Symbol	Page	Symbol	Page	Symbol	Page
$[x, y]$	44	$<_Z$	98	E	182
$\bigcup_{i \in I}$	51	$\mathbb{Q}$	102	$\cong$	184
$\bigcap_{i \in I}$	51	0_Q	102	$<_L$	185
F_i	51	1_Q	102	$<^\circ$	189
${}^A B$	52	$+_Q$	103	Ω	194
$\bigtimes_{i \in I}$	54	$\cdot_Q$	105	$\aleph_1$	199
$[x]_R$	57	r^{-1}	107	V_α	200
$[x]$	57	$s \div r$	107	rank S	204
A/R	58	$<_Q$	108	$\aleph_\alpha$	212
0	67	$\lvert r \rvert$	109	$\beth_\alpha$	214
1	67	$\mathbb{R}$	113	sup S	216
2	67	$<_R$	113	it$\langle A, R \rangle$	220
3	67	$+_R$	114	kard S	222
a^+	68	0_R	115	$\oplus$	222
ω	69	$-x$	117	$\rho + \sigma$	222
σ	71	$\lvert x \rvert$	118	$\bar{\alpha}$	223
$\mathbb{Z}$	75	$\cdot_R$	118	η	223
$m + n$	79	1_R	119	ρ^*	224
A1	79	$\approx$	129	$R * S$	224
A2	79	(x, y)	130	$\rho \cdot \sigma$	224
$m \cdot n$	80	card S	136	$\alpha + \beta$	228
M1	80	$\aleph_0$	137	$\alpha \cdot \beta$	228
M2	80	$\kappa + \lambda$	139	$<_H$	228
m^n	80	$\kappa \cdot \lambda$	139	A3	229
E1	80	κ^λ	139	M3	229
E2	80	$\kappa!$	144	α^β	232
$\in$	83	$\preccurlyeq$	145	E3	232
$\in_S$	83	$\kappa \leq \lambda$	145	ε_0	240
$\subset$	85	$\kappa < \lambda$	146	R^ι	243
$\mathbb{Z}$	91	Sq(A)	160	σ^M	249
$+_Z$	92	$\subset_S$	167	ZFC	253
0_Z	92	$\leq$	168	HF	256
$-a$	95	seg	173	cf λ	257
$\cdot_Z$	95	${}^{<A}B$	175	ssup S	262
1_Z	97	TC S	178	$\ulcorner \varphi \urcorner$	263

CHAPTER 1

INTRODUCTION

BABY SET THEORY

We shall begin with an informal discussion of some basic concepts of set theory. In these days of the "new math," much of this material will be already familiar to you. Indeed, the practice of beginning each mathematics course with a discussion of set theory has become widespread, extending even to the elementary schools. But we want to review here elementary-school set theory (and do it in our notation). Along the way we shall be able to point out some matters that will become important later. We shall not, in these early sections, be particularly concerned with rigor. The more serious work will start in Chapter 2.

A *set* is a collection of things (called its *members* or *elements*), the collection being regarded as a single object. We write "$t \in A$" to say that t is a member of A, and we write "$t \notin A$" to say that t is not a member of A.

For example, there is the set whose members are exactly the prime numbers less than 10. This set has four elements, the numbers 2, 3, 5, and 7. We can name the set conveniently by listing the members within braces (curly brackets):

$$\{2, 3, 5, 7\}.$$

relation ($\in$). If we want to know whether $A \in B$, we look at the set A as a single object, and we check to see if this single object is among the members of B. By contrast, if we want to know whether $A \subseteq B$, then we must open up the set A, examine its various members, and check whether its various members can be found among the members of B.

Examples 1. $\varnothing \subseteq \varnothing$, but $\varnothing \notin \varnothing$.

2. $\{\varnothing\} \in \{\{\varnothing\}\}$ but $\{\varnothing\} \nsubseteq \{\{\varnothing\}\}$. $\{\varnothing\}$ is not a subset of $\{\{\varnothing\}\}$ because there is a member of $\{\varnothing\}$, namely $\varnothing$, that is not a member of $\{\{\varnothing\}\}$.

3. Let Us be the set of all people in the United States, and let Un be the set of all countries belonging to the United Nations. Then

$$\text{John Jones} \in Us \in Un.$$

But John Jones $\notin Un$ (since he is not even a country), and hence $Us \nsubseteq Un$.

