it, we say X is *cofinal in* Y (sometimes called "unbounded") iff there is no element of Y that is strictly above every element of X, or equivalently, for every element of Y, there is some element of X above it. Thus, the primes are cofinal in the natural numbers, and the natural numbers are cofinal in the reals. A set of natural numbers is cofinal in the natural numbers iff it is infinite, but for a set of reals to be cofinal in the reals is a bit more complicated.

We are interested in cofinal subsets of ordinals. If $\alpha = S(\beta)$, then a subset of α is cofinal in α iff it contains β. But if α is an infinite limit ordinal, then $X \subseteq \alpha$ is cofinal in α iff $\bigcup X = \alpha$.

If α and β are two ordinals, and $f\colon \alpha \longrightarrow \beta$ is a function, we say that f is cofinal in β iff the range of f is cofinal in the above definition. The *cofinality* of β, notated '$\mathrm{cf}(\beta)$' is the least ordinal α for which there exists such a cofinal function. Equivalently, α is the least ordinal that is the order-type of a cofinal subset of β.

Exercise 5.52 $\mathrm{cf}(\omega) = \omega$.

Exercise 5.53 $\mathrm{cf}(\omega + \omega) = \omega$.

Exercise 5.54 $\mathrm{cf}(\omega \cdot \omega) = \omega$.

Exercise 5.55 (AC) $\mathrm{cf}(\aleph_1) = \aleph_1$.

Exercise 5.56 $\mathrm{cf}(\aleph_1 + \omega) = \omega$.

Exercise 5.57 $\mathrm{cf}(\alpha) \leq \alpha$.

Exercise 5.58 $\mathrm{cf}(\mathrm{cf}(\alpha)) = \mathrm{cf}(\alpha)$.

Exercise 5.59 $\mathrm{cf}(\alpha)$ is always a cardinal. (Hint: Show that if α is not a cardinal, then there is a cofinal function from some smaller ordinal, so that $\mathrm{cf}(\alpha) < \alpha$, and use the previous exercise.)

Cardinals that are the cofinality of some ordinal will also be their own cofinality, and such cardinals are said to be *regular*. Otherwise, $\mathrm{cf}(\kappa) < \kappa$, and κ is said to be *singular*.

Exercise 5.60 $\aleph_0$ is regular.

Exercise 5.61 $\aleph_\omega$ is singular.

Exercise 5.62 (AC) $\aleph_{S(\alpha)}$ is regular. (Hint: Use Exercise 5.34.)

One might conjecture that if $\kappa = \aleph_\lambda$, where λ is a limit ordinal, then it is singular. This is clearly true if $\lambda < \kappa$ (as in Exercise 5.61). But as we showed in Exercise 5.27, there are cardinals $\kappa = \aleph_\kappa$, and we haven't yet established whether such a cardinal is regular or singular.

Exercise 5.63 There is a singular cardinal κ with $\kappa = \aleph_\kappa$. (Hint: Consider the limit of the sequence $\aleph_0, \aleph_{\aleph_0}, \aleph_{\aleph_{\aleph_0}}, \ldots.$)

However, we will show in Exercise 6.51 that there is in some sense an open question whether there could be a regular cardinal that is indexed by a limit ordinal (which would have to be itself, though it couldn't be the first fixed point of the $\aleph$ function).

Exercise 5.64 There is a singular cardinal κ with $\kappa = \beth_\kappa$.

A regular limit cardinal is said to be "inaccessible", and it is said to be "strongly inaccessible" iff it is also a fixed point of the $\beth$ function.

Exercise 5.65 *(AC) (König's Theorem) If every member of X has cardinality less than κ, then $|\bigcup X| < \kappa^{|X|}$. (Hint: If this were false, then there would be a surjective function $f\colon \bigcup X \longrightarrow {}^X\kappa$. But for any such f, we can construct a member of ${}^X\kappa$, that is not in the range of f. A member of ${}^X\kappa$ is a function $g\colon X \longrightarrow \kappa$. For g not to be in the range of f, it suffices if, for each $y \in X$, and each $z \in y$, $g(y) \neq f(z)(y)$.)

Exercise 5.66 $\kappa^{\mathrm{cf}(\kappa)} > \kappa$. (Hint: There is $X \subseteq \kappa$ with $\bigcup X = \kappa$ and $|X| = \mathrm{cf}(\kappa)$.)

