# 4 Theories

In this chapter, we look at first-order theories: sets of sentences in a formal, first-order language that are assumed to describe a particular domain of objects. They could describe the behaviour of physical or social systems, but we'll focus on mathematical theories – specifically, arithmetic and set theory.

## 4.1 Arithmetic

For most of history, people did maths in an informal manner, relying on a loose collection of techniques for solving different types of problems. In proofs, assumptions that seemed obviously true – for example, that $0 \neq 1$ – were simply taken for granted.

In the 19th century, mathematics went through a phase of formalization. Weierstrass, Dedekind, and others gave precise definitions of limits, differentiation, and other mathematical concepts. They formalized the exact assumptions that were needed to derive well-known results (such as the intermediate value theorem), giving rigorous proofs without any appeals to obviousness or intuition. It now became possible to show that the techniques mathematicians had been using were indeed valid.

At the same time, more powerful mathematical theories were developed – such as the set theory of Cantor (formalized by Zermelo, Fraenkel, and others), or the type theory of Russell and Whitehead. All known branches of maths, it seemed, could be unified in such a theory, allowing for new results to be derived from the emerging connections between previously separate domains. Theorems from topology could be used to prove results in algebra.

Formally, a *theory* is a set of sentences that is closed under entailment, so that it contains everything that is entailed by it. In this chapter, we'll be concerned with theories in a formal, first-order language. We often write $\vdash_T A$ or $T \vdash A$ (rather than $A \in T$) to say that $A$ is in the theory $T$.

Note that this fits our earlier use of the turnstile: If $A$ is in $T$, then $T \vdash A$ by Mon and Id. Conversely, if $T \vdash A$, then $T \models A$ by the completeness of first-order logic, and then $A$ is in $T$ because $T$ is closed under entailment. We could write $T \models A$ instead of

$T \vdash A$. Conceptually, however, theories belong to the "syntax" or "proof theory" side of logic. A theory is simply a set of sentences. This set is usually specified by laying down some non-logical axioms. The theory then contains all and only the sentences that can be derived from these axioms. We say that a theory $T$ is *axiomatized* by a set of sentences $\Gamma$ if it contains exactly the sentences that are derivable from $\Gamma$.

**Exercise 4.1** Let $\mathfrak{L}$ be some first-order language (with identity). Let $T_1$ be the theory axiomatized by the set of all $\mathfrak{L}$-sentences, $T_2$ the theory axiomatized by the empty set of sentences, and $T_3$ the theory axiomatized by $\{\forall x\, x \neq x\}$. Which of $T_1$, $T_2$, and $T_3$ are the same?

Let's take a closer look at formal theories of arithmetic. Arithmetic is the study of the natural numbers 0, 1, 2, 3, etc. A lot is known about these numbers. We know, for example, that $1 + 2 = 3$, that there are infinitely many primes, or that the factorial function $x!$ grows faster than any polynomial $x^n$. There are also things we don't know. We don't know if every even number greater than 2 is the sum of two primes (Goldbach's conjecture). The aim of an axiomatized theory of arithmetic is to capture all truths about the natural numbers, showing exactly which assumptions (or axioms) are needed to derive which results.

An important part of the axiomatic project is to reduce the number of primitive concepts. In section **??**, I already mentioned that we don't need separate individual constants '$n$' for each number $n$: we can instead use a single constant '0' for the number 0, and a function symbol '$s$' for the successor function, so that the number 1 is denoted by '$s(0)$', the number 2 by '$s(s(0))$', and so on. This is useful because it means that we don't need special axioms for each number: having defined 1, 2, and 3 as $s(0)$, $s(s(0))$, and $s(s(s(0)))$, respectively, we may hope to derive that $1 + 2 = 3$ from general assumptions about zero and the successor function. If '1', '2', and '3' were primitive symbols, it is hard to see how '$1 + 2 = 3$' could be derived from more basic principles.

In section **??**, I suggested that a first-order theory of arithmetic might use primitive symbols for 0, the successor function, addition, multiplication, and the less-than relation. In fact, the less-than relation can be defined in terms of the other concepts and logical expressions:

$$x < y \text{ iff } \exists z(x + s(z) = y).$$

If $t_1$ and $t_2$ are any terms, we can therefore treat '$t_1 < t_2$' as a metalinguistic abbreviation of '$\exists x(t_1 + s(x) = t_2)$', where $x$ is a variable not occurring in $t_1$ or $t_2$.

Let's see how we could define the concept of a prime number. Remember that a

number is prime if it is greater than 1 and divisible only by 1 and itself. We can define divisibility:

$$x \text{ divides } y \text{ iff } \exists z(z \times x = y).$$

With this, we can define primality:

$$x \text{ is prime iff } s(0) < x \wedge \forall y(y \text{ divides } x \to (y = s(0) \vee y = x)).$$

**Exercise 4.2** Define the concepts of (a) an even number and (b) a square number.

Other concepts are harder to define. It is not obvious how one could define exponentiation $x^y$ or the factorial $x!$ in terms of 0, $s$, $+$, and $\times$. We'll see in chapter **??** how it can be done. Indeed, we'll see that all computable functions and relations on the natural numbers can be defined in terms of our four primitives. That is, whenever there is an algorithm for computing a function, or for determining whether a relation holds between some numbers, then the function or relation can be defined in terms of 0, $s$, $+$, and $\times$.

**Exercise 4.3** Can you find another primitive that we could use instead of '$s$'? (That is, can you find a primitive symbol $\varphi$ so that $s(t)$ can be defined from 0, $\varphi$, $+$, and $\times$?)

