Chapter 1

Axiomatic set theory

1.1 Why axiomatic set theory?

Essentially all mathematical theories deal with sets in one way or another. In most cases, however, the use of set theory is limited to its basics: elementary operations on sets, fundamental facts about functions, and, in some cases, rudimentary elements of cardinal arithmetic. This basic part of set theory is very intuitive and can be developed using only our "good" intuition for what sets are. The theory of sets developed in that way is called "naive" set theory, as opposed to "axiomatic" set theory, where all properties of sets are deduced from a fixed set of axioms.

Clearly the "naive" approach is very appealing. It allows us to prove a lot of facts on sets in a quick and convincing way. Also, this was the way the first mathematicians studied sets, including Georg Cantor, a "father of set theory." However, modern set theory departed from the "paradise" of this simple-minded approach, replacing it with "axiomatic set theory," the highly structured form of set theory. What was the reason for such a replacement?

Intuitively, a set is any collection of all elements that satisfy a certain given property. Thus, the following **axiom schema of comprehension,** due to Frege (1893), seems to be very intuitive.

> If φ is a property, then there exists a set $Y = \{X : \varphi(X)\}$ of all elements having property φ.

This principle, however, is false! It follows from the following theorem of Russell (1903) known as Russell's antinomy or Russell's paradox.

Russell's paradox There is no set $S = \{X : X \notin X\}$.

The axiom schema of comprehension fails for the formula $\varphi(X)$ defined as "$X \notin X$." To see it, notice that if S had been a set we would have had for every Y

$$Y \in S \Leftrightarrow Y \notin Y.$$

Substituting S for Y we obtain

$$S \in S \Leftrightarrow S \notin S,$$

which evidently is impossible.

This paradox, and other similar to it, convinced mathematicians that we cannot rely on our intuition when dealing with abstract objects such as arbitrary sets. To avoid this trouble, "naive" set theory has been replaced with axiomatic set theory.

The task of finding one "universal" axiomatic system for set theory that agrees with our intuition and is free of paradoxes was not easy, and was not without some disagreement. Some of the disagreement still exists today. However, after almost a century of discussions, the set of ten axioms/schemas, known as the Zermelo–Fraenkel axioms (abbreviated as ZFC, where C stands for the axiom of choice), has been chosen as the most natural. These axioms will be introduced and explained in the next chapters. The full list of these axioms with some comments is also included in Appendix A.

It should be pointed out here that the ZFC axioms are far from "perfect." It could be expected that a "perfect" set of axioms should be *complete*, that is, that for any statement φ expressed in the language of set theory (which is described in the next section) either φ or its negation is a consequence of the axioms. Also, a "good" set of axioms should certainly be *consistent*, that is, should not lead to a contradiction. Unfortunately, we cannot prove either of these properties for the ZFC axioms. More precisely, we do *believe* that the ZFC axioms are consistent. However, if this belief is correct, we can't prove it using the ZFC axioms alone. Does it mean that we should search for a better system of set-theory axioms that would be without such a deficiency? Unfortunately, there is no use in searching for it, since no "reasonable" set of axioms of set theory can prove its own consistency. This follows from the following celebrated theorem of Gödel.

Theorem 1.1.1 (Gödel's second incompleteness theorem) *Let T be a set of axioms expressed in a formal language $\mathcal{L}$ (such as the language of set theory described in Section 1.2) and assume that T has the following "reasonable" properties.*

(1) *T is consistent.*

(2) *There is an effective algorithm that decides for an arbitrary sentence of the language $\mathcal{L}$ whether it is in T or not.*

(3) *T is complicated enough to encode simple arithmetic of the natural numbers.*

Then there is a sentence φ of the language $\mathcal{L}$ that encodes the statement "T is consistent." However, φ is not a consequence of the axioms T.

In other words, Theorem 1.1.1 shows us that for whatever "reasonable" systems of axioms of set theory we choose, we will always have to rely on our intuition for its consistency. Thus, the ZFC axioms are as good (or bad) in this aspect as any other "reasonable" system of axioms.

So what about the completeness of the ZFC axioms? Can we prove at least that much? The answer is again negative and once again it is a common property for all "reasonable" systems of axioms, as follows from another theorem of Gödel.

Theorem 1.1.2 (Gödel's first incompleteness theorem) *Let T be a set of axioms expressed in a formal language $\mathcal{L}$ (such as the language of set theory described in Section 1.2) and assume that T has the following "reasonable" properties.*

(1) *T is consistent.*

(2) *There is an effective algorithm that decides for an arbitrary sentence of the language $\mathcal{L}$ whether it is in T or not.*

(3) *T is complicated enough to encode simple arithmetic of the natural numbers.*

Then there is a sentence φ of the language $\mathcal{L}$ such that neither φ nor its negation $\neg\varphi$ can be deduced from the axioms T.