Any set A will have one or more subsets. (In fact, if A has n elements, then A has 2^n subsets. But this is a matter we will take up much later.) We can gather all of the subsets of A into one collection. We then have the set of all subsets of A, called the *power*[1] *set* $\mathscr{P}A$ of A. For example,

$$\mathscr{P}\varnothing = \{\varnothing\},$$
$$\mathscr{P}\{\varnothing\} = \{\varnothing, \{\varnothing\}\},$$
$$\mathscr{P}\{0, 1\} = \{\varnothing, \{0\}, \{1\}, \{0, 1\}\}.$$

A very flexible way of naming a set is the method of *abstraction*. In this method we specify a set by giving the condition—the entrance requirement—that an object must satisfy in order to belong to the set. In this way we obtain the set of all objects x such that x meets the entrance requirement. The notation used for the set of all objects x such that the condition $\_\_\, x \,\_\_$ holds is

$$\{x \mid \_\_\, x \,\_\_\}.$$

For example:

1. $\mathscr{P}A$ is the set of all objects x such that x is a subset of A. Here "x is a subset of A" is the entrance requirement that x must satisfy in order to belong to $\mathscr{P}A$. We can write

$$\mathscr{P}A = \{x \mid x \text{ is a subset of } A\}$$
$$= \{x \mid x \subseteq A\}.$$

[1] The reasons for using the word "power" in this context are not very convincing, but the usage is now well established.

2. $A \cap B$ is the set of all objects y such that $y \in A$ and $y \in B$. We can write

$$A \cap B = \{y \mid y \in A \text{ and } y \in B\}.$$

It is unimportant whether we use "x" or "y" or another letter as the symbol (which is used as a pronoun) here.

3. The set $\{z \mid z \neq z\}$ equals $\varnothing$, because the entrance requirement "$z \neq z$" is not satisfied by any object z.

4. The set $\{n \mid n \text{ is an even prime number}\}$ is the same as the set $\{2\}$.

There are, however, some dangers inherent in the abstraction method. For certain bizarre choices of the entrance requirement, it may happen that there is no set containing exactly those objects meeting the entrance requirement. There are two ways in which disaster can strike.

One of the potential disasters is illustrated by

$$\{x \mid x \text{ is a positive integer definable in one line of type}\}.$$

The tricky word here is "definable." Some numbers are easy to define in one line. For example, the following lines each serve to define a positive integer:

12,317,
the millionth prime number,
the least number of the form $2^{2^n} + 1$ that is not prime,
the 23rd perfect[2] number.

Observe that there are only finitely many possible lines of type (because there are only finitely many symbols available to the printer, and there is a limit to how many symbols will fit on a line). Consequently

$$\{x \mid x \text{ is a positive integer definable in one line of type}\}$$

is only a finite set of integers. Consider the least positive integer not in this set; that is, consider

the least positive integer not definable in one line of type.

The preceding line defines a positive integer in one line, but that number is, by its construction, not definable in one line! So we are in trouble, and the trouble can be blamed on the entrance requirement of the set, i.e., on the phrase "is a positive integer definable in one line of type." While it may have

[2] A positive integer is *perfect* if it equals the sum of its smaller divisors, e.g., $6 = 1 + 2 + 3$. It is *deficient* (or *abundant*) if the sum of its smaller divisors is less than (or greater than, respectively) the number itself. This terminology is a vestigial trace of numerology, the study of the mystical significance of numbers. The first four perfect numbers are 6, 28, 496, and 8128.

appeared originally to be a meaningful entrance requirement, it now appears to be gravely defective. (This example was given by G. G. Berry in 1906. A related example was published in 1905 by Jules Richard.)

There is a second disaster that can result from an overly free-swinging use of the abstraction method. It is exemplified by

$$\{x \mid x \notin x\},$$

this is, by the set of all objects that are not members of themselves. Call this set A, and ask "is A a member of itself?" If $A \notin A$, then A meets the entrance requirement for A, whereupon $A \in A$. But on the other hand, if $A \in A$, then A fails to meet the entrance requirement and so $A \notin A$. Thus both "$A \in A$" and "$A \notin A$" are untenable. Again, we are in trouble. The phrase "is not a member of itself" appears to be an illegal entrance requirement for the abstraction method. (This example is known as Russell's paradox. It was communicated by Bertrand Russell in 1902 to Gottlob Frege, and was published in 1903. The example was independently discovered by Ernst Zermelo.)

These two sorts of disaster will be blocked in precise ways in our axiomatic treatment, and less formally in our nonaxiomatic treatment. The first sort of disaster (the Berry example) will be avoided by adherence to entrance requirements that can be stated in a totally unambiguous form, to be specified in the next chapter. The second sort of disaster will be avoided by the distinction between *sets* and *classes*. Any collection of sets will be a *class*. Some collections of sets (such as the collections $\varnothing$ and $\{\varnothing\}$) will be sets. But some collections of sets (such as the collection of all sets not members of themselves) will be too large to allow as sets. These oversize collections will be called *proper classes*. The distinction will be discussed further presently.

