Chapter 6
Set Theory

The ontological decision concerning infinity can then simply be phrased as: an infinite natural multiplicity exists.
Alain Badiou *Being and Event* [1]

Abstract In previous chapters we introduced mathematical structures, and we followed with a detailed description of basic number structures. Now it is time to look at structures in general. The classical number structures fit very well the definition: a set with a set of relations on it. But what about other structures? Are they all sets? Can a set of relations always be associated with them? Clearly not. Not everything in this world is a set. I am a structured living organism, but I am definitely not a set. Nevertheless, once a serious investigation of set theory got underway, it revealed a fantastically rich universe of sets, and it showed that, in a certain sense, every structure can be thought of a set with a set of relations on it. To explain how it is possible, we need to get a closer look at sets. As we saw in the previous chapter, the deceptively simple intuitive concept of set (collection) leads to unexpected consequences when we apply the well understood properties of finite sets to infinite collections. The role of axiomatic set theory is to provide basic and commonly accepted principles from which all other knowledge about infinity should follow in a formal fashion. There are many choices for such theories. In this chapter we will discuss the commonly used axioms of Zermelo and Fraenkel.

Keywords Axiomatic set theory · Axioms of ZF · Unordered pairs · Actual infinity · Power set axiom

6.1 What to Assume About Infinite Sets?

Sets are collections of objects. To construct a mathematical universe of sets, we will start with very little, strictly speaking with nothing—an empty set. Then we will determine how other sets can be constructed from it. What initially creates a difficulty in thinking about sets is that in the universe of sets, there is nothing but

R. Kossak, *Mathematical Logic*, Springer Graduate Texts in Philosophy 3,
https://doi.org/10.1007/978-3-319-97298-5_6

sets. There are only sets and nothing else. No numbers. No points. No triangles and circles. How can the rich world of mathematical objects and structures be recreated in such a setting? This is what we will see in this chapter.

Set theory as a mathematical discipline did not exist until the second half of the nineteenth century. Around 1870, Georg Cantor began a systematic study of certain sets of points on the number line, and, to do that, he developed a mechanism of counting beyond the finite. Cantor defined the notion of the size (cardinality) of a set and in the process of comparing sizes of various infinite sets he discovered a vast hierarchy of infinities. The new discipline attracted a lot of attention and one of the outcomes was the creation of the axiomatic system, known today as Zermelo-Fraenkel Set Theory, abbreviated by ZF. This formal theory does not explain what sets are. Instead, the axioms of ZF state that sets with certain properties exist, and that the universe of all sets is closed under certain operations, which means that those operations, when applied to sets, yield sets.

In formal theories, axioms can be chosen arbitrarily; they are often formulated just to see what their formal consequences are. For a theory as fundamental as set theory, the choice of axioms is a more demanding task. The axioms are all supposed to be evident. Some undoubtedly are, but certainly not all. A fuller coverage of all related issues would take us too far. Our present aim is just to get familiar with sets, set operations, and the set-theoretic language.

All axioms of ZF can be easily expressed without logical symbolism, but we will write them in a more formal way. The reason is to see how all statements about sets can be expressed in first-order logic. Some of the formal axioms are not easy to digest at first. We will be clarifying their meaning as we move along.

We will use the language of first-order logic with only one binary relation symbol $\in$. We will refer to this language as the *language of set theory*.

Since the axioms will express properties of sets, you should think of the variables $x, y, z, \ldots$ as representing arbitrary sets. The choice of variables is not important. Later on, we will use other variables such as $A, B, \ldots X, Y, \ldots$. One can also use more descriptive names, like Set_1, Set_2, and the like. The membership relation $\in$ is binary, but for greater clarity, atomic formulas of the form $\in(x, y)$ will be written as $x \in y$. We read $x \in y$ as x is an element (or member) of y.

The symbol $\Longrightarrow$ is the implications symbol. In logic formalism, we write "if φ then ψ" as $\varphi \Longrightarrow \psi$. The symbol $\Longleftrightarrow$ is used for equivalence: $\varphi \Longleftrightarrow \psi$ abbreviates the statement $(\varphi \Longrightarrow \psi) \wedge (\psi \Longrightarrow \varphi)$.