Let's turn to the second part of the axiomatic project. Having reduced the number of primitive concepts, we need to lay down axioms that describe how these concepts behave. The aim is to reduce all truths about the natural numbers to a small number of basic principles.

It's useful to divide the axioms into two parts. The first concern just 0 and $s$, the second $+$ and $\times$. What do we know about 0 and $s$? We know, for example, that every number has a successor. But we don't need to postulate this as an axiom: all function symbols in first-order logic denote total functions. What isn't guaranteed is that '$s$' denotes an injective function: we need to postulate that no two numbers have the same successor.

Q1 $\quad \forall x \forall y\,(s(x) = s(y) \to x = y)$

We also know that 0 is not the successor of any number:

Q2 $\quad \forall x\, 0 \neq s(x)$

These two axioms are already quite powerful. Let's think about what a model of them must look like. There must be at least one object, denoted by 0. There must also be an object $s(0)$. Can this be the same as 0? No: otherwise 0 would be the successor of itself, which contradicts Q2. So $s(0)$ is another object. What about $s(s(0))$? This can't be 0, by Q2. And so it can't be $s(0)$ either, by Q1: if $s(s(0)) = s(0)$, then $s(0)$ and 0 would have the same successor. So $s(s(0))$ is a third object. By iterating this reasoning, we can see that any model of Q1 and Q2 must have a chain of infinitely many objects $0, s(0), s(s(0)), \ldots$, connected by the successor function.

**Exercise 4.4** Can you find a model in which Q1 and Q2 are true, but $\forall x(s(x) \neq x)$ is false?

Exercise 4.4 shows that Q1 and Q2 don't suffice to capture all truths about 0 and $s$. The problem is that they don't rule out the existence of other objects, outside the chain $0, s(0), s(s(0)), \ldots$. On these other objects, the successor relation must still be injective, but it can go in a loop, or it can form a second infinite chain $a, s(a), s(s(a)), \ldots$. We might use the following axiom to rule out such chains:

Q3 $\quad \forall x\,(x \neq 0 \rightarrow \exists y\, x = s(y))$

This says that there is no object other than 0 that is not a successor. This doesn't help with the looping case, however. Intuitively, we'd like to have a postulate saying that every number can eventually be reached from 0 by repeated application of $s$. But there's no way to express this in first-order logic (as we proved in section **??**). Still, we can get close by adding the following axiom schema, called the *induction schema*:

Ind $\quad (A(0) \wedge \forall x\,(A(x) \rightarrow A(s(x))) \rightarrow \forall x\, A(x))$

Here, $A(x)$ is any formula with one free variable. Think of every such formula as expressing a property. Ind then says that if some (expressible) property holds of 0, and if it is inherited from any number to its successor, then it holds of all numbers. The schema is obviously related to the method of inductive proof, where we show that all numbers have a property by showing that 0 has it and that it is inherited from any number to its successor.

Ind rules out the looping case. Consider the simplest version, where there's an object $a$ outside the chain $0, s(0), s(s(0)), \ldots$ that is its own successor. In this model, $\forall x(s(x) \neq x)$ is false. But $\forall x(s(x) \neq x)$ follows from Q1, Q2, and Ind, as follows. Let $A(x)$ be the

formula $s(x) \neq x$. Then $A(0)$ is $s(0) \neq 0$. This is entailed by Q2. $\forall x(A(x) \to A(s(x)))$ is $\forall x(s(x) \neq x \to s(s(x)) \neq s(x))$. This is entailed by Q1. By Ind, we can derive $\forall x(s(x) \neq x)$.

**Exercise 4.5** How does Ind rule out loops with two elements? That is, why isn't there a model of Q1, Q2, and Ind with two objects $a$ and $b$ outside the chain $0, s(0), s(s(0)), \ldots$ that are successors of each other?

Ind also rules out a second chain $a, s(a), s(s(a)), \ldots$, as we can see from the fact that it entails Q3:

**Proposition 4.1**

Ind entails Q3.

*Proof.* let $A(x)$ be the formula $x \neq 0 \to \exists y\, x = s(y)$. Q3 is $\forall x A(x)$. To derive this via Ind, we need to derive

(i) $A(0)$, and
(ii) $\forall x(A(x) \to A(s(x)))$.

Both of these are logical truths. (i) holds because $0 = 0$; so the antecedent of $A(0)$ is false. For (ii), note that $A(s(x))$ is trivial: its consequent says that $\exists y(s(x) = s(y))$. So $A(x) \to A(s(x))$ holds for all $x$. □

Let's turn to addition and multiplication. A common way to define functions on $\mathbb{N}$ is to describe how they apply to 0 and then define their value for any successor number in terms of their value for the previous number. For example, the factorial function $n!$ that maps every number $n$ to the product $1 \times 2 \times \ldots \times n$ can be defined by the following two clauses:

(i) $0! = 1$
(ii) $s(n)! = n! \times s(n)$

This is called a definition by *(primitive) recursion*. It may at first look circular, but it is not. Take, for example, the input 2 to the factorial function. By clause (ii) of the definition, $2!$ is $1! \cdot 2$. To evaluate this, we need to know $1!$. By clause (ii) again, $1!$ is

$0! \cdot 1$. By the first clause, $0!$ is 1. Putting all this together, we have

$$2! = (1 \times 1) \times 2 = 2.$$

We can similarly define the addition function by primitive recursion on its second argument:

(i) $x + 0 = x$
(ii) $x + s(y) = s(x + y)$

These two claims are easily translated into the language of arithmetic, which gives us our next two axioms:

Q4 $\forall x(x + 0 = x)$
Q5 $\forall x \forall y(x + s(y) = s(x + y))$

The same trick works for multiplication, which we can define as repeated addition:

Q6 $\forall x(x \times 0 = 0)$
Q7 $\forall x \forall y(x \times s(y) = (x \times y) + x)$

**Exercise 4.6** Explain how the primitive recursive definition of addition determines the value of $3 + 2$.

The theory axiomatized by Q1–Q7 is called *Robinson Arithmetic*, or Q. It will play an important role in chapter **??**. The standard first-order theory of arithmetic, called *Peano Arithmetic*, or PA, replaces Q3 by Ind: its axioms are Q1, Q2, Ind, and Q4–Q7. (The theory is named after Giuseppe Peano, although Peano points out that essentially the same theory was proposed earlier by Dedekind).

Are all truths in the language arithmetic entailed by the axioms of PA? For a while, this seemed plausible. Gödel's first *incompleteness* theorem revealed that the answer is no: there are arithmetical truths that aren't provable in PA. So PA $\neq$ Th($\mathfrak{N}$). We'll prove this in ch. **??**. As we'll see, the problem can't be fixed by adding another axiom or axiom schema. PA isn't just incomplete; there's a good sense in which it is *incompletable*.

**Exercise 4.7** Show that (a) PA $\vdash \forall x(x \times s(0) = x)$; (b) PA $\vdash x \neq 0 \to \exists y(x = s(y))$.

**Exercise 4.8** We've seen that Q1–Q3 don't rule out structures in which the successor function goes in a loop for some objects outside $0, s(0), s(s(0)), \ldots$.

(a) Show that adding Q4–Q7 doesn't help: there are models of Q1–Q7 with two objects $a$ and $b$ that are successors of each other.

(b) Using the definition of '<' from earlier in this section, determine whether $a < b$, $a < a$, and $0 < a$ (in your model).

(c) Show that Q $\nvdash \forall x(x + y = y + x)$.

## 4.2 Set theory

In the 19th century, set-theoretic concepts were increasingly used by mathematicians to make their theories and definitions more precise. For example, Dedekind defined the real numbers in terms of sets of rational numbers, which allowed for new, more rigorous proofs of many results in real analysis.

The concept of a set was initially regarded not as belonging to a separate mathematical theory, which we now know as set theory. Rather, it was treated as a logical concept. To speak of the set of such-and-suchs, it was assumed, is just to speak of the such-and-suchs taken together. As Georg Cantor put it in 1895: a set is 'a collection of definite, well-differentiated objects [...] into a whole'. It was assumed that, as a matter of logic, there whenever there is a collection of (definite, well-differentiated) objects, there is also a set of these objects.

Dedekind had defined the real numbers in terms of sets of rational numbers, which can easily be defined as pairs of integers (numerator and denominator), which can be defined as pairs of natural numbers ($a - b$). Frege realized that it is possible to define the natural numbers entirely in terms of sets. (See section **??** below.) Familiar properties of the natural numbers (and, by extension, of the integers, rationals, and reals) could then be derived from apparently logical properties of sets. Hence there emerged the philosophical project of *logicism*: the idea that all of maths could be reduced to logic and definitions.

This was the life project of Frege, who invented the calculus of predicate logic in order to show that all of arithmetic could be derived from purely logical axioms by simple

logical rules like MP and Gen. Frege's "logical axioms" included one assumption about sets: his "axiom V". This is a second-order axiom, but we can express it as a first-order schema:

$$\text{V} \quad \{x \mid A(x)\} = \{x \mid B(x)\} \leftrightarrow \forall x(A(x) \leftrightarrow B(x)).$$

$A(x)$ and $B(x)$ are arbitrary formulas with one free variable. Axiom V says that different sets never have the very same members. This makes sense if a set of things is just those things "considered as a whole". But Axiom V also implies that for any formula $A(x)$ there is a corresponding set $\{x : A(x)\}$. This is known as the *naive comprehension principle*. Remember that singular terms in predicate logic can never be empty, and Axiom V allows us to form a term $\{x \mid A(x)\}$ for any formula $A(x)$.

Unfortunately for Frege, the naive comprehension principle turns out to be inconsistent, as Bertrand Russell pointed out to him in a letter in 1902. Consider the formula $x \notin x$, saying that $x$ is not a member of itself. Assume that there is a set of all things to which this formula applies. Call this set $R$. We can ask whether $R$ is a member of itself. If it is, then by the definition of $R$, it must not be a member of itself. If it isn't, then by the definition of $R$, it must be a member of itself. Either way, we get a contradiction. So $x \notin x$ is a formula for which there is no corresponding set $\{x : x \notin x\}$.

In hindsight, there is something odd about the idea that a set might contain itself. One imagines sets as abstract "containers", and a container can hardly contain itself. Ernst Zermelo, who had independently noticed Russell's paradox, developed this intuition into a paradox-free formal theory.

According to Zermelo, we should think of the sets as built in stages or layers. We start with things that are not sets, called *individuals* or *urelemente*. At the next stage, we form all sets of these individuals. We may now have sets of rocks and cities, like {Athens, Berlin}, but we don't have any sets containing other sets. At the next stage, we form all sets whose elements are either individuals or sets of individuals. This includes all sets from the first stage, but it also includes sets like {{Athens, Berlin}, Athens}, with sets from the previous stage as elements. We continue in this manner. Whenever a set occurs at some stage, it can be used as an element of sets at later stages. As a consequence, a set first appears at a stage only after all its elements have appeared. So we never get a set that contains itself. Nor do we get a set of all sets that don't contain themselves: this would be the set of all sets; such a set would contain itself, which is impossible.