In practice, avoidance of disaster will not really be an oppressive or onerous task. We will merely avoid ambiguity and avoid sweepingly vast sets. A prudent person would not want to do otherwise.

Exercises

1. Which of the following become true when "$\in$" is inserted in place of the blank? Which become true when "$\subseteq$" is inserted?

(a) $\{\varnothing\}$ __ $\{\varnothing, \{\varnothing\}\}$.
(b) $\{\varnothing\}$ __ $\{\varnothing, \{\{\varnothing\}\}\}$.
(c) $\{\{\varnothing\}\}$ __ $\{\varnothing, \{\varnothing\}\}$.
(d) $\{\{\varnothing\}\}$ __ $\{\varnothing, \{\{\varnothing\}\}\}$.
(e) $\{\{\varnothing\}\}$ __ $\{\varnothing, \{\varnothing, \{\varnothing\}\}\}$.

2. Show that no two of the three sets $\varnothing$, $\{\varnothing\}$, and $\{\{\varnothing\}\}$ are equal to each other.

3. Show that if $B \subseteq C$, then $\mathscr{P}B \subseteq \mathscr{P}C$.

4. Assume that x and y are members of a set B. Show that $\{\{x\}, \{x, y\}\} \in \mathscr{P}\mathscr{P}B$.

SETS—AN INFORMAL VIEW

We are about to present a somewhat vague description of how sets are obtained. (The description will be repeated much later in precise form.) None of our later work will actually depend on this informal description, but we hope it will illuminate the motivation behind some of the things we will do.

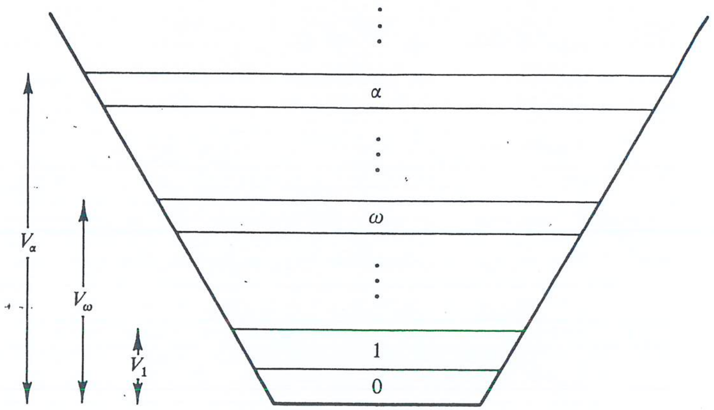


Fig. 2. V_0 is the set A of atoms.

First we gather together all those things that are not themselves sets but that we want to have as members of sets. Call such things *atoms*. For example, if we want to be able to speak of the set of all two-headed coins, then we must include all such coins in our collection of atoms. Let A be the set of all atoms; it is the first set in our description.

We now proceed to build up a hierarchy

$$V_0 \subseteq V_1 \subseteq V_2 \subseteq \cdots$$

of sets. At the bottom level (in a vertical arrangement as in Fig. 2) we take $V_0 = A$, the set of atoms. The next level will also contain all sets of atoms:

$$V_1 = V_0 \cup \mathscr{P}V_0 = A \cup \mathscr{P}A.$$

The third level contains everything that is in a lower level, plus all sets of things from lower levels:

$$V_2 = V_1 \cup \mathscr{P}V_1.$$

And in general

$$V_{n+1} = V_n \cup \mathscr{P}V_n.$$

Thus we obtain successively V_0, V_1, V_2, ... But even this infinite hierarchy does not include enough sets. For example, $\varnothing \in V_1$, $\{\varnothing\} \in V_2$, $\{\{\varnothing\}\} \in V_3$, etc., but we do not yet have the infinite set

$$\{\varnothing, \{\varnothing\}, \{\{\varnothing\}\}, \ldots\}.$$

To remedy this lack, we take the infinite union

$$V_\omega = V_0 \cup V_1 \cup \cdots,$$

and then let $V_{\omega+1} = V_\omega \cup \mathscr{P}V_\omega$, and we continue. In general for any α,

$$V_{\alpha+1} = V_\alpha \cup \mathscr{P}V_\alpha,$$

and this goes on "forever." Whenever you might think that the construction is finished, you instead take the union of all the levels obtained so far, take the power set of that union, and continue.

