

Chapter 4

Set Theory

As a branch of mathematics, set theory is less than one hundred years old, yet it occupies a unique and critical position. Set-theoretic principles and methods pervade mathematics. Set-theoretic results have shaken the worlds of analysis, algebra, and topology. Simple questions about sets have split the mathematical community into hostile camps, and the romance of its infinite sets have charmed and challenged philosophers as nothing else in mathematics.

JIM HENLE

Set theory and set-theoretic notation were born out of the nineteenth century struggle in mathematics to give a clear account of the real number system. In this chapter we will investigate an axiom system known as ZF (for Zermelo-Fraenkel).

An investigation of set theory begins by asking the question: *What is a set?* In school you were probably taught that a set is a collection of objects. While this intuition is important, this approach won't get us very far toward our goal of a rigorous foundation for set theory. It is not very precise. (After all, what is a *collection*? For that matter, what is an *object*?) Furthermore, this description makes the (for us) unnecessary distinction between sets and objects.

Instead, we will take the approach of stipulating how sets *behave*. The fundamental property of any set has to do with what “objects” are “in” it. So it seems natural that our language will need to have some way of talking about membership of one set in another set (our objects will all be sets themselves). Let $\mathcal{L} = \{\in\}$ be the language with one binary predicate (the intended meaning of which is to indicate membership of one set in another). It turns out that this simple language will suffice to express axioms rich enough to do an amazing amount mathematics not only about sets but also about many other familiar mathematical objects, like numbers, functions, etc.

4.1 Naive Set Theory

Our first attempt at specifying how sets behave, naive set theory had a certain elegance about it. It was based on only two fundamental principals (the axiom of extension and the axiom scheme of comprehension), both of which seemed intuitively obvious, or at least reasonable assumptions based on the way we naively think about sets.

Our study of naive set theory revealed two things:

1. Naive set theory seemed to be useful for doing mathematics.

We were able to define many useful mathematical objects like intersections, unions, ordered pairs, functions, relations, etc. and prove things about them.

2. Naive set theory is *inconsistent*.

We used a form of Russel's paradox to show that naive set theory was able to prove the sentence

$$\forall x(\mathcal{P}(x) \not\subseteq x)$$

is both true and false.

We would like to remedy this by revising our axioms in such a way that the resulting theory of sets is still useful for doing mathematics, but no longer able to prove contradictory statements.

4.2 A Second Attempt at Set Thoery

We won't actually quite succeed in meeting the goals listed above. In fact, it is inherently impossible to do so. But we will come close. We will introduce a new set of axioms called ZF that will not longer be susceptible to the Russell attack on its consistency. Unfortunately, we won't know with certainty that there is not some other inconsistency that no one has yet detected.

We also won't take time to fully develop "all of mathematics" in ZF, but we will give some indication that this might be possible by doing the following:

- We will show that $\text{ZF} \vdash \text{Con}(\text{PA})$.

That is, assuming the axioms of Zermelo-Fraenkel set theory, we can build a model for the axioms of Peano Arithmetic. Although we will not do it here, a similar thing can be done to construct models for the integers, the rationals and the reals. In fact, one can attempt in this way to "do all of ordinary mathematics" (like calculus) and see how much truth there is to the slogan "All mathematics is set theory."

- We will do some infinite arithmetic.

Actually there are two types of infinite arithmetic, ordinal and cardinal. Hopefully, we will have a chance to talk a little about each one.

But before we can make progress on these goals, we need to take a closer look at the axioms of ZF.

4.3 The Axioms of ZF

Our first axiom will formalize our statement that a set is determined by what is “in” it. It is identical to the Axiom of Extensionality from naive set theory.

Axiom (Extensionality): Membership determines the set.

$$\forall a \forall b [\forall z [z \in a \leftrightarrow z \in b] \leftrightarrow a = b].$$

In particular, this says that it doesn’t matter how we describe a set, how we denote a set, or how we construct a set, only what ends up belonging to the set (as determined by the relation \in). Same members, same set. Different members, different sets.

Our second axiom provides us with our first example of a set.