Oddly, this hierarchical construction works even if there are no individuals. Starting with no individuals, we can construct one set of individuals: the empty set $\emptyset$. From this, we can form another set: $\{\emptyset\}$. And once we have $\emptyset$ and $\{\emptyset\}$, we can form $\emptyset$, $\{\emptyset\}$,

$\{\emptyset, \{\emptyset\}\}$, and $\{\{\emptyset\}\}$. And off we go. For purely mathematical applications, it turns out that this *pure* hierarchy is often enough. Every set-theoretic structure with individuals is isomorphic to a structure of pure sets.

Let's make the structure of this hierarchy, called the *cumulative hierarchy*, or simply $V$, more precise. The hierarchy divides into levels or stages. Each stage is a set of sets. The first stage, $V_0$, is the set of individuals. In the pure hierarchy, this is the empty set:

$$V_0 = \emptyset.$$

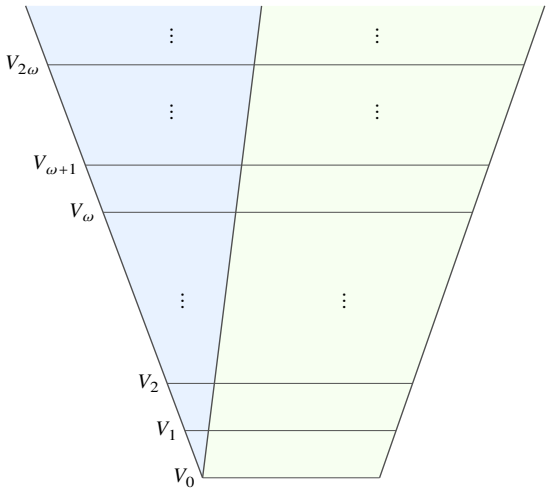
From any stage $V_k$, we recursively define the next stage $V_{k+1}$ as the set of all sets whose elements are in $V_k$. This is just the power set of $V_k$:

$$V_{k+1} = \mathcal{P}(V_k).$$

So $V_1 = \mathcal{P}(\emptyset) = \{\emptyset\}$, $V_2 = \mathcal{P}(\{\emptyset\}) = \{\emptyset, \{\emptyset\}\}$, and so on. Thus we get an infinite sequence $V_0, V_1, V_2, \ldots$ of ever-larger sets, all ultimately built from the empty set.

But we don't stop there. After all the stages $V_0, V_1, V_2, \ldots$, there is another stage $V_\omega$ ("$V$ omega"). $V_\omega$ contains all sets that have appeared at any earlier stage. That is, $V_\omega$ is the union of all earlier stages:

$$V_\omega = \bigcup_{k<\omega} V_k.$$




While all sets in the sequence $V_0, V_1, V_2, \ldots$ are finite, $V_\omega$ has infinitely many elements.

From $V_\omega$, we can form yet further sets by repeating the previous recipes. At stage $V_{\omega+1}$, we collect all the subsets of $V_\omega$. (Many of these are infinite and thus didn't appear at any earlier stage.) That is, $V_{\omega+1} = \mathcal{P}(V_\omega)$. We then form $V_{\omega+2} = \mathcal{P}(V_{\omega+1})$, and so on. After all the stages $V_\omega, V_{\omega+1}, V_{\omega+2}, \ldots$, there is another stage $V_{\omega+\omega}$, or $V_{\omega\cdot 2}$, containing all sets that have appeared at any earlier stage: $V_{\omega\cdot 2} = \bigcup_{k<\omega} V_{\omega+k}$. From $V_{\omega\cdot 2}$, we construct $V_{\omega\cdot 2+1}$, $V_{\omega\cdot 2+2}$, …by taking power sets. Then we construct $V_{\omega\cdot 3}$ by taking the union of all earlier stages. And so on and on.

And we don't stop there. After all the stages $V_\omega, \ldots, V_{\omega\cdot 2}, \ldots, V_{\omega\cdot 3}, \ldots$, there is another stage $V_{\omega\cdot\omega}$, or $V_{\omega^2}$, where we take the union of all previous stages. From this,

we construct further stages by taking power sets and unions. Eventually, we reach $V_{\omega^3}$, then $V_{\omega^4}$, etc. Then we take the union of all these stages to get $V_{\omega^\omega}$, and so on and on, through $V_{\omega^{\omega^\omega}}$, through stages with infinitely high towers of $\omega$, and much, much further. The cumulative hierarchy is *vast*.

**Exercise 4.9** How many sets are in $V_3$ How many are in $V_4$?

**Exercise 4.10** Is the cardinality of $V_{\omega+1}$ greater than the cardinality of $V_\omega$?

Let's now try to axiomatize this conception of sets. That is, we'll try to find a set of sentences in a suitable first-order language that describes the structure of the cumulative hierarchy. The description I just gave, with its 'and so on's and 'after all these stages' can't be directly translated into first-order logic. We have to take a more indirect approach.

The most popular axiomatization of set theory is *ZFC*, for 'Zermelo-Fraenkel set theory with the Axiom of Choice'. Its only primitive concept is the membership relation. So we have a single non-logical symbol: the binary predicate symbol '$\in$'. From this, other concepts are defined. For example, we can define the subset relation $\subseteq$ as follows:

$$x \subseteq y \text{ iff } \forall z(z \in x \to z \in y).$$

(I now use variables $x, y$ rather than $t_1, t_2$ because the only singular terms in the language of set theory are variables.)

