1
Preliminaries

In one sense, set theory is the study of mathematics using the tools of mathematics. After millennia of doing mathematics, mathematicians started trying to write down the rules of the game. Since mathematics had already fanned out into many subareas, each with its own terminology and concerns, the first task was to find a reasonable common language. It turns out that everything mathematicians do can be reduced to statements about sets, equality and membership. These three concepts are so fundamental that we cannot define them; we can only describe them. About equality alone, there is little to say other than "two things are equal if and only if they are the same thing." Describing sets and membership has been trickier. After several decades and some false starts, mathematicians came up with a system of laws that reflected their intuition about sets, equality and membership, at least the intuition that they had built up so far. Most importantly, all of the theorems of mathematics that were known at the time could be derived from just these laws. In this context, it is common to refer to laws as *axioms*, and to this particular system as *Zermelo–Fraenkel Set Theory with the Axiom of Choice*, or *ZFC*. In the first unit of the course, through Chapter 4, we examine this system and get some practice using it to build up the theory of infinite numbers.

In another sense, set theory is a part of mathematics like any other, rich in ideas, techniques and connections to other areas. This perspective is emphasized more than the foundational aspects of set theory throughout the course but especially in the second half, Chapters 5–7. There, our choice of topics within set theory is designed to give the reader an impression of the depth and breadth of the subject and where it fits within the whole of mathematics.

To get started, we review some basic notation and terminology. We expect that the reader is familiar with the following notions but perhaps has not seen them expressed in exactly the same way.

Ordered pairs are used everywhere in mathematics, for example, to refer to points on the plane in geometry. The precise meaning of (x, y) is left to the imagination in most other courses but we need to be more specific.

Definition 1.1 $(x, y) = \{\{x\}, \{x, y\}\}$ is the *ordered pair with first coordinate* x *and second coordinate* y.

It is convenient that (x, y) is defined in terms of sets. After all, this is set theory, so everything should be a set! The main point of the definition is that from looking at $\{\{x\}, \{x, y\}\}$ we can tell which is the first coordinate and which is the second coordinate. Namely, if $\{\{x\}, \{x, y\}\}$ has exactly two elements, then the first coordinate is

$$x = \text{the unique } z \text{ such that } \{z\} \in \{\{x\}, \{x, y\}\}$$

and the second coordinate is

$$y = \text{the unique } z \neq x \text{ such that } \{x, z\} \in \{\{x\}, \{x, y\}\}.$$

And, if $\{\{x\}, \{x, y\}\}$ has just one element, which can only happen if $x = y$, then the first and second coordinates are both

$$x = \text{the unique } z \text{ such that } \{z\} \in \{\{x\}\}.$$

To understand this formula, keep in mind that

$$\{x, y\} = \{y, x\}$$

and

$$\{x, x\} = \{x\}.$$

In particular,

$$\{\{x\}, \{x, x\}\} = \{\{x\}, \{x\}\} = \{\{x\}\}$$

and $\{x\}$ is the only element of $\{\{x\}\}$.

Definition 1.2 $A \times B = \{(x, y) \mid x \in A \text{ and } y \in B\}$ is the *Cartesian product of* A *and* B.

Definition 1.3 R is a *relation from A to B* iff R is a subset of $A \times B$, that is

$$R \subseteq A \times B.$$

Sometimes, if we know that R is a relation, then we write xRy instead of $(x, y) \in R$. For example, we write

$$\sqrt{2} < \pi$$

not

$$(\sqrt{2}, \pi) \in <$$

because the latter is confusing.

Definition 1.4 Let R be a relation from A to B and $S \subseteq A$.

1. The *domain of R* is

$$\text{dom}(R) = \{x \in A \mid \text{there exists } y \text{ such that } xRy\}.$$

2. The *image of S under R* is

$$R[S] = \{y \in B \mid \text{there exists } x \in S \text{ such that } xRy\}.$$

3. The *range of R* is

$$\text{ran}(R) = \{y \in B \mid \text{there exists } x \text{ such that } xRy\}.$$

Notice that $\text{ran}(R) = R[\text{dom}(R)]$.