Axiom (Empty Set): There is a set with no members.

$$\exists a \forall x x \notin a$$

Notice that since we won’t have the powerful comprehension scheme of naive set theory, we need to have a separate axiom to build this set. (How could one show that it does not follow from extensionality alone that there is an empty set? Must there be a set at all?)

Also notice that we have introduced an abbreviation here. Whenever we use the “phrase” $y \notin x$, officially we mean $\neg y \in x$. In fact, we will use many abbreviations in our study of set theory. This will make our expressions much easier to read and understand. But in principal, every such statement with abbreviations could be written down as a first order wff in the language $\mathcal{L} = \{\in\}$. In fact, for any relation defined by a formula φ , we will allow ourselves the convenience of introducing an abbreviation.

Examples.

1. The binary relation $a \subseteq b$ is defined by the wff (with 2 free variables) $\varphi(a, b) = \forall x [x \in a \rightarrow x \in b]$.
2. The 3-ary relation $c = a \cap b$ is defined by the wff (with 3 free variables) $\varphi(a, b, c) = \forall x [x \in c \leftrightarrow [x \in a \wedge x \in b]]$.

This example deserves a bit more discussion. Typically, we think of intersection (\cap) as being an operation that builds a new set from two sets, rather than as a relation among three sets. Once we have shown that for any a and b there is a set c such that $c = a \cap b$ (this will require some more axioms) and that it must be unique (that much we can already do from Extensionality), then we will be free to treat \cap as a operator (or function) in this way. But officially, when we make some claim $\Psi(a \cap b)$ we really mean

$$\exists c [c = a \cap b \wedge \Psi(c)] ,$$

which is of course just an abbreviation for

$$\exists c [\forall x [x \in c \leftrightarrow [x \in a \wedge x \in b]] \wedge \Psi(c)] .$$

You can begin to see why abbreviations are going to be necessary to do anything complicated.

3. $a = \emptyset$ is an abbreviation for $\forall x x \notin a$, so the Axiom of Extensionality can be rewritten as

$$\exists a a = \emptyset .$$

Exercise 4.1. Write down a wff that defines each of the following relations: $c = a \cup b$; $c = \{a, b\}$; $c = a \setminus b$.¹ ◁

Exercise 4.2. There is another kind of abbreviation that will be useful to us. Write wffs that are abbreviated by $\exists x \in a \varphi$, $\forall x \in a \varphi$, $\exists!x \varphi$, and $\exists!x \in a \varphi$. ($\exists!x$ is intended to mean ‘there is a unique x such that’). ◁

We can’t do much with just these first two axioms. In fact,

Exercise 4.3. Show that there is a model for Extensionality and Empty Set that has only one element in its universe. ◁

What we need is some way to generate more sets. The next few axioms of ZF will provide us with ways to build new sets from old sets. We know from the paradoxes we have already discussed that we will need to be at least a little bit careful as we do this. In general, the guiding principal is not to allow sets that are “too big” (like the set of all sets or some other “bad” thing). The four most important of these “set-building axioms” are

¹ $a \setminus b$ is the set of all elements of a that are not elements of b . In set theory, we often use the symbol \setminus instead of the usual subtraction symbol since everything is a set and we want to distinguish between, for example, $5 - 3$ (which equals 2) and $5 \setminus 3$ (which equals $\{3, 4\}$).

Axiom (Pairing): If a and b are sets, so is $\{a, b\}$.

$$\forall a \forall b \exists c \forall x [x \in c \leftrightarrow [x = a \vee x = b]]$$

Axiom (Union): If a is a set, then $\cup a$ is a set.

$$\forall a \exists c \forall x [x \in c \leftrightarrow \exists b [b \in a \wedge x \in b]]$$

Axiom (Powerset): If a is a set, so is the set of all subsets of a .

$$\forall a \exists b \forall x [x \in b \leftrightarrow x \subseteq a]$$

Axiom (Separation): If a is a set and φ is a wff, then $\{x \in a \mid \varphi(x)\}$ is a set.

$$\forall a \exists b \forall x [x \in b \leftrightarrow [x \in a \wedge \varphi(x)]]$$

Exercise 4.4. We used an abbreviation (\subseteq) in the Powerset Axiom. Rewrite that axiom as an \mathcal{L} -wff without any abbreviations. \triangleleft

The Pairing and Powerset Axioms are fairly straightforward. The Union Axiom deserves a little explanation. By $\cup a$ we mean the set $\cup a = \{x \mid \exists b (b \in a \wedge x \in b)\}$. In this notation, the more familiar $A \cup B$ is $\cup\{A, B\}$. Since by Pairing, $\{A, B\}$ is a set if A and B are sets, we see that ZF allows us to construct $A \cup B$ whenever A and B are sets.

