CHAPTER 4

NATURAL NUMBERS

There are, in general, two ways of introducing new objects for mathematical study: the axiomatic approach and the constructive approach. The axiomatic approach is the one we have used for sets. The concept of set is one of our primitive notions, and we have adopted a list of axioms dealing with the primitive notions.

Now consider the matter of introducing the natural numbers[1]

$$0, \quad 1, \quad 2, \ldots$$

for further study. An axiomatic approach would consider "natural number" as a primitive notion and would adopt a list of axioms. Instead we will use the constructive approach for natural numbers. We will define natural numbers in terms of other available objects (sets, of course). In place of axioms for numbers we will be able to prove the necessary properties of numbers from known properties of sets.

[1] There is a curious point of terminology here. Is 0 a natural number? With surprising consistency, the present usage is for school books (through high-school level) to exclude 0 from the natural numbers, and for upper-division college-level books to include 0. Freshman and sophomore college books are in the transition zone. In this book we include 0 among the natural numbers.

Constructing the natural numbers in terms of sets is part of the process of "embedding mathematics in set theory." The process will be continued in Chapter 5 to obtain more exotic numbers, such as $\sqrt{2}$.

INDUCTIVE SETS

First we need to define natural numbers as suitable sets. Now numbers do not at first glance appear to be sets. Not that it is an easy matter to say what numbers *do* appear to be. They are abstract concepts, which are slippery things to handle. (See, for example, the section on "Two" in Chapter 5.) Nevertheless, we can construct specific sets that will serve perfectly well as numbers. In fact this can be done in a variety of ways. In 1908, Zermelo proposed to use

$$\varnothing, \quad \{\varnothing\}, \quad \{\{\varnothing\}\}, \ldots$$

as the natural numbers. Later von Neumann proposed an alternative, which has several advantages and has become standard. The guiding principle behind von Neumann's construction is to make each natural number be the set of all smaller natural numbers. Thus we define the first four natural numbers as follows:

$$\begin{aligned} 0 &= \varnothing, \\ 1 &= \{\varnothing\} = \{\varnothing\}, \\ 2 &= \{0, 1\} = \{\varnothing, \{\varnothing\}\}, \\ 3 &= \{0, 1, 2\} = \{\varnothing, \{\varnothing\}, \{\varnothing, \{\varnothing\}\}\}. \end{aligned}$$

We could continue in this way to define 17 or any chosen natural number. Notice, for example, that the set 3 has three members. It has been selected from the class of all three-member sets to represent the size of the sets in that class.

This construction of the numbers as sets involves some extraneous properties that we did not originally expect. For example,

$$0 \in 1 \in 2 \in 3 \in \cdots$$

and

$$0 \subseteq 1 \subseteq 2 \subseteq 3 \subseteq \cdots.$$

But these properties can be regarded as accidental side effects of the definition. They do no harm, and actually will be convenient at times.

Although we have defined the first four natural numbers, we do not yet have a definition of what it means in general for something to be a natural number. That is, we have not defined the set of all natural numbers. Such a definition cannot rely on linguistic devices such as three dots or phrases like "and so forth." First we define some preliminary concepts.

Definition For any set a, its *successor* a^+ is defined by

$$a^+ = a \cup \{a\}.$$

A set A is said to be *inductive* iff $\varnothing \in A$ and it is "closed under successor," i.e.,

$$(\forall a \in A)\, a^+ \in A.$$

In terms of the successor operation, the first few natural numbers can be characterized as

$$0 = \varnothing, \quad 1 = \varnothing^+, \quad 2 = \varnothing^{++}, \quad 3 = \varnothing^{+++}, \ldots.$$

These are all distinct, e.g., $\varnothing^+ \neq \varnothing^{+++}$ (Exercise 1). And although we have not yet given a formal definition of "infinite," we can see informally that any inductive set will be infinite.

We have as yet no axioms that provide for the existence of infinite sets. There are indeed infinitely many distinct sets whose existence we could establish. But there is no one set having infinitely many members that we can prove to exist. Consequently we cannot yet prove that any inductive set exists. We now correct that fault.

Infinity Axiom There exists an inductive set:

$$(\exists A)[\varnothing \in A \;\&\; (\forall a \in A)\, a^+ \in A].$$

Armed with this axiom, we can now define the concept of natural number.

Definition A *natural number* is a set that belongs to every inductive set.

We next prove that the collection of all natural numbers constitutes a set.

Theorem 4A There is a set whose members are exactly the natural numbers.

Proof Let A be an inductive set; by the infinity axiom it is possible to find such a set. By a subset axiom there is a set w such that for any x,

$$\begin{aligned} x \in w \quad &\Leftrightarrow \quad x \in A \;\&\; x \text{ belongs to every other inductive set} \\ &\Leftrightarrow \quad x \text{ belongs to every inductive set.} \end{aligned}$$

(This proof is essentially the same as the proof of Theorem 2B.) ⊣

The set of all natural numbers is denoted by a lowercase Greek omega:

$$\begin{aligned} x \in \omega \quad &\Leftrightarrow \quad x \text{ is a natural number} \\ &\Leftrightarrow \quad x \text{ belongs to every inductive set.} \end{aligned}$$

In terms of classes, we have

$$\omega = \bigcap \{A \mid A \text{ is inductive}\},$$

but the class of all inductive sets is not a set.

Theorem 4B ω is inductive, and is a subset of every other inductive set.

Proof First of all, $\varnothing \in \omega$ because $\varnothing$ belongs to every inductive set. And second,

$$\begin{aligned} a \in \omega \quad &\Rightarrow \quad a \text{ belongs to every inductive set} \\ &\Rightarrow \quad a^+ \text{ belongs to every inductive set} \\ &\Rightarrow \quad a^+ \in \omega. \end{aligned}$$

Hence ω is inductive. And clearly ω is included in every other inductive set. ⊣

Since ω is inductive, we know that $0\ (=\varnothing)$ is in ω. It then follows that $1\ (=0^+)$ is in ω, as are $2\ (=1^+)$ and $3\ (=2^+)$. Thus 0, 1, 2, and 3 are natural numbers. Unnecessary extraneous objects have been excluded from ω, since ω is the *smallest* inductive set. This fact can also be restated as follows.

Induction Principle for ω Any inductive subset of ω coincides with ω.

Suppose, for example, that we want to prove that for every natural number n, the statement __ n __ holds. We form the set

$$T = \{n \in \omega \mid \_\_\ n\ \_\_\}$$

of natural numbers for which the desired conclusion is true. If we can show that T is inductive, then the proof is complete. Such a proof is said to be a proof *by induction*. The next theorem gives a very simple example of this method.

Theorem 4C Every natural number except 0 is the successor of some natural number.

Proof Let $T = \{n \in \omega \mid \text{either } n = 0 \text{ or } (\exists p \in \omega)\ n = p^+\}$. Then $0 \in T$. And if $k \in T$, then $k^+ \in T$. Hence by induction, $T = \omega$. ⊣

CHAPTER 6

CARDINAL NUMBERS AND THE AXIOM OF CHOICE

EQUINUMEROSITY

We want to discuss the *size* of sets. Given two sets A and B, we want to consider such questions as:

(a) Do A and B have the same size?
(b) Does A have more elements than B?

Now for finite sets, this is not very complicated. We can just count the elements in the two sets, and compare the resulting numbers. And if one of the sets is finite and the other is infinite, it seems conservative enough to say that the infinite set definitely has more elements than does the finite set.

