1-2016

Discrete Mathematics: Chapter 5, Set Theory & Infinity

Calvin Jongsma
Dordt College, calvin.jongsma@dordt.edu

Discrete Mathematics: Chapter 5, Set Theory & Infinity

Abstract
In this chapter we will explore the notion of cardinality or numerosity from a more theoretical perspective. And our focus here will be on infinite sets: what distinguishes them from finite sets, and what distinguishes them from one another. This is the central topic that initiated the development of Set Theory by the German mathematicians Cantor and Dedekind in the last quarter of the nineteenth century, from 1872 – 1897.

Keywords
set theory, infinity, Georg Cantor, Richard Dedekind, logic

Disciplines
Christianity | Computer Sciences | Mathematics

Comments

- From Discrete Mathematics: An Integrated Approach, a self-published textbook for use in Math 212
- © 2016 Calvin Jongsma

✠ Chapter 5 ✠

SET THEORY
&
INFINITY

5.1 Countably Infinite Sets

Chapter 4 introduced elementary *Set Theory* and then used it as a platform on which to investigate a number of combinatorial counting principles. Our main focus there was finding efficient ways to count the cardinality of different finite sets.

In this chapter we will explore the notion of cardinality or numerosity from a more theoretical perspective. And our focus here will be on infinite sets: what distinguishes them from finite sets, and what distinguishes them from one another. This is the central topic that initiated the development of *Set Theory* by the German mathematicians Cantor and Dedekind in the last quarter of the nineteenth century, from 1872 – 1897.

Historically, though *Set Theory* was a branch of mathematics, in Cantor's thought it also had close connections with Roman Catholic philosophy and theology, which he became quite familiar with during his research. In fact, Cantor viewed himself as God's prophet of the infinite, revealing necessary truths of mathematics and theology grounded in the Mind of God.*

Because of the philosophical character of Cantor's set theory, a number of mathematicians at the time were antagonistic toward it, believing mathematics had lost its moorings and wandered into the field of religion. However, infinity enters nearly every part of mathematics in essential ways, so the topic can't be avoided. Cantor's virtue lay in his persistent belief that the notion of infinity could be consistently explicated if one used clear definitions to treat the topic. In so doing, Cantor exposed the sources of certain conceptual difficulties in earlier philosophical treatments of infinity, and he provided a way to think about infinity that many contemporary philosophers now find both attractive and definitive.

Mathematical logicians since Cantor's time have continued to expand the field. In David Hilbert's seminal 1900 address to the International Congress of Mathematicians on the 23 most important unsolved problems of the time, problems connected to transfinite set theory received prominent attention. *Set Theory* is still an area of active research today by people interested in the foundations of mathematics. It contains technical complications and subtleties that we will not be able to explore here, but we will open the door on the field and peek in. In the final section devoted to this topic, we will discuss the standard axiomatization of *Set Theory*, and we will look at one of the ways in which this area has impacted theoretical computer science.

Early on, Cantor discovered that in order to unravel some of the perplexities associated with infinity he had to refine his (and others') intuitions about sets and numerosity. The key mathematical notion that permitted him to do this was that of a one-to-one correspondence. We'll begin there ourselves as well, showing how it relates to numerosity and cardinality.

One-to-One Correspondence and Numerosity

How can you tell when two finite sets are exactly the same size? There are two ways this can be done, one yielding more information than the other. First of all, you can just count the sets. When you're done, you'd know whether they're the same size by comparing the two numbers. If they're equal, the sets have the same numerosity; otherwise not. You not only know whether the two sets are the same size; you also know what sizes they are.

A second way to compare the numerosity of sets is less informative: you only know in the end whether the sets are the same size, not what size they are. This occurs when you match up the two sets' elements one by one. If there are no unmatched elements in either set, they're the same size; otherwise the one with unmatched elements is more numerous.

Because this second method (matching) is more elementary than the first (counting), it can be used by children even before they're very familiar with numbers. They can tell that there

* For more on the theological and metaphysical context of Cantor's work in set theory, see the article by his biographer, Joseph Dauben, *Georg Cantor and the Battle for Transfinite Set Theory*, online at the ACMS website: http://www.acmsonline.org/journal/2004/Dauben-Cantor.pdf.

are just as many glasses as plates on the dinner table even if they can't tell you how many there are. The primitive character of the matching technique thus makes it an excellent method when you're beginning to learn about comparing quantities of things, but Cantor discovered that it is also the right method for an advanced theoretical treatment of cardinality. In fact, this approach works even when counting fails; namely, in the case where the sets are infinite. Some care has to be exercised here, but matching yields significant results. Cantor therefore made the idea of a one-to-one correspondence central to his whole approach.

DEFINITION* 5.1 - 1: *One-to-One Correspondence Between Sets

S can be put into one-to-one correspondence with T iff all the elements of S can be matched with all the elements of T in a one-to-one fashion.

This treats one-to-one correspondence informally, defining it in terms of a certain kind of matching. Later in the book (Sections 6.1 and 6.3), once we have formally introduced functions and relations, we will define one-to-one correspondence more precisely, both as a special sort of function and as an important type of relation. However, introducing such formal precision at this stage would make our work on the cardinality of infinite sets more complex, so we will stick with our more informal definition. The notion of infinity has enough conceptual complexity of its own without adding technical complications.

Note that the definition of one-to-one correspondence talks about S being put into one-to-one correspondence with T (in that order), but the relation of *being in one-to-one correspondence* is actually an equivalence relation (see Exercise 1): in particular, if S can be put into one-to-one correspondence with T, then T can also be put into one-to-one correspondence with S. This fact justifies our talking about one-to-one correspondences in a non-directed way, saying things like "S and T can be put into one-to-one correspondence with one another" or "there is a one-to-one correspondence between S and T."

Numerosity Relations Among Sets

One-to-one correspondences provide us with a tool for deciding about the relative sizes of sets: are two given sets the same size, or is one of them bigger than the other? The following definition spells out how this can be decided by means of one-to-one correspondences.

DEFINITION* 5.1 - 2: *Equinumerous Sets, Less Numerous Sets (***Cantor 1878***)

a) *S* ***is equinumerous with*** *T, written $S \sim T$, iff there is a one-to-one correspondence between S and T.*

b) *S* ***is less numerous than or equinumerous to*** *T, written $S \preceq T$, iff S is equinumerous with some subset of T.*

c) *S* ***is less numerous than*** *T, written $S \prec T$, iff $S \preceq T$ but $S \not\sim T$.*

We have taken $\preceq$ here as primary and $\prec$ as derivative because the definition of $\preceq$ in terms of equinumerosity is simpler. However, the obvious expected result between these relations holds: $S \preceq T$ iff $S \prec T$ or $S \sim T$ (see Exercise 8). This justifies the way we read $S \preceq T$. A number of other results also hold, though their proofs aren't always as simple as one might expect. The reason for this is the fact that we are basing everything upon one-to-one correspondences, which behave a bit differently for infinite sets than they do for finite sets (see, for example, Exercise 3).

One result that is central for working with these relations but which turns out to be far from trivial is the following theorem. The theorem was first stated by Cantor, who was unable to prove it. Dedekind was the first to give a proof, communicating it in an unpublished letter of 1887 to Cantor. In 1896 Schröder gave a flawed proof for the result; a year later Bernstein published the first valid proof. The theorem is usually referred to as the *Schröder-Bernstein Theorem*, though a more accurate name would probably be the *Cantor-Bernstein Theorem*.

THEOREM* 5.1 - 1: *The Schröder-Bernstein Theorem (***1897***)
If $S \preceq T$ and $T \preceq S$, then $S \sim T$.

***Proof*:**
Given the proof's complexity, we will not present one here. A proof can be rather easily constructed, however, given the following lemma (see Exercise 4a):

Lemma*: *Nested Equinumerous Sets
If $S_2 \subseteq S_1 \subseteq S_0$ and $S_0 \sim S_2$, then $S_0 \sim S_1 \sim S_2$.

Of course, this merely concentrates the theorem's difficulty in the lemma; now proving the lemma requires some ingenuity. Though this result may seem obvious, its proof is not trivial. We will leave this as an exploration (see Exercise 4b). ■

Finite and Infinite Sets: Some Distinctions and Background

What makes one set finite and another infinite? According to the most basic way of thinking about it, this is easy: a set S is finite iff it can be counted off by some natural number n. If $|S| = n$ for some $n \in \mathbb{N}$, S is finite; otherwise it's infinite. Thus $\mathbb{N}$ itself is not finite; there is no largest natural number. The set of all points on a line segment is likewise infinite. Infinite sets are larger than finite sets: they belong to a *transfinite* realm.