Small finite sets can be given by a list of their members. The list of elements are traditionally written within curly brackets {, }. For example, the set whose only elements are a and b is $\{a, b\}$. We can give this set a name, say x. Then $x = \{a, b\}$. This notation is not part of our first-order formalism, but notice that the statement $(a \in x) \wedge (b \in x) \wedge \forall y[y \in x \Longrightarrow (x = a \vee x = b)]$ expresses that $x = \{a, b\}$, hence we can use the latter as an abbreviation.

Recall that if the same free[1] variable is used more than once in a formula, it stands in for the same set; however, when we use different variables, it does not automatically mean that the sets they refer to are different. For example, when we say that something holds "for all sets x and all sets y," we really mean "for all" and this includes the case when the x is the same as y.

We will start with the following three axioms. The first axiom declares that there is a set that is empty; it has no elements. The axiom of the empty set is not the most natural axiom that comes to mind when one thinks about sets. In fact, one could create a formal set theory without it, but, as we will see, it is convenient. The second axiom says that if x and y are sets, then there is a set z which contains all elements of the set x, all elements of the set y, and nothing more. We call such a set z the *union* of x and y and denote it by $x \cup y$. For example if $x = \{a, b\}$ and $y = \{a, c, d\}$, then $x \cup y = \{a, b, c, d\}$. The third axiom is the *axiom of extensionality*; we will comment on its importance below. Remember that all variables $x, y, z, \ldots$ represent sets.

Axiom 1 There is an empty set

$$\exists x \forall y \neg(y \in x)$$

Axiom 2 Unions of sets are sets

$$\forall x \forall y \exists z \forall t \; [t \in z \iff ((t \in x) \vee (t \in y))]$$

Axiom 3 Extensionality

$$\forall x \forall y \; [\forall z \; (z \in x \iff z \in y) \implies x = y]$$

Axiom 1 is a set existence axiom; it declares that a certain set exists. Axiom 2 is a closure axiom; it says that if certain sets exist, some other sets must exist as well. Axiom 3 is of a different nature. It declares that in the world of sets only the membership relation matters. If two sets x and y have exactly the same members, they must be equal. This axiom has important consequences. For example, it implies that there is only one empty set (a somewhat curious fact). Since there is only one empty set we can introduce a symbol for it; the one commonly used is $\varnothing$. Another consequence of Axiom 3 is that the sets $\{a, b\}$ and $\{b, a\}$ are equal. In specific examples of sets, linearity of notation forces us to list elements in a particular order. It follows from Axiom 3 that this order is irrelevant. Also, since the sets $\{a, a\}$ and $\{a\}$ have exactly the same elements, they must be equal. The expressions in curly brackets are not sets, they are lists of elements of sets for our informal use; if an element of a set is listed twice, it does not make the set any bigger.

[1] A free variable in a formula, is a variable that is not within a scope of a quantifier. For example, in $\forall x \exists y [x \in y \wedge \neg(z \in z)]$ only z is free.

There are more axioms. In fact there are infinitely many, and this is because some of them are in the form of axiom schemas. This means that a single statement is used to represent infinitely many statements, all of the same shape. This is the case of the next axiom, which is known as the *comprehension axiom*. The axiom says that if x is a set, and $\varphi(z)$ is a first-order formula with the free variable z, then there is a set y which contains exactly those elements of the set x which have the property expressed by $\varphi(z)$.[2]

Axiom 4 The axiom schema of comprehension

For every first-order formula $\varphi(z)$ *in the language of set theory*
$$\forall x \exists y\ [\forall z(z \in y \iff (z \in x \wedge \varphi(z))]$$

Axiom 4 is an *axiom schema* because we have one axiom for each formula $\varphi(z)$. It is known that we cannot reduce Axiom 4 to a single first-order statement in the language of set theory. Of course, as written it is a single sentence and it has the flavor of a first-order statement. It begins with "for all formulas $\varphi(z), \ldots$," so we are quantifying over formulas not sets. This obstacle can be overcome. Since (almost) all mathematical objects can be represented as sets, so can be formulas of set theory. There is no problem with treating the syntax of a first-order language in set-theoretic fashion. There is a much more serious problem though. To make sense of Axiom 4 in this way we would also need to be able to express set-theoretically the semantic part, i.e. we would need a set-theoretic definition of truth for formulas of set theory. In the second part of the book we will discuss the celebrated theorem of Alfred Tarski that implies that there are no such definitions.