The Separation Axiom is really an axiom scheme. By this we mean that it is a description of countably many axioms, one for each possible wff φ .² Separation allows us to form “definable” subsets. That is, we can select out from any set all of the members of that set which satisfy some property that can be expressed with a wff. This is weaker than saying that any subset can be formed, since there may be subsets that cannot be defined in this manner. For doing everyday mathematics, however, this is usually sufficient, since usually it is not difficult to express the kinds of subsets we need in this manner. And Separation is much weaker than Comprehension, which allowed us to make any definable set, even if it wasn’t a subset of some other set. This is the sense in which we have tried to avoid sets that are “too large”.

²We have been a bit imprecise about what kind of wff φ may be. Clearly φ will usually have x free. It may also have a free, but c should not appear free in φ . Also, φ may have additional free variables, in which case we need to preface the axiom with universal quantifiers over all the additional variables – what we have denoted as $\forall z$ in the past. This is called the *universal closure* of a wff with free variables. The separation scheme actually includes all universal closures of the wffs just described, but we will suppress the listing of the free variables and their universal quantifiers to keep the notation manageable.

Note that once we have Separation, then weaker versions of the Pairing, Union, and Powerset Axioms suffice:

$$\forall x \forall y \exists z [x \in z \wedge y \in z]$$

$$\forall x \exists z \forall y \forall u [y \in x \wedge u \in y] \rightarrow u \in z]$$

$$\forall x \exists y \forall u [u \subseteq x \rightarrow u \in z]$$

If one's goal is to study models of set theory, then these weaker versions are somewhat easier to establish for a given model. If we are only interested in consequences of ZF, then it doesn't matter which version we assume, since the stronger versions follow from the weaker versions and Separation.

The Empty Set Axiom is also unnecessary provided we know that there is some set (like the one guaranteed by the Infinity Axiom discussed below), since we can now define the empty set as

$$\{x \in a \mid x \neq x\}$$

for any set a .

The remaining axioms of ZF are Infinity, Regularity, and Replacement. The Infinity Axiom is like the Empty Set Axiom in that it guarantees the existence of one particular set, rather than giving a general way for building new sets from old.

Axiom (Infinity): There is an infinite set.

$$\exists a [\emptyset \in a \wedge \forall x [x \in a \rightarrow x \cup \{x\} \in a]]$$

Without this axiom, it is possible to have a model in which every set is finite (although the model itself must be infinite).

Exercise 4.5. Show that any model of ZF – Inf, the axioms of ZF without the Infinity Axiom, must be infinite. Hint: How large is a powerset? Show that there must be sets of infinitely many different sizes. \triangleleft

With this axiom, there must be a set that contains \emptyset , $\emptyset \cup \{\emptyset\} = \{\emptyset\}$, $\{\emptyset\} \cup \{\{\emptyset\}\} = \{\emptyset, \{\emptyset\}\}$, ... This axiom will be important in our definition of a model for PA, and in fact, we will interpret the successor function of PA using $S(x) = x \cup \{x\}$.

Axiom (Regularity):

$$\forall a [a \neq \emptyset \rightarrow \exists b [b \in a \wedge b \cap a = \emptyset]]$$

The Axiom of Regularity is less intuitive than some of the others, and in fact, much can be done without it, but is useful to avoid certain pathological sets and relationships between sets. ZF^- is sometimes used to denote ZF with the Regularity Axiom deleted.

Lemma 4.3.1 *There is no set a such that $a \in a$.*

Proof. First note that $\{a\}$ exists since it is the same as $\{a, a\}$, which exists by Pairing. Now apply Regularity to $\{a\}$. Since $\{a\}$ is non-empty and a is the only element in $\{a\}$, it must be that $a \cap \{a\} = \emptyset$. Since clearly $a \in \{a\}$, this means that $a \notin a$, else the intersection would be non-empty. \square

Exercise 4.6. Show that if $a \in b$, then $b \notin a$. Hint: Apply Regularity to $\{a, b\}$. \triangleleft

We will defer discussion of Replacement until we need it.