But now consider the case of two infinite sets. Our first need is for a definition: What exactly should "A has the same size as B" mean when A and B are infinite? After we select a reasonable definition, we can then ask, for example, whether any two infinite sets have the same size. (We have not yet officially defined "finite" or "infinite," but we will soon be in a position to define these terms in a precise way.)

An Analogy In order to find a solution to the above problem, we can first consider an analogous problem, but one on a very simple level.

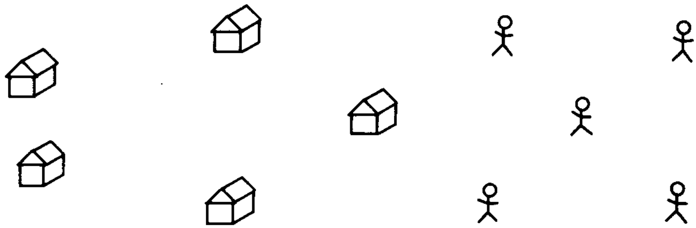

Fig. 30. Are there exactly as many houses as people?

Imagine that your mathematical education is just beginning—that you are on your way to nursery school. You are apprehensive about going, because you have heard that they have mathematics lessons and you cannot count past three. Sure enough, on the very first day they show you Fig. 30 and ask, "Are there exactly as many houses as people?" Your heart sinks. There are too many houses and too many people for you to count. This is just the predicament described earlier, where we had sets A and B that, being infinite, had too many elements to count.

But wait! All is not lost. You take your crayon and start pairing people with houses (Fig. 31). You soon discover that there are indeed exactly as many houses as people. And you did not have to count past three. You get a gold star and go home happy. We adopt the same solution.

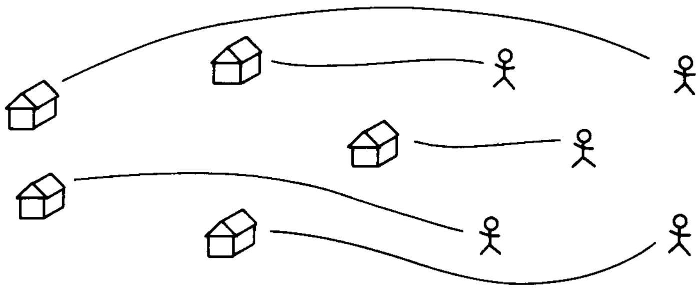

Fig. 31. How to answer the question without counting.

Definition A set A is *equinumerous* to a set B (written $A \approx B$) iff there is a one-to-one function from A onto B.

A one-to-one function from A onto B is called a *one-to-one correspondence* between A and B. For example, in Fig. 30 the set of houses is equinumerous to the set of people. A one-to-one correspondence between the sets is exhibited in Fig. 31.

Example The set $\omega \times \omega$ is equinumerous to ω. There is a function J mapping $\omega \times \omega$ one-to-one onto ω, shown in Fig. 32, where the value of

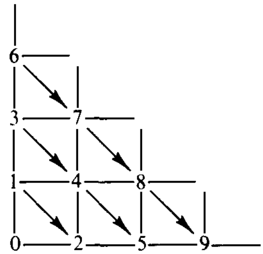

Fig. 32. $\omega \times \omega \approx \omega$.

$J(m, n)$ is written at the point with coordinates $\langle m, n \rangle$. In fact we can give a polynomial expression for J:

$$J(m, n) = \tfrac{1}{2}[(m + n)^2 + 3m + n],$$

as you are asked to verify in Exercise 2.

Example The set of natural numbers is equinumerous to the set $\mathbb{Q}$ of rational numbers, i.e., $\omega \approx \mathbb{Q}$. The method to be used here is like the one used in the preceding example. We arrange $\mathbb{Q}$ in an orderly pattern, then thread a path through the pattern, pairing natural numbers with the rationals as we go. The pattern is shown in Fig. 33. We define $f: \omega \to \mathbb{Q}$, where $f(n)$ is the rational next to the bracketed numeral for n in Fig. 33. To ensure that f is one-to-one, we skip rationals met for the second (or third or later) time. Thus $f(4) = -1$, and we skip $-2/2$, $-3/3$, and so forth.

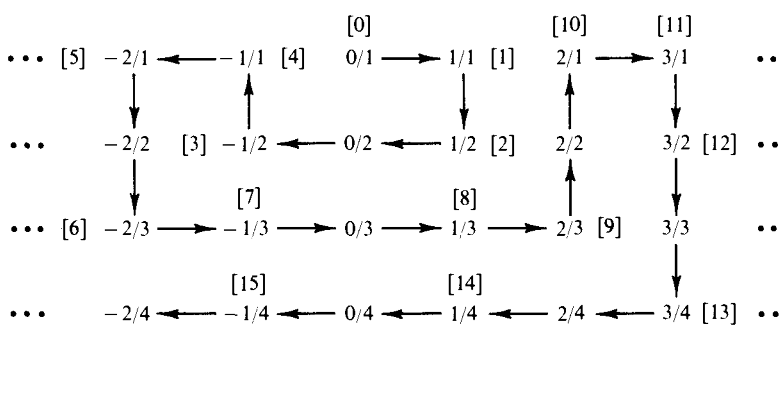

Fig. 33. $\omega \approx \mathbb{Q}$.

Example The open unit interval

$$(0, 1) = \{x \in \mathbb{R} \mid 0 < x < 1\}$$

is equinumerous to the set $\mathbb{R}$ of all real numbers. A geometric construction of a one-to-one correspondence is shown in Fig. 34. Here (0, 1) has been bent into a semicircle with center P. Each point in (0, 1) is paired with its projection (from P) on the real line.

There is also an analytical proof that $(0, 1) \approx \mathbb{R}$. Let $f(x) = \tan \pi(2x - 1)/2$. Then f maps (0, 1) one-to-one (and continuously) onto $\mathbb{R}$.

As the above example shows, it is quite possible for an infinite set, such as $\mathbb{R}$, to be equinumerous to a proper subset of itself, such as (0, 1). (For finite sets this never happens, as we will prove shortly.) Galileo remarked in 1638 that ω was equinumerous to the set $\{0, 1, 4, 9, \ldots\}$ of squares of natural numbers, and found this to be a curious fact. The

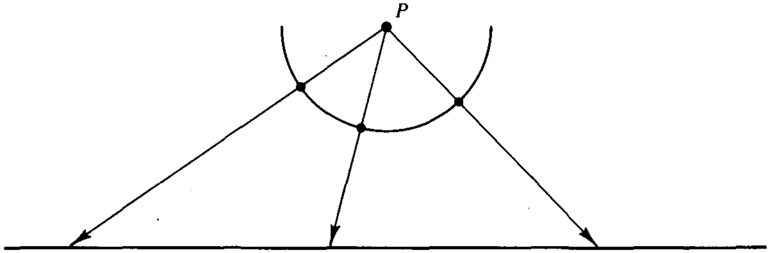

Fig. 34. $(0, 1) \approx \mathbb{R}$.

squares are in some sense a small part of the natural numbers, e.g., the fraction of the natural numbers less than n that are squares converges to 0 as n tends to infinity. But when viewed simply as two abstract sets, the set of natural numbers and the set of squares have the same size. Similarly the set of even integers is equinumerous to the set of all integers. If we focus attention on the way in which even integers are placed among the others, then we are tempted to say that there are only half as many even integers as there are integers altogether. But if we instead view the two sets as two different abstract sets, then they have the same size.