Exercise 5.67 $2^{\aleph_0} \neq \aleph_\omega$. (Hint: Use Exercise 5.51.)

Somewhat surprisingly, although ZFC suffices to settle all facts about cardinal addition and multiplication, we have now proven essentially *everything* that can be proven about cardinal exponentiation. We can prove that $2^{\aleph_0}$ is not precisely equal to any singular cardinal whose cofinality is equal to $\aleph_0$, like $\aleph_\omega$, or $\aleph_{\omega+\omega}$, but by using Paul Cohen's method of "forcing", we can make it equal to any other cardinal.

6 Models of Set Theory

In this section we start to show that certain axioms of set theory are independent of each other, and give consistency proofs of large fragments of ZF. However, by Gödel's theorems, we can't prove ZF itself to be consistent unless we assume something stronger than ZF.

6.1 Simple models

If φ is any sentence, and M is any non-empty set, then let φ^M be the sentence that one gets when all quantifiers are restricted to range only over elements of M. That is, every instance of '$\forall x$' is replaced by '$\forall x{\in}M$' (recall that this is an abbreviation for '$\forall x(x \in m \rightarrow \quad)$'), and every instance of '$\exists x$' is replaced by '$\exists x{\in}M$' (which is itself an abbreviation for '$\exists x(x \in M$ and $\quad)$').

We say a non-empty M *satisfies* a sentence φ (in symbols, '$M \models \varphi$') iff φ^M is true. Roughly, to say that M satisfies φ is to say that if the only objects that existed were the ones in the set M, then φ would be true.[13]

If M satisfies "enough" axioms, then we think of it as a *model* of set theory. If a set of sentences are all satisfied by the same model, then anything that is provable from them will also be satisfied by that same model. Since no model satisfies a contradiction, this also means that any set of sentences satisfied by a model is consistent.

Recall that in the Exercises leading up to 2.73, we showed that ω is a model for the axioms of Peano arithmetic. We will now start to consider various sets as models of smaller or larger amounts of ZF set theory itself.

Exercise 6.1 ω satisfies Axiom 3 (Union) but not Axiom 2 (Pairing).

Exercise 6.2 Show that $\{\emptyset, \{\{\emptyset\}\}\}$ does not satisfy Axiom 1. (Hint: Show that this universe has two distinct "empty" sets.)

Exercise 6.3 For any non-empty M, M satisfies Axiom 4 (Null Set). (Hint: Apply Axiom 6 (Foundation) to M.)

The last exercise shows that "satisfies" is often quite cheap — the objects that play the role in the "model" don't really do the right thing. But this is related to the roughness in the claim that M satisfying φ means that if the only objects that existed were the ones in M, then φ would be true.

Recall the definition of a 'transitive set' from Section 3 — it is a set that contains anything contained in any of its elements. If M is not transitive, then the difficulty with 'if the only objects that existed were the ones in M' goes beyond the difficulty of evaluating counterfactuals about mathematical objects — if M is not transitive, then some elements of objects in M are not themselves in M, so that it's hard to say what it would mean for the object to exist without some of its elements existing. Thus, we will tend to restrict consideration of potential models to transitive sets. (For each, there will be an exercise to prove that it is transitive.)

[13] If you are familiar with model theory, then you will note that we are not here considering *all* possible models for the language, but only models where '$\in$' is interpreted by the $\in$-relation within the set. In model theory, we can not only let M be an arbitrary set, but also interpret '$\in$' with any relation on that set, to get even more exotic models than the ones that we consider here. For instance, although Exercise 6.3 shows that every model with the "real" '$\in$' relation satisfies the Null Set Axiom, we can construct a model in the broader sense that does not, for instance, by letting the domain of the model be the integers $\mathbb{Z}$, and interpreting '$\in$' as the predecessor relation.

Exercise 6.4 If M is transitive, then M satisfies Axiom 1 (Extensionality).

Exercise 6.5 For any M (transitive or not), M satisfies Axiom 6 (Foundation). (Hint: To show that the instance of Foundation involving the set x is true in M, apply Foundation in the real universe to $M \cap x$.)