4.4 The axioms of PA

Peano Arithmetic is an axiomatization of basic arithmetic on the natural numbers (non-negative integers). Peano Arithmetic works in the language $\mathcal{L} = \{0, 1, S, +, \times\}$, where 0 and 1 are constants, S is a unary function, and $+$ and \times are binary functions. S stands for *successor*, and the intended meaning is that $S(x)$ should be the “next” natural number after x . The goal in choosing these axioms is to choose statements that are “obviously true” about arithmetic on natural numbers, powerful enough to form the basis of our reasoning about and with the natural numbers, yet as simple and few as possible. Over time, the axioms proposed by Guiseppe Peano, an Italian mathematician, have become the standard choice.

Peano Arithmetic has seven regular axioms and one axiom scheme. The axiom scheme is intended to capture how induction works.

1. $\forall x \forall y [S(x) = S(y) \rightarrow x = y]$
2. $\forall x [S(x) \neq 0]$
3. $S(0) = 1$
4. $\forall x [x + 0 = x]$
5. $\forall x \forall y [x + S(y) = S(x + y)]$
6. $\forall x [x \times 0 = 0]$
7. $\forall x \forall y [x \times S(y) = (x \times y) + x]$
8. For any wff φ with free variable x (and possibly other free variables, too), the universal closure of

$$[\varphi(0) \wedge \forall x (\varphi(x) \rightarrow \varphi(S(x)))] \rightarrow \forall x \varphi(x)$$

is an axiom. This axiom is called induction on the wff φ .

You might have expected some other familiar statements to be in this list, things like $\forall x \forall y x + y = y + x$, and other sentences saying that addition and multiplication have the usual associative, commutative and distributive properties. It turns out that all these (and much more) are consequences of PA, so for simplicity's sake, we leave them out of the axioms and instead make them theorems (statements we prove with the axioms as premises). Here are two examples. The parenthesized PA before each statement indicates that these are consequences of the Peano axioms.

Theorem 4.4.1 (PA) $\forall x S(x) = x + 1$

Proof. Let x be arbitrary (i.e., an arbitrary natural number). By (3) (and the indiscernibility of identicals), $x + 1 = x + S(0)$. By (5), $x + S(0) = S(x + 0)$. By (4), $S(x + 0) = S(x)$. So by transitivity and symmetry of $=$ (or by several uses of indiscernibility of identicals, if you want to go all the way back to first principles), $S(x) = x + 1$. Since x was arbitrary, this is true for all x . \square

Theorem 4.4.1 may leave you wondering why one bothers to introduce S into the language at all, since everything can be expressed using $+$. Indeed, *Language, Proof and Logic* does not introduce S into the language. (See page 456 for the axioms there.) Nevertheless, there are reasons to do so. In particular, there is a subtle distinction between adding 1 (successor) and adding an arbitrary number. The successor is the foundation on which all the other addition is built, as you can see from the axioms. Furthermore, what $S(x)$ is depends in a much more significant way on x than on 1, so it proper to think of it as a unary operator. This corresponds in some sense the difference between $x+1$ and $x++$ in the C programming language. In any case, we will need to think carefully about the successor when we build our model of PA in ZF, so the distinction is a useful one for us.

Theorem 4.4.2 (PA) $\forall x x + 1 = 1 + x$

Proof. This proof is more involved, since it requires the use of induction. Our wff $\varphi(x)$ will be the following:

$$x + 1 = 1 + x$$

We must show that $\varphi(0)$ holds (base case) and that for any x , $\varphi(x) \rightarrow \varphi(x + 1)$ (inductive step). (Now that we know that $S(x) = x + 1$, we will freely choose which one we use in a particular situation.) We begin showing the base case, $0 + 1 = 1 + 0$. By axiom (3), $0 + 1 = 1$. By axiom (4), $1 + 0 = 1$. So $0 + 1 = 1 + 0$.