Example For any set A, we have $\mathscr{P}A \approx {}^{A}2$. To prove this, we define a one-to-one function H from $\mathscr{P}A$ onto ${}^{A}2$ as follows: For any subset B of A, $H(B)$ is the characteristic function of B, i.e., the function f_B from A into 2 for which

$$f_B(x) = \begin{cases} 1 & \text{if } x \in B, \\ 0 & \text{if } x \in A - B. \end{cases}$$

Then any function $g \in {}^{A}2$ is in ran H, since

$$g = H(\{x \in A \mid g(x) = 1\}).$$

The following theorem shows that equinumerosity has the property of being reflexive (on the class of all sets), symmetric, and transitive. But it cannot be represented by an equivalence relation, because it concerns *all* sets. In von Neumann–Bernays set theory, one can form the class

$$\mathsf{E} = \{\langle A, B\rangle \mid A \approx B\}.$$

Then E is an "equivalence relation on V," in the sense that it is a class of ordered pairs that is reflexive on V, symmetric, and transitive. But E is not a set, lest its field V be a set. In Zermelo–Fraenkel set theory, we have only the "equivalence concept" of equinumerosity.

Theorem 6A For any sets A, B, and C:

(a) $A \approx A$.
(b) If $A \approx B$, then $B \approx A$.
(c) If $A \approx B$ and $B \approx C$, then $A \approx C$.

Proof See Exercise 5. ⊣

In light of the examples presented up to now, you might well ask whether any two infinite sets are equinumerous. Such is not the case; some infinite sets are much larger than others.

Theorem 6B (Cantor 1873) (a) The set ω is not equinumerous to the set $\mathbb{R}$ of real numbers.

(b) No set is equinumerous to its power set.

Proof We will show that for any function $f\colon \omega \to \mathbb{R}$, there is a real number z not belonging to $\operatorname{ran} f$. Imagine a list of the successive values of f, expressed as infinite decimals:

$$\begin{aligned} f(0) &= 236.001\ldots, \\ f(1) &= -7.777\ldots, \\ f(2) &= 3.1415\ldots, \\ &\ \ \vdots \end{aligned}$$

(In Chapter 5 we did not go into the matter of decimal expansions, but you are surely familiar with them.) We will proceed to construct the real z. The integer part is 0, and the $(n+1)$st decimal place of z is 7 unless the $(n+1)$st decimal place of $f(n)$ is 7, in which case the $(n+1)$st decimal place of z is 6. For example, in the case shown,

$$z = 0.767\ldots.$$

Then z is a real number not in $\operatorname{ran} f$, as it differs from $f(n)$ in the $(n+1)$st decimal place.

The proof of (b) is similar. Let $g\colon A \to \mathscr{P}A$; we will construct a subset B of A that is not in $\operatorname{ran} g$. Specifically, let

$$B = \{x \in A \mid x \notin g(x)\}.$$

Then $B \subseteq A$, but for each $x \in A$,

$$x \in B \quad \Leftrightarrow \quad x \notin g(x).$$

Hence $B \neq g(x)$. ⊣

The set $\mathbb{R}$ happens to be equinumerous to $\mathscr{P}\omega$, as we will soon be able to prove. A larger set is $\mathscr{P}\mathbb{R}$, and $\mathscr{P}\mathscr{P}\mathbb{R}$ is larger still.

Before continuing our consideration of infinite sets, we will study the other alternative: the sets that are "small" at least to the extent of being finite.

Exercises

1. Show that the equation

$$f(m, n) = 2^m(2n + 1) - 1$$

defines a one-to-one correspondence between $\omega \times \omega$ and ω.

2. Show that in Fig. 32 we have:

$$\begin{aligned} J(m, n) &= [1 + 2 + \cdots + (m + n)] + m \\ &= \tfrac{1}{2}[(m + n)^2 + 3m + n]. \end{aligned}$$

3. Find a one-to-one correspondence between the open unit interval (0, 1) and $\mathbb{R}$ that takes rationals to rationals and irrationals to irrationals.

4. Construct a one-to-one correspondence between the closed unit interval

$$[0, 1] = \{x \in \mathbb{R} \mid 0 \leq x \leq 1\}$$

and the open unit interval (0, 1).

5. Prove Theorem 6A.

FINITE SETS

Although we have long been using the words "finite" and "infinite" in an informal way, we have not yet given them precise definitions. Now is the time.

Definition A set is *finite* iff it is equinumerous to some natural number. Otherwise it is *infinite*.

Here we rely on the fact that in our construction of ω, each natural number is the set of all smaller natural numbers. For example, any natural number is itself a finite set.

We want to check that each finite set S is equinumerous to a *unique* number n. The number n can then be used as a count of the elements in S.

We first need the following theorem, which implies that if n objects are placed into fewer than n pigeonholes, then some pigeonhole receives more than one object. Recall that a set A is a *proper subset* of B iff $A \subseteq B$ and $A \neq B$.

Pigeonhole Principle No natural number is equinumerous to a proper subset of itself.

Proof Assume that f is a one-to-one function from the set n into the set n. We will show that ran f is all of the set n (and not a proper subset of n). This suffices to prove the theorem.

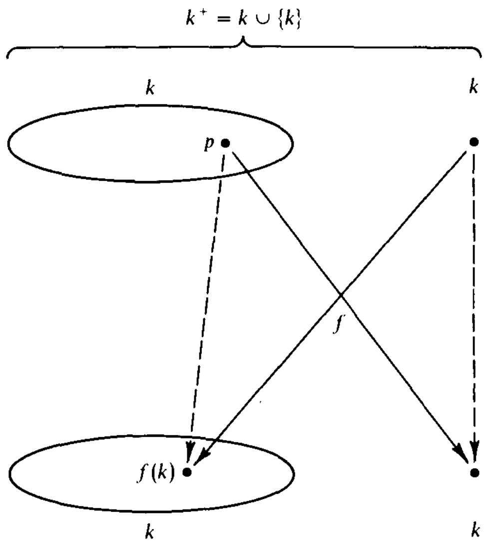

Fig. 35. In f we interchange two values to obtain $\hat{f}$.

We use induction on n. Define:

$$T = \{n \in \omega \mid \text{any one-to-one function from } n \text{ into } n \text{ has range } n\}.$$

Then $0 \in T$; the only function from the set 0 into the set 0 is $\varnothing$ and its range is the set 0. Suppose that $k \in T$ and that f is a one-to-one function from the set k^+ into the set k^+. We must show that the range of f is all of the set k^+; this will imply that $k^+ \in T$. Note that the restriction $f \upharpoonright k$ of f to the set k maps the set k one-to-one into the set k^+.

Case I Possibly the set k is closed under f. Then $f \upharpoonright k$ maps the set k into the set k. Then because $k \in T$ we may conclude that ran $(f \upharpoonright k)$ is all of the set k. Since f is one-to-one, the only possible value for $f(k)$ is the number k. Hence ran f is $k \cup \{k\}$, which is the set k^+.