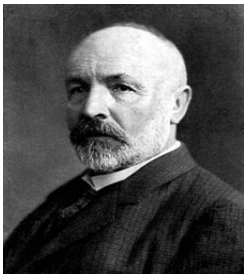
This approach takes the notion of *being finite* as primary. Infinity is treated as a purely negative concept: infinite means not-finite. This is traditionally how infinity was conceptualized. Furthermore, before the late nineteenth century, nearly everyone adhered to Aristotle's dictum on infinity: a quantity is infinite only in a potential sense, in that it can be continually added to. You can never consider a magnitude to be infinitely extended, only infinitely extendible. A similar result was thought true of a collection of objects: an infinite set that is present altogether is inconceivable; it leads to conceptual problems. Infinity is related to an *unending process*, not to a *completed totality* or entity. Potential infinity was acceptable, but not actual infinity.

Aristotle's dictum reigned among orthodox mathematicians concerned with logical rigor for over 2000 years. In support of this doctrine, mathematicians observed that if you considered certain infinite totalities as bona fide mathematical objects, then paradoxical contradictions arose. We will illustrate this with a historical example having to do with infinite sets.

In the seventeenth century Galileo noted that while the set of all perfect squares is smaller than the full set of positive integers, being a proper part of it, it is nevertheless equinumerous with the full set: each positive integer can be matched with its square. This paradoxical state of affairs was taken as substantiating Aristotle's dictum: completed infinities are contradictory.

The contradiction occurs, however, only if you assume that being smaller in the sense of being a subset entails being smaller in the sense of cardinality. This notion seems captured by *Axiom 5* of Euclid's ***Elements***: "The whole is greater than the part." This accords with common sense based on our experience with finite collections and finite continuous magnitudes, such as line lengths, so it seems self-evident. However, "common sense" fails us in the realm of the infinite; there we are faced with a choice that must be consciously made. What criterion should be used for comparing the numerosity of infinite sets?

After reflecting on this issue, Cantor chose to treat transfinite sets as completed totalities. Cantor rejected the Aristotelian viewpoint on infinite collections and adopted a more "theological" attitude, as some mathematicians characterized it: actual infinity is a legitimate notion for mathematics to use. Other mathematicians followed him in this, though some balked at importing what they considered "metaphysical ideas" into mathematics.



Georg Cantor

Cantor made the notion of one-to-one correspondence the basis of cardinality comparisons, as we did above. On this foundation it is possible to prove for finite sets that the cardinality of the whole is indeed greater than the cardinality of the proper part. But when sets are infinite, this property fails dramatically, as our example from Galileo illustrates. In Cantor's opinion, this is not due to treating infinite sets as completed totalities but to making an invalid generalization from the finite case. Self-evident properties that hold for finite sets may not simply be assumed to be true for sets in general. A set can be a proper part of the whole and still be equinumerous with it. In fact, this is the case for all infinite sets, as we will prove below (see *Corollary* 1 to *Theorem* 10). No contradiction automatically arises, so long as we don't postulate Euclid's axiom as holding universally for sets. We may still be left, however, with the sense of an uncomfortable strangeness in the face of learning how differently infinite sets behave in comparison with finite ones. This should eventually dissipate, however, after you work with infinite sets for a while.

Countably Infinite Sets: Definition and Basic Numerosity Results

We will begin our expedition into the realm of the infinite by exploring *countably infinite* sets. These are sets whose elements can be counted off, as it were, proceeding one by one without ever stopping. Since we can also count off finite sets, we will include them in the class of *countable* sets.

DEFINITION* 5.1 - 3: *Countably Infinite and Countable Sets

a) *A set S is* ***countably infinite*** *iff $\mathbb{N}^+ \sim S$, where $\mathbb{N}^+ = \{1, 2, \dots\}$.*

b) *A set S is* ***countable*** *iff it is countably infinite or finite.*

To demonstrate that a set is countably infinite, therefore, you must match up its elements with the entire sequence of positive natural numbers. Such a correspondence induces a sequential order on the set (first, second, etc.). We are thus led to the following definition and proposition. Note that the definition does not assume sets have enumerations; it just tells you what such a thing is if a set has one.

DEFINITION* 5.1 - 4: *Enumeration of a Set

An enumeration of a set S is a non-repetitious listing of its elements as a finite or infinite sequence; in which case the sequence is said to enumerate the set.

PROPOSITION* 5.1 - 1: *Countably Infinite Sets and Enumeration

A set S is countably infinite iff it can be enumerated by an infinite sequence.

***Proof*:**

Suppose S is countably infinite. Then it can be put into one-to-one correspondence with $\mathbb{N}^+$. Denoting by x_n the element of S matched with n, we have an enumeration of S: $x_1, x_2, \dots$. On the other hand, suppose S can be enumerated by an infinite sequence $x_1, x_2, \dots$. This sets up a one-to-one correspondence between S and $\mathbb{N}^+$, so S is countably infinite. ■

***COROLLARY*: *Countably Infinite Subsets of* $\mathbb{N}$**

a) *$\mathbb{N}^+$ is countably infinite.*

b) *$\mathbb{N}$ is countably infinite.*

***Proof*:**

Both parts of this corollary are immediate consequences of the proposition: enumerate the sets in their naturally occurring order. ■

The following proposition and its corollary give us an alternative characterization of being countably infinite.

PROPOSITION* 5.1 - 2: *Equinumerosity of Countably Infinite Sets
If T is countably infinite, then S is countably infinite iff $S \sim T$.

***Proof*:**
This depends on the fact that $\sim$ is an equivalence relation. See Exercise 6. ■

COROLLARY*: *Countably Infinite Sets: Alternative Characterization
S is countably infinite iff $\mathbb{N} \sim S$.

***Proof*:**
This follows immediately from the last two results. ■

There are times when it is handier to compare a set with $\mathbb{N}$, and there are times when $\mathbb{N}^+$ is the one to use. It all depends upon the circumstances. As we will see in some proofs below, for some occasions it is even convenient to use both characterizations. In addition, we have the option of using *Proposition* 1's enumeration characterization for countably infinite sets.

Countably Infinite Sets and Their Subsets

Both $\mathbb{N}^+$ and $\mathbb{N}$ are countably infinite. What about other sets of numbers, such as the integers, the rational numbers, and the real numbers? They're certainly infinite; are they also countable? In the next subsection we will investigate some familiar sets of numbers to determine whether they are countably infinite. At the same time, we will extend those results to obtain more general theorems. We will begin here, however, by making some important observations about the size of countably infinite sets in comparison to other sets.

Countably infinite sets are infinite and are thus larger than finite sets. On the other hand, they are the smallest of infinite sets: sets strictly smaller than countably infinite sets are finite. This is the intuitive content of the next few results. We begin with the theorem that was earlier thought to be paradoxical.

THEOREM* 5.1 - 2: *Countably Infinite Sets Contain Equinumerous Proper Subsets
If S is countably infinite, then S contains a countably infinite proper subset S^*; that is, S is equinumerous with a proper subset of itself.

***Proof*:**
This holds for $\mathbb{N}$ according to the *Corollary* to *Proposition* 1: $\mathbb{N} \sim \mathbb{N}^+$.
We can show a similar thing in general.
Since S is countably infinite, let $x_0, x_1, x_2, \ldots$ be an enumeration of its elements (here we're making use of S being equinumerous with $\mathbb{N}$).
Then $S^* = S - \{x_0\}$ can be enumerated by $x_1, x_2, \ldots$, so it is also countably infinite and thus equinumerous with S by *Proposition* 2. ■

THEOREM* 5.1 - 3: *Existence of Finite Subsets of Countably Infinite Sets
If S is a countably infinite set, then S contains finite subsets of all sizes.

***Proof*:**
Let S be a countably infinite set and let $\{x_1, x_2, \ldots, x_n, \ldots\}$ be an enumeration of S.
Let $S_n = \{x_1, x_2, \ldots, x_n\}$ for any $n \in \mathbb{N}$. This is a finite subset of S having n elements. ■

COROLLARY*: *Comparing Countably Infinite Sets and Finite Sets
If S is countably infinite and F is finite, then $F \prec S$.

***Proof*:**
See Exercise 16. ■

THEOREM* 5.1 - 4: *Infinite Subsets of Countably Infinite Sets
If S is an infinite subset of a countably infinite set T, then S is also countably infinite.

***Proof*:**
Since T is countably infinite, it can be listed by an infinite sequence $x_1, x_2, x_3, \ldots$.
Deleting all terms that are not elements of S leaves a subsequence $x_{k_1}, x_{k_2}, x_{k_3}, \ldots$.
This gives an enumeration of S.
Furthermore, it will be an infinite sequence because S is infinite.
Thus, S is countably infinite by *Proposition* 1. ■

COROLLARY* 1: *Subsets of Countably Infinite Sets
If $S \subseteq T$ and T is countably infinite, then S is finite or countably infinite.

***Proof*:**
See Exercise 24a. ■

Using the terminology of *countable sets* from *Definition* 3, the last corollary says that *subsets of countably infinite sets are countable.*

COROLLARY* 2: *Sets Less Numerous than Countably Infinite Sets
If $S \prec T$ and T is countably infinite, then S is finite.