Let's go though the axioms of ZFC. The intended domain are the pure sets. So all quantifiers range over pure sets. Our first axiom is known as the axiom of *extensionality*.

$$\text{Z1} \quad \forall x \forall y((\forall z(z \in x \leftrightarrow z \in y)) \to x = y).$$

This says that a set is determined by its elements: no two sets have the same elements. Unlike Frege's Axiom V, Z1 doesn't imply that for any formula $A(x)$ there is a corresponding set $\{x : A(x)\}$. Instead of this unrestricted comprehension principle, we have a more restricted principle, called the *separation axiom*. It's actually a schema:

$$\text{Z2} \quad \forall y \exists z \forall x(x \in z \leftrightarrow (x \in y \wedge A(x)))$$

This says that for any set $y$ and any formula $A(y)$, there is a set $x$ that contains just those elements of $y$ that satisfy $A(x)$. That is, provided that we already have a set $y$, we can use any formula to carve out a subset of $y$ from the elements of $y$ of which the formula is true.

The next axiom postulates the existence of the empty set, the base level of the hierarchy.

Z3 $\quad \exists x \forall y (y \notin x)$.

This says that there is something (a set) that has no elements. By the extensionality axiom, there is only one such thing. It's convenient to have a name for it: '$\emptyset$'. But '$\emptyset$' isn't officially part of the language. The only singular terms in the language of set theory are variables. So we can't say that '$\emptyset$' is shorthand for some more complex term in the language, in the way we could treat '3' as shorthand for 's(s(s(0)))'. What we can do instead is give a *contextual* or *syncategorematic* definition of '$\emptyset$', as follows:

$A(\emptyset)$ abbreviates $\exists x(\forall y(y \notin x) \wedge A(x))$.

Here, $A(\emptyset)$ is a string of symbols containing '$\emptyset$', $x$ and $y$ are two variables that don't occur in $A$, and $A(x)$ is the result of replacing all occurrences of '$\emptyset$' in $A(\emptyset)$ by $x$.

For example, consider the sentence

$$\forall x(\emptyset \subseteq x),$$

saying that the empty set is a subset of every set. By the convention for $\emptyset$, it is shorthand for

$$\exists z(\forall y(y \notin z) \wedge \forall x(z \subseteq x)),$$

which in turn (by the convention for $\subseteq$) is shorthand for

$$\exists z(\forall y(y \notin z) \wedge \forall x \forall v(v \in z \to v \in x)),$$

(You may note that this is a logical truth.)

The same trick is needed to talk about operations on sets. To define the union $\cup$ operation, for example, we need to find a formula that is true of sets $x$, $y$, and $z$ iff $z$ is the union of sets $x$ and $y$. Such a formula is not hard to find:

$$\forall v(v \in x \vee v \in y \leftrightarrow v \in z).$$

With this, we can give a contextual definition of '$\cup$':

$$A(t_1 \cup t_2) \text{ abbreviates } \exists x(\forall y(y \in t_1 \vee y \in t_2 \leftrightarrow y \in x) \wedge A(x)),$$

where $x$ and $y$ do not occur in $A$.

We can similarly define the intersection operation $\cap$:

$$A(t_1 \cap t_2) \text{ abbreviates } \exists x(\forall y(y \in t_1 \wedge y \in t_2 \leftrightarrow y \in x) \wedge A(x)).$$

**Exercise 4.11** Give contextual definitions of $\bigcup t$ and $\mathcal{P}(t)$. $\bigcup t$ is the union of all sets in $t$; $\mathcal{P}(t)$ is the set of all subsets of $t$.

The next two axioms guarantee that for every set $x$, there is a set $\bigcup x$ comprising all elements of elements of $x$, and a set $\mathcal{P}(x)$ comprising all subsets of $x$. Z4 is the *union axiom*, Z5 the *powerset axiom*.

Z4 $\forall x \exists u \forall y(y \in u \leftrightarrow \exists z(z \in x \wedge y \in z))$.
Z5 $\forall x \exists p \forall y(y \in p \leftrightarrow y \subseteq x)$.

Z3 guarantees the existence of the empty set $\emptyset$. By Z4 and Z5, we also have the union $\bigcup \emptyset$ and the power set $\mathcal{P}(\emptyset)$ of the empty set. But this gets us nowhere: $\bigcup \emptyset$ and $\mathcal{P}(\emptyset)$ are just the empty set again. To get the hierarchy off the ground, we need the next axiom, the *pairing axiom*:

Z6 $\forall v \forall w \exists x \forall y(y \in x \leftrightarrow (y = v \vee y = z))$.

This says that for any sets $v, w$ there is a set $\{v, w\}$ that contains exactly $v$ and $w$. $v$ and $w$ can be the same thing. For this case, the axiom says that frinfor every set $v$ there is a set $\{v, v\} = \{v\}$ that contains exactly $v$. This is called the *singleton* set of $v$. We'll help ourselves to $\{v\}$ as a contextually defined term.

Given the empty set $\emptyset$ from Z3, the pairing axiom gives us the singleton of the empty set, $\{\emptyset\}$. From this, we can form further sets like $\{\emptyset, \{\emptyset\}\}$ by applying the union and powerset operations. Z1–Z6 thus give us the finite levels $V_0, V_1, V_2, \ldots$ of the cumulative hierarchy. But they don't guarantee the existence of $V_\omega$, For this, we need the *axiom of infinity*:

Z7 $\exists x(\emptyset \in x \wedge \forall y(y \in x \to y \cup \{y\} \in x))$.