Case II Otherwise $f(p) = k$ for some number p less than k. In this case we interchange two values of the function. Define $\hat{f}$ by

$$\begin{aligned} \hat{f}(p) &= f(k), \\ \hat{f}(k) &= f(p) = k, \\ \hat{f}(x) &= f(x) \qquad \text{for other } x \in k^+ \end{aligned}$$

(see Fig. 35). Then $\hat{f}$ maps the set k^+ one-to-one into the set k^+, and the set k is closed under $\hat{f}$. So by Case I, ran $\hat{f} = k^+$. But ran $\hat{f}$ = ran f.

Thus in either case, ran $f = k^+$. So T is inductive and equals ω. ⊣

Corollary 6C No finite set is equinumerous to a proper subset of itself.

Proof This is the same as the pigeonhole principle, but for an arbitrary finite set A instead of a natural number. Since A is equinumerous to a natural number n, we can use the one-to-one correspondence g between A and n to "transfer" the pigeonhole principle to the set A.

Suppose that, contrary to our hopes, there is a one-to-one correspondence f between A and some proper subset of A. Consider the composition $g \circ f \circ g^{-1}$, illustrated in Fig. 36. This composition maps n into n, and it is one-to-one by Exercise 17 of Chapter 3. Furthermore its range C is a proper subset of n. (Consider any a in A − ran f; then $g(a) \in n - C$.) Thus n is equinumerous to C, in contradiction to the pigeonhole principle. ⊣

The foregoing proof uses an argument that is useful elsewhere as well. We have sets A and n that are "alike" in that $A \approx n$, but different in that they have different members. Think of the members of A as being red, the members of n as being blue. Then the function $g\colon A \to n$ paints red

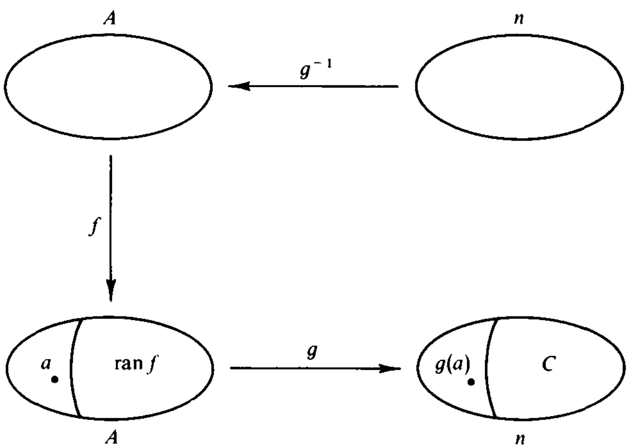

Fig. 36. What f does to A, $g \circ f \circ g^{-1}$ does to n.

things blue, and the function $g^{-1}: n \to A$ paints blue things red. The composition $g \circ f \circ g^{-1}$ paints things red, applies f, and then restores the blue color.

Corollary 6D (a) Any set equinumerous to a proper subset of itself is infinite.

(b) The set ω is infinite.

Proof The preceding corollary proves part (a). Part (b) follows at once from part (a), since the function σ whose value at each number n is n^+ maps ω one-to-one onto $\omega - \{0\}$. ⊣

Corollary 6E Any finite set is equinumerous to a *unique* natural number.

Proof Assume that $A \approx m$ and $A \approx n$ for natural numbers m and n. Then $m \approx n$. By trichotomy and Corollary 4M, either $m = n$ or one is a proper subset of the other. But the latter alternative is impossible since $m \approx n$. Hence $m = n$. ⊣

For a finite set A, the unique $n \in \omega$ such that $A \approx n$ is called the *cardinal number* of A, abbreviated card A. For example, card $n = n$ for $n \in \omega$. And if a, b, c, and d are all distinct objects, then card$\{a, b, c, d\} = 4$. This is because $\{a, b, c, d\} \approx 4$; selecting a one-to-one correspondence is the process called "counting." Observe that for any finite sets A and B, we have $A \approx$ card A and

$$\text{card } A = \text{card } B \quad \text{iff} \quad A \approx B.$$

What about infinite sets? The number card A measures the size of a finite set A. We want "numbers" that similarly measure the size of infinite sets. Just what sets these "numbers" are is not too crucial, any more than it was crucial just what set the number 2 was. The essential demand is that we define card A for arbitrary A in such a way that

$$\text{card } A = \text{card } B \quad \text{iff} \quad A \approx B.$$

Now it turns out that there is no way of defining card A that is really simple. We therefore postpone until Chapter 7 the actual definition of the set card A. The information we need for the present chapter is embodied in the following promise.

Promise For any set A we will define a set card A in such a way that:

(a) For any sets A and B,

$$\text{card } A = \text{card } B \quad \text{iff} \quad A \approx B.$$

(b) For a finite set A, card A is the natural number n for which $A \approx n$.

(In making good on this promise, we will use in Chapter 7 additional axioms, namely the replacement axioms and the axiom of choice. If you plan to omit Chapter 7, then regard card A as an additional primitive notion and the promise as an additional axiom.)

We define a cardinal number to be something that is card A for some set A. By part (b) of the promise, any natural number n is also a cardinal number, since $n = \text{card } n$. But card ω is not a natural number (card $\omega \neq n = \text{card } n$, since ω is not equinumerous to n). Just what set card ω is will not be revealed until Chapter 7. Meanwhile we give it the name that Cantor gave it:

$$\text{card } \omega = \aleph_0 .$$

The symbol $\aleph$ is aleph, the first letter of the Hebrew alphabet.

In general, for a cardinal number κ, there will be a great many sets A of cardinality κ, i.e., sets with card $A = \kappa$. (The one exception to this occurs when $\kappa = 0$.) In fact, for any nonzero cardinal κ, the class

$$\mathsf{K}_\kappa = \{X \mid \text{card } X = \kappa\}$$

of sets of cardinality κ is too large to be a set (Exercise 6). But all of the sets of cardinality κ look, from a great distance, very much alike—the elements of two such sets may differ but the number of elements is always κ. In particular, if one set X of cardinality κ is finite, then all of them are; in this case κ is a *finite cardinal*. And if not, then κ is an *infinite cardinal*. Thus the finite cardinals are exactly the natural numbers. $\aleph_0$ is an infinite cardinal, as are card $\mathbb{R}$, card $\mathscr{P}\omega$, card $\mathscr{P}\mathscr{P}\omega$, etc.

Before leaving this section on finite sets, we will verify a fact that, on an informal level, appears inevitable: Any subset of a finite set is finite. (Later we will find another proof of this.)

Lemma 6F If C is a proper subset of a natural number n, then $C \approx m$ for some m less than n.

Proof We use induction. Let

$$T = \{n \in \omega \mid \text{any proper subset of } n \text{ is equinumerous to a member of } n\}.$$

Then $0 \in T$ vacuously, 0 having no proper subsets. Assume that $k \in T$ and consider a proper subset C of k^+.

Case I $C = k$. Then $C \approx k \in k^+$.

Case II C is a proper subset of k. Then since $k \in T$, we have $C \approx m$ for $m \in k \in k^+$.

Case III Otherwise $k \in C$. Then $C = (C \cap k) \cup \{k\}$ and $C \cap k$ is a proper subset of k. Because $k \in T$, there is $m \in k$ with $C \cap k \approx m$. Say f is a one-to-one correspondence between $C \cap k$ and m; then $f \cup \{\langle k, m \rangle\}$ is a one-to-one correspondence between C and m^+. Since $m \in k$, we have $m^+ \in k^+$. Hence $C \approx m^+ \in k^+$ and $k^+ \in T$.