Definition 1.5 f is a *function from A to B* iff f is a relation from A to B and, for every $x \in A$, there exists a unique y such that $(x, y) \in f$.

If we happen to know that f is a function, then we write

$$f(x) = y$$

instead of $(x, y) \in f$. When we write $f : A \to B$, it is implicit that f is a function from A to B. In certain situations, we refer to a function f by writing $x \mapsto f(x)$ or $\langle f(x) \mid x \in A \rangle$. There are times when we write f_x instead of $f(x)$; this is when we are thinking of elements x of A as *indices* and $\langle f_x \mid x \in A \rangle$ as an *indexed family*. If the domain of f consists of ordered pairs, then it is common to write $f(x, x')$ instead of $f((x, x'))$. Functions are also called *operations* and *maps*. Some people distinguish between a function $f : A \to B$ and its graph,

$$\mathrm{graph}(f) = \{(x, f(x)) \mid x \in A\},$$

but we do not. To us they are the same, that is, $f = \mathrm{graph}(f)$, as we see from Definition 1.5.

Definition 1.6 If $f : A \to B$ is a function and $S \subseteq A$, then the *restriction of f to S* is

$$f \restriction S = \{(x, f(x)) \mid x \in S\}.$$

Definition 1.7 Let $f : A \to B$ be a function.

1. f is an *injection* iff for all $x, x' \in A$, if $x \neq x'$, then $f(x) \neq f(x')$.
2. f is a *surjection* iff for every $y \in B$, there exists $x \in A$ such that $f(x) = y$.
3. f is a *bijection* iff f is both an injection and a surjection.

Injections are also called *one-to-one* functions. Surjections from A to B are also called functions from A *onto* B. Bijections are also called *one-to-one correspondences*.

Definition 1.8 If f is an injection from A to B, then we write f^{-1} for the unique injection $g : f[A] \to A$ with the property that $g(f(x)) = x$ for every $x \in A$. In other words,

$$f^{-1} = \{(f(x), x) \mid x \in A\}.$$

Finally, we assume that the reader has good intuition about the set of integers,

$$\mathbb{Z} = \{\ldots, -2, -1, 0, 1, 2, \ldots\},$$

the set of rational numbers,

$$\mathbb{Q} = \{m/n \mid m, n \in \mathbb{Z} \text{ and } n \neq 0\}$$

and the set of real numbers, $\mathbb{R}$. One thing we will do in this course is define all these kinds of numbers, starting from the natural numbers 0, 1, 2, 3, 4, etc. Each natural number will be the set of natural numbers that precedes it. Thus $0 = \emptyset$, where $\emptyset$ is the set with no members. After that, $1 = \{0\}$, $2 = \{0, 1\}$, $3 = \{0, 1, 2\}$,

$4 = \{0, 1, 2, 3\}$, etc. This happens to be very convenient because then

$$m < n \iff m \in n.$$

In other words, the usual ordering on the natural numbers coincides with membership.

We use natural numbers to denote cardinality, for example, when we say, "Lance Armstrong won the Tour de France seven times." And we use natural numbers to denote order, for example, when we say, "the attorney general is seventh in the presidential line of succession." Another thing we will do in this course is extend the notions of cardinality and order into the infinite. Finite cardinal and ordinal numbers are basically the same thing; one could say that the difference between "seven" and "seventh" is just grammatical. However, the difference between infinite cardinal and ordinal numbers is more profound, as we will explain in Chapters 3 and 4.

Exercises

Exercise 1.1 If R is a relation, then we define

$$R^{-1} = \{(y, x) \mid xRy\}.$$

Give an example where R is a function but R^{-1} is not.

Exercise 1.2 How many functions whose domain is the empty set are there? In other words, given a set B, how many functions $f : \emptyset \to B$ are there?

Exercise 1.3 Explain why $(x, y, z) = (x, (y, z))$ is a reasonable definition of an *ordered triple*.