A better explanation of the "forever" idea must be delayed until we discuss (in Chapter 7) the "numbers" being used as subscripts in the preceding paragraphs. These are the so-called "ordinal numbers." The ordinal numbers begin with 0, 1, 2, ...; then there is the infinite number ω, then $\omega + 1$, $\omega + 2$, ...; and this goes on "forever."

A fundamental principle is the following: Every set appears somewhere in this hierarchy. That is, for every set a there is some α with $a \in V_{\alpha+1}$. That is what the sets are; they are the members of the levels of our hierarchy.

Examples Suppose that a and b are sets. Say that $a \in V_{\alpha+1}$ and $b \in V_{\beta+1}$ and suppose that $V_{\beta+1}$ is "higher" in the hierarchy than $V_{\alpha+1}$. Then both a and b are in $V_{\beta+1}$, since each level includes all lower levels. Consequently in $V_{\beta+2}$ we have the pair set $\{a, b\}$. On the other hand at no point do we obtain a set of all sets, i.e., a set having all sets as members. There simply is no such set.

There is one way in which we can simplify our picture. We were very indefinite about just what was in the set A of atoms. The fact of the matter is that the atoms serve no mathematically necessary purpose, so we banish them; we take $A = \varnothing$. In so doing, we lose the ability to form sets of flowers or sets of people. But this is no cause for concern; we do not need set theory to talk about people and we do not need people in our set theory. But we definitely do want to have sets of numbers, e.g., $\{2, 3 + i\pi\}$. Numbers do not appear at first glance to be sets. But as we shall discover (in Chapters 4 and 5), we can find sets that serve perfectly well as numbers.

Our theory then will ignore all objects that are not sets (as interesting and real as such objects may be). Instead we will concentrate just on "pure" sets that can be constructed without the use of such external objects. In

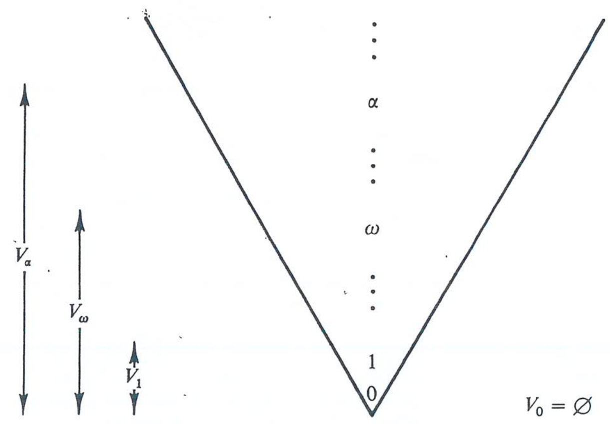


Fig. 3. The ordinals are the backbone of the universe.

particular, any member of one of our sets will itself be a set, and each of *its* members, if any, will be a set, and so forth. (This does not produce an infinite regress, because we stop when we reach $\varnothing$.)

Now that we have banished atoms, the picture becomes narrower (Fig. 3). The construction is also simplified. We have defined $V_{\alpha+1}$ to be $V_\alpha \cup \mathscr{P}V_\alpha$. Now it turns out that this is the same as $A \cup \mathscr{P}V_\alpha$ (see Exercise 6). With $A = \varnothing$, we have simply $V_{\alpha+1} = \mathscr{P}V_\alpha$.

Exercises

5. Define the rank of a set c to be the least α such that $c \subseteq V_\alpha$. Compute the rank of $\{\{\varnothing\}\}$. Compute the rank of $\{\varnothing, \{\varnothing\}, \{\varnothing, \{\varnothing\}\}\}$.
6. We have stated that $V_{\alpha+1} = A \cup \mathscr{P}V_\alpha$. Prove this at least for $\alpha < 3$.
7. List all the members of V_3. List all the members of V_4. (It is to be assumed here that there are no atoms.)

CLASSES

There is no "set of all sets," i.e., there is no set having all sets as members. This is in accordance with our informal image of the hierarchical way sets are constructed. Later, the nonexistence of a set of all sets will become a theorem (Theorem 2A), provable from the axioms.

Nonetheless, there is some mild inconvenience that results if we forbid ourselves even to speak of the collection of all sets. The collection cannot be a set, but what status can we give it? Basically there are two alternatives:

The Zermelo-Fraenkel alternative The collection of all sets need have no ontological status at all, and we need never speak of it. When tempted to speak of it, we can seek a rephrasing that avoids it.

The von Neumann–Bernays alternative The collection of all sets can be called a *class*. Similarly any other collection of sets can be called a class. In particular, any set is a class, but some classes are too large to be sets. Informally, a class A is a set if it is included in some level V_α of our hierarchy (and then is a member of $V_{\alpha+1}$). Otherwise it is not a set, and can never be a member of a set.