Thus T is inductive and coincides with ω. ⊣

Corollary 6G Any subset of a finite set is finite.

Proof Consider $A \subseteq B$ and let f be a one-to-one correspondence between B and some n in ω. Then $A \approx f[\![A]\!] \subseteq n$ and by the lemma $f[\![A]\!] \approx m$ for some $m \underline{\in} n$. Hence $A \approx m \underline{\in} n \in \omega$. ⊣

Exercises

6. Let κ be a nonzero cardinal number. Show that there does not exist a set to which every set of cardinality κ belongs.

7. Assume that A is finite and $f: A \to A$. Show that f is one-to-one iff $\operatorname{ran} f = A$.

8. Prove that the union of two finite sets is finite (Corollary 6K), without any use of arithmetic.

9. Prove that the Cartesian product of two finite sets is finite (Corollary 6K), without any use of arithmetic.

ORDERING CARDINAL NUMBERS

We can use the concept of equinumerosity to tell us when two sets A and B are of the same size. But when should we say that B is larger than A?

Definition A set A is *dominated* by a set B (written $A \preccurlyeq B$) iff there is a one-to-one function from A *into* B.

For example, any set dominates itself. If $A \subseteq B$, then A is dominated by B, since the identity function on A maps A one-to-one into B. More generally we have: $A \preccurlyeq B$ iff A is equinumerous to some subset of B. This is just a restatement of the definition, since f is a function from A *into* B iff it is a function from A *onto* a subset of B (see Fig. 38).

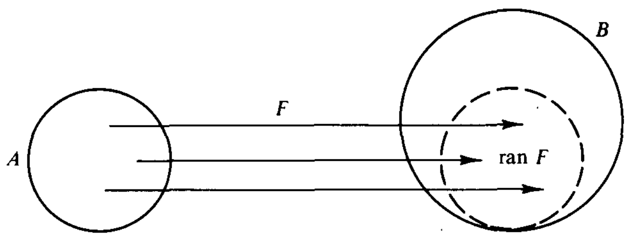

Fig. 38. F shows that $A \preccurlyeq B$.

We define the ordering of cardinal numbers by utilizing the concept of dominance:

$$\text{card } A \leq \text{card } B \quad \text{iff} \quad A \preccurlyeq B.$$

As with the operations of cardinal arithmetic, it is necessary to check that the ordering relation is well defined. For suppose we start with two cardinal numbers, say κ and λ. In order to determine whether or not $\kappa \leq \lambda$, our definition demands that we employ selected representatives K and L for which $\kappa = \text{card } K$ and $\lambda = \text{card } L$. Then

$$\kappa \leq \lambda \quad \text{iff} \quad K \preccurlyeq L.$$

But the truth or falsity of "$\kappa \leq \lambda$" must be independent of *which* selected representatives are chosen. Suppose also that $\kappa = \text{card } K'$ and $\lambda = \text{card } L'$. To avoid embarrassment, we must be certain that

$$K \preccurlyeq L \quad \text{iff} \quad K' \preccurlyeq L'.$$

To prove this, note first that $K \approx K'$ and $L \approx L'$ (because card K = card K' and card L = card L'). If $K \preccurlyeq L$, then we have one-to-one maps (i) from K' onto K, (ii) from K into L, and (iii) from L onto L'. By composing the three functions, we can map K' one-to-one into L', and hence $K' \preccurlyeq L'$.

We further define

$$\kappa < \lambda \quad \text{iff} \quad \kappa \leq \lambda \quad \text{and} \quad \kappa \neq \lambda.$$

Thus in terms of sets we have

$$\text{card } K < \text{card } L \quad \text{iff} \quad K \preccurlyeq L \quad \text{and} \quad K \not\approx L.$$

Notice that this condition is stronger than just saying that K is equinumerous to a proper subset of L. After all, ω is equinumerous to a proper subset of itself, but we certainly do not want to have card $\omega <$ card ω. The definition of "$<$" has the expected consequence that

$$\kappa \leq \lambda \quad \text{iff} \quad \text{either} \quad \kappa < \lambda \quad \text{or} \quad \kappa = \lambda.$$

Examples 1. If $A \subseteq B$, then card $A \leq$ card B. Conversely, whenever $\kappa \leq \lambda$, then there exist sets $K \subseteq L$ with card $K = \kappa$ and card $L = \lambda$. To prove this, start with any sets C and L of cardinality κ and λ, respectively. Then $C \preccurlyeq L$, so there is a one-to-one function f from C into L. Let $K = \operatorname{ran} f$; then $C \approx K \subseteq L$.

2. For any cardinal κ, we have $0 \leq \kappa$.

3. For any finite cardinal n, we have $n < \aleph_0$. (Why?) For any two finite cardinals m and n, we have

$$m \in n \quad \Rightarrow \quad m \subseteq n \quad \Rightarrow \quad m \leq n.$$

Furthermore the converse implications hold. For if $m \leq n$, then $m \preccurlyeq n$ and there is a one-to-one function $f\colon m \to n$. By the pigeonhole principle, it is impossible to have n less than m, so by trichotomy $m \in n$. Thus our new ordering on finite cardinals agrees with the epsilon ordering of Chapter 4.

4. $\kappa < 2^\kappa$ for any cardinal κ. For if A is any set of cardinality κ, then $\mathscr{P}A$ has cardinality 2^κ. Then $A \preccurlyeq \mathscr{P}A$ (map $x \in A$ to $\{x\} \in \mathscr{P}A$), but by Cantor's theorem (Theorem 6B) $A \not\approx \mathscr{P}A$. Hence $\kappa \leq 2^\kappa$ but $\kappa \neq 2^\kappa$, i.e., $\kappa < 2^\kappa$. In particular, there is no largest cardinal number.

The first thing to prove about the ordering we have defined for cardinals is that it actually behaves like something we would be willing to call an ordering. After all, just using the symbol "$\leq$" does not confer any special properties, but it does indicate the expectation that special properties will be forthcoming. For a start, we ask if the following are valid for all cardinals κ, λ, and μ:

1. $\kappa \leq \kappa$.
2. $\kappa \leq \lambda \leq \mu \Rightarrow \kappa \leq \mu$.
3. $\kappa \leq \lambda \;\&\; \lambda \leq \kappa \Rightarrow \kappa = \lambda$.
4. Either $\kappa \leq \lambda$ or $\lambda \leq \kappa$.

The first is obvious, since $A \preccurlyeq A$ holds for any set A. The second item follows at once from the fact that

$$A \preccurlyeq B \;\&\; B \preccurlyeq C \quad \Rightarrow \quad A \preccurlyeq C.$$

(We prove this by taking the composition of two functions.) The third item is nontrivial. But the assertion is correct, and is called the *Schröder–Bernstein theorem*. We will also prove the fourth item, but that proof will require the axiom of choice.

First we will prove the Schröder–Bernstein theorem, which will be a basic tool in calculating the cardinalities of sets. Typically when we want to calculate card S for a given set S, we try to squeeze card S between upper and lower bounds. If possible, we try to get these bounds to coincide,

$$\kappa \leq \text{card } S \leq \kappa,$$

whereupon the Schröder–Bernstein theorem asserts that card $S = \kappa$. We will see examples of this technique after proving the theorem.