A property $\varphi(x)$ is said to be *absolute* iff for every transitive set M, $\forall x \in M\ ((\varphi(x))^M \leftrightarrow \varphi(x))$. Similarly, a relation $\varphi(x, y)$ is said to be absolute iff for every transitive set M, $\forall x{\in}M\ \forall y{\in}M\ (\varphi(x, y)^M \leftrightarrow \varphi(x, y))$, and so on for formulas with more free variables.[14]

The next several exercises involve proving that a given formula is absolute. For each of these, there are *non*-transitive sets M containing x and y for which $\phi(x, y)^M$ is true, but $\phi(x, y)$ is in fact false (or vice versa). It is useful to consider what these counterexamples look like to help prove why this *can't* happen if M is transitive.

Exercise 6.6 '$x \subseteq y$' is absolute (but there is a non-transitive M with $x, y \in M$ and $(x \subseteq y)^M$ even though $(x \not\subseteq y)$).

Exercise 6.7 '$x = \{y, z\}$' is absolute (but there is a non-transitive M with $x, y, z \in M$ and $(x = \{y, z\})^M$ even though $(x \neq \{y, z\})$).

Exercise 6.8 '$x = \bigcup y$' is absolute (but there is a non-transitive M with $x, y \in M$ and $(x = \bigcup y)^M$ even though $(x \neq \bigcup y)$).

Exercise 6.9 '$x = \emptyset$' is absolute (but there is a non-transitive M with $x \in M$ and $(x = \emptyset)^M$ even though $(x \neq \emptyset)$).

Exercise 6.10 Any property that can be defined from absolute properties using just logical connectives and bounded quantifiers is itself absolute.

Exercise 6.11 'x is transitive' is absolute.

Exercise 6.12 'x is an ordinal' is absolute. (Consider this as 'x is a transitive set, all of whose members are transitive.'

Exercise 6.13 '$x = y \cup \{y\}$' ('$x = S(y)$') is absolute.

Exercise 6.14 'x is a limit ordinal' is absolute.

Exercise 6.15 '$x = \omega$' is absolute. (Consider this as 'x is a limit ordinal that has no limit ordinals as members', rather than 'x is the intersection of all sets containing $\emptyset$ and closed under S'.)

Recall that we have defined the finite ordinals as $0 = \emptyset$, $1 = S(0)$, $2 = S(1)$, etc. We can use these to construct a strange model showing that the power set operation is *not* absolute:

[14] Note that some authors use the word 'absolute' in a slightly stricter or weaker sense, either requiring this to hold for a broader class of models than the transitive sets, or only requiring it for a specific collection of transitive sets.

Exercise 6.16 $(3 = \mathcal{P}(2))^4$. (That is, even though 3 is not *actually* the power set of 2, when quantifiers are restricted to range only over objects that are contained within the ordinal $4 = \{0, 1, 2, 3\}$, 3 is the set that contains all the things in this range that are subsets of 2, and nothing else.)

Intuitively, the power set operation is *not* absolute, because it depends not just on the *elements* of the set involved, but also *all* sets that might be a subset of it. While a transitive set must contain all the elements of the sets it contains, there might be actual subsets that it can't see, and so a transitive set might think of the 'wrong' set as being the power set of something that it contains.

A similar consideration applies to cardinalities. Recall that '$|X| = |Y|$' means 'there exists a bijection between X and Y'. If we let $M = \{\emptyset, \{\emptyset\}, \{\{\emptyset\}\}\} = \{0, 1, \{1\}\}$, then even though $|1| = |\{1\}|$, it is not the case that $(|1| = |\{1\}|)^M$, because M doesn't contain *any* sets of ordered pairs, let alone a set that is a function that is a bijection between 1 and $\{1\}$.

Tracking which properties and relations are or are not absolute will be helpful in what follows. When a term is not absolute, we will use notation like '$(\mathcal{P}(x))^M$' and '$(\aleph_1)^M$' to denote the objects that M 'thinks' are the power set of x, or the first uncountable ordinal.

6.2 The Cumulative Hierarchy

Recall that 'V' is the symbol that we *would* use for the universe of set theory, if it were a set. One reason for the choice of 'V' for the universe rather than 'U' is the construction we are about to begin.

We define V_α by transfinite recursion, with $V_0 = \emptyset$, $V_{S(\alpha)} = \mathcal{P}(V_\alpha)$, and $V_\lambda = \bigcup\{V_\beta \mid \beta < \lambda\}$ for limit ordinals λ. We can think of each V_α as composed of a small base point (the empty set) with various wider and wider layers on top of it, that we might naturally draw as a big letter 'V'.