Next, we have the axiom of *foundation* (or *regularity*) to ensure that every set (every object in the domain) is part of the cumulative hierarchy. Consider any nonempty set $x$. The elements of $x$ are other sets. If $x$ is in the cumulative hierarchy, then its elements must have appeared at earlier stages in the construction, and there must be some stage at which the first of them appeared. Let $y$ be one of these earliest elements. Since all elements of $y$ appear strictly before $y$, it follows that none of the elements of $y$ can be in $x$. That is, every nonempty set $x$ of sets must have an element $y$ that is disjoint from $x$:

Z8 $\quad \forall x[x \neq \emptyset \to \exists y \in x(x \cap y = \emptyset)]$

There are two more axioms. Next is the *axiom of replacement*, due to Fraenkel. To see what it does, remember that the axiom of infinity guarantees the existence of an infinite set. As we'll see in section **??**, we may think of this set as the set $\mathbb{N}$ of natural numbers $0, 1, 2, \ldots$. Now suppose that we have a construction that defines a set $s_n$ for each number $n \in \mathbb{N}$. We would like to say that these sets $s_0, s_1, s_2, \ldots$ themselves form a set. After all, there aren't any more of the $s_n$ than there are members of $\mathbb{N}$. Zermelo's original axioms, from 1908, didn't guarantee this. The axiom of replacement does. It is called 'replacement' because it allows replacing all members $i$ of a known set by other things $f(i)$.

To state the replacement axiom (which is actually a schema), I use the uniqueness quantifier $\exists!$. '$\exists! x A(x)$' is short for '$\exists x(A(x) \wedge \forall y(A(y) \to y{=}x))$'.

Z9 $\quad \forall v[(\forall x \in v \exists! y A(x, y)) \to \exists w \forall x \in v \exists y \in w A(x, y)]$

Replacement ensures that higher levels of the cumulative hierarchy exist. With the previous axioms, we get $V_0, V_1, V_2, \ldots, V_\omega, V_{\omega+1}, V_{\omega+2}, \ldots$, but we don't get $V_{\omega+\omega}$. Replacement fills the gap: since there are only countably many stages below $V_{\omega+\omega}$, we can define a mapping from $\mathbb{N}$ to these stages. Since we already know that $\mathbb{N}$ is a set, replacement guarantees that there's a set containing all the stages below $V_{\omega+\omega}$. The union of this set is $V_{\omega+\omega}$.

Finally, we have Zermelo's Axiom of Choice. This says that if we have a set $x$ of non-empty sets, then there is a set $c$ that contains exactly one element from each set in $x$.

Z10 $\quad \forall x[\forall z(z \in x \to z \neq \emptyset) \to \exists y \forall z(z \in x \to \exists! v(v \in z \wedge v \in y))]$

Unlike the other axioms, the Axiom of Choice states that a certain set exists without describing how it can be constructed: we are not told *which* element of each set in $x$ is in $c$. For this reason (as well as others), the axiom has long been controversial. Nowadays, it is generally accepted, as many important mathematical results depend on it.

**Exercise 4.12** Infinity says that there is a set $x$ of a certain kind. List 3 members of this set.

**Exercise 4.13** Prove from the axioms of ZFC that for any three things a,b,c, there is a set {a,b,c}.

**Exercise 4.14** Explain why the separation axiom implies that there is no set of all sets.

## 4.3 Sets and numbers

The Axiom of Infinity draws attention to an infinite sequence of sets, called the *finite von Neumann ordinals*, or simply the *finite ordinals*:

- $\emptyset$
- $\{\emptyset\}$
- $\{\emptyset, \{\emptyset\}\}$
- $\{\emptyset, \{\emptyset\}, \{\emptyset, \{\emptyset\}\}\}$
- ...

This sequence has the structure of the natural numbers. We can think of $\emptyset$ as 0, $\{\emptyset\}$ as 1, $\{\emptyset, \{\emptyset\}\}$ as 2, and so on. The successor of any number $n$ is the $n \cup \{n\}$. (Conveniently, each number $n$ has exactly $n$ elements.)

More formally, we can use the finite ordinals to define a model of arithmetical theories like $\mathrm{Th}(\mathfrak{N})$ and PA. Recall that a model of a theory is a structure consisting of a domain and an interpretation of the non-logical symbols in which all axioms of the theory are true. For a model of $\mathrm{Th}(\mathfrak{N})$, we can choose as the domain the set $\omega$ of finite ordinals. The interpretation function maps the '0' symbol to $\emptyset$ and the successor symbol '$s$' to the function that maps each set $x \in \omega$ to $x \cup \{x\}$. The standard recursive definitions of addition and multiplication then determine the interpretation of '+' and '×'. (If $n$ and $m$

are in $\omega$, $n + m$ will be the unique set in $\omega$ that has exactly $n + m$ elements, and $n \times m$ the unique set with $n \times m$ elements.) This shows that the natural number structure can be embedded in the structure of sets. The same is true for almost every mathematical structure.

There is more. Suppose we read

- '0' as an abbreviation of '$\emptyset$',
- '$s(t)$' as an abbreviation of '$t \cup \{t\}$',
- '$t_1 + t_2$' and '$t_1 \times t_2$' as abbreviations of the corresponding operations on sets,

and we restrict all quantifiers in PA to range over $\omega$. Then all axioms of PA are provable in ZFC. This means that PA is *interpretable* in ZFC. In general, a theory $T$ is interpretable in ZFC is there is a translation scheme of the kind I've sketched under which all axioms of $T$ are provable in ZFC.