Schröder–Bernstein Theorem (a) If $A \preccurlyeq B$ and $B \preccurlyeq A$, then $A \approx B$.
(b) For cardinal numbers κ and λ, if $\kappa \leq \lambda$ and $\lambda \leq \kappa$, then $\kappa = \lambda$.

Proof It is done with mirrors (see Fig. 39). We are given one-to-one functions $f: A \rightarrow B$ and $g: B \rightarrow A$. Define C_n by recursion, using the formulas

$$C_0 = A - \text{ran } g \qquad \text{and} \qquad C_{n^+} = g[\![f[\![C_n]\!]]\!].$$

Thus C_0 is the troublesome part that keeps g from being a one-to-one correspondence between B and A. We bounce it back and forth, obtaining $C_1, C_2, \ldots$. The function showing that $A \approx B$ is the function $h: A \rightarrow B$ defined by

$$h(x) = \begin{cases} f(x) & \text{if } \; x \in C_n \; \text{ for some } n, \\ g^{-1}(x) & \text{otherwise.} \end{cases}$$

Note that in the second case ($x \in A$ but $x \notin C_n$ for any n) it follows that $x \notin C_0$ and hence $x \in \text{ran } g$. So $g^{-1}(x)$ makes sense in this case.

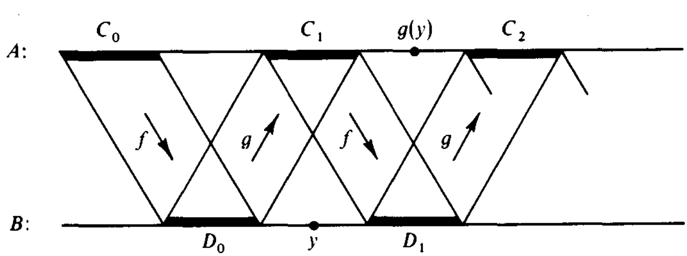

Fig. 39. The Schröder–Bernstein theorem.

Does it work? We must verify that h is one-to-one and has range B. Define $D_n = f[\![C_n]\!]$, so that $C_{n^+} = g[\![D_n]\!]$. To show that h is one-to-one, consider distinct x and x' in A. Since both f and g^{-1} are one-to-one, the only possible problem arises when, say, $x \in C_m$ and $x' \notin \bigcup_{n \in \omega} C_n$. In this case,

$$h(x) = f(x) \in D_m,$$

whereas

$$h(x') = g^{-1}(x') \notin D_m,$$

lest $x' \in C_{m^+}$. So $h(x) \neq h(x')$.

Finally we must check that ran h exhausts B. Certainly each $D_n \subseteq$ ran h, because $D_n = h[\![C_n]\!]$. Consider then a point y in $B - \bigcup_{n \in \omega} D_n$. Where is $g(y)$? Certainly $g(y) \notin C_0$. Also $g(y) \notin C_{n^+}$, because $C_{n^+} = g[\![D_n]\!]$, $y \notin D_n$, and g is one-to-one. So $g(y) \notin C_n$ for any n. Therefore $h(g(y)) = g^{-1}(g(y)) = y$. This shows that $y \in$ ran h, thereby proving part (a).

Part (b) is a restatement of part (a) in terms of cardinal numbers. ⊣

The Schröder–Bernstein theorem is sometimes called the "Cantor–Bernstein theorem." Cantor proved the theorem in his 1897 paper, but his proof utilized a principle that is equivalent to the axiom of choice. Ernst Schröder announced the theorem in an 1896 abstract. His proof, published in 1898, was imperfect, and he published a correction in 1911. The first fully satisfactory proof was given by Felix Bernstein and was published in an 1898 book by Borel.

Examples The usefulness of the Schröder–Bernstein theorem in calculating cardinalities is indicated by the following examples.

1. If $A \subseteq B \subseteq C$ and $A \approx C$, then all three sets are equinumerous. To prove this, let κ = card A = card C and let λ = card B. Then by hypothesis $\kappa \leq \lambda \leq \kappa$, so by the Schröder–Bernstein theorem $\kappa = \lambda$.

2. The set $\mathbb{R}$ of real numbers is equinumerous to the closed unit interval $[0, 1]$. For we have

$$(0, 1) \subseteq [0, 1] \subseteq \mathbb{R},$$

and (as noted previously) $\mathbb{R} \approx (0, 1)$. Thus by the preceding example, all three sets are equinumerous. (For a more direct construction of a one-to-one correspondence between $\mathbb{R}$ and $[0, 1]$, we suggest trying Exercise 4.)

3. If $\kappa \leq \lambda \leq \mu$, then, as we observed before, $\kappa \leq \mu$. We can now give an improved version:

$$\kappa \leq \lambda < \mu \quad \Rightarrow \quad \kappa < \mu,$$
$$\kappa < \lambda \leq \mu \quad \Rightarrow \quad \kappa < \mu.$$

For by the earlier observation we obtain $\kappa \leq \mu$; if equality held, then (as in the first example) all three cardinal numbers would coincide.

4. $\mathbb{R} \approx {}^{\omega}2$, and hence $\mathbb{R} \approx \mathscr{P}\omega$. Thus the set of real numbers is equinumerous to the power set of ω. To prove this it suffices, by the Schröder–Bernstein theorem, to show that $\mathbb{R} \preccurlyeq {}^{\omega}2$ and ${}^{\omega}2 \preccurlyeq \mathbb{R}$.

To show that $\mathbb{R} \preccurlyeq {}^{\omega}2$, we construct a one-to-one function from the open unit interval $(0, 1)$ into ${}^{\omega}2$. The existence of such a function, together with the fact that $\mathbb{R} \approx (0, 1)$, gives us

$$\mathbb{R} \approx (0, 1) \preccurlyeq {}^{\omega}2.$$

The function is defined by use of binary expansions of real numbers; map the real whose binary expansion is 0.1100010... to the function in ${}^{\omega}2$ whose successive values are 1, 1, 0, 0, 0, 1, 0, ... In general, for a real number z in $(0, 1)$, let $H(z)$ be the function $H(z)\colon \omega \to 2$ whose value at n equals the $(n + 1)$st bit (binary digit) in the binary expansion of z. Clearly H is one-to-one. (But H does not have all of ${}^{\omega}2$ for its range. Note that $0.1000\ldots = 0.0111\ldots = \frac{1}{2}$. For definiteness, always select the nonterminating binary expansion.)

To show that ${}^{\omega}2 \preccurlyeq \mathbb{R}$ we use decimal expansions. The function in ${}^{\omega}2$ whose successive values are 1, 1, 0, 0, 0, 1, 0, ... is mapped to the real number with decimal expansion 0.1100010.... This maps ${}^{\omega}2$ one-to-one into the closed interval $[0, \frac{1}{9}]$.

5. By virtue of the above example,

$$\text{card } \mathbb{R} = 2^{\aleph_0}.$$

Consequently the plane $\mathbb{R} \times \mathbb{R}$ has cardinality

$$2^{\aleph_0} \cdot 2^{\aleph_0} = 2^{\aleph_0 + \aleph_0} = 2^{\aleph_0}.$$

Thus the line $\mathbb{R}$ is equinumerous to the plane $\mathbb{R} \times \mathbb{R}$. This will not come as a surprise if you have heard of "space-filling" curves.