Now we do the inductive step. Let x be any number such that $\varphi(x)$. We must show $\varphi(x + 1)$, i.e., that $(x + 1) + 1 = 1 + (x + 1)$. By the inductive

hypothesis $(\varphi(x))$, we know that $x + 1 = 1 + x$, so

$$\begin{aligned} (x + 1) + 1 &= (1 + x) + 1 \text{ by inductive hypothesis} \\ &= 1 + (x + 1) \text{ by axiom (5)} \end{aligned}$$

□

The other familiar properties of arithmetic follow by similar (but longer) sorts of proofs. Each one will use induction. This is because Peano's axiom scheme gives inductive (or recursive) definitions of addition and multiplication.

One last note on PA. We have constant symbols for 0 and for 1, but not for the other natural numbers. If necessary, we can consider 2, 3, etc. to be abbreviations for $S(S(0))$, $S(S(S(0)))$, etc. In this way, we are able to talk about any of our old favorite natural numbers we like even though they do not have constant symbol names.

4.5 A Model for PA: The Universe

We want to show that in ZF we can construct a model of PA. This will show us that if ZF is consistent, then so is PA, since it has a model. (Of course, if ZF is not consistent, then it can prove that there is a model for PA and it can prove there is no such model.) We take this as evidence of two things: ZF is useful for doing mathematics, and PA is a reasonable axiom set for arithmetic.

Let's call the model we are going to construct \mathcal{M} . We are part of the way there already. We will let $0^{\mathcal{M}}$ be \emptyset , and $S^{\mathcal{M}}$ be defined by $S^{\mathcal{M}}(x) = x \cup \{x\}$. But we are getting a little bit ahead of ourselves. We haven't even said what the universe of our model is supposed to be. Furthermore, we need to say something about functions, since $S^{\mathcal{M}}$ is supposed to be a function defined on the universe. We will deal with the universe in this section and the functions – including $+^{\mathcal{M}}$ and $*^{\mathcal{M}}$ as well as $S^{\mathcal{M}}$ – in the next section.

We want our universe to be exactly $\omega = \{0, S(0), S(S(0)), \dots\}$. (In set theory this set is usually denoted by ω rather than \mathbb{N} , although it is essentially the same object.) So we are building the “standard model” of PA, which until now we have just been assuming existed. All we need to do is show that ZF implies that the ω exists. The Infinity Axiom almost says this, but we need to combine it with Separation to get just the set we want.

Let $\varphi(z)$ be the wff

$$\varphi(z) = \emptyset \in z \wedge \forall x [x \in z \rightarrow x \cup \{x\} \in z].$$

Then the Infinity Axiom is just $\exists z \varphi(z)$. For some such z , we use Separation to build the set

$$\omega = \{x \in z \mid \forall u [[u = x \vee u \in x] \rightarrow [u = \emptyset \vee \exists v S(v) = u]]\}.$$

The intuition for this definition is that we want to include in ω only those things which are built up from \emptyset using successor. We want to remove from z any other types of sets. We will say more about how to formalize $S(x) = x \cup \{x\}$ as a function from ω to ω in the next section. For now, let's plow on and show that S has the desired properties.

Lemma 4.5.1 (ZF) *The following statements are true about S and ω :*

1. $\forall x \ 0 \neq S(x)$;
2. $\forall x \forall y \ [S(x) = S(y) \rightarrow x = y]$;
3. $\forall x \in \omega \ S(x) \in \omega$;
4. $\forall \vec{y} \ [\varphi(0) \wedge \forall x \ [\varphi(x) \rightarrow \varphi(S(x))] \rightarrow \forall x \in \omega \ \varphi(x)]$.

Exercise 4.7. Prove Lemma 4.5.1. Hint: For (2) use Exercise 6. For (4) apply Regularity to $X = \{x \in \omega \mid \neg \varphi(x)\}$ to show that X must be empty. \triangleleft

By Lemma 4.5.1, once we have shown that S is a function on ω , our model \mathcal{M} will satisfy axioms 1, 2 and Induction of PA. Statement (4) also justifies our use of informal induction on ω , which we can use to prove the following useful facts about the way \in behaves on ω .