***Proof*:**
See Exercise 24b. ■

Countably Infinite Sets and Set Theoretic Operations

The results of the last subsection compare the numerosity of countably infinite sets with that of other sets, both finite and infinite. We will postpone doing any more along this line (what is it we haven't done yet?) until the end of this section because of a technical complication. Here we will instead look at how countably infinite sets interact with others using set theoretic operations. We'll begin by noting that we can adjoin/remove a finite number of elements to/from such a set and the resulting set will still be countably infinite.

THEOREM* 5.1 - 5: *Disjoint Union of Finite and Countably Infinite Sets
If S is a finite set and T is a countably infinite set disjoint from S, then $S \cup T$ is countably infinite.

***Proof*:**
To prove this, we merely enumerate the finite set $S = \{x_1, x_2, \ldots, x_n\}$ prior to beginning our listing for $T = \{y_1, y_2, y_3, \ldots\}$.
This gives us an infinite enumeration for $S \cup T$: $x_1, x_2, \ldots, x_n, y_1, y_2, y_3, \ldots$. ■

COROLLARY* 1: *Union of Finite and Countably Infinite
If S is finite and T is countably infinite, then $S \cup T$ is countably infinite.

***Proof*:**
See Exercise 25a. ■

COROLLARY* 2: *Deletion of Finite from Countably Infinite
If S is finite and T is countably infinite, then $T - S$ is countably infinite.

***Proof*:**
See Exercise 25b. ■

The import of *Corollary* 2 is the following: you cannot make countably infinite sets finite by chipping off finite parts. The reverse is thus also true: countably infinite sets cannot be gotten by combining finite sets, not matter how large. There is an immense gulf between the finite and the infinite that cannot be finitely spanned. Ordinary speech about very large numbers being "nearly infinite" is a picturesque way to describe the size of such a natural number, but that's all it is. Amazingly, such numbers are really no closer to being infinite than 0 is.

The particular countably infinite sets we have encountered so far all exhibit the required enumerations in their natural orders. They have a first element, then a second, and so on. Giving an enumeration is essential to showing that a set is countably infinite, but nothing requires the enumeration to match the way the elements of the set are naturally ordered. In fact, the elements will often need to be rearranged in order to enumerate the set. The set of integers, for instance, has no first element, yet it, too, is countable.

PROPOSITION* 5.1 - 3: *The Integers are Countably Infinite

$\mathbb{Z} = \{\ldots, -2, -1, 0, 1, 2, \ldots\}$ *is countably infinite.*

***Proof*:**

We will give what has been called a *zigzag argument*.

To enumerate the integers, we must disrupt the natural order, since $\mathbb{Z}$ is infinite in two directions.

The zigzag enumeration $0, 1, -1, 2, -2, 3, -3, \ldots, n, -n, \ldots$ establishes the claim. ■

This proposition easily generalizes, as the next theorem shows. In fact, it can be generalized further, as the following results demonstrate.

THEOREM* 5.1 - 6: *Union of Countably Infinite Sets

If S and T are countably infinite, then $S \cup T$ is countably infinite.

***Proof*:**

Enumerate S and T and then merge their lists in a zigzag fashion, omitting any elements already in the list. ■

The last two theorems can be combined into a single statement: *If S is countable and T is countably infinite, then $S \cup T$ is countably infinite.* We can illustrate this with a cute example of dubious practicality.*

✠ **EXAMPLE 5.1 - 1**

Hilbert's Hotel is an imaginary spacious inn containing a countable infinity of rooms. Show that even if the inn is already full for the night, it can still accommodate a countable group of additional guests.

Solution

Since the full group of all guests, those present and those arriving, is still only countably infinite, reassign rooms according to the enumeration provided by the last two theorems.

Of course, reassigning rooms will be rather a nuisance for those already down for the night, for infinitely many people will have to get up and change rooms when each new group arrives, even if it's only one person. A set-theoretically knowledgeable innkeeper would instead house all her guests at the outset in such a way as to *always* leave a countable infinity of rooms open for further occupancy (see Exercise 27).

---

* Hilbert introduced this example without fanfare in a 1925 lecture to illustrate the difference between finite and infinite sets. Sometime later it was popularized in works discussing infinity, and it has now also been used in debates about cosmology and theology.

Coming back down to mathematical reality, the last theorem can be easily generalized to obtain the following corollary. Going even further with our generalization, *Theorem* 7 may surprise you; it requires a whole new proof strategy to justify.

COROLLARY*: *Finite Union of Countably Infinite Sets

If $S_1, S_2, \ldots, S_n$ are countably infinite sets, then $\bigcup_{i=1}^{n} S_i$ is countably infinite.

***Proof*:**

This result follows from *Theorem* 6, using mathematical induction. See Exercise 26. ■

THEOREM* 5.1 - 7: *Countably Infinite Union of Countably Infinite Sets

If each set S_i is countably infinite for $i \in \mathbb{N}^+$, then $\bigcup_{i=1}^{\infty} S_i$ is countably infinite.

***Proof*:**

A simple *diagonal argument* establishes this result. Here we will zigzag back and forth through all the sets, as it were, not just through two of them as before.

Let x_{ij} list the elements of each S_i for $j = 1, 2, 3, \ldots$.

Make an infinite array out of these lists, putting the elements of S_i in row i:

$$\begin{array}{ccccc} x_{11} & x_{12} & x_{13} & x_{14} & \cdots \\ x_{21} & x_{22} & x_{23} & x_{24} & \cdots \\ x_{31} & x_{32} & x_{33} & x_{34} & \cdots \\ x_{41} & x_{42} & x_{43} & x_{44} & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{array}$$

Now list all the elements in the union by moving through the array diagonally, starting at the top left corner and going down along the downward-left-sloping minor diagonals:

$$x_{11}; x_{12}, x_{21}; x_{13}, x_{22}, x_{31}; \ \ldots$$

If the sets S_i are pairwise disjoint, this enumeration will include every element in the array exactly once. If the sets overlap, elements will be repeated, so we should omit listing any element that is already in the list to avoid duplication. The resulting enumeration established the union as countably infinite. ■

Returning to consider specific sets of numbers, we've already shown that $\mathbb{N}$ and $\mathbb{Z}$ are countably infinite. What about $\mathbb{Q}$? $\mathbb{Z}$ was not terribly difficult to enumerate, because while it had no first element, at least its elements were all successors of others. But $\mathbb{Q}$ is not sparsely populated like $\mathbb{Z}$; its elements are densely packed together. Given any two rational numbers, it is possible to find a third one between them, so no rational number directly follows or precedes any other. On the face of it, then, it seems impossible to enumerate $\mathbb{Q}$. However, having proved *Theorem* 7, the job is relatively easy, as we now show.

PROPOSITION* 5.1 - 4: *The Rational Numbers are Countably Infinite

$\mathbb{Q} = \left\{ \frac{m}{n} : m, n \in \mathbb{Z},\ n \neq 0 \right\}$ is countably infinite.

***Proof*:**

A diagonal argument exactly like the one above shows that the positive rationals are countably infinite: list m/n for $m, n \in \mathbb{N}^+$ as x_{mn} to set up the array.

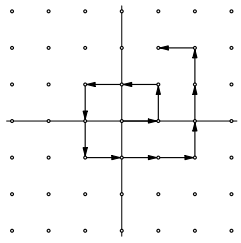
A similar argument shows that the negative rational numbers are countably infinite.

Thus the set of all non-zero rational numbers is countably infinite.

Adjoining 0 at the head of such a list shows that $\mathbb{Q}$ is countably infinite. ■

For a second, more geometric way of demonstrating this same result, first identify each rational number m/n with the point (m, n) in the plane.

These integral lattice points may then be enumerated in a counterclockwise spiral, beginning with $(0, 0)$ and continuing with $(1, 0)$, $(1, 1)$, $(0, 1)$, $(-1, 1)$, $(-1, 0)$, and so on.
To obtain a list of the rational number representatives (m, n), we omit all pairs ending in a 0 and all pairs that are integral multiples of earlier ones; for example, we omit $(-1, -1)$ because $(1, 1)$ is already in the list. This sequential listing also shows $\mathbb{Q}$ is countably infinite. ■



The Cartesian product of two countably infinite sets is often listed in array form as above. This is the basis of the next theorem.

THEOREM* 5.1 - 8: *Cartesian Product of Countably Infinite Sets
If S and T are countably infinite, then $S \times T$ is countably infinite.

***Proof*:**
See Exercise 29a. ■

COROLLARY*: *Finite Cartesian Product of Countably Infinite Sets

If S_i is countably infinite for each i, then $S_1 \times S_2 \times \cdots \times S_n = \{(x_1, x_2, \ldots, x_n) : x_i \in S_i\}$ is countably infinite.