AXIOM OF CHOICE

At several points in this book we have already encountered the need for a principle asserting the possibility of selecting members from nonempty sets. We can no longer postpone a systematic discussion of such a principle. There are numerous equivalent formulations of the axiom of choice. The following theorem lists six of them. Others will be found in the exercises.

Theorem 6M The following statements are equivalent.

(1) Axiom of choice, I. For any relation R, there is a function $F \subseteq R$ with dom $F =$ dom R.

(2) Axiom of choice, II; multiplicative axiom. The Cartesian product of nonempty sets is always nonempty. That is, if H is a function with domain I and if $(\forall i \in I)\, H(i) \neq \varnothing$, then there is a function f with domain I such that $(\forall i \in I)\, f(i) \in H(i)$.

(3) Axiom of choice, III. For any set A there is a function F (a "choice function" for A) such that the domain of F is the set of nonempty subsets of A, and such that $F(B) \in B$ for every nonempty $B \subseteq A$.

(4) Axiom of choice, IV. Let $\mathscr{A}$ be a set such that (a) each member of $\mathscr{A}$ is a nonempty set, and (b) any two distinct members of $\mathscr{A}$ are disjoint. Then there exists a set C containing exactly one element from each member of $\mathscr{A}$ (i.e., for each $B \in \mathscr{A}$ the set $C \cap B$ is a singleton $\{x\}$ for some x).

(5) Cardinal comparability. For any sets C and D, either $C \preccurlyeq D$ or $D \preccurlyeq C$. For any two cardinal numbers κ and λ, either $\kappa \leq \lambda$ or $\lambda \leq \kappa$.

(6) Zorn's lemma. Let $\mathscr{A}$ be a set such that for every chain $\mathscr{B} \subseteq \mathscr{A}$, we have $\bigcup \mathscr{B} \in \mathscr{A}$. ($\mathscr{B}$ is called a *chain* iff for any C and D in $\mathscr{B}$, either $C \subseteq D$ or $D \subseteq C$.) Then $\mathscr{A}$ contains an element M (a "maximal" element) such that M is not a subset of any other set in $\mathscr{A}$.

Statements (1)–(4) are synoptic ways of saying that there exist uniform methods for selecting elements from sets. On the other hand, statements (5) and (6) appear to be rather different.

COUNTABLE SETS

The definition below applies the word "countable" to those sets whose elements can, in a sense, be counted by means of the natural numbers. A "counting" of a set can be taken to be a one-to-one correspondence between the members of the set and the natural numbers (or the natural numbers less than some number n). This requires that the set be no larger than ω.

Definition A set A is *countable* iff $A \preccurlyeq \omega$, i.e., iff card $A \leq \aleph_0$.

Since we have recently found that

$$\kappa < \aleph_0 \quad \Leftrightarrow \quad \kappa \text{ is finite,}$$

we can also formulate the definition as follows: A set A is countable iff either A is finite or A has cardinality $\aleph_0$.

For example, the set ω of natural numbers, the set $\mathbb{Z}$ of integers, and the set $\mathbb{Q}$ of rational numbers are all infinite countable sets. But the set $\mathbb{R}$ of real numbers is uncountable (Theorem 6B).

Any subset of a countable set is obviously countable. The union of two countable sets is countable, as is their Cartesian product. (The union has at most cardinality $\aleph_0 + \aleph_0$, the product at most $\aleph_0 \cdot \aleph_0$. But both of these numbers equal $\aleph_0$.) On the other hand, $\mathscr{P}A$ is uncountable for any infinite set A. (If $2^\kappa \leq \aleph_0$, then $\kappa < \aleph_0$.)

Recall (from the example preceding Theorem 6N) that a nonempty set B is countable iff there is a function from ω onto B. This fact is used in the proof of the next theorem.

Theorem 6Q A countable union of countable sets is countable. That is, if $\mathscr{A}$ is countable and if every member of $\mathscr{A}$ is a countable set, then $\bigcup\mathscr{A}$ is countable.

Proof We may suppose that $\varnothing \notin \mathscr{A}$, for otherwise we could simply remove it without affecting $\bigcup\mathscr{A}$. We may further suppose that $\mathscr{A} \neq \varnothing$, since $\bigcup\varnothing$ is certainly countable. Thus $\mathscr{A}$ is a countable (but nonempty) from $\omega \times \omega$ onto $\bigcup\mathscr{A}$. We already know of functions from ω onto $\omega \times \omega$, and the composition will map ω onto $\bigcup\mathscr{A}$, thereby showing that $\bigcup\mathscr{A}$ is countable.

Since $\mathscr{A}$ is countable but nonempty, there is a function G from ω onto $\mathscr{A}$. Informally, we may write

$$\mathscr{A} = \{G(0), G(1), \ldots\}.$$

(Here G might not be one-to-one, so there may be repetitions in this enumeration.) We are given that each set $G(m)$ is countable and nonempty.

Hence for each m there is a function from ω onto $G(m)$. We must use the axiom of choice to select such a function for each m.

Because the axiom of choice is a recent addition to our repertoire, we will describe its use here in some detail. Let $H\colon \omega \to {}^{\omega}(\bigcup \mathscr{A})$ be defined by

$$H(m) = \{g \mid g \text{ is a function from } \omega \text{ onto } G(m)\}.$$

We know that $H(m)$ is nonempty for each m. Hence there is a function F with domain ω such that for each m, $F(m)$ is a function from ω onto $G(m)$.

To conclude the proof we have only to let $f(m, n) = F(m)(n)$. Then f is a function from $\omega \times \omega$ onto $\bigcup \mathscr{A}$. ⊣

Example For any set A, define a *sequence* in A to be a function from some natural number into A. Let $\mathrm{Sq}(A)$ be the set of all sequences in A:

$$\begin{aligned}\mathrm{Sq}(A) &= \{f \mid (\exists n \in \omega)\ f \text{ maps } n \text{ into } A\}\\ &= {}^{0}A \cup {}^{1}A \cup {}^{2}A \cup \cdots.\end{aligned}$$

The *length* of a sequence is simply its domain.

In order to verify that $\mathrm{Sq}(A)$ is a legal set, note that if $f\colon n \to A$, then

$$f \subseteq n \times A \subseteq \omega \times A,$$

so that $f \in \mathscr{P}(\omega \times A)$. Hence $\mathrm{Sq}(A) \subseteq \mathscr{P}(\omega \times A)$.

We now list some observations that establish the existence of transcendental real numbers.

1. $\mathrm{Sq}(\omega)$ has cardinality $\aleph_0$. This can be proved by using primarily the fact that $\omega \times \omega \approx \omega$. Another very direct proof is the following. Consider any $f \in \mathrm{Sq}(\omega)$; say the length of f is n. Then define

$$H(f) = 2^{f(0)+1} \cdot 3^{f(1)+1} \cdot \ldots \cdot p_{n-1}^{f(n-1)+1},$$

where p_i is the $(i + 1)$st prime. (If the length of f is 0, then $H(f) = 1$.) Thus $H\colon \mathrm{Sq}(\omega) \to \omega$ and by the fundamental theorem of arithmetic (which states that prime factorizations are unique) H is one-to-one. Hence card $\mathrm{Sq}(\omega) \le \aleph_0$, and the opposite inequality is clear. (In Chapter 4 we did not actually develop the theory of prime numbers. But there are no difficulties in embedding any standard development of the subject into set theory.)