Lemma 4.5.2 (ZF) *Properties of \in on ω :*

1. $\forall x \in \omega \ \forall y \in \omega \ [x \in y \rightarrow y \notin x]$.
2. $\forall x \in \omega \ \forall y \ [y \in x \rightarrow y \subsetneq x]$;
3. $\forall x \in \omega \ \forall y \in \omega \ x \in y \leftrightarrow x \subsetneq y$
4. $\forall x \in \omega \ \forall y \in \omega \ \forall z \in \omega \ [x \in y \wedge y \in z \rightarrow x \in z]$;
5. $\forall x \in \omega \ \forall y \in \omega \ [x = y \vee x \in y \vee y \in x]$.

Proof. (1): This follows immediately from Exercise 6.

(2): Induct on x . If x is 0, then the statement is vacuously true. If $x = k \cup \{k\}$ and the statement is true for k in place of x , then $y \in x = k \cup \{k\}$ implies that either $y \in k$ so by induction $y \subsetneq k \subsetneq x$, or else $y = k \subsetneq x$.

(3): This is proved by induction on y .

- Base Case: $y = 0$.

It is vacuously true if $y = 0$, since there are no x such that $x \in 0$ and there are no x such that $x \subsetneq 0$.

- Inductive Step: Suppose that $y = S(z) = z \cup \{z\}$ for some z and that $x \in z \leftrightarrow x \subsetneq z$.

First, we show that if $x \in y$, then $x \subsetneq y$. Since $x \in y = z \cup \{z\}$, there are two cases to consider:

- Case 1: $x = z$. In this case, $x = z \subsetneq z \cup \{z\} = y$, so $x \subsetneq y$.
- Case 2: $x \in z$.

In this case, by the inductive hypothesis, $x \subsetneq z \subsetneq z \cup \{z\} = y$.

So in either case $x \subsetneq y$.

On the other hand, suppose that $x \subsetneq y = z \cup \{z\}$. Again there are two cases to consider.

- $x \subseteq z$

If $x \subseteq z$, then either $x \subsetneq z$ or $x = z$. If $x \subseteq z$, then $x \in z$ (by the inductive hypothesis). Either way, $x \in z \cup \{z\} = y$.

- $z \in x$

In this case, by the inductive hypothesis, $z \subsetneq x$. So $z \subsetneq x \subsetneq z \cup \{z\}$, which is a contradiction, since z is missing only one element from $z \cup \{z\}$. So this case doesn't happen.

(4): If $x \in y$ and $y \in z$, then by (2) $x \subsetneq y$ and $y \subsetneq z$, so $x \subsetneq z$, hence by (2) again, $x \in z$.

(5): First notice that by (3), (5) is equivalent to $\forall x \in \omega \forall y \in \omega [x \subseteq y \vee y \subseteq x]$. Now induct on x .

- Base case: If $x = 0$ the result is obvious since $0 \subseteq y$ for any y .
- Inductive step: $x = k \cup \{k\}$, and for all $y \in \omega$, $y \subsetneq k$, $k \subsetneq y$ or $y = k$. Let's look at each case.

- If $y \subseteq k$, then $y \subsetneq k \subseteq x$, so $y \subsetneq x$.
- If $y = k$, then $y \subseteq x$.
- If $k \subsetneq y$, then (by (3)) $k \in y$, so $\{k\} \subseteq y$, so $x = k \cup \{k\} \subseteq y$.

□

Notice that (1), (3) and (4) imply that \in is a strict linear order on ω . You may remember that when we introduced PA, we mentioned that one can define $x < y$ by $\exists z x + S(z) = y$. Of course, once we have defined addition on ω , we could do the same thing here. It turns out that both orders are the same, that is

$$\forall x \in \omega \forall y \in \omega [x \in y \leftrightarrow \exists z \in \omega x + S(z) = y]$$

We will write $x < y$ if $x \in y$, and $x \leq y$ if $x \in y \vee x = y$.

4.6 A Model for PA: The Functions

In the last section we postponed the question of what a function is. Now we need to answer it. So what is a function? Well, in set theory, everything is a set, so a function must be some sort of set. And sets are determined by their members, so what we are really asking is what belongs to (the set representing) a function.