A sentence φ as in Theorem 1.1.2 is said to be *independent* of the axioms T. It is not difficult to prove that a sentence φ is independent of the consistent set of axioms T if and only if both $T \cup \{\varphi\}$ and $T \cup \{\neg\varphi\}$ are consistent too. Part of this course will be devoted to studying the sentences of set theory that are independent of the ZFC axioms.

The preceding discussion shows that there is no way to find a good complete set of axioms for set theory. On the other hand, we can find a set of axioms that reach far enough to allow encoding of all set-theoretic operations and all classical mathematical structures. Indeed, the ZFC axioms do satisfy this requirement, and the rest of Part I will be devoted to describing such encodings of all structures of interest.

1.2 The language and the basic axioms

Any mathematical theory must begin with undefined concepts. In the case of set theory these concepts are the notion of a "set" and the relation "is an element of" between the sets. In particular, we write "$x \in y$" for "x is an element of y."

The relation $\in$ is primitive for set theory, that is, we do not define it. All other objects, including the notion of a set, are described by the axioms. In particular, all objects considered in formal set theory are sets. (Thus, the word "set" is superfluous.)

In order to talk about any formal set theory it is necessary to specify first the language that we will use and in which we will express the axioms. The correct expressions in our language are called *formulas*. The basic formulas are "$x \in y$" and "$z = t$," where x, y, z, and t (or some other variable symbols) stand for the sets. We can combine these expressions using the basic logical connectors of negation $\neg$, conjunction &, disjunction $\vee$, implication $\rightarrow$, and equivalence $\leftrightarrow$. Thus, for example, $\neg\varphi$ means "not φ" and $\varphi\rightarrow\psi$ stands for "φ implies ψ." In addition, we will use two quantifiers: existential $\exists$ and universal $\forall$. Thus, an expression $\forall x\varphi$ is interpreted as "for all x formula φ holds." In addition, the parentheses "(" and ")" are used, when appropriate.

Formally, the formulas can be created only as just described. However, for convenience, we will also often use some shortcuts. For example, an expression $\exists x \in A\varphi(x)$ will be used as an abbreviation for $\exists x(x \in A \ \&\ \varphi(x))$, and we will write $\forall x \in A\varphi(x)$ to abbreviate the formula $\forall x(x \in A\rightarrow\varphi(x))$. Also we will use the shortcuts $x \neq y$, $x \notin y$, $x \subset y$, and $x \not\subset y$, where, for example, $x \subset y$ stands for $\forall z(z \in x\rightarrow z \in y)$.

Finally, only variables, the relations $=$ and $\in$, and logical symbols already mentioned are allowed in formal formulas. However, we will often use some other constants. For example, we will write $x = \emptyset$ (x is an empty set) in place of $\neg\exists y(y \in x)$.

We will discuss ZFC axioms throughout the next few sections as they are needed. Also, in most cases, we won't write in the main text the formulas representing the axioms. However, the full list of ZFC axioms together with the formulas can be found in Appendix A.

Let us start with the two most basic axioms.

Set existence axiom There exists a set: $\exists x(x = x)$.

Extensionality axiom If x and y have the same elements, then x is equal to y.

The set existence axiom follows from the others. However, it is the most basic of all the axioms, since it ensures that set theory is not a trivial theory. The extensionality axiom tells us that the sets can be distinguish only by their elements.

Comprehension scheme (or schema of separation) For every formula $\varphi(s,t)$ with free variables s and t, set x, and parameter p there exists a set $y = \{u \in x \colon \varphi(u,p)\}$ that contains all those $u \in x$ that have the property φ.

Notice that the comprehension scheme is, in fact, the scheme for infinitely many axioms, one for each formula φ. It is a weaker version of Frege's axiom schema of comprehension. However, the contradiction of Russell's paradox can be avoided, since the elements of the new set $y = \{u \in x \colon \varphi(u,p)\}$ are chosen from a fixed *set* x, rather than from an undefined object such as "the class of all sets."

From the set existence axiom and the comprehension scheme used with the formula "$u \neq u$," we can conclude the following stronger version of the set existence axiom.

Empty set axiom There exists the *empty set* $\emptyset$.

To see the implication, simply define $\emptyset = \{y \in x \colon y \neq y\}$, where x is a set from the set existence axiom. Notice that by the extensionality axiom the empty set is unique.

An interesting consequence of the comprehension scheme axiom is the following theorem.