***Proof*:**
Apply *PMI* to the number of sets involved. See Exercise 29b. ■

On the basis of the results proved so far, you might be ready to leap to the following ***conjecture:*** *the countably infinite Cartesian product of countably infinite sets is countably infinite.* However, as you recall, generalizing on the basis of earlier results doesn't always work out. Just because one set is a proper part of another set doesn't mean it's smaller, for instance. So if this result is true, it will need a careful proof. At this point, due to some complications, we'll leave this result as a conjecture for you to think about, taking it up in Section 5.2. In the meantime we will consider one more standard countably infinite set of numbers. In order to formulate a result about it, we must first give a definition.

DEFINITION* 5.1 - 5: *Algebraic Numbers
A complex number a is algebraic iff $p(a) = 0$ for some polynomial $p(x) = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0$ whose coefficients a_i are all integers.

Examples of algebraic numbers include rational numbers and radicals such as $\sqrt{2}$ and $i = \sqrt{-1}$, for each of these is the zero of some polynomial having integer coefficients (see Exercise 31ab). We will use the notation $\mathbb{A}$ for the set of all algebraic complex numbers.

PROPOSITION* 5.1 - 5: *The Algebraic Numbers are Countably Infinite (***Cantor 1874***)
$\mathbb{A}$, the set of all algebraic numbers, is countably infinite.

***Proof*:**
We will sketch the main points here, leaving the details as an exercise (see Exercise 31c).
For each $n \in \mathbb{N}^+$, there are countably many polynomials with integer coefficients of degree n.
From algebra, we know that each of these polynomials has at most n distinct zeros or roots.
Thus there are countably many zeros associated with polynomials of degree n.
Taking the union of all of these zeros for each positive integer n, we obtain a countably infinite set $\mathbb{A}$. ■

This proposition obviously entails the countability of $\mathbb{Z}$ and $\mathbb{Q}$ as well, since all rational numbers are algebraic.

Transfinite Cardinal Arithmetic and Countable Sets (Optional)

Finite sets have familiar cardinalities: the natural numbers. What about infinite sets? Don't they also have a definite "size" that can be denoted by means of the same cardinality symbol $|S|$? We will assume that the concept of cardinality is a primitive (undefined) notion capturing the idea of *how many/numerosity* for sets and that this notion makes some sense for all sets.

Before restating some of the above results in the language of cardinality, we will first connect the notion of cardinality to numerosity relations. The following definition asserts how cardinality order-relations relate to being equinumerous or less numerous.*

DEFINITION* 5.1 - 6: *Cardinality Comparisons and Numerosity

a) $|S| = |T| \leftrightarrow S \sim T$.
b) $|S| < |T| \leftrightarrow S \prec T$.
c) $|S| > |T| \leftrightarrow |T| < |S|$.

In earlier mathematics courses you probably used ∞ to stand for infinity, but this symbol often indicates an infinite process (shades of Aristotle's dictum), not a completed infinity of elements. So we will introduce new symbols to stand for the numerosity of infinite sets. The symbol Cantor used for the *cardinality of countably infinite sets* is the first letter of the Hebrew alphabet, $\aleph$, with the subscript 0: $\aleph_0$ is read "aleph nought" or "aleph null."

Given this notation, we can reformulate the above propositions simply as follows:

$$|\mathbb{N}| = \aleph_0 \quad |\mathbb{Z}| = \aleph_0 \quad |\mathbb{Q}| = \aleph_0 \quad |\mathbb{A}| = \aleph_0$$

Cardinal numbers denote the numerosity of sets, but we can also compare them with one another and even do arithmetic with them. Two numerosity comparison results given above translate into the following cardinality claims:

$$(\forall n \in \mathbb{N})(n < \aleph_0)$$
$$m \leq \aleph_0 \rightarrow (\exists n \in \mathbb{N})(m = n) \vee m = \aleph_0, \text{ where } m \text{ is any cardinal number.}$$

Various theorems on numerosity and set theoretic operations, on the other hand, translate into the following cardinal arithmetic results, given appropriate definitions for cardinal number operations of addition, multiplication, and exponentiation (which we won't pursue here):

$$(\forall n \in \mathbb{N})(\aleph_0 \pm n = \aleph_0)$$
$$\aleph_0 + \aleph_0 = \aleph_0$$
$$(\forall n \in \mathbb{N}^+)(n \cdot \aleph_0 = \aleph_0)$$
$$\sum_{i=1}^{\infty} \aleph_0 = \aleph_0 \cdot \aleph_0 = \aleph_0$$
$$(\forall n \in \mathbb{N}^+)(\aleph_0^n = \aleph_0)$$

These results require some time to digest and may seem strange, given what you know about ordinary finite arithmetic, but they provide cardinality formulations of what we have shown above. They give us the first results in transfinite cardinal arithmetic and are only the tip of the iceberg. In Section 5.2 we will obtain a few additional results, but we will not follow these ideas out any further. If you are interested in learning more about transfinite arithmetic, you can consult a standard textbook on *Set Theory.*

* A more rigorous set-theoretic treatment of cardinality would begin by independently defining the concept of cardinality; then this definition turns out to be a theorem.

Countably Infinite Sets and Infinite Sets

The theorems we have so far give us a partial determination of the position of countably infinite sets in the numerosity order-hierarchy for sets. We know that such sets are more numerous than finite sets and that they are infinite. Given that every subset of a countably infinite set is either finite or infinite, it seems clear that countably infinite sets are the smallest of all infinite sets. But this fact still needs to be argued: maybe different kinds of infinite sets can't even be compared with respect to size.

We left this result until now because it requires a result known as the *Axiom of Choice*. Informally, this axiom says that given any pairwise-disjoint infinite collection of sets, choice sets exist that contain one selected element for each set in the collection. This axiom seems intuitively clear, at least given the kinds of collections we're familiar with, but in fact it is one of the more controversial results of set theory and leads to some peculiar consequences, such as the *Banach-Tarski Paradox*, results that in their popular forms are highly counterintuitive. We won't go into more detail on this here; you can explore this as well by consulting a textbook in *Set Theory* or by looking up "axiom of choice" on-line (see Exercise 34).

THEOREM* 5.1 - 9: *Infinite Sets Contain Countably Infinite Subsets

If T is an infinite set, then T contains a countably infinite subset S.

***Proof*:**

Let T be as given. We construct S by selecting and enumerating its elements in stages.

Since T is infinite, $T \neq \emptyset$ and so contains some element x_0.

Let $S_0 = \{x_0\}$ and $T_0 = T - S_0$. Since S_0 is finite, its complement T_0 must be infinite and hence non-empty.

Continuing recursively, suppose that $S_n = \{x_0, x_1, \ldots, x_n\}$ is a set of distinct elements chosen from T. Then since S_n is finite, its complement $T_n = T - S_n$ is non-empty.

Choose $x_{n+1} \in T_n$ and let $S_{n+1} = S_n \cup \{x_{n+1}\}$. This gives an enlarged set of distinct elements of T.

Now take $S = \{x_0, x_1, \ldots, x_n, \ldots\}$. Since all x_i are distinct, given the way we choose them, S is a countably infinite subset of T. ■

COROLLARY*: *Countably Infinite Subsets of Infinite Sets are Less- or Equi-Numerous

If T is an infinite set and S is a countably infinite subset of T, then $S \preceq T$.

***Proof*:**

See Exercise 32. ■

Given *Theorem* 2 and *Theorem* 9, we can go on to prove an even stronger result: all infinite sets are equinumerous with proper subsets of themselves.

THEOREM* 5.1 - 10: *Infinite Sets Contain Equinumerous Proper Subsets

If T is an infinite set, then T is equinumerous with some proper subset of itself.

***Proof*:**

Let T be any infinite set, and let $S = \{x_0, x_1, x_2, \ldots\}$ denote a countably infinite subset.

Following the method of *Theorem* 2, we first put S into one-to-one correspondence with a proper subset of *itself* before tackling the full set.

Match each element x_n in S with its successor x_{n+1} in the list. This places S in one-to-one correspondence with the subset $S^* = S - \{x_0\}$.

We now show how to extend this matching to get a one-to-one correspondence between T and $T^* = T - \{x_0\}$.

If an element is in S, match it in the way just indicated. If an element is outside S, match it with itself. This matches everything in T with elements in T^* in the way required.

Thus $T \sim T^*$, so T is equinumerous with a proper subset. ■

COROLLARY* 1: *Infinite Sets: Alternate Characterization

A set T is infinite iff T is equinumerous to a proper subset of itself.

***Proof*:**

See Exercise 33a. ■

Richard Dedekind

This corollary stands Euclid's *Axiom* 5 on its head: for infinite sets, the whole is not necessarily greater than its part. Recognizing this fact led Richard Dedekind to adopt this positive characterization of infinite sets as their defining property. This definition of being infinite has the advantage that it does not require any prior knowledge of the natural numbers, as our earlier definition does. Being finite can then be defined, according to Dedekind, as being not infinite.