Let's suppose $f : A \rightarrow B$ is a function. What set should it be? Typically, we think of the function f as a rule telling us how to assign to each element $a \in A$ a unique element $b \in B$. In set theory, we will assume that that rule is expressed as a list (i.e., a set) of all such ordered pairs a and b .

But what is an ordered pair? Once again, it must be some set (everything is a set). Let's denote an ordered pair by $\langle a, b \rangle$. The key property of an ordered pair is that $\langle a, b \rangle = \langle c, d \rangle$ if and only if $a = c$ and $b = d$. The Pairing Axiom lets us build sets like $\{a, b\}$, but this set does not distinguish the order of a and b and is the same as $\{b, a\}$. After a little experimenting, we find that there is a reasonable set to call $\langle a, b \rangle$ and that the axioms of ZF imply that this set exists whenever a and b are sets.

Exercise 4.8. Here is a list of possibilities for $\langle a, b \rangle$. Only one of them has the desired properties. Find it and prove that it works. For the others, show why they fail to work:

$$\{a, b\} \quad \{a, \{b\}\} \quad \{\{a\}, \{b\}\} \quad \{\{a\}, \{a, b\}\} \quad \{a, b, \{a, b\}\}$$

◁

Now we can define $A \times B$ to be the set of all ordered pairs $\langle a, b \rangle$ where $a \in A$ and $b \in B$.

Exercise 4.9. Prove that if A and B are sets, then $A \times B$ is a set. Hint: Find a set big enough to contain $A \times B$ and then use Separation to get exactly $A \times B$. You will need the answer to Exercise 8 to do this. ◁

A function from A to B is now just a set with the following properties:

- $f \subseteq A \times B$,
- $\forall a \in A \exists b \in B \langle a, b \rangle \in f$,
- $\forall a \forall y \forall z [\langle a, y \rangle \in f \wedge \langle a, z \rangle \in f] \rightarrow y = z$.

Notice that the last two properties can be combined into the following wff (with abbreviations):

- $\forall a \in A \exists! b \in B \langle a, b \rangle \in f$,

Now we introduce a number of abbreviations for functions. $f : A \rightarrow B$ is an abbreviation for “ f is a function from A to B ” (i.e., for the conjunction of the three wffs in the definition above); “ f is a function” abbreviates

$\exists A \exists B f : A \rightarrow B; f(x) = y$ abbreviates $\langle x, y \rangle \in f$; $\text{range}(f) = \{y \mid \exists x f(x) = y\}$; and $f \upharpoonright D = \{\langle x, y \rangle \in f \mid x \in D\}$.

Exercise 4.10. Write wffs (you may use other abbreviations) that define “ f is one-to-one”, “ f is onto B ”, and “ f is a function from A to B and $g = f^{-1}$ ”. \triangleleft

These abbreviations will allow us to use our usual notation for functions when it is convenient to do so. There are times, however, when knowing that f is really just a set with certain properties is also handy.

The following properties of functions are easy to prove in ZF:

Lemma 4.6.1 (Function Lemma)

1. If f is a function, then $\text{range}(f)$ is a set.
2. If f is a function and $X \subseteq f$, then f is a function.
3. If f is a function and D is a set, then $f \upharpoonright D$ is a function.
4. If $f : A \rightarrow B$ is one-to-one and $g = f^{-1}$, then g is a one-to-one function.

Exercise 4.11. Prove Lemma 4.6.1. Hint: For (2), remember that “ f is a function” means $f : A \rightarrow B$ for some sets A and B . The trick is to show that the appropriate sets A and B exist. Use Union and Separation for this. For (3) and (4), use (2). \triangleleft

Now we are ready to finish our model \mathcal{M} for PA. First let’s deal with successor:

Lemma 4.6.2 *There is a function $f : \omega \rightarrow \omega$ such that for every $x \in \omega$, $f(x) = S(x)$.*

Proof. Let $f = \{\langle x, y \rangle \in \omega \times \omega : S(x) = y\}$.

Note that by Lemma 4.5.1 if we let $S^{\mathcal{M}}$ be the function from the previous lemma, then axioms 1, 2 and the Induction Scheme of PA are satisfied by our model.