A wide range of mathematical theories are interpretable in ZFC. In that sense, ZFC is *at least as strong* as these other theories: whatever they can prove, ZFC can prove as well (if only under the appropriate translation scheme).

I'm not going to prove that PA is interpretable in ZFC. The proof isn't hard, but a little fiddly. To get a sense of what needs to be shown, consider the second axiom of PA:

Q2 $\quad \forall x\, 0 \neq s(x)$

Under the above translation scheme, this turns into $\forall(x \in \omega \to (\emptyset \neq x \cup \{x\}))$. And that's easily provable in ZFC.

**Exercise 4.15** Sketch a proof of the translated Q2 axiom (from the axioms of ZFC).

Let's have a closer look at the finite ordinals. They have some interesting properties. For one, every member of a finite ordinal is also a subset of it. Sets of this kind are called *transitive*. That's because a transitive set $z$ is a set such that whenever $x \in y$ and $y \in z$ then $x \in z$. Another special property of the finite ordinals is that they are $\in$-*well-ordered*. A set $x$ is $\in$-*well-ordered* if any two members of $x$ are related one way or the other by $\in$.

The finite ordinals can be defined as the transitive, $\in$-well-ordered sets with finitely many elements. Now suppose we drop the finiteness condition. Let's define an *ordinal* as a transitive, $\in$-well-ordered set. The finite ordinals are ordinals, but they are not the only ones. For example, $\omega$, the set of finite ordinals, is itself an ordinal. (You can easily

confirm that it is transitive and $\in$-well-ordered.) $\omega$ is an infinite ordinal. So is $\omega \cup \{\omega\}$: the set we get from $\omega$ by adding $\omega$ itself as an element. By our earlier definition of the successor relation, $\omega \cup \{\omega\}$ is the successor of $\omega$. The successor of $\omega \cup \{\omega\}$ is $\omega \cup \{\omega, \omega \cup \{\omega\}\}$, and so on.

The ordinals form a *transfinite* sequence. If we identify the finite ordinals with the natural numbers, the early parts of the sequence look like this:

$$0, 1, 2, \ldots, \omega, \omega + 1, \omega + 2, \ldots \omega + \omega, \ldots$$

Like 0, $\omega$ is not the successor of any ordinal. Infinite ordinals of this kind are called *limit ordinals*. The next limit ordinal after $\omega$ is called $\omega + \omega$, or $\omega \cdot 2$. It is the union of all ordinals $\omega + n$, where $n$ is a finite ordinal. The next limit ordinal after $\omega \cdot 2$ is $\omega \cdot 3$, and so on. After all these comes $\omega \cdot \omega$, or $\omega^2$. It is the union of all ordinals $\omega \cdot n$, where $n$ is a finite ordinal. Much later we reach $\omega^\omega$, $\omega^{\omega^\omega}$, and so on.

The ordinals provide a way of "counting" beyond the finite. This is often useful in mathematics. Above, I've used the ordinals to label stages in the cumulative hierarchy. If there were only finitely many stages, I could have labelled them with natural numbers: $V_0, V_1, V_2, \ldots$. With the ordinals, we can use limit ordinals to label stages at which we take the union of the earlier stages, and successor ordinals to label stages at which we take power sets.

**Exercise 4.16** Show that $\omega$ is transitive and $\in$-well-ordered.

**Exercise 4.17** Is the set of all ordinals an ordinal?

Another use of ordinals is to interpret the theory of cardinals that I outlined in the previous chapter. Remember that two sets have the same cardinality iff they are equinumerous, meaning that there is a bijection between them. For finite sets, cardinalities are naturally identified with natural numbers. The set {Athens, Berlin, Cairo} has cardinality 3. But what kind of thing is the cardinality of an infinite set? In section **??**, we gave them names: we called them $\aleph_0$, $\aleph_1$, etc. But we didn't say more about what these things might be.

The standard answer in set theory identifies the cardinals with certain ordinals: the cardinality of any set $x$ is defined as *the least ordinal that is equinumerous with* $x$.

For finite sets, this yields the expected results. {Athens, Berlin, Cairo} is equinumerous with $\{\emptyset, \{\emptyset\}, \{\emptyset, \{\emptyset\}\}\}$, which is the ordinal 3. So the cardinality of {Athens, Berlin, Cairo}

is 3. What about the cardinality of $\omega$?

The cardinality of $\omega$ is $\omega$. That's because $\omega$ is the least ordinal that is equinumerous with $\omega$. Since $\omega$ is countably infinite, and we defined $\aleph_0$ as the cardinality of any countably infinite set, this means that $\omega = \aleph_0$.