Exercise 1.4 Equivalence relations play an important role in this book. We assume that the reader has studied them before but this exercise reviews all the necessary definitions and facts. Let A be a set and R be a *relation on* A, that is, $R \subseteq A \times A$. Then:

- R is a *reflexive relation on* A iff for every $x \in A$, xRx.
- R is a *symmetric relation on* A iff for all $x, y \in A$, if xRy, then yRx.

- R is a *transitive relation on* A iff for all $x, y, z \in A$, if xRy and yRz, then xRz.[1]
- R is an *equivalence relation on* A iff R is a reflexive, symmetric and transitive relation on A.

Assuming that R is an equivalence relation on A, for every $x \in A$, we define the *equivalence class of* x to be

$$[x]_R = \{y \in A \mid xRy\}.$$

It is also standard to write

$$A/R = \{[x]_R \mid x \in A\}.$$

A *partition of* A is a family $\mathcal{F}$ of non-empty subsets of A such that

- A is the union of $\mathcal{F}$, that is,

  $$A = \bigcup \mathcal{F} = \{x \mid \text{there exists } X \in \mathcal{F} \text{ such that } x \in X\}$$

  and
- the elements of $\mathcal{F}$ are pairwise disjoint, that is, for all $X, Y \in \mathcal{F}$, if $X \neq Y$, then $X \cap Y = \emptyset$.

Now here are the exercises:

1. Let R be an equivalence relation on A. Prove that A/R is a partition of A.
2. Let $\mathcal{F}$ be a partition of A. Prove that there exists a unique equivalence relation R such that $\mathcal{F} = A/R$.

[1] Later in the book we will define *transitive set*, which is different from *transitive relation*. Unfortunately, it will be important to pay attention to this subtle difference in terminology.

Cambridge University Press & Assessment
978-1-107-00817-5 — A Course on Set Theory
Ernest Schimmerling
Excerpt
More Information

2

ZFC

In the most general terms, when we talk about a mathematical *theory*, we have in mind a collection of *axioms* in a certain *language*. The *language of set theory* has two symbols, $=$ and $\in$, although sometimes we add symbols that are defined in terms of these two to make things easier to read. For example, we write $A \subseteq B$ when we mean that, for every x, if $x \in A$, then $x \in B$.

Zermelo–Fraenkel Set Theory with the Axiom of Choice, or *ZFC* for short, is a certain theory in the language of set theory that we will describe in this chapter. There are infinitely many axioms of ZFC, each of which says something rather intuitive about sets, equality and membership. In our list below, some axioms of ZFC are presented individually whereas others are presented as *schemes* for generating infinitely many axioms. One last comment about terminology before we begin: throughout the course,

$$\text{set} = \text{collection} = \text{family}$$

and

$$\text{member} = \text{element}.$$

Also, the three phrases,

- x belongs to A,
- x is an element of A and
- x is a member of A,

all mean the same thing, namely $x \in A$.

Empty Set Axiom

This axiom says that there is a unique set without members. Formally, it is written

$$\exists! A \; \forall x \;\; (x \notin A)\,.$$

In plain English, this says:

There exists a unique A such that, for every x, x is not an element of A.

The unique set without elements is written $\emptyset$.

Extensionality Axiom

This axiom says that two sets are equal if they have the same members. Formally, it is written

$$\forall A \; \forall B \;\; [\; \forall x \;\; (x \in A \iff x \in B) \implies A = B \;]\,.$$

Because we defined

$$A \subseteq B \iff \forall x \;\; (x \in A \implies x \in B)\,,$$

another way to write the Extensionality Axiom is

$$\forall A \; \forall B \;\; [\; (A \subseteq B \text{ and } B \subseteq A) \implies A = B \;]\,.$$

In other words, two sets are equal if each is a subset of the other.