Theorem 1.2.1 *There is no set of all sets.*

Proof If there were a set S of all sets then the following set

$$Z = \{X \in S \colon X \notin X\}$$

would exist by the comprehension scheme axiom. However, with S being the set of all sets, we would have that $Z = \{X \colon X \notin X\}$, the set from Russell's paradox. This contradiction shows that the set S of all sets cannot exist. □

By the previous theorem all sets do not form a set. However, we sometimes like to talk about this object. In such a case we will talk about a *class* of sets or the set-theoretic *universe*. We will talk about classes only on an intuitive level. It is worth mentioning, however, that the theory of classes can also be formalized similarly to the theory of sets. This, however, is far beyond the scope of this course. Let us mention only that there are other proper classes of sets (i.e., classes that are not sets) that are strictly smaller than the class of all sets.

The comprehension scheme axiom is a conditional existence axiom, that is, it describes how to obtain a set (subset) from another set. Other basic conditional existence axioms are listed here.

Pairing axiom For any a and b there exists a set x that contains a and b.

Union axiom For every family $\mathcal{F}$ there exists a set U containing the *union* $\bigcup \mathcal{F}$ of all elements of $\mathcal{F}$.

Power set axiom For every set X there exists a set P containing the set $\mathcal{P}(X)$ (the *power set*) of all subsets of X.

In particular, the pairing axiom states that for any a and b there exists a set x such that $\{a,b\} \subset x$. Although it does not state directly that there exists a set $\{a,b\}$, the existence of this set can easily be concluded from the existence of x and the comprehension scheme axiom:

$$\{a,b\} = \{u \in x \colon u = a \vee u = b\}.$$

Similarly, we can conclude from the union and power set axioms that for every sets $\mathcal{F}$ and X there exist the union of $\mathcal{F}$

$$\bigcup \mathcal{F} = \{x \colon \exists F \in \mathcal{F}\ (x \in F)\} = \{x \in U \colon \exists F \in \mathcal{F}\ (x \in F)\}$$

and the power set of X

$$\mathcal{P}(X) = \{z \colon z \subset X\} = \{z \in P \colon z \subset X\}.$$

It is also easy to see that these sets are defined uniquely. Notice also that the existence of a set $\{a,b\}$ implies the existence of a singleton set $\{a\}$, since $\{a\} = \{a,a\}$.

The other basic operations on sets can be defined as follows: the union of two sets x and y by

$$x \cup y = \bigcup \{x,y\};$$

the difference of sets x and y by

$$x \setminus y = \{z \in x \colon z \notin y\};$$

the arbitrary intersections of a family $\mathcal{F}$ by

$$\bigcap \mathcal{F} = \left\{z \in \bigcup \mathcal{F} \colon \forall F \in \mathcal{F}\ (z \in F)\right\};$$

and the intersections of sets x and y by

$$x \cap y = \bigcap \{x,y\}.$$

The existence of sets $x \backslash y$ and $\bigcap \mathcal{F}$ follows from the axiom of comprehension.

We will also sometimes use the operation of *symmetric difference* of two sets, defined by

$$x \triangle y = (x \setminus y) \cup (y \setminus x).$$

Its basic properties are listed in the next theorem. Its proof is left as an exercise.

Theorem 1.2.2 *For every x, y, and z*

(a) $x \triangle y = y \triangle x$,

(b) $x \triangle y = (x \cup y) \setminus (x \cap y)$,

(c) $(x \triangle y) \triangle z = x \triangle (y \triangle z)$.

We will define an ordered pair $\langle a, b \rangle$ for arbitrary a and b by

$$\langle a, b \rangle = \{\{a\}, \{a, b\}\}. \tag{1.1}$$

It is difficult to claim that this definition is natural. However, it is commonly accepted in modern set theory, and the next theorem justifies it by showing that it maintains the intuitive properties we usually associate with the ordered pair.

Theorem 1.2.3 *For arbitrary a, b, c, and d*

$$\langle a, b \rangle = \langle c, d \rangle \textit{ if and only if } a = c \textit{ and } b = d.$$

Proof Implication $\Leftarrow$ is obvious.

To see the other implication, assume that $\langle a, b \rangle = \langle c, d \rangle$. This means that $\{\{a\}, \{a, b\}\} = \{\{c\}, \{c, d\}\}$. In particular, by the axiom of extensionality, $\{a\}$ is equal to either $\{c\}$ or $\{c, d\}$.

If $\{a\} = \{c\}$ then $a = c$. If $\{a\} = \{c, d\}$, then c must belong to $\{a\}$ and we also conclude that $a = c$. In any case, $a = c$ and we can deduce that $\{\{a\}, \{a, b\}\} = \{\{a\}, \{a, d\}\}$. We wish to show that this implies $b = d$.

But $\{a, b\}$ belongs to $\{\{a\}, \{a, d\}\}$. Thus we have two cases.

Case 1: $\{a, b\} = \{a, d\}$. Then $b = a$ or $b = d$. If $b = d$ we are done. If $b = a$ then $\{a, b\} = \{a\}$ and so $\{a, d\} = \{a\}$. But d belongs then to $\{a\}$ and so $d = a$. Since we had also $a = b$ we conclude $b = d$.