For advanced works in set theory, the Zermelo–Fraenkel alternative seems to be the better of the two. It profits from the simplicity of having to deal with only one sort of object (sets) instead of two (classes and sets). And the circumlocutions it requires (in order to avoid reference to classes that are not sets) are things to which set-theorists learn early to adapt.

For introductory works in set theory (such as this book), the choice between the two alternatives is less clear. The prohibition against mentioning any class that fails to be a set seems unnatural and possibly unfair. Desiring to have our cake and eat it too, we will proceed as follows. We officially adopt the Zermelo–Fraenkel alternative. Consequently the axioms and theorems shall make no mention of any class that is not a set. But in the expository comments, we will not hesitate to mention, say, the class of all sets if it appears helpful to do so. To avoid confusion, we will reserve upper-case sans serif letters ($\mathsf{A}, \mathsf{B}, \ldots$) for classes that are not guaranteed to be sets.

AXIOMATIC METHOD

In this book we are going to state the axioms of set theory, and we are going to show that our theorems are consequences of those axioms. The great advantage of the axiomatic method is that it makes totally explicit just what our initial assumptions are.

It is sometimes said that "mathematics can be embedded in set theory." This means that mathematical objects (such as numbers and differentiable functions) can be defined to be certain sets. And the theorems of mathematics (such as the fundamental theorem of calculus) then can be viewed as statements about sets. Furthermore, these theorems will be provable from our axioms. Hence our axioms provide a sufficient collection of assumptions for the development of the whole of mathematics—a remarkable fact. (In Chapter 5 we will consider further the procedure for embedding mathematics in set theory.)

The axiomatic method has been useful in other subjects as well as in set theory. Consider plane geometry, for example. It is quite possible to talk about lines and triangles without using axioms. But the advantages of axiomatizing geometry were seen very early in the history of the subject.

The nonaxiomatic approach to set theory is often referred to as "naive set theory," a terminology that does not hide its bias. Historically, set theory originated in nonaxiomatic form. But the paradoxes of naive set theory (the most famous being Russell's paradox) forced the development of axiomatic set theory, by showing that certain assumptions, apparently plausible, were inconsistent and hence totally untenable. It then became mandatory to give explicit assumptions that could be examined by skeptics for possible inconsistency. Even without the foundational crises posed by the paradoxes of naive set theory, the axiomatic approach would have been developed to cope with later controversy over the truth or falsity of certain principles, such as the axiom of choice (Chapter 6). Of course our selection of axioms will be guided by the desire to reflect as accurately as possible our informal (preaxiomatic) ideas regarding sets and classes.

It is nonetheless quite possible to study set theory from the nonaxiomatic viewpoint. We have therefore arranged the material in the book in a "two-tier" fashion. The passages dealing with axioms will henceforth be marked by a stripe in the left margin. The reader who omits such passages and reads only the unstriped material will thereby find a nonaxiomatic development of set theory. Perhaps he will at some later time wish to look at the axiomatic underpinnings. On the other hand, the reader who omits nothing will find an axiomatic development. Notice that most of the striped passages appear early in the book, primarily in the first three chapters. Later in the book the nonaxiomatic and the axiomatic approaches largely converge.

Our axiom system begins with two *primitive notions*, the concepts of "set" and "member." In terms of these concepts we will define others, but the primitive notions remain undefined. Instead we adopt a list of *axioms* concerning the primitive notions. (The axioms can be thought of as divulging partial information regarding the meaning of the primitive notions.)

Having adopted a list of axioms, we will then proceed to derive sentences that are *logical consequences* (or *theorems*) of the axioms. Here a sentence σ is said to be a logical consequence of the axioms if any assignment of meaning to the undefined notions of set and member making the axioms true also makes σ true.

We have earlier sketched, in an informal and sometimes vague way, what "set" and "member" are *intended* to mean. But for a sentence to be a logical consequence of the axioms, it must be true *whatever* "set" and "member" mean, provided only that the axioms are true. The sentences that appear, on the basis of our informal viewpoint, as if they ought to be true, must still be shown to be logical consequences of the axioms before being accepted as theorems. In return for adopting this restriction, we escape any drawbacks of the informality and vagueness of the nonaxiomatic viewpoint.

(There is an interesting point here concerning the foundations of mathematics. If σ is a logical consequence of a list of axioms, is there then a finitely long *proof* of σ from the axioms? The answer is affirmative, under a very reasonable definition of "proof." This is an important result in mathematical logic. The topic is treated, among other places, in our book *A Mathematical Introduction to Logic*, Academic Press, 1972.)