This approach is legitimate, but it is abstract and non-intuitive, so we did not adopt it as our main approach. Moreover, by what we noted above, the validity of this result depends on a version of the *Axiom of Choice*. Without the *Axiom of Choice*, there can be sets that are infinite in the sense of our original definition that are not Dedekind-infinite (infinite in Dedekind's sense). Equivalently, there can be sets that are Dedekind-finite (finite in Dedekind's sense) that are not finite according to our definition.

COROLLARY* 2: *Infinite Sets Absorb Finite or Countably Infinite Sets

a) *If S is finite or countably infinite and T is infinite, then $S \cup T \sim T$.*

b) *If S is finite or countably infinite and T is uncountably infinite, then $T - S \sim T$.*

***Proof*:**

This is proved by modifying the proof of *Theorem* 10. See Exercise 33bc. ■

EXERCISE SET 5.1

Problems 1 - 4: Equinumerosity

Prove the following results concerning equinumerosity.

*1. Show that the relation of equinumerosity is an equivalence relation on sets. That is, show it satisfies the following properties, using the definition of $\sim$. (*Note*: do not use properties of $|S|$ here.)

a. *Reflexive Property*: $S \sim S$.

*b. *Symmetric Property*: If $S \sim T$, then $T \sim S$.

*c. *Transitive Property*: If $R \sim S$ and $S \sim T$, then $R \sim T$.

*2. Show by example that it is possible for two sets S and T to be matched up in a one-to-one fashion by two different matchings so that in the first matching all the elements of S are matched up with elements of T and T has elements with no mates, while in the second matching just the opposite is true (the roles of S and T are reversed). *Hint*: could this occur for finite sets?

*3. Prove that $S \times T \sim T \times S$.

4. *The Schröder-Bernstein Theorem*

a. Given the *Lemma on Nested Equinumerous Sets*, prove the *Schröder-Bernstein Theorem*: If $S \preceq T$ and $T \preceq S$, then $S \sim T$.

b. Prove the *Lemma on Nested Equinumerous Sets*: *If $S_2 \subseteq S_1 \subseteq S_0$ and $S_0 \sim S_2$, then $S_0 \sim S_1 \sim S_2$.* (This result is much more difficult to prove.)

5. True or false? Prove or disprove your claims, using results about $\sim$.

a. If $R \sim T$ and $S \sim V$, then $R \cup S \sim T \cup V$.

b. If $R \sim T$ and $S \sim V$, then $R \cup S \sim T \cup V$, provided $R \cap S = \emptyset$ and $T \cap V = \emptyset$.

c. If $R \sim T$ and $S \sim V$, then $R \cap S \sim T \cap V$.

6. Prove *Proposition* 2: *If T is countably infinite, then S is countably infinite iff $S \sim T$.* Use your knowledge of logic to design a proof strategy.

Problems 7 - 10: Numerosity Order Comparisons

Using the definitions and results in this section about $\preceq$ (but not Axiom 1 and results about $\leq$), prove the following numerosity order results.

*7. *Monotonicity Property*

*a. If $S \subseteq T$, then $S \preceq T$.

b. Why isn't the following true: if $S \subset T$, then $S \prec T$?

*c. Using part a, argue why $S \preceq S$, $S \cap T \preceq S$, $S \preceq S \cup T$, and $S - T \preceq S$.

*8. $S \preceq T$ iff $S \prec T$ or $S \sim T$.

9. $S \prec T$ iff $S \preceq T$ and $T \not\preceq S$.

10. If $S \preceq T$and $T \preceq U$, then $S \preceq U$.

Problems 11 - 16: Strict Numerosity Order Comparisons

Using the definitions and results in this section about $\prec$ (but not Axiom 1 and results about $<$), prove the following strict numerosity order results.

11. $S \not\prec S$.

12. If $S \prec T$, then $T \not\preceq S$.

13. If $S \prec T$, then $T \not\prec S$.

*14. If $R \prec S$ and $S \prec T$, then $R \prec T$.

15. If $S_1 \sim S_2$ and $T_1 \sim T_2$, then $(S_1 \prec T_1 \leftrightarrow S_2 \prec T_2)$.

16. Prove the *Corollary* to *Theorem* 3: *If S is countably infinite and F is finite, then $F \prec S$.*

Problems 17 - 20: True or False

Are the following statements true or false? Explain your answer.

17. $S \sim T$ iff for each element $x \in S$ there is a unique element $y \in T$ that can be associated with it.

*18. $S \prec T$ iff S can be put into one-to-one correspondence with a proper subset of T.

19. A set is countable if its elements can be listed off by a finite or infinite sequence.

*20. Since the set of rational numbers are dense (between any two of them there is a third one) while the set of integers is not, there are more rational numbers than integers.

Problems 21 - 31: Countably Infinite Sets

Work the following problems involving countably infinite sets.

*21. *Countable Subsets of* $\mathbb{N}$

a. Prove that the set of natural numbers greater than 100 is countably infinite by enumerating its elements. Then determine an explicit formula for the n^{th} element in your list, starting with $n = 0$.

*b. Show that the set of all positive integral powers of ten is countably infinite, both by enumerating its elements and by finding a formula $f(n)$ that gives the n^{th} element, starting with $n = 0$.

EC 22. *The Zigzag Argument*

Determine an explicit formula $f(n)$ for the n^{th} element in the enumeration set up by the zigzag argument in *Proposition* 5.1-3. Take $f(0) = 0, f(1) = 1, f(2) = -1$, etc. Hint: use the floor function $\lfloor x \rfloor$ = the largest integer less than or equal to x.

23. *The Odd Integers are Countably Infinite*

a. Prove that the set of all odd integers (positive and negative) is countably infinite by listing its elements.

b. Give an explicit formula $f(n) = x_n$ for the n^{th} element in your listing of part a.

24. *Theorem 4*
    a. Prove *Corollary* 1 to *Theorem* 4: *If $S \subseteq T$ and T is countably infinite, then S is finite or countably infinite.*
    b. Prove *Corollary* 2 to *Theorem* 4: *If $S \prec T$ and T is countably infinite, then S is finite.*

*25. *Theorem 5*
    *a. Prove *Corollary* 1 to *Theorem* 5: *If S is finite and T is countably infinite, then $S \cup T$ is countably infinite.*
    b. Prove *Corollary* 2 to *Theorem* 5: *If S is finite and T is countably infinite, then $T - S$ is countably infinite.*

26. *Theorem 6*
    Prove the *Corollary* to *Theorem* 6: *If $S_1, S_2, \ldots, S_n$ are each countably infinite sets, then $\bigcup_{i=1}^{n} S_i$ is countably infinite.*

*27. *Hilbert's Hotel*
    *a. Imagine that you're the inn-keeper of *Hilbert's Hotel* (see Example 1). Explain how you can house a countably infinite number of guests and still have room for countably many late arrivals.
EC b. Suppose you've housed a countably infinite number of guests as in part a, and then two more such groups arrive. Explain how you will house these in succession without making anyone move to a new room. If still more groups arrive, will you be able to accommodate them in the same way?

EC 28. Determine a formula for enumerating the infinite rectangular array given in first part of the proof of *Theorem* 7. This may be easier if you find a formula for $f(m, n)$, where x_{mn} is listed in the $f(m, n)^{\text{th}}$ place in the enumeration.

29. *Theorem 8*
    a. Prove *Theorem* 8: *If S and T are countably infinite, then $S \times T$ is countably infinite.*
    b. Prove the *Corollary* to *Theorem* 8: *If S_i is countably infinite for each i, then $S_1 \times S_2 \times \cdots \times S_n$ is countably infinite.*

*30. *Finite Sequences*
    *a. How many finite sequences of 0's and 1's are there? Prove your answer.
    *b. How many finite subsets does $\mathbb{N}$ have? Hint: show how to model any given finite subset using a string of 0's and 1's and then use part a.
    c. How many different finite sequences of natural numbers (finite lists of natural numbers) are there? Prove your answer.

31. *Algebraic Numbers*
    a. Using *Definition* 5, prove that any rational number m/n is an algebraic number.
    b. Using *Definition* 5, prove that both $\sqrt{2}$ and $i = \sqrt{-1}$ are algebraic numbers.
    c. Fill in the set-theoretic details for the proof of *Proposition 5*.
    d. Prove that the set of all real algebraic numbers ($\mathbb{A} \cap \mathbb{R}$) is countably infinite.

Problems 32 - 33: Infinite Sets and Countably Infinite Sets

Work the following problems relating infinite sets and countably infinite sets.

32. *Theorem 9*
    Prove the *Corollary* to *Theorem* 9: *If T is an infinite set and S is a countably infinite subset of T, then $S \prec T$ or $S \sim T$.*

33. *Theorem 10*
    a. Prove *Corollary* 1 to *Theorem* 10: *A set T is infinite iff T is equinumerous to a proper subset of itself.*
    b. Prove *Corollary* 2a to *Theorem* 10: *If S is finite or countably infinite and T is infinite, then $S \cup T \sim T$.*
    c. Prove *Corollary* 2b to *Theorem* 10: *If S is finite or countably infinite and T is uncountably infinite, then $T - S \sim T$.*

34. Look up the terms *Axiom of Choice* and *Banach-Tarski paradox* using an internet search-engine such as Googol. Explain why some feel the *Axiom of Choice* is a questionable principle when used in full generality (there are restricted versions of the axiom that are less controversial).