2. $\mathrm{Sq}(A)$ is countable for any countable set A. By the countability of A there is a one-to-one function g from A into ω. This function naturally induces a one-to-one map from $\mathrm{Sq}(A)$ into $\mathrm{Sq}(\omega)$. Hence card $\mathrm{Sq}(A) \le$ card $\mathrm{Sq}(\omega) = \aleph_0$. (An alternative proof writes $\mathrm{Sq}(A) = \bigcup\{{}^{n}A \mid n \in \omega\}$, a countable union of countable sets.)

We can think of this set A as an alphabet, and the elements of $\mathrm{Sq}(A)$ as being *words* on the alphabet A. In this terminology, the present example can be stated: On any countable alphabet, there are countably many words.

3. There are $\aleph_0$ algebraic numbers. (Recall that an algebraic number is a real number that is the root of some polynomial with integer coefficients. For this purpose we exclude from the polynomials the function that is identically equal to 0.) As a first step in counting the algebraic numbers, note that the set $\mathbb{Z}$ of integers has cardinality $\aleph_0 + \aleph_0 = \aleph_0$. Next we calculate the cardinality of the set P of polynomials with integer coefficients. We can assign to each polynomial (of degree n) its sequence (of length $n + 1$) of coefficients, e.g., $1 + 7x - 5x^2 + 3x^4$ is assigned the sequence of length 5 whose successive values are 1, 7, -5, 0, 3. This defines a one-to-one map from P into $\text{Sq}(\mathbb{Z})$, a countable set. Hence P is countable. Since each polynomial in P has only finitely many roots, the set of algebraic numbers is a countable union of finite sets. Hence it is countable, by Theorem 6Q. Since the set of algebraic numbers is certainly infinite, it has cardinality $\aleph_0$.

4. There are uncountably many transcendental numbers. (Recall that a transcendental number is defined to be a real number that is not algebraic.) Since the set of algebraic numbers is countable, the set of transcendental numbers cannot also be countable lest the set $\mathbb{R}$ be countable. (Soon we will be able to show that the set of transcendental numbers has cardinality $2^{\aleph_0}$.)

Exercises

26. Prove the following generalization of Theorem 6Q: If every member of a set $\mathscr{A}$ has cardinality κ or less, then

$$\text{card} \bigcup \mathscr{A} \leq (\text{card}\, \mathscr{A}) \cdot \kappa.$$

27. (a) Let A be a collection of circular disks in the plane, no two of which intersect. Show that A is countable.

(b) Let B be a collection of circles in the plane, no two of which intersect. Need B be countable?

(c) Let C be a collection of figure eights in the plane, no two of which intersect. Need C be countable?

28. Find a set $\mathscr{A}$ of open intervals in $\mathbb{R}$ such that every rational number belongs to one of those intervals, but $\bigcup \mathscr{A} \neq \mathbb{R}$. [*Suggestion*: Limit the sum of the lengths of the intervals.]

29. Let A be a set of positive real numbers. Assume that there is a bound b such that the sum of any finite subset of A is less than b. Show that A is countable.

30. Assume that A is a set with at least two elements. Show that $\text{Sq}(A) \preccurlyeq {}^{\omega}A$.

NOTATION

For sets A and B we can form the collection of functions F from A into B. Call the set of all such functions ${}^{A}B$:

$$ {}^{A}B = \{F \mid F \text{ is a function from } A \text{ into } B\}. $$

If $F: A \to B$, then $F \subseteq A \times B$, and so $F \in \mathscr{P}(A \times B)$. Consequently we can apply a subset axiom to $\mathscr{P}(A \times B)$ to construct the set of all functions from A into B.

The notation ${}^{A}B$ is read "B-pre-A." Some authors write B^A instead; this notation is derived from the fact that if A and B are finite sets and the number of elements in A and B is a and b, respectively, then ${}^{A}B$ has b^a members. (To see this, note that for each of the a elements of A, we can choose among b points in B into which it could be mapped. The number of ways of making all a such choices is $b \cdot b \cdot \cdots \cdot b$, a times.) We will return to this point in Chapter 6.

Example Let $\omega = \{0, 1, 2, \ldots\}$. Then ${}^{\omega}\{0, 1\}$ is the set of all possible functions $f: \omega \to \{0, 1\}$. Such an f can be thought of as an infinite sequence $f(0), f(1), f(2), \ldots$ of 0's and 1's.

Example For a nonempty set A, we have ${}^{A}\varnothing = \varnothing$. This is because no function could have a nonempty domain and an empty range. On the other hand, ${}^{\varnothing}A = \{\varnothing\}$ for any set A, because $\varnothing: \varnothing \to A$, but $\varnothing$ is the only function with empty domain. As a special case, we have ${}^{\varnothing}\varnothing = \{\varnothing\}$.

CONTINUUM HYPOTHESIS

We have in this chapter given some examples of countable sets and uncountable sets. But every uncountable set examined thus far has had cardinality $2^{\aleph_0}$ or more. This raises the question: Are there any sets with cardinality between $\aleph_0$ and $2^{\aleph_0}$? The "continuum hypothesis" is the assertion that the answer is negative, i.e., that there is no κ with $\aleph_0 < \kappa < 2^{\aleph_0}$. Or equivalently, the continuum hypothesis can be stated: Every uncountable set of real numbers is equinumerous to the set of all real numbers.

Cantor conjectured that the continuum hypothesis was true. And David Hilbert later published a purported proof. But the proof was incorrect, and more recent work has cast doubt on the continuum hypothesis. In 1939 Gödel proved that on the basis of our axioms for set theory (which we here assume to be consistent) the continuum hypothesis could not be *disproved*. Then in 1963 Paul Cohen showed that the continuum hypothesis could not be *proved* from our axioms either.

But since the continuum hypothesis is neither provable nor refutable from our axioms, what can we say about its truth or falsity? We have some informal ideas about what sets are like, but our intuition might not assign a definite answer to the continuum hypothesis. Indeed, one might well question whether there is any meaningful sense in which one can say that the continuum hypothesis is either true *or* false for the "real" sets. Among those set-theorists nowadays who feel that there is such a meaningful sense, the majority seems to feel that the continuum hypothesis is false.

The "generalized continuum hypothesis" is the assertion that for every infinite cardinal κ, there is no cardinal number between κ and 2^κ. Gödel's 1939 work shows that even the generalized continuum hypothesis cannot be disproved from our axioms. And of course Cohen's result shows that it cannot be proved from our axioms (even in the special case $\kappa = \aleph_0$).

There is the possibility of extending the list of axioms beyond those in this book. And the new axioms might conceivably allow us to prove or to refute the continuum hypothesis. But to be acceptable as an *axiom*, a statement must be in clear accord with our informal ideas of the concepts being axiomatized. It would not do, for example, simply to adopt the generalized continuum hypothesis as a new axiom. It remains to be seen whether any acceptable axioms will be found that settle satisfactorily the correctness or incorrectness of the continuum hypothesis.

The work of Gödel and Cohen also shows that the axiom of choice can neither be proved nor refuted from the other axioms (which we continue to assume are consistent). But unlike the continuum hypothesis, the axiom of choice conforms to our informal view of how sets should behave. For this reason, we have adopted it as an axiom.

Results such as those by Gödel and Cohen belong to the *metamathematics* of set theory. That is, they are results that speak of set theory itself, in contrast to theorems within set theory that speak of sets.