For example, the first of our axioms is the axiom of extensionality, which is almost as follows: Whenever A and B are sets such that exactly the same things are members of one as are members of the other, then $A = B$. Imagine for the moment that this were our *only* axiom. We can then consider some of the logical consequences of this one axiom.

For a start, take the sentence: "There cannot be two different sets, each of which has no members." This sentence *is* a logical consequence of extensionality, for we claim that any assignment of meaning to "set" and "member" making extensionality true also makes the above sentence true. To prove this, we argue as follows. Let A and B be any sets, each of which has no members. Then exactly the same things belong to A as to B, since none belong to either. Hence by extensionality, $A = B$. (The validity of this argument, while independent of the meaning of "set" or "member," does depend on the meaning of the logical terms such as "each," "no," "equal," etc.)

On the other hand, consider a simple sentence such as σ: "There are two sets, one of which is a member of the other." This sentence is *not* a logical consequence of extensionality. To see this, let the word "set" mean "a number equal to 2," and let "member of" mean "unequal to." Under this interpretation, extensionality is true but σ is false. Of course we will soon add other axioms of which σ will be a logical consequence.

NOTATION

To denote sets we will use a variety of letters, both lowercase (a, b, ...), uppercase (A, B, ...), and even script letters and Greek letters. Where feasible, we will attempt to have the larger and fancier letters denote sets higher in our hierarchy of levels than those denoted by smaller and plainer letters. In addition, letters can be embellished with subscripts, primes, and the like. This assures us of an inexhaustible supply of symbols for sets. In other words, when it comes to naming sets, anything goes.

It will often be advantageous to exploit the symbolic notation of mathematical logic. This symbolic language, when used in judicious amounts to replace the English language, has the advantages of both conciseness (so that expressions are shorter) and preciseness (so that expressions are less ambiguous).

The following symbolic expressions will be used to abbreviate the corresponding English expressions:

$\forall x$	for every set x
$\exists x$	there exists a set x such that
$\neg$	not
&	and
or	or (in the sense "one or the other *or both*")
$\Rightarrow$	implies ("__ $\Rightarrow$ __" abbreviates "if __, then __")
$\Leftrightarrow$	if and only if, also abbreviated "iff"

You have probably seen these abbreviations before; they are discussed in more detail in the Appendix.

We also have available the symbols $\in$ and $=$ (and $\notin$ and $\neq$, although we could economize by eliminating, e.g., "$a \notin B$" in favor of "$\neg a \in B$"). With all these symbols, variables (a, b, ...), and parentheses we *could* avoid the English language altogether in the statement of axioms and theorems. We will not actually do so, or at least not all at once. But the fact that we have this splendid formal language available is more than a theoretical curiosity. When we come to stating certain axioms, the formal language will be a necessity.

Notice that we read $\forall x$ as "for all *sets* x," rather than "for all *things* x." This is due to our decision to eliminate atoms from our theory. A result of this elimination is that everything we consider is a set; e.g., every member of a set will itself be a set.

Example The principle of extensionality can be written as

$$\forall A\ \forall B[(A \text{ and } B \text{ have exactly the same members}) \Rightarrow A = B].$$

Then "A and B have exactly the same members" can be written as

$$\forall x(x \in A \quad \Leftrightarrow \quad x \in B)$$

so that extensionality becomes

$$\forall A\ \forall B[\forall x(x \in A \quad \Leftrightarrow \quad x \in B) \quad \Rightarrow \quad A = B].$$

Example "There is a set to which nothing belongs" can be written

$$\exists B\ \forall x\ x \notin B.$$

These two examples constitute our first two axioms. It is not really necessary for us to state the axioms in symbolic form. But we will seize the opportunity to show how it *can* be done, on the grounds that the more such examples you see, the more natural the notation will become to you.

HISTORICAL NOTES

The concept of a set is very basic and natural, and has been used in mathematical writings since ancient times. But the theory of abstract sets, as objects to be studied for their own interest, was originated largely by Georg Cantor (1845–1918). Cantor was a German mathematician, and his papers on set theory appeared primarily during the period from 1874 to 1897.

Cantor was led to the study of set theory in a very indirect way. He was studying trigonometric series (Fourier series), e.g., series of the form

$$a_1 \sin x + a_2 \sin 2x + a_3 \sin 3x + \cdots.$$

Such series had been studied throughout the nineteenth century as solutions to differential equations representing physical problems. Cantor's work on Fourier series led him to consider more and more general sets of real numbers. In 1871 he realized that a certain operation on sets of real numbers (the operation of forming the set of limit points) could be iterated more than a finite number of times; starting with a set P_0 one could form $P_0, P_1, P_2, \ldots, P_\omega, P_{\omega+1}, \ldots, P_{\omega+\omega}, \ldots$. In December of 1873 Cantor proved that the set of all real numbers could not be put into one-to-one correspondence with the integers (Theorem 6B); this result was published in 1874. In 1879 and subsequent years, he published a series of papers setting forth the general concepts of abstract sets and "transfinite numbers."