HINTS TO STARRED EXERCISES 5.1

1. b. Use the definition of $\sim$.
   c. To begin your proof, suppose that $R \sim S$ and $S \sim T$, then use the definition of $\sim$.
2. Take a pair of familiar infinite sets and match them up as required.
3. What's an obvious ordered pair in $T \times S$ to match $(x, y) \in S \times T$ to? Explain why your choice gives a one-to-one correspondence.
7. a. Use the fact that S can be put into one-to-one correspondence with itself (Exercise 1a).
   b. [No hint.]
8. Use *EO* to prove one direction and *Cases* to prove the other.
14. Work this *without* using what you know about *finite* cardinalities. Properties of the transfinite realm are established using results about $\sim$, $\preceq$, and $\prec$. Use the definition of $\prec$ here.
18. [No hint.]
20. [No hint.]
21. b. Use exponential notation for your formula.
25. a. To show that $S \cup T$ is countably infinite, find a way to enumerate its elements.
27. a. After you figure this out, try part b for EC.
30. a. There is a countable infinity of finite strings of 0s and 1s. Find a way to enumerate these strings.
    b. Use the result of part a, but show how it's relevant.

5.2 Uncountably Infinite Sets

Prior to Cantor's work on *Transfinite Set Theory* in the last quarter of the nineteenth century, mathematicians and philosophers had distinguished finite quantities from infinite ones, but they had little inkling that there might be different orders of infinity. Being infinite meant being not-finite. When Cantor discovered there were different size infinities, he began talking about *transfinite* cardinal numbers, cardinalities beyond those of finite numbers. He found that the realm of *Transfinite Set Theory* has some definite contours, that one could calculate with and compare such cardinal numbers, and that the transfinite realm was richer and far stranger than the realm of finite arithmetic. One of the problems he bequeathed to mathematics was determining the exact size of the linear continuum. This made it to the very top of Hilbert's famous list of 23 open problems delivered in his 1900 address to the second International Congress of Mathematicians. The *Continuum Problem's* surprising two-stage solution during the mid-twentieth century brought fame to those involved.

In this section we will explore the most elementary portions of all this, just enough to present a few key results and understand what Cantor's *Continuum Problem* is all about. Much more has been investigated on these topics and in great depth since the time of Cantor. There is also the related topic of transfinite *ordinals* (numbers corresponding to well-ordered sets), which we won't touch at all.

Transfinite Set Theory has very important connections to logic and philosophy as well as to certain foundational portions of various branches of mathematics, but at this time it remains somewhat remote from rudimentary topics in *Discrete Mathematics*. We will point out in Section 5.3, however, that some proof techniques (diagonalization) associated with the results below have been applied in a number of disciplines, including areas of theoretical computer science that deal with computability.

Infinity and Countable Infinity

In Section 5.1 we began our exploration of infinite sets. Using the criterion that countably infinite sets are the ones that can be put into an infinite sequence or placed in one-to-one correspondence with $\mathbb{N}^+$, we were able to prove a number of results about such sets. The familiar sets of numbers $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{Q}$, and $\mathbb{A}$ were all shown to be countably infinite, and the set-theoretic operations of union and Cartesian product applied to two (and thus finitely many) countably infinite sets were shown to yield countably infinite sets.

Given the evidence amassed on countably infinite sets, one might wonder whether all the hoopla about countably infinite sets isn't just so much sophisticated hot air, whether infinite sets can't *all* be shown to be countably infinite, given enough ingenuity. Sets like $\mathbb{Q}$ didn't look like they could be listed at the outset, but given an inventive rearrangement of the elements' natural order, it was found to be possible. Maybe this can always be done. Can't you just choose some element as the first one, then pick another as the second, and so on, until eventually everything in the entire set is listed?

The answer to this is mostly "no". It is a little bit "yes," in the highly qualified sense that every set can be *well-ordered* or sequentialized, provided the *Axiom of Choice* is accepted. This is the result that originally got Zermelo started on axiomatizing *Set Theory* in 1908. But even given this very surprising result, which is too involved to get into here, it is still "no" with respect to numerosity. A listing procedure can be used to demonstrate that every infinite set contains a countably infinite subset (*Theorem* 5.1-10), but unless you are able to show that your countably infinite sequence eventually catches every element of the set, you do not know whether the whole set can be enumerated. And, in fact, not all infinite sets are countably infinite, as we will show in a powerful way below. Some infinite sets are so big that they are *uncountable*, in the strong, human-independent sense that they cannot be listed by any

countably infinite sequence, no matter how ingeniously devised. Such sets are more numerous than the set of natural numbers. The notion of infinity thus gets refracted into a spectrum of different transfinite cardinalities when it is passed through the prism of Cantor's *Set Theory*.

Uncountably Infinite Sets and Geometric Point Sets

We will now give repeated demonstration of the mathematical existence of uncountably infinite sets. In fact, there are infinitely many different-sized uncountable sets, though that will not be immediately apparent, just as it wasn't to Cantor. We begin with the obvious definition.

DEFINITION* 5.2 - 1: *Uncountably Infinite Sets
A set S is ***uncountably infinite/uncountable*** *iff S is infinite but not countably infinite.*

The problem with uncountable sets is not that they can't be matched up with other sets. *Theorem* 5.1-10 shows that any infinite set can be put into one-to-one correspondence with some other set (a proper subset of itself). The problem is that uncountable sets are simply too big to match up with $\mathbb{N}^+$, even though that, too, is infinite. On the other hand, uncountable sets may or may not be equinumerous with one another. The following proposition gives some basic comparison results for uncountably infinite sets.

PROPOSITION* 5.2 - 1: *Numerosity and Uncountably Infinite Sets
a) *If S is uncountably infinite and $S \sim T$, then T is uncountably infinite;*
b) *If S is uncountably infinite and $S \subseteq T$, then T is uncountably infinite;*
c) *If S is countably infinite and T is uncountably infinite, then $S \prec T$.*

***Proof*:**
See Exercise 16. ■

You probably noticed in our discussion of countably infinite sets that we failed to treat either the set of real numbers or the set of complex numbers, even though we treated the set of algebraic numbers. There was good reason for this: it turns out that these sets are not countably infinite. We will show this, starting with demonstrating that the real line is uncountable. We will then consider higher dimensional spaces to see what new complications they might bring into the picture. Our train of consequences rests upon the fundamental fact that the set of real numbers between 0 and 1 (the open unit interval) is uncountable.

***THEOREM* 5.2 - 1: *Numerosity of* $(0,1)$ (*Cantor 1874*)**
The open interval $(0,1) = \{x \in \mathbb{R} : 0 < x < 1\}$ is uncountably infinite.

***Proof*:**
We will give Cantor's second, famous diagonal argument (1891), which involves proof by contradiction.

We can *uniquely represent* any number in the interval $(0,1)$ by choosing its non-terminating decimal representation $.a_1a_2a_3\cdots$; for example, .5 gets represented by $.4\overline{9}$, $1/7$ by $.\overline{142857}$, and $\pi/4$ by $.78539816339\cdots$.
Suppose now that the following sequence lists all the elements of $(0,1)$, and let x_{ij} denote the j^{th} digit in the decimal representation of the i^{th} number in the list.

$$
\begin{array}{l}
.\,x_{11}x_{12}x_{13}\cdots \\
.\,x_{21}x_{22}x_{23}\cdots \\
.\,x_{31}x_{32}x_{33}\cdots \\
\cdots\cdots\cdots\cdots \\
\cdots\cdots\cdots\cdots
\end{array}
$$

Such a list is *inherently incomplete*: it *must* miss a decimal between 0 and 1. We will manufacture one such number by proceeding along the main diagonal of the listing.
Construct a non-terminating decimal $d = .d_1d_2d_3\cdots$ in the following way (the choice of numerals is arbitrary, except you must steer clear of 0 to avoid a terminating decimal).
Let d_n be any definite number different from x_{nn}; say,

$$d_n = \begin{cases} 1: & x_{nn} \neq 1, \\ 2: & x_{nn} = 1. \end{cases}$$

Since d differs in the n^{th} place from the n^{th} number listed, it cannot be in the list.
Thus, there is no way to list all the real numbers in the open interval $(0, 1)$. ■

Theorem 1 has an immediate consequence via *Proposition* 1b: the set of real numbers $\mathbb{R}$ or indeed any superset of $(0, 1)$ is uncountable.

COROLLARY* 1: $\mathbb{R}$ *is Uncountably Infinite

If $S \supseteq (0, 1)$, then S is uncountable. Hence, $\mathbb{R}$ is uncountable.