These V_α will supply some useful models of much of set theory, and we will eventually show that V itself (the full universe, which is a proper class) is in a certain sense just the union of all the V_α.

Exercise 6.17 For any α, V_α is transitive.

Define a function such that $f(0, x) = x$ and $f(n + 1, x) = \bigcup f(n, x)$ for all natural numbers n. Then we call $\bigcup\{f(n, x) \mid n \in \omega\}$ the *transitive closure* of x, or $\mathrm{tc}(x)$. The next two exercises show that $\mathrm{tc}(x)$ is the smallest transitive set that has x as a subset.

Exercise 6.18 $\mathrm{tc}(x)$ is transitive.

Exercise 6.19 If y is a transitive set and $x \subseteq y$, then $f(n, x) \subseteq y$. (Hint: work by induction on n.)

Using a process similar to transfinite recursion, we can define $Rank(x) = \{Rank(y) \mid y \in \mathrm{tc}(x)\}$. The main difference is that instead of requiring the domain to be an ordinal, we should require it to be a transitive set.

Roughly, the rank of a set is one less than the number of layers of nested brackets needed when writing it out (though of course, we can't *actually* write out infinite sets with nested brackets). The next several exercises allow us to demonstrate the meaning of this rough statement, and evaluate ranks for all the set building operations described in the axioms.

Exercise 6.20 Calculate the ranks of the following sets: $\emptyset$; $\{\emptyset\}$; $\{\{\emptyset\}\}$; $\{\emptyset, \{\emptyset\}\}$; $\{\{\{\emptyset\}\}\}$; $\{\{\{\emptyset\}\}, \emptyset\}$. (It will be helpful to write each of these ranks as a number — we will show in Exercise 6.22 that rank is in fact always an ordinal.)

Exercise 6.21 $Rank(x)$ is defined for all x. (Hint: Use Axiom 6 (the Axiom of Foundation) on the set $\{y \in \mathrm{tc}(x) \mid Rank(y) \text{ is undefined}\}$.)

Exercise 6.22 $Rank(x)$ is always an ordinal. (Hint: Recall that a transitive set of ordinals is itself an ordinal. Show that $Rank(x)$ is always transitive, and apply Foundation again to the set $\{y \in \mathrm{tc}(x) \mid Rank(y) \text{ is not an ordinal}\}$.)

Exercise 6.23 $Rank(\{x, y\}) = \max(Rank(x), Rank(y)) + 1$.

Exercise 6.24 $Rank(\bigcup x) = \bigcup\{Rank(y) \mid y \in x\}$. (Note that this means that $Rank(x)$ is the least ordinal greater than the rank of every element of x. Some texts define rank this way.)

Exercise 6.25 $Rank(\bigcup x) \leq Rank(x)$. (As a more challenging exercise, prove that $Rank(x) = S(Rank(\bigcup x))$ unless $Rank(x)$ is 0 or a limit ordinal, in which case it is equal to $Rank(\bigcup x)$.)

Exercise 6.26 If $x \in y$ then $Rank(x) < Rank(y)$.

Exercise 6.27 If $x \subseteq y$ then $Rank(x) \leq Rank(y)$.

Exercise 6.28 $Rank(\mathcal{P}(x)) = Rank(x) + 1$.

Exercise 6.29 For every ordinal α, $Rank(\alpha) = \alpha$ and $Rank(V_\alpha) = \alpha$. (Hint: Prove these by induction on α.)

Exercise 6.30 $x \in V_\alpha$ iff $Rank(x) < \alpha$. (Hint: Prove this by induction on α.)

The sets V_α are called the *cumulative hierarchy*, and we have just shown that every set appears at some specific stage in the hierarchy (in particular, it appears as a member of the stage immediately after its rank). In addition, each stage of the hierarchy is itself a set that exists at the next stage of the hierarchy. The cumulative hierarchy can be used to give an understanding of the universe of set theory, as argued by George Boolos in "The Iterative Conception of Set".

It turns out that this claim is equivalent to Axiom 6 (Foundation) in the presence of the other axioms, though formulating the precise statement that every set is a member of some V_α is somewhat complicated.

Exercise 6.31 $|V_0| = 0$ and $|V_{n+1}| = 2^{|V_n|}$ for finite n.

Exercise 6.32 $|V_{\omega+\alpha}| = \beth_\alpha$ for all α, and thus if $\alpha > \omega \cdot \omega$, then $|V_\alpha| = \beth_\alpha$.