Lemma 4.6.3 *There are functions $\alpha : \omega \times \omega \rightarrow \omega$ and $\mu : \omega \times \omega \rightarrow \omega$ such that*

1. For all $n \in \omega$, $\alpha(\langle n, 0 \rangle) = n$.
2. For all $n, k \in \omega$, $\alpha(\langle n, S(k) \rangle) = S(\alpha(\langle n, k \rangle))$.
3. For all $n \in \omega$, $\mu(\langle n, 0 \rangle) = 0$.
4. For all $n, k \in \omega$, $\mu(\langle n, S(k) \rangle) = \alpha(\langle \mu(\langle n, k \rangle), n \rangle)$.

Note that once we have proven the lemma we will let $+^{\mathcal{M}} = \alpha$ and $*^{\mathcal{M}} = \mu$.

Proof. We will only prove the result for α , the proof for μ is similar.

It turns out that this result is trickier than it might first appear. In fact, we will need to introduce the axiom scheme of Replacement in order to accomplish it. Here is the basic idea of the proof. Suppose we want to show that such a set/function α exists. We would like to build it up in stages. For example, we know how to add when the second addend is 0: $m + 0 = m$. So let's let

$$A_0 = \langle m, n, r \rangle \in \omega \times \omega \times \omega \mid n = 0 \wedge m = r .$$

A_0 exists by Separation.

Now we want to let A_1 be the part of α that tells us what to do when we add 1:

$$A_1 = \langle m, S(n), S(r) \rangle \in \omega \times \omega \times \omega \mid \langle m, n, r \rangle \in A_0 .$$

And of course, we want to have

$$A_{i+1} = A_{S(i)} = \langle m, S(n), S(r) \rangle \in \omega \times \omega \times \omega \mid \langle m, n, r \rangle \in A_i .$$

Finally we let $\alpha = \cup_{i=0}^{\infty} A_i$. In this way we build up α stage by stage.

So where is the rub? We need to justify the existence of all the A_i 's and $\alpha = \cup_{i=0}^{\infty} A_i$. For the latter, we need a set $A = \{A_i \mid i \in \omega\}$. Then we can use the Union Axiom to define $\alpha = \cup A$. (Remember, that is the way we formalize $\cup_{i=0}^{\infty} A_i$.)

The Axiom of Replacement allows us to build just this sort of set. Notice the form of this set: for each $i \in \omega$, we want to put A_i into A . That is we want to replace each i with A_i . More generally, Replacement allows us to build sets of the form

$$\{F(x) \mid x \in A\}$$

where A is a set and F behaves like a function (for each $x \in A$ there must be exactly one $F(x)$) but is defined in terms of wffs rather than sets (since we want to use it to prove the existence of the sets involved).

Why should we allow such an axiom in ZF? The intuition is that if we map each element x of the set A to $F(x)$, then the resulting collection of all these images under F is no larger than A was, and no more complicated, so it should be a set, too.

Here is the formal axiom (scheme):

Axiom (Replacement): For each wff φ with free variables x and \vec{z} , we have the axiom

$$\forall a \forall \vec{z} [\forall x \in a \exists! y \varphi(x, y, \vec{z}) \rightarrow \exists b \forall x \in a \exists y \in b P(x, y, \vec{z})]$$

By combining this with separation, we can change the conclusion to be

$$\forall y [y \in b \leftrightarrow P(x, y, \vec{z})] .$$

That is, if we can prove that for each x in some set A there is a unique y such that $\varphi(x, y)$, then we may justify the formation of

$$\{y \mid x \in A \wedge \varphi(x, y)\} .$$

As with our axiom schemes of Comprehension and Separation, the \vec{z} can be thought of as allowing parameterization.

Now we have just what we need. We can use Replacement (and Separation) to define the A'_i 's.

By Extension, each A_i will be a unique set, so we can apply Replacement (and Separation) to build $A = A_i \mid i \in \omega$.

Finally, we let $\alpha = \cup A = \cup_{i \in \omega} A_i$. □

Putting everything together, we get

Theorem 4.6.4 $ZF \vdash \text{Con}(\text{PA})$. *That is, assuming the axioms of Zermelo-Fraenkel set theory, we can show that there is a model for PA, and hence that PA cannot prove a contradiction.* □

Note that we are implicitly using Soundness and Completeness here to say that there is such a connection between proof and truth.