This result still leaves open the question whether all such real number supersets of $(0, 1)$ are uncountable *in the same way* as $(0, 1)$. On first thought you might suspect the answer is no; after all, $(0, 1)$ is only a short line segment, while $\mathbb{R} = (-\infty, \infty)$ can be thought of as an infinite line. The next results answer this question. We will motivate them by an example.

✠ **EXAMPLE 5.2 - 1**

Show that the following intervals are equinumerous to $(0, 1)$.
a) $(1, 2)$
b) $(0, 2)$
c) $(1, 4)$

Solution

The following formulas define one-to-one correspondences between $(0, 1)$ and the above sets and so establish the equinumerosity claimed. A geometric interpretation is also suggested for each of these matchings. Convince yourself that each of these matchings are one-to-one correspondences.

a) $y = x + 1$; this matches points by translating those in $(0, 1)$ one unit right.
Alternatively, think about this in two-dimensional terms: the graph of $y = x + 1$ above the x-interval $(0, 1)$ is directly to the right of the y-interval $(1, 2)$. Going straight up from an x-value to the graph and then across to the y-axis sets up a one-to-one matching between x-values in $(0, 1)$ and y-values in $(1, 2)$.

b) $y = 2x$; this stretches the unit interval to twice its length.
The graph of $y = 2x$ over the interval $(0, 1)$ can be used to establish the one-to-one correspondence between it and $(0, 2)$, just as in part *a*.

c) $y = 3x + 1$; this stretches the unit interval to three times its length and then translates that one unit right.
The graph of this line in 2-D can be used to establish a matching between $(0, 1)$ and $(1, 4)$, as in parts *a* and *b*.

PROPOSITION* 5.2 - 2: *All Open Intervals Are Equinumerous and Uncountable

All finite open intervals are equinumerous and are uncountably infinite: $(a, b) \sim (c, d)$ for any real numbers $a < b$, $c < d$.

***Proof*:**

See Exercise 8. ■

So stretching or shrinking an interval by a finite magnification factor does not affect its numerosity: every finite open interval contains the same number of points. This also holds if the interval is closed or half-open and half-closed (see Exercise 11). In particular, the intervals $(0, 1]$ and $[0, 1]$ are equinumerous with $(0, 1)$: adjoining one or two more points don't change the infinite cardinality of a set (see *Corollary* 2a to *Theorem* 5.1-10). The next result shows that magnifying the interval's length by an infinite factor, as it were, doesn't affect the interval's numerosity, either.

PROPOSITION* 5.2 - 3: $\mathbb{R}$ *is Equinumerous with (0,1)

$\mathbb{R} \sim (0, 1)$

***Proof*:**

The formula $y = \dfrac{x - .5}{x(x - 1)}$ [alternatively, $y = \tan(\pi x - \pi/2)$] sets up a one-to-one correspondence between the open interval $(0, 1)$ and the full set of real numbers (see Exercise 9). Thus $\mathbb{R} \sim (0, 1)$. ■

We will now look at an important consequence of the fact that while the rational numbers and even the algebraic numbers are countable, the set of all real numbers is uncountable. We first state another definition.

DEFINITION* 5.2 - 2: *Transcendental Numbers

A real number is transcendental iff it is not algebraic.

All transcendental real numbers are irrational, but not conversely. Many irrational numbers, such as $\sqrt{2}$, are algebraic. Examples of familiar transcendental real numbers are π and e. A natural question to investigate is, which real numbers are transcendental, and how many are there? Is the set of all transcendental numbers finite? Countably infinite? Uncountably infinite? The next proposition answers this question. It is an immediate corollary of the work done by Cantor in 1874 in showing that algebraic numbers are countable while real numbers are uncountable. Dedekind saw no practical consequence of $\mathbb{A}$ being countable; Cantor did.

PROPOSITION* 5.2 - 4: *Transcendental Numbers Are Uncountable

There are uncountably many transcendental real numbers.

***Proof*:**

We use *Proof by Contradiction*.

Since the set $\mathbb{A}$ of algebraic numbers is countably infinite, its restriction to the reals, $\mathbb{A} \cap \mathbb{R}$, is countable.

Thus if the transcendental real numbers were countably infinite, the union of this set with the set of algebraic real numbers, namely $\mathbb{R}$, would also be countable; but it's not.

So the set of transcendental real numbers is uncountably infinite. ■

This result gives a slick answer to a difficult problem. In 1844 Liouville had demonstrated how to construct infinitely many transcendental reals, such as $0.11000100000000000000000001\cdots$, where 1s occur in the $n!$ decimal places. In 1873 Hermite managed to show that e is transcendental. In the following year Cantor showed that every interval of real numbers contains infinitely many transcendentals, constructively showing how to generate such numbers. The proof given above, published by Felix Klein in 1894, fails to construct or identify any specific transcendental number, but it nevertheless demonstrates that there are uncountably many of them. In essence, then, real numbers are typically (almost always) transcendental numbers, even though we are not familiar with many of them in particular. In 1882 Lindemann showed

that π is transcendental. Work done in the twentieth century has investigated more generally when a^b is transcendental and has shown that numbers like $2^{\sqrt{2}}$ are transcendental.*

Once it is clear from the examples of $\mathbb{N}$ and $\mathbb{R}$ that there are infinite sets of two different sizes, it is natural to query whether all other known infinite sets are one of these two types or whether there are still other sizes of infinity. We found many sets that were countably infinite/the same size as $\mathbb{N}$, and we discovered that every infinite set is at least as large as $\mathbb{N}$ because it contains a countably infinite subset. Are all uncountably infinite sets the same size as $\mathbb{R}$? There are two or three subquestions to this query:

1. Are there other well-known sets the same size as $\mathbb{R}$?
2. Are there any sets whose cardinality is strictly larger than that of $\mathbb{R}$?
3. Are there any sets whose cardinality lies strictly between that of $\mathbb{N}$ and $\mathbb{R}$?

We will take up these questions in the order just stated. Actually, we will begin by looking at sets that have the potential to be larger than $\mathbb{R}$ and show (like we did earlier for $\mathbb{N}$) that in fact they are still the same size as $\mathbb{R}$. Then we will go on to show, however, that there are lots of sets (infinitely many) strictly larger than $\mathbb{R}$. In conclusion, we will look briefly at what is known about sets of intermediate size.

Sets Equinumerous with the Continuum

An obvious thing to explore next is what happens to cardinality when the dimensionality of the system is increased. $\mathbb{R}$ quantifies the linear/one-dimensional continuum; what if we consider $\mathbb{R}^2 = \mathbb{R} \times \mathbb{R}$, which quantifies two-dimensional space? if we take $\mathbb{R}^3$ (three-dimensional space)? or $\mathbb{R}^n$?

Using the Cartesian-product operator on $\mathbb{N}$ didn't produce anything new: we found that all $\mathbb{N}^n$ are still countably infinite. According to the next theorem, the same thing happens here: moving from one dimension to two or three or even any finite dimension fails to increase the total number of points. In other words, the size of the collection of points in space is largely independent of the dimension. This fact greatly surprised Cantor, who like others at the time thought higher dimensional space must contain a greater multiplicity of points than the linear continuum. In writing to Dedekind about the result, Cantor exclaimed, "I see it, but I don't believe it!"** Some mathematicians at the time likewise saw it but didn't like it, finding Cantor's brand of set-theoretic mathematics unsavory.

THEOREM* 5.2 - 2: *The Plane is Equinumerous with the Line (***Cantor 1878***)
$\mathbb{R}^2 \sim \mathbb{R}$

***Proof*:**
Since $\mathbb{R} \sim (0, 1]$, $\mathbb{R}^2 \sim (0, 1] \times (0, 1]$.
To prove our result, then, it suffices to show $(0, 1] \times (0, 1] \sim (0, 1]$.
The basic intuition behind matching these two sets is the following.
Given any point (x, y) in the unit square, we use the numbers' non-terminating decimal representations and write its components as $(. x_1x_2x_3\cdots, . y_1y_2y_3\cdots)$.
We match this ordered pair of real numbers to the real number z constructed in zigzag fashion: $z = . x_1y_1x_2y_2x_3y_3\cdots$.
This matching very nearly works. It matches different ordered pairs with different numbers, but unfortunately some numbers will not be matched up with any ordered pair. (For instance, both $z = .1101010\cdots$ and $z = .101010\cdots$ will have no ordered pair (x, y)

* This was featured as Problem 7 in Hilbert's 1900 list: is a^b transcendental if a is algebraic (but not 0 or 1) and b is irrational? The answer by Russian mathematician Alexander Gelfond in 1934 was: yes, when b is an algebraic irrational. Characterizing when a^b is transcendental remains open.