Beyond $\aleph_0$, things get interesting. The cardinality of $\omega + 1$ is still $\aleph_0$. Remember that $\omega + 1$ is $\omega \cup \{\omega\}$. So $\omega + 1$ is just like $\omega$, except that it has one extra element. If you add a single element to a countably infinite set, you always get another countably infinite set. So $\omega + 1$ is equinumerous with $\omega$. After $\omega$, the *ordinal numbers* and the *cardinal numbers* diverge. $\omega$ is both an ordinal and a cardinal. But $\omega + 1$ only an ordinal. The next cardinal after $\omega$ (a.k.a. $\aleph_0$) is $\aleph_1$. This is, by definition, an ordinal: it is the first ordinal in the transfinite sequence that is not equinumerous with $\omega$. It comes surprisingly late in the sequence. It's not $\omega + 1$, or $\omega \cdot 2$, or $\omega^2$, or $\omega^\omega$, or $\omega^{\omega^\omega}$. These can all be shown to be equinumerous with $\omega$. $\aleph_1$ comes much later. And yet we know, from Cantor's theorem, that there are infinitely many different cardinalities. Indeed, for every ordinal $\kappa$, there is a distinct cardinal $\aleph_\kappa$, which is itself an ordinal!

xxx exercise?

By Cantor's theorem, the cardinality of $\mathcal{P}(\omega)$ is greater than $\aleph_0$. How much greater? Cantor conjectured, but was unable to prove, that the cardinality of $\mathcal{P}(\omega)$ is $\aleph_1$. We know that $\mathcal{P}(\omega)$ is equinumerous with the set of real numbers $\mathbb{R}$, which is also known as the *continuum*. So Cantor's conjecture was that there is no set whose cardinality is strictly between that of the natural numbers and that of the real numbers. This became known as the *continuum hypothesis*.

In 1938, Gödel proved that the continuum hypothesis is consistent with ZFC (assuming ZFC itself is consistent): it can't be disproved from the axioms of ZFC. In 1963, Paul Cohen showed that the negation of the continuum hypothesis is also consistent with ZFC (assuming ZFC is consistent). So the continuum hypothesis can be neither proved nor disproved in ZFC.

This is odd. Take the set of real numbers $\mathbb{R}$. We know that this set is uncountable. Obviously, we can get a countable set by removing sufficiently many numbers from $\mathbb{R}$, Can we also remove numbers from $\mathbb{R}$ so that we get a set that's still uncountable, but smaller than $\mathbb{R}$? This straightforward question seems to have no definite answer. It can't be answered from the standard axioms of set theory. We could, of course, add the continuum hypothesis as a further axiom. But we could equally add its negation. Neither leads to a contradiction.

Early set theorists assumed that all questions about pure sets have definite answers that can be established by an extended kind of logic. The status of the continuum hypothesis casts doubt on this picture. By now, hundreds of other statements are known that can

neither be proved nor disproved in ZFC. We can investigate structures in which they hold and structures in which they fail. Perhaps there is no "true" structure of sets after all. When we describe the cumulative hierarchy, we seem to describe a unique structure. We say that $V_{\omega+2}$ contains *all subsets* of $V_{\omega+1}$. But we can't tell whether these subsets include sets with a cardinality between $\aleph_0$ and the continuum. If the concept of 'all subsets' has a definite meaning, this meaning seems impossible to pin down.

## 4.4 Unintended models, again

In section 4.1, we looked at non-standard models of Q: models in which all axioms of Q are true but whose structure is clearly not that of the natural numbers. I didn't emphasize it at the time, but Peano Arithmetic also has non-standard models. These are much harder to construct directly. But we know that they exist, from the compactness theorem.

**Theorem 4.1**
There are non-standard models of Peano Arithmetic.

*Proof.* Let $c$ be an individual constant other than 0. Let $\Gamma$ be the set of sentences consisting of the axioms of PA together with all the sentences

$$c \neq 0, c \neq s(0), c \neq s(s(0)), \ldots.$$

Every finite subset of $\Gamma$ is true in the standard model of arithmetic: just interpret $c$ as a sufficiently large natural number. By the compactness theorem, $\Gamma$ has a model. All axioms of PA are true in this model. But the object denoted by $c$ (in this model) can't be a natural number: it lies outside the number sequence 0,1,2,.... □

Intuitively, Peano Arithmetic doesn't "know" that there are no numbers besides 0,1,2,...: its axioms are compatible with the existence of further numbers. And we know from theorem **??** that there's no way to add the missing information to PA, in the form of further axioms: even the set of all truths in the language of arithmetic has non-standard models.

**Exercise 4.18** PA rules out structures in which the "non-standard numbers" form either a loop or a second chain $a, s(a), s(s(a)), \ldots$. What else could a non-standard model look like?

What about ZFC? Does it have non-standard models?

What does a model of ZFC look like? It consists of a set $D$ of objects and an interpretation function that assigns some relation on $D$ to the symbol '$\in$'. In the *intended* model, $D$ is the set of all sets, and '$\in$' denotes the membership relation. Wait…There is no set of all sets!

In a sense, every model of ZFC is a non-standard model. For every model has a set as its domain, but the sets don't fit into any set. They form a *proper class*, as people say. A *proper class* is a collection that is too big to be a set. (The concept of a proper class is formalized in certain extensions of ZFC.)

The problem is that we formalized our semantic concepts in set theoretic terms. We've define models as set-theoretic structures. The intended interpretation of ZFC can't be formalized in this way.

You may wonder how ZFC can have models at all. In any model of ZFC, the domain is a set, and it is interpreted as containing all sets, but ZFC entails that there is no set of all sets. We can strengthen this puzzle. By the (downward) Löwenheim-Skolem theorem, ZFC has a countable model. In this model, the domain contains only countably many objects, but the axioms and theorems of ZFC are all true, and among these theorems are sentences saying that there are uncountably many sets. This is known as "Skolem's Paradox".

It's not a real paradox. A set is countable if there is an injective function from it to the finite ordinals. According to ZFC, there are sets for which there is no such function. In the countable non-standard models, *we* can find such functions, but they are not in the domain of the model.