Exercise 6.33 The Generalized Continuum Hypothesis (the statement that for all ordinals α, $\beth_\alpha = \aleph_\alpha$) implies the Well-Ordering Principle, and therefore the Axiom of Choice. (Hint: Note that every set is a member of some V_α and thus has cardinality less than some $\beth_\alpha$, and $\aleph_\alpha$ is well-ordered.)

The fact that every set appears at some specific spot in the cumulative hierarchy also allows us to define 'Scott's trick' for identifying cardinals with sets even *without* assuming the Axiom of Choice (or anything stronger that implies it, like the Generalized Continuum Hypothesis). If S is a non-well-orderable set, then there is no ordinal that is equinumerous with it, let alone a minimal such ordinal. But the idea of identifying cardinals with ordinals was somewhat artificial to begin with. The natural idea, which Frege, Russell, and others originally considered, was to identify cardinals with an equivalence class of the equinumerosity relation. The problem is that, other than the cardinality of the empty set, every such equivalence class is too big to be a set. (The 'set' of all singletons would contain the singleton of each object in the universe, and so its union would be the whole universe, so it can't exist as a set. The same is true for any other non-empty cardinality.)

Under the Axiom of Choice, we note that each such equivalence class contains an ordinal, so we can represent the equivalence class with the least ordinal in it. (The other members are defined by being equinumerous with this ordinal.) But Dana Scott noted that even without the Axiom of Choice, for every S, there is an ordinal α such that $S \in V_\alpha$. For any set S, we can find the least β such that there is a member of V_β equinumerous with S, and then define $|S|$ to be the set $\{X \in V_\beta \mid |X| = |S|\}$. Instead of picking a single set of a given cardinality to stand for that cardinality, we get a set of sets that are all of that cardinality. But in the absence of the Axiom of Choice, that is often what happens to things that we thought we could identify uniquely.

Now we use the V_α to construct models of parts of set theory.

Exercise 6.34 For any α, V_α satisfies Axioms 3 (Union) and 7 (Separation). (Hint: The methods required for these two axioms are different. For the Union axiom, we use the fact that the sentence '$y = \bigcup x$' is absolute and argue that this set is in M. But for the Separation axiom, we note that '$y = \{z \in x \mid \varphi(z)\}$' is *not* absolute (since different objects might satisfy $(\varphi(z))^M$ than $\varphi(z)$), but we note that *whatever* subset of x is given by separating with φ^M, there is *some* subset corresponding to this in the full universe, and argue that this subset will itself be in M.)

Exercise 6.35 If X is a set of non-empty sets that are pairwise disjoint, and Z contains exactly one member of each member of X, then $Rank(Z) \leq Rank(X)$.

Exercise 6.36 (AC) For any α, V_α satisfies Axiom 12 (The Axiom of Disjoint Choice). (Hint: all the conditions involved in the Axiom of Disjoint Choice are absolute.)

Exercise 6.37 V_α satisfies Axioms 2 (Pairing) and 5 (Power Set) iff α is not a successor ordinal. (Hint: Being a pair set is absolute, and while being a power set is not, there are no problems within V_λ.)

Exercise 6.38 V_α satisfies Axiom 10 (Infinity) iff $\alpha > \omega$.

Exercise 6.39 If $\lambda > \omega$ is a limit ordinal, then V_λ satisfies all the axioms of ZF except possibly 8 and 9 (the two versions of Replacement).

Exercise 6.40 $V_{\omega+\omega}$ fails to satisfy Axiom 8. (Hint: Recall that the definition of ordinal addition required Exercise 3.52, which in turn required the Axiom of Replacement. Note that $\omega \in V_{\omega+\omega}$ but $\omega + \omega$ is not.)

Note that at this point we have shown that several of the axioms are independent of each other. In particular, $V_{S(\alpha)}$ satisfies Axioms 1 (Extensionality), 3 (Union), 4 (Empty Set), 6 (Foundation) and 7 (Separation) without satisfying Axioms 2 (Pairing), or 5 (Power Set), so that the latter can't be proven from the former. Since α can be chosen to be greater or less than ω, this means that these axioms can't prove Axiom 10 (Infinity), but also that Infinity still doesn't change the previous independence. $V_{\omega+\omega}$ shows that even all of these axioms together can't prove Replacement.