To appreciate the strength of Axiom 4 we need to know what kind of properties can be expressed in the first-order language of set theory. Later we will see that almost everything about mathematics can expressed that way. For now let us consider one example. The property "the set z has at least two elements" can be expressed by the following formula with the free variable z: $\exists t \exists u[(t \in z) \wedge (u \in z) \wedge \neg(t = u)]$. Recall, that all variables represent sets; hence if z is an element of a set x, then z itself is a set, and it may or may not have at least two elements. Axiom 4 assures us that for every set x there exists a set y composed exactly of those sets z in x which have at least two elements.

The key step in Cantor's not-yet-formalized set theory, and in subsequent formalizations, was an open act of faith. This act is expressed by the *axiom of infinity*.

Axiom 5 The axiom of infinity

Infinite sets exist.

We are not saying here that "infinity exists," meaning that some processes can be potentially continued forever. The axiom says that there are sets that are actually infinite. There are many ways in which Axiom 5 can be expressed formally. We will

[2]Here, and in other axioms, we assume that formulas, such as $\varphi(z)$, can have other free variables.

examine how it is done in more detail, but before that we must introduce two more axioms.

Axiom 6 The axiom of pairing

$$\forall x \forall y \exists z \forall t \; [t \in z \Longleftrightarrow ((t = x) \vee (t = y))]$$

Axiom 6 declares that any two sets x and y can be combined to form another set z, whose only elements are x and y. The set $z = \{x, y\}$ is called a *pair*, and sometimes we refer to is as an *unordered pair*. An important particular case is when $x = y$. Then the set $\{x, y\} = \{x, x\} = \{x\}$. Such a set is called a *singleton*.

The axiom of pairing looks innocuous. In the world of mathematics, pairing objects is one of the most elementary operations. In fact, we have already used it several times in our discussion of the number systems. From a philosophical viewpoint though, the self-evidence of the axiom becomes less clear. The universe of all sets is vast. One may think of it as a universe of all there is. What right do I have to take *any* object x and combine it with *any other* object y? What makes them a pair? Do x and y need to have something in common? According to the axiom, they don't. The axiom of pairing is a very peculiar axiom.

Axioms 6 looks very similar to Axiom 2. Compare the two axioms. The difference is not difficult to see. It can be proven that neither of the axioms implies the other.

Using the axioms listed so far, we can begin a reconstruction of familiar mathematical objects as sets. To make natural numbers, we will use a representation due to John von Neumann. Let us start with the number 0. In set theory, 0 can be represented by the empty set $\varnothing$. Notice that $\varnothing$ has 0 elements. The number 1 is defined as the set $\{\varnothing\}$. This set is not empty, it has exactly one element—the empty set. The next number 2 is $\{\varnothing, \{\varnothing\}\}$, and 3 is $\{\varnothing, \{\varnothing\}, \{\varnothing, \{\varnothing\}\}\}$. This rather mysterious process can be explained as follows. If we have a set x representing a natural number n, then the next number $n+1$ is represented $x \cup \{x\}$. This construction is designed so that the set representing a number n has exactly n elements. For each natural number n, we defined a unique set that represents it. We use the word "represents" metaphorically. We do not assume that the natural numbers *exist* somewhere else and here we reconstruct those numbers as sets. No, in set theory there is nothing but sets, and we *define* numbers as sets. Using the axioms we already listed, one can prove that all sets $\varnothing$, $\{\varnothing\}$, $\{\varnothing, \{\varnothing\}\}$, ... exist. Once this is done, we can give them the familiar names 0, 1, 2, ... To generate all those numbers (step-by-step) all we need is the empty set to begin with, and the rules that allow us to construct $x \cup \{x\}$ given x. This way we get set-theoretic representations of all natural numbers. We will need a bit more to define the set of all natural numbers.

A one way to express the axiom of infinity is to declare that there is a set containing all natural numbers as defined above. This is the meaning of the following formal statement

$$\exists x[\varnothing \in x \wedge \forall y(y \in x \Longrightarrow y \cup \{y\} \in x].$$

In this form, the axiom of infinity says that there is a set so large that it contains all natural numbers. To conclude that the set of natural numbers itself exists, more axioms of ZF are needed, but we will not go into the details. Let us just accept that the set of all natural numbers exists. As before, we will use $\mathbb{N}$ to denote the set of all natural numbers.