Cantor's work was well received by some of the prominent mathematicians of his day, such as Richard Dedekind. But his willingness to regard infinite sets as objects to be treated in much the same way as finite sets was bitterly attacked by others, particularly Kronecker. There was no objection to a "potential infinity" in the form of an unending process, but an "actual infinity" in the form of a completed infinite set was harder to accept.

About the turn of the century, attempts were made to present the principles of set theory as being principles of logic—as self-evident truths of deductive thought. The foremost work in this direction was done by Gottlob Frege. Frege was a German mathematician by training, who contributed to both mathematics and philosophy. In 1893 and 1903 he published a two-volume work in which he indicated how mathematics could be developed from principles that he regarded as being principles of logic. But just as the second volume was about to appear, Bertrand Russell informed Frege of a contradiction derivable from the principles (Russell's paradox).

Russell's paradox had a tremendous impact on the ideas of that time regarding the foundations of mathematics. It was not, to be sure, the first paradox to be noted in set theory. Cantor himself had observed that some collections, such as the collection of all sets, had to be regarded as "inconsistent totalities," in contrast to the more tractable "consistent totalities," such as the set of numbers. In this he foreshadowed the distinction between proper classes and sets, which was introduced by John von Neumann in 1925. Also in 1897, Cesare Burali-Forti had observed a paradoxical predicament in Cantor's theory of transfinite ordinal numbers. But the simplicity and the directness of Russell's paradox seemed to destroy utterly the attempt to base mathematics on the sort of set theory that Frege had proposed.

The first axiomatization of set theory was published by Ernst Zermelo in 1908. His axioms were essentially the ones we state in Chapters 2 and 4. (His *Aussonderung* axioms were later made more precise by Thoralf Skolem and others.) It was observed by several people that for a satisfactory theory of ordinal numbers, Zermelo's axioms required strengthening. The axiom of replacement (Chapter 7) was proposed by Abraham Fraenkel (in 1922) and others, giving rise to the list of axioms now known as the "Zermelo–Fraenkel" (ZF) axioms. The axiom of regularity or foundation (Chapter 7) was at least implicit in a 1917 paper by Dmitry Mirimanoff and was explicitly included by von Neumann in 1925.

An axiomatization admitting proper classes as legitimate objects was formulated by von Neumann in the 1925 paper just mentioned. Some of his ideas were utilized by Paul Bernays in the development of a more satisfactory axiomatization, which appeared in a series of papers published in 1937 and later years. A modification of Bernays's axioms was used by Kurt Gödel in a 1940 monograph. This approach is now known as "von Neumann–Bernays" (VNB) or "Gödel–Bernays" (GB) set theory.

The use of the symbol $\in$ (a stylized form of the Greek epsilon) to denote membership was initiated by the Italian mathematician Giuseppe Peano in 1889. It abbreviates the Greek word ἐστὶ, which means "is." The underlying rationale is illustrated by the fact that if B is the set of all blue objects, then we write "$x \in B$" in order to assert that x *is* blue.

Present-day research in set theory falls for the most part into two branches. One branch involves investigating the consequences of new and stronger axioms. The other branch involves the "metamathematics" of set theory, which is the study not of sets but of the workings of set *theory* itself: its proofs, its theorems, and its nontheorems.

CHAPTER 2

AXIOMS AND OPERATIONS

In this chapter we begin by introducing the first six of our ten axioms. Initially the axiomatization might appear to be like cumbersome machinery to accomplish simple tasks. But we trust that it will eventually prove itself to be powerful machinery for difficult tasks. Of course the axioms are not chosen at random, but must ultimately reflect our informal ideas about what sets are.

In addition to the introduction of basic concepts, this chapter provides practice in using the symbolic notation introduced in Chapter 1. Finally the chapter turns to the standard results on the algebra of sets.

AXIOMS

The first of our axioms is the principle of extensionality.

Extensionality Axiom If two sets have exactly the same members, then they are equal:

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

Call this set A. And let B be the set of all solutions to the polynomial equation

$$x^4 - 17x^3 + 101x^2 - 247x + 210 = 0.$$

Now it turns out (as the industrious reader can verify) that the set B has exactly the same four members, 2, 3, 5, and 7. For this reason A and B are the same set, i.e., $A = B$. It matters not that A and B were defined in different ways. Because they have exactly the same elements, they are equal; that is, they are one and the same set. We can formulate the general principle:

Principle of Extensionality If two sets have exactly the same members, then they are equal.