6.3 Hereditary cardinalities, and models of most of ZF

For any infinite, well-ordered, regular cardinal κ, let H_κ be $\{x \mid |\mathrm{tc}(x)| < \kappa\}$, the collection of sets whose transitive closure is of cardinality less than κ. H_ω is called the set of "hereditarily finite sets", $H_{\aleph_1}$ is called the set of "hereditarily countable sets", and so on. (Strictly speaking, we won't have proved that these are actually *sets* until Exercise 6.42.)

Exercise 6.41 Let κ be a well-ordered regular cardinal, and consider $x \in V_{\kappa+1} \setminus V_\kappa$. Show that $\kappa \leq |x|$. (Hint: consider the least upper bound of the ranks of members of x.)

Exercise 6.42 For any infinite, well-ordered, regular cardinal κ, $H_\kappa \subseteq V_\kappa$.

Exercise 6.43 $V_\omega = H_\omega$.

Exercise 6.44 $V_{\aleph_1} \neq H_{\aleph_1}$. (Hint: Find some member of $V_{\omega+2}$ that is not in $H_{\aleph_1}$.)

Exercise 6.45 For any infinite, well-ordered, regular cardinal κ, H_κ is transitive, and thus satisfies Axioms 1 (Extensionality), 4 (Empty Set), and 6 (Foundation).

Exercise 6.46 For any infinite, well-ordered, regular cardinal κ, H_κ satisfies Axioms 2 (Pairing), 3 (Union), and 7 (Separation). (Hint: Show that the operations of pairing, union, and separation can't take sets whose hereditary size is below κ to a set whose hereditary size is κ or above.)

Exercise 6.47 If κ is an infinite, well-ordered, regular cardinal, then H_κ satisfies Axiom 8 (Replacement).

Exercise 6.48 If κ is an infinite, well-ordered, regular cardinal, then H_κ satisfies all the axioms of ZF except possibly Axiom 5 (Power Set) and 10 (Infinity).

Exercise 6.49 V_ω satisfies all the axioms of ZFC except 10 (Infinity). (Note that we can prove AC^{V_ω} without actually assuming AC! For any particular *finite* set of non-empty sets, we can prove the existence of a choice function without using AC, and every element of V_ω is finite.)

If λ is a limit ordinal and $\aleph_\lambda$ is regular, then $\aleph_\lambda$ is called a *weakly inaccessible cardinal*. If $\beth_\lambda$ is regular, then $\beth_\lambda$ is called a *strongly inaccessible cardinal*.

Exercise 6.50 If κ is (strongly or weakly) inaccessible, then $\kappa = \aleph_\kappa$, and if κ is strongly inaccessible, then $\kappa = \beth_\kappa$. (Compare this to Exercise 5.63.)

Exercise 6.51 If κ is strongly inaccessible, then $V_\kappa = H_\kappa$, and thus satisfies *all* the axioms of ZF (as well as Choice, if it is true in the full universe). (Hint: Recall Exercise 6.32.)

By Gödel's Incompleteness Theorem, it is not possible for a set of axioms to prove the existence of a model satisfying all those same axioms. Thus, ZF does not prove the existence of a strongly inaccessible cardinal. In his other work, Gödel constructed a family of sets L_α such that if κ is weakly inaccessible, then L_κ actually satisfies ZFC+GCH. Thus, ZF does not prove the existence of a weakly inaccessible cardinal either, but it also does not refute AC or GCH.

These weakly and strongly inaccessible cardinals are the first 'large cardinals'. Further large cardinal axioms are independent in the same way — each large cardinal axiom proves the existence of a model that satisfies ZFC as well as the existence of the smaller large cardinals, and thus none of them can even be proven to be consistent, except by assuming something at least as strong.

Paul Cohen's method of 'forcing' proves independence statements by constructing new models in a different way. He starts with one model M and constructs another model $M[G]$ from it. To prove the consistency of CH, he makes sure that $(\mathcal{P}(\omega))^M = (\mathcal{P}(\omega))^{M[G]}$, but for every cardinal $\kappa < (\beth_1)^M$ in M, he adds a bijection between this cardinal and ω in $M[G]$, so that $(\beth_1)^M = (\beth_1)^{M[G]} = (\aleph_1)^{M[G]}$. To prove the consistency of the negation of CH, he makes sure that every cardinal in M is also a cardinal in $M[G]$ (there are no new bijections between cardinals and smaller sets), but adds additional subsets of ω, so that $(\beth_1)^{M[G]} = (\aleph_\alpha)^{M[G]} = (\aleph_\alpha)^M$. This latter construction only works if $\aleph_\alpha$ has cofinality greater than ω.