$\mathbb{N}$ is just a set, but it follows from axioms that the ordering relation on $\mathbb{N}$ is definable in set theory. In other words, if follows from the axioms that the set of ordered pairs $\{(m, n) : m \in \mathbb{N} \wedge n \in \mathbb{N} \wedge m < n\}$ exists. To this end, the relation $m < n$ must be defined in set-theoretic terms. This is done by observing that m is less than n if and only if there is a one-to-one function $f : m \longrightarrow n$ that is not onto. Here we take advantage of the fact that the set-theoretic natural numbers are sets, and that the axioms give us enough power to express statements about functions. All this development is a bit technical, but fairly routine. Once all of this is done we can define the ordered structure $(\mathbb{N}, <)$ and then declare that it is the first infinite counting number ω.[3]

The set-theoretic counting numbers have an interesting property: each number is the set of numbers smaller than itself. Thus $1 = \{0\}$, $2 = \{0, 1\}$, $3 = \{0, 1, 2\}$, and so on,[4] and ω can be identified with the ordered set of all finite counting numbers (i.e. all natural numbers). The great beauty of this approach is that there is no reason to stop at ω. The next counting number is $\omega + 1$, the one after that is $\omega + 2$ and so on. What is $\omega + 1$? According to the definition it is $\omega \cup \{\omega\} = \{0, 1, 2, \ldots, \omega\}$, and we extend the ordering relation by declaring that $m < \omega$ for all finite m.[5]

Starting with nothing (the empty set) and basic axioms that declare existence of other sets, one can generate a rich world of objects, which, while being "just sets" can be used to represent mathematical entities which do not, at first glance, look like sets at all. In fact, it turns out that, with very specific exceptions, most objects of modern mathematics can be interpreted as sets, and most theorems of mathematics can be derived as consequences of the axioms of set theory. This is absolutely remarkable. If you do not have a good idea of what "mathematical entities" other than numbers are, think of mathematics as the language of physics, which in turn provides models of our actual physical universe both in the micro and macro scales. All of those mathematical models can be interpreted as sets.

We have not listed all the axioms of ZF but the list is almost complete. A few axioms are missing, some are of a more technical nature and we will not discuss them here, but we do need one more—the Power Set Axiom—without which would

[3] Cantor defined ω to be the *order type* of $(\mathbb{N}, <)$.

[4] To see this, you need to look at the set theoretic representation of natural numbers. For numbers so constructed, it makes perfect sense to write expression like $3 \in 5$.

[5] A word of caution: What we described here is the process of generating transfinite counting numbers. They are known in set theory as *ordinal numbers* or just *ordinals*. Ordinal numbers are used to count steps in infinite processes, but they are not used to measure sizes of infinite sets. Cantor's *cardinal numbers* serve that second purpose.

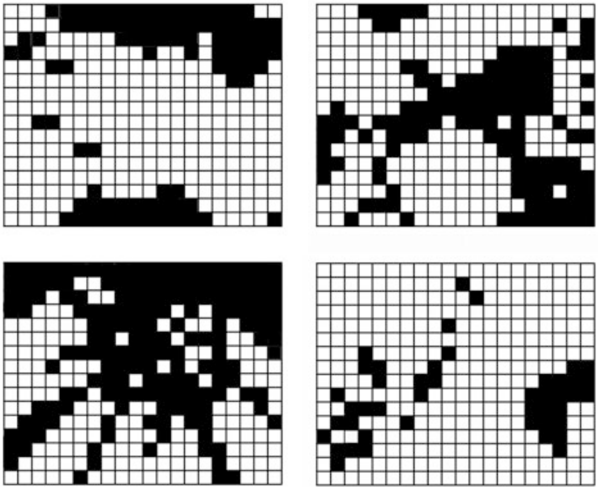

Fig. 6.1 Four of many subsets of a 320-element set

not be able to reconstruct the objects such as the number line, a plane, or the 3-dimensional space as sets.[6]

Let A be a set. Any collection of elements of A is called a *subset* of A. In particular, the empty set and the whole set A are considered subsets of A. Let A be a set of three different elements, let us call them a, b, and c. Here is a list of all subsets of A: $\varnothing$, $\{a\}$, $\{b\}$, $\{c\}$, $\{a, b\}$, $\{a, c\}$,$\{b, c\}$,$\{a, b, c\}$. The set A has three elements and eight subsets.