Here and elsewhere, we can state things more concisely and less ambiguously by utilizing a modest amount of symbolic notation. Also we abbreviate the phrase "if and only if" as "iff." Thus we have the restatement:

Principle of Extensionality If A and B are sets such that for every object t,

$$t \in A \quad \text{iff} \quad t \in B,$$

then $A = B$.

For example, the set of primes less than 10 is the same as the set of solutions to the equation $x^4 - 17x^3 + 101x^2 - 247x + 210 = 0$. And the set $\{2\}$ whose only member is the number 2 is the same as the set of even primes.

Incidentally, we write "$A = B$" to mean that A and B are the same object. That is, the expression "A" on the left of the equality symbol names the same object as does the expression "B" on the right. If $A = B$, then automatically (i.e., by logic) anything that is true of the object A is also true of the object B (it being the same object). For example, if $A = B$, then it is automatically true that for any object t, $t \in A$ iff $t \in B$. (This is the converse to the principle of extensionality.) As usual, we write "$A \neq B$" to mean that it is not true that $A = B$.

A small set would be a set $\{0\}$ having only one member, the number 0. An even smaller set is the empty set $\varnothing$. The set $\varnothing$ has no members at all. Furthermore it is the only set with no members, since extensionality tells us that any two such sets must coincide. It might be thought at first that the empty set would be a rather useless or even frivolous set to mention, but, in fact, from the empty set by various set-theoretic operations a surprising array of sets will be constructed.

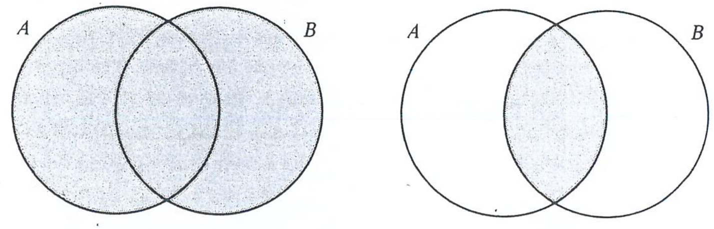
For any objects x and y, we can form the pair set $\{x, y\}$ having just the members x and y. Observe that $\{x, y\} = \{y, x\}$, as both sets have exactly the same members. As a special case we have (when $x = y$) the set $\{x, x\} = \{x\}$.

For example, we can form the set $\{\varnothing\}$ whose only member is $\varnothing$. Note that $\{\varnothing\} \neq \varnothing$, because $\varnothing \in \{\varnothing\}$ but $\varnothing \notin \varnothing$. The fact that $\{\varnothing\} \neq \varnothing$ is reflected in the fact that a man with an empty container is better off than a man with nothing—at least he has the container. Also we can form $\{\{\varnothing\}\}$, $\{\{\{\varnothing\}\}\}$, and so forth, all of which are distinct (Exercise 2).

Similarly for any objects x, y, and z we can form the set $\{x, y, z\}$. More generally, we have the set $\{x_1, \ldots, x_n\}$ whose members are exactly the objects $x_1, \ldots, x_n$. For example,

$$\{\varnothing, \{\varnothing\}, \{\{\varnothing\}\}\}$$

is a three-element set.



(a) (b)

Fig. 1. The shaded areas represent (a) $A \cup B$ and (b) $A \cap B$.

Two other familiar operations on sets are union and intersection. The *union* of sets A and B is the set $A \cup B$ of all things that are members of A or B (or both). Similarly the *intersection* of A and B is the set $A \cap B$ of all things that are members of both A and B. For example,

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

and

$$\{2, 3, 5, 7\} \cap \{1, 2, 3, 4\} = \{2, 3\}.$$

Figure 1 gives the usual pictures illustrating these operations. Sets A and B are said to be *disjoint* when they have no common members, i.e., when $A \cap B = \varnothing$.

A set A is said to be a *subset* of a set B (written $A \subseteq B$) iff all the members of A are also members of B. Note that any set is a subset of itself. At the other extreme, $\varnothing$ is a subset of every set. This fact (that $\varnothing \subseteq A$ for any A) is "vacuously true," since the task of verifying, for every member of $\varnothing$, that it also belongs to A, requires doing nothing at all.

If $A \subseteq B$, then we also say that A is *included* in B or that B *includes* A. The inclusion relation ($\subseteq$) is not to be confused with the membership