7 The Constructible Hierarchy, Gödel's L

None of the exercises in this section depend on the Axiom of Choice, so it should not be used at any point.

For each formula $\varphi(y, x_1, \ldots, x_n)$, and each set X, let $\mathcal{D}_\varphi(X) = \{Y \mid (\forall y \exists x_1 \ldots \exists x_n (y \in Y \leftrightarrow \varphi(y, x_1, \ldots, x_n))^X)\}$. This is the set of all subsets of X that are definable in X by a particular formula φ using parameters in X.

Exercise 7.1 $\mathcal{D}_\varphi(X) \subseteq \mathcal{P}(X)$.

Exercise 7.2 If X is an infinite well-orderable set then $|\mathcal{D}_\varphi(X)| = |X|$. (Hint: Use Exercise 5.33.)

Exercise 7.3 '$Z = \mathcal{D}_\varphi(X)$' is absolute.

Let $\mathcal{D}(X) = \bigcup\{\mathcal{D}_\varphi(X) \mid \varphi \text{ is a well-formed formula of the language of set theory}\}$. This is the set of all subsets of X definable just using parameters in X. Note that there are only countably many formulas φ.

In order to properly define this within the language of set theory, we need a way to enumerate all the well-formed formulas of the language. Showing that this enumeration can be done inside the language is beyond the scope of this document.

Exercise 7.4 If X is finite then $\mathcal{D}(X) = \mathcal{P}(X)$.

Exercise 7.5 If X is an infinite well-orderable set then $|\mathcal{D}(X)| = |X|$.

Exercise 7.6 If X is transitive, then $X \subseteq \mathcal{D}(X)$.

Exercise 7.7 '$Y = \mathcal{D}(X)$' is absolute.

Define $L_0 = V_0 = \emptyset$, $L_{S(\alpha)} = \mathcal{D}(L_\alpha)$, and $L_\lambda = \bigcup\{L_\beta \mid \beta < \lambda\}$ for limit ordinals λ.

Exercise 7.8 For any α, L_α is transitive.

Exercise 7.9 $L_\alpha = V_\alpha$ for $\alpha \leq \omega$.

Exercise 7.10 $L_{\omega+1} \neq V_{\omega+1}$.

Exercise 7.11 $L_\alpha \in L_\beta$ iff $\alpha < \beta$.

Exercise 7.12 For any ordinal α, L_α satisfies Axioms 1 (Extensionality), 3 (Union), 4 (Empty Set), and 6 (Foundation).

Exercise 7.13 For any limit ordinal λ, L_λ satisfies Axioms 2 (Pairing) and 7 (Separation).

Exercise 7.14 '$X = L_\alpha$' is absolute.

Show that if $|S| = \kappa$ then no new subsets of S appear after L_{κ^+}. This might require defining the Gödel operations. Or maybe it proceeds by showing that every new set in $L_{S(\alpha)}$ must have transitive closure with cardinality equal to that of α.

Also, we should prove that L_α is absolute, so that L is the smallest model containing all the ordinals, and that L is absolute, and that $V = L$ is consistent.

Exercise 7.15 If λ is a limit ordinal, then $L_{\aleph_\lambda}$ satisfies Axiom 5 (Power Set).

Exercise 7.16 If κ is an infinite, regular, well-orderable cardinal, then L_κ satisfies Axiom 8 (Replacement).

Exercise 7.17 For any ordinal α, L_α is well-orderable. (Hint: This is only difficult for α a limit ordinal. In this case, the difficulty is to construct a single procedure that gives a well-ordering for every L_β simultaneously, where $\beta < \alpha$.)

Exercise 7.18 For any infinite ordinal α, $|L_\alpha| = |\alpha|$.

Exercise 7.19 For any limit ordinal λ, L_λ satisfies the Well-Ordering Principle.

Exercise 7.20 If κ is a weakly inaccessible cardinal then L_κ satisfies all of ZFC, as well as the Generalized Continuum Hypothesis.

We can write down a formula $\varphi(x)$ that says 'There is an ordinal α such that $x \in L_\alpha$.' The sentence $\forall x \varphi(x)$ is usually written as $V = L$.