Imagine now the set M of pixels on a computer monitor contained in a rectangle of dimensions 20 pixels by 16 pixels. This rectangle is made of $20 \times 16 = 320$ pixels, so it is rather small (especially if your screen resolution is large). Think of pictures that are configurations of black and white pixels inside the square (see Fig. 6.1). Each picture can be identified with the set of black pixels (with the other pixels remaining white). Also, each subset of M can be identified with a picture by assigning black to all pixels in the subset and keeping all other pixels white. We see that the number of all possible black and white pictures inside the rectangle is the same as the number of subsets of the set M.

[6]Another word of caution: to say that we would not be able to reconstruct the real number line and other similar objects without the Power Set Axiom is not quite precise. We need that axiom to do it more or less naturally within the Zermelo-Fraenkel set theory. There are other axiomatic systems that do not have such an axiom, in which one can formalize much of modern mathematics, one of the more prominent ones being the second-order arithmetic. Those other systems are interesting, and they are studied for many reasons, but none of them has the status of ZF and its extensions that became a lingua franca of mathematics.

The set M has many subsets. For any number n, a set with n elements has exactly 2^n subsets. For example, the set of letters $\{a, b, c, d, e, f, g, h, i, j\}$ has $2^{10} = 1024$ subsets. It could take a while, but one can make a list all of those subsets. Since our set M above has 320 elements, the number of its subsets is 2^{320}. That number is huge. The number of molecules in the observable universe is currently estimated to be below 10^{81} and this number is much smaller than 2^{320}. Any list all subsets of M, no matter how small a font one would use, would have to fill the entire universe many times over. With this is mind, it becomes a serious philosophical problem to explain the status of the statement: "There exists the set of all subsets of M." If it does exist then where and how? Still, one can freely click here and there to select any set of pixels, so it would be hard to say that some subsets of M do exist, but some others don't. Since none can be excluded, we are inclined to accept that they all somehow "exist," and therefore there must "exist" the set of all of them. This is what the power set axiom declares:

Axiom 7 The power set axiom

For every set x, there exists the set of all subsets of x.

Following the route we outlined in Chap. 4, one can use the axioms to build the sets $\mathbb{N}$, $\mathbb{Z}$ and $\mathbb{Q}$ and the arithmetic operations as relations on them, but still need to see how relations are dealt with in set theory, and this is postponed until the next chapter.

Once we define $\mathbb{Q}$ as a set, the Power Set Axiom allows us to create the set of all its subsets. Then, using Axiom 4, one can show that the set of all Dedekind cuts exists. Then one can go on to define the set of all real numbers and their arithmetic operations, all in the language of set theory. This way, the real line, a classical object of mathematical analysis, gets reconstructed (from the empty set!) in set-theoretic terms. This is an important stepping stone in formalized mathematics, that opens the door to many further developments.

Exercises

Exercise 6.1 *Write down the set theoretic representations of* 4 *and* 5.

Exercise 6.2 *The first-order statement expressing that y is a subset of x is*

$$\forall z(z \in y \implies z \in x).$$

Use this to write a first-order statement expressing the power set axiom.

Exercise 6.3 ** In set theory,* 1 *is defined as* $\{\varnothing\}$ *and* 2 *as* $\{\varnothing, \{\varnothing\}\}$. *Use the axioms of ZF to prove that* 1 *is not equal to* 2.

Exercise 6.4 *The set-theoretic representation of natural numbers in this chapter is due to John von Neumann. Earlier, Ernst Zermelo defined natural numbers this way:* $\varnothing$, $\{\varnothing\}$, $\{\{\varnothing\}\}$, $\{\{\{\varnothing\}\}\}$, $\ldots$. *Write the axiom of infinity using Zermelo's definition.*

Exercise 6.5 ** This exercise requires a lot of patience and some familiarity with formalized set theory. Write down formal set theoretic definitions of* $\mathbb{Z}$, $\mathbb{Q}$ *and* $\mathbb{R}$.