Case 2: $\{a, b\} = \{a\}$. Then b belongs to $\{a\}$ and so $b = a$. Hence we conclude that $\{a, d\} = \{a\}$, and as in case 1 we can conclude that $b = d$. $\square$

Now we can define an ordered triple $\langle a, b, c\rangle$ by identifying it with $\langle\langle a, b\rangle, c\rangle$ and, in general, an ordered n-tuple by

$$\langle a_1, a_2, \ldots, a_{n-1}, a_n\rangle = \langle\langle a_1, a_2, \ldots, a_{n-1}\rangle, a_n\rangle.$$

The agreement of this definition with our intuition is given by the following theorem, presented without proof.

Theorem 1.2.4 *$\langle a_1, a_2, \ldots, a_{n-1}, a_n\rangle = \langle a'_1, a'_2, \ldots, a'_{n-1}, a'_n\rangle$ if and only if $a_i = a'_i$ for all $i = 1, 2, \ldots, n$.*

Next we will define a *Cartesian product* $X \times Y$ as the set of all ordered pairs $\langle x, y\rangle$ such that $x \in X$ and $y \in Y$. To make this definition formal, we have to use the comprehension axiom. For this, notice that for every $x \in X$ and $y \in Y$ we have

$$\langle x, y\rangle = \{\{x\}, \{x, y\}\} \in \mathcal{P}(\mathcal{P}(X \cup Y)).$$

Hence, we can define

$$X \times Y = \{z \in \mathcal{P}(\mathcal{P}(X \cup Y))\colon \exists x \in X\ \exists y \in Y\ (z = \langle x, y\rangle)\}. \tag{1.2}$$

The basic properties of the Cartesian product and its relation to other set-theoretic operations are described in the exercises.

The last axiom we would like to discuss in this section is the infinity axiom. It states that there exists at least one infinite set. This is the only axiom that implies the existence of an infinite object. Without it, the family $\mathcal{F}$ of all finite subsets of the set of natural numbers would be a good "model" of set theory, that is, $\mathcal{F}$ satisfies all the axioms of set theory except the infinity axiom.

To make the statements of the infinity axiom more readable we introduce the following abbreviation. We say that y is a successor of x and write $y = S(x)$ if $y = x \cup \{x\}$, that is, when

$$\forall z[z \in y \leftrightarrow (z \in x \vee z = x)].$$

Infinity axiom (Zermelo 1908) There exists an infinite set (of some special form):

$$\exists x \left[\forall z(z = \emptyset \rightarrow z \in x)\ \&\ \forall y \in x \forall z(z = S(y) \rightarrow z \in x)\right].$$

Notice that the infinity axiom obviously implies the set existence axiom.

EXERCISES

1 Prove that if $F \in \mathcal{F}$ then $\bigcap \mathcal{F} \subset F \subset \bigcup \mathcal{F}$.

2 Show that for every family $\mathcal{F}$ and every set A

(a) if $A \subset F$ for every $F \in \mathcal{F}$ then $A \subset \bigcap \mathcal{F}$, and

(b) if $F \subset A$ for every $F \in \mathcal{F}$ then $\bigcup \mathcal{F} \subset A$.

3 Prove that if $\mathcal{F} \cap \mathcal{G} \neq \emptyset$ then $\bigcap \mathcal{F} \cap \bigcap \mathcal{G} \subset \bigcap(\mathcal{F} \cap \mathcal{G})$. Give examples showing that the inclusion cannot be replaced by equality and that the assumption $\mathcal{F} \cap \mathcal{G} \neq \emptyset$ is essential.

4 Prove Theorem 1.2.2.

5 Show that $\langle\langle a, b\rangle, c\rangle = \langle\langle a', b'\rangle, c'\rangle$ if and only if $\langle a, \langle b, c\rangle\rangle = \langle a', \langle b', c'\rangle\rangle$ if and only if $a = a'$, $b = b'$, and $c = c'$. Conclude that we could define an ordered triple $\langle a, b, c\rangle$ as $\langle a, \langle b, c\rangle\rangle$ instead of $\langle\langle a, b\rangle, c\rangle$.

6 Prove that $X \times Y = \emptyset$ if and only if $X = \emptyset$ or $Y = \emptyset$.

7 Show that for arbitrary sets X, Y, and Z the following holds.

(a) $(X \cup Y) \times Z = (X \times Z) \cup (Y \times Z)$.

(b) $(X \cap Y) \times Z = (X \times Z) \cap (Y \times Z)$.

(c) $(X \setminus Y) \times Z = (X \times Z) \setminus (Y \times Z)$.

8 Prove that if $X \times Z \subset Y \times T$ and $X \times Z \neq \emptyset$ then $X \subset Y$ and $Z \subset T$. Give an example showing that the assumption $X \times Z \neq \emptyset$ is essential.