By logic alone, if $A = B$, then A and B have the same members. Combining this fact with the Extensionality Axiom, we have that

$$\forall A \; \forall B \;\; [\; \forall x \;\; (x \in A \iff x \in B) \iff A = B \;]\,.$$

Equivalently,

$$\forall A \; \forall B \;\; [\; (A \subseteq B \text{ and } B \subseteq A) \iff A = B \;]\,.$$

Pairing Axiom

This axiom allows us to form singletons and unordered pairs. Its formal statement is

$$\forall x \; \forall y \; \exists! A \; \forall z \;\; [z \in A \iff (z = x \text{ or } z = y)]\,.$$

If $x \neq y$, then we write $\{x, y\}$ for the unique set whose only members are x and y and call it an *unordered pair*. We always write $\{x\}$ instead of $\{x, x\}$ and call it a *singleton*. At this point, it makes sense to define the first three natural numbers $0 = \emptyset$, $1 = \{0\}$ and $2 = \{0, 1\}$. We can also justify defining *ordered pairs* by setting

$$(x, y) = \{\{x\}, \{x, y\}\}$$

whenever we are given x and y as we did in Definition 1.1. As a reminder, when $x = y$, what we really have is

$$(x, x) = \{\{x\}\}.$$

Notice that, based on this definition, when we write (x, y), we can tell that x is the *first coordinate* and y is the *second coordinate*. Formally, this means we can prove that for all x, y, x' and y',

$$(x, y) = (x', y') \iff (x = x' \text{ and } y = y').$$

Union Axiom

This axiom allows us to form unions. Its formal statement is

$$\forall \mathcal{F}\ \exists! A\ \forall x\ \ [x \in A \iff \exists Y \in \mathcal{F}\ \ (x \in Y)].$$

We write $\bigcup \mathcal{F}$ for the unique set whose members are exactly the members of the members of $\mathcal{F}$. In other words,

$$\bigcup \mathcal{F} = \{x \mid \text{there exists } Y \in \mathcal{F} \text{ such that } x \in Y\}.$$

It is important to note that, in the Union Axiom, the family $\mathcal{F}$ is allowed to be infinite. We often use different notation when $\mathcal{F}$ is finite. For example, we define

$$A \cup B = \bigcup \{A, B\}$$

and

$$A \cup B \cup C = \bigcup \{A, B, C\}.$$

At this point, we can define the remaining natural numbers

$$3 = 2 \cup \{2\} = \{0, 1, 2\},$$

$$4 = 3 \cup \{3\} = \{0, 1, 2, 3\},$$

$$5 = 4 \cup \{4\} = \{0, 1, 2, 3, 4\}$$

and, in general,

$$n + 1 = n \cup \{n\} = \{0, \ldots, n\}.$$

Power Set Axiom

This axiom allows us to form the set of all subsets of a given set. Its formal statement is

$$\forall A \; \exists! \mathcal{F} \; \forall X \; (X \in \mathcal{F} \iff X \subseteq A).$$

We write $\mathcal{P}(A)$ for the unique set of subsets of A. In other words,

$$\mathcal{P}(A) = \{X \mid X \subseteq A\}.$$

We call $\mathcal{P}(A)$ the *power set* of A. As an example, let us see what happens when we start with the empty set and take power sets over and over. Define

$$V_0 = \emptyset,$$

$$V_1 = \mathcal{P}(V_0) = \{\emptyset\},$$

$$V_2 = \mathcal{P}(V_1) = \{\emptyset, \{\emptyset\}\},$$

$$V_3 = \mathcal{P}(V_2) = \{\emptyset, \{\emptyset\}, \{\{\emptyset\}\}, \{\emptyset, \{\emptyset\}\}\}$$

and, in general,

$$V_{n+1} = \mathcal{P}(V_n).$$

The sets V_n will come up again later.

Comprehension Scheme

This axiom scheme gives us a way to form specific subsets of a given set. It says the following.

> For each "property" $P(x)$, the following is an axiom:
>
> $$\forall A \; \exists! B \; \forall x \; [x \in B \iff (x \in A \text{ and } P(x))].$$

Notice that the word "property" appears in quotes. There are infinitely many properties, which is why ZFC has infinitely many axioms. We will not give a formal definition of "property" because it involves first-order logic, which is not a prerequisite. It is enough