** Today's mathematics students may be less surprised by this phenomenon, particularly if they have seen space filling curves generated as the limit of a sequence of curves defined by an iterative process.

related to them. In the first case we would need to have $x = .100\cdots$, which is not a non-terminating decimal, while in the second case $y = .000\cdots$, which is neither positive nor non-terminating.)
A clever modification of the basic procedure, however, using *blocks of digits* instead of single digits, avoids these problems: first list a block of x's digits, continuing until a non-zero digit is reached [why must this happen?], then do the same for y's digits, then go back to x and list another block of its digits, etc. This procedure will now yield a one-to-one correspondence between the unit square and the unit interval (see Exercise 13a). ■

COROLLARY* 1: *Complex Numbers are Equinumerous with Real Numbers
$\mathbb{C} \sim \mathbb{R}$

***Proof*:**
$\mathbb{C} \sim \mathbb{R}^2$ because $a + bi$ can be paired off with (a, b) in $\mathbb{R}^2$ (this is the standard geometric representation for $\mathbb{C}$).
Since $\mathbb{R}^2 \sim \mathbb{R}$, our result follows by transitivity of $\sim$. ■

COROLLARY* 2: *N-Dimensional Space is Equinumerous with the Linear Continuum
$\mathbb{R}^n \sim \mathbb{R}$*; i.e., n-dimensional space is equinumerous with one dimensional space.*

***Proof*:**
Use *Proof by Mathematical Induction*. See Exercise 13b. ■

Sets Larger than the Continuum: Cardinality and the Power Set

Finite Cartesian products didn't increase the cardinality of $\mathbb{R}$; what about other operations? Intersections and set differences make sets smaller, so they won't be any help. Unions might – sets get enlarged here – but it turns out they don't increase cardinality, either: cardinality remains the same, just as for countably infinite sets. Surprisingly, infinite sets absorb smaller sets without affecting their size (cf. *Corollary* 2 to *Theorem* 5.1-10 and Exercise 20).

The only other operator we have to work with is the power-set operator. We saw that this greatly increased the size of finite sets: if $|S| = n$, then $|\mathcal{P}(S)| = 2^n > n$. Of course by now, your intuitions about cardinality have been sufficiently jumbled so that you probably don't trust *any* generalization from the finite to the infinite case. If so, that's good; you shouldn't just generalize. But in this case the generalization does go through. *Cantor's Theorem* says that every set is strictly less numerous than its power set. We'll look at this situation first via an example, which is important in its own right.

***PROPOSITION* 5.2-5: $\mathcal{P}(\mathbb{N})$ *is Uncountably Infinite/Equinumerous with* $\mathbb{R}$**
$\mathcal{P}(\mathbb{N}) \sim \mathbb{R}$

***Proof*:**
Real numbers in the set $(0, 1]$ can be uniquely represented in binary fraction notation by non-terminating sequences of 0s and 1s following a "binary point", where the n^{th} bit now has the place value of 2^{-n} instead of 10^{-n}. For example, $1/2 = .0111\cdots$ because $1/2 = 1/4 + 1/8 + 1/16 + \cdots$. All other reals in $(0, 1]$ can likewise be represented; all that differs from the usual decimal representation is the base being used.
Let B denote the associated set of all non-terminating binary sequences – just drop the prefixed binary point. This set is obviously equinumerous with $(0, 1]$ by what was just said and hence also with $\mathbb{R}$.

Now enlarge B by adjoining all "terminating" binary sequences, that is, all the binary sequences having finitely many 0s and 1s followed by all 0s. Let T denote this set of terminating binary sequences. T is essentially the set of all finite 0-1 sequences (lop off the ending 0s): it is countably infinite (see Exercise 5.1-28a). Let S stand for the set of all possible sequences of 0s and 1s, whether terminating or not. S is thus the disjoint union of the uncountable set B with the countable set T.
By *Corollary* 2 to *Theorem* 5.1-10, S is uncountable and of the same cardinality as B, the set of non-terminating sequences. But $B \sim \mathbb{R}$, so $S \sim \mathbb{R}$, too.
To finish the proof, it suffices to exhibit a one-to-one correspondence between S and the power set $\mathcal{P}(\mathbb{N})$: then $\mathcal{P}(\mathbb{N})$ and $\mathbb{R}$ are equinumerous.
Each sequence in S uniquely determines a subset P of positive integers, in the following way: $n \in P$ iff there is a 1 in the n^{th} term of the sequence (consider the sequences as starting with the zeroth term). This construction provides a one-to-one correspondence between the set of all binary sequences S and $\mathcal{P}(\mathbb{N})$.
Thus $S \sim \mathcal{P}(\mathbb{N})$, and therefore $\mathcal{P}(\mathbb{N}) \sim \mathbb{R}$. ∎

Since $\mathbb{R}$ is uncountable, $\mathcal{P}(\mathbb{N})$ is, too: $\mathbb{N} \prec \mathcal{P}(\mathbb{N})$. The cardinality of the real line $\mathbb{R}$ used to be denoted by the German letter c (standing for *continuum*), but in light of the last proposition and our earlier notation, it can also be symbolized by $2^{\aleph_0}$. Restating our set comparison in terms of the associated cardinal numbers, we have $\aleph_0 < 2^{\aleph_0}$.

Having seen that $S \prec \mathcal{P}(S)$ holds in the finite case and also in the case of a countably infinite set, we're ready to generalize. This is the content of *Cantor's Theorem*. Once we have this theorem, the population of the transfinite realm explodes.

THEOREM* 5.2 - 3: *Cantor's Theorem: Numerosity of Power Sets (***1892***)
$S \prec \mathcal{P}(S)$

***Proof*:**
Let S be any set. We will show that S is strictly less numerous than $\mathcal{P}(S)$ using a type of diagonalization procedure (see Exercise 21).
It's obvious that $S \preceq \mathcal{P}(S)$: match each element x of S with the singleton $\{x\}$ in $\mathcal{P}(S)$. This shows S is equinumerous with a subset of $\mathcal{P}(S)$.
We show that $S \not\sim \mathcal{P}(S)$ by showing that *any* attempted matching (not only the one just mentioned) misses some subset of S; i.e., it necessarily misses some element of $\mathcal{P}(S)$.
Suppose, then, that each x in S is matched up in some way with a subset $M_x \in \mathcal{P}(S)$.
Now consider $N = \{x \in S : x \notin M_x\}$.
N is obviously a subset of S. However, N is *not* matched with any $x \in S$, so our matching fails to be a one-to-one correspondence.
For suppose $N = M_a$ for some $a \in S$. Where does a lie with respect to N?
Well, $a \in N \leftrightarrow a \notin M_a$ according to the definition of N.
Substituting from our supposition, $a \in N \leftrightarrow a \notin N$, which is a contradiction.
Hence there is no a such that $N = M_a$.
Consequently $x \mapsto M_x$ is not a one-to-one correspondence.
Since M was a perfectly general assignment of subsets to elements, no matching can be found between S and $\mathcal{P}(S)$.
Thus $S \not\sim \mathcal{P}(S)$, and so $S \prec \mathcal{P}(S)$. ∎

Let's summarize where we are now with respect to infinite cardinalities and sets of numbers. We know that $\mathbb{N}$ is the smallest type of infinite set, and we know that many other sets are also countably infinite. In addition we know that $\mathbb{R}$ is strictly bigger than $\mathbb{N}$, and that all n-dimensional real spaces are the same size as $\mathbb{R}$. So far, though, we have no sets of larger cardinality than c. But now, — *ka-boom!! Cantor's Theorem* changes everything in one fell

swoop. The power set of any set is a still larger set. Thus, $\mathcal{P}(\mathbb{N})$ is bigger than $\mathbb{N}$, though it turns out to be the same size as $\mathbb{R}$. Not to worry: $\mathcal{P}(\mathbb{R})$ is strictly larger than $\mathbb{R}$. And $\mathcal{P}(\mathcal{P}(\mathbb{R}))$ is bigger yet, and ... and ... . A sequence of successive power sets gives us a list of sets of ever increasing size. Infinite cardinalities go on and on just like the finite ones do! In fact, the realm of transfinite sets is far richer and more bizarre than even this suggests; it is very different from the finite realm. Our discussion here touches only the beginning of what has been investigated by twentieth-century set theorists. You can maybe start to see why some mathematicians were questioning: is this still mathematics, or have we entered the domains of speculative philosophy and theology? Cantor, who realized that the transfinite realm had no largest number, also held that there was an absolute infinity, God with his divine attributes, an infinity that remained set apart from the actual infinities present in creation.

Intermediate Cardinalities and the Continuum Hypothesis

Our final topic will reveal some of the seemingly built-in ambiguity of the transfinite realm. We now know that sets continue to grow in size if we apply the power set operator. But do they grow successively or by leaps and bounds? In the finite realm, the power set of a set gives us something much larger than we started out with; there are many cardinal numbers between n and 2^n. Is this the same for the realm of transfinite sets? Or is the power set the next size set, with no intervening size? This is the most general formulation of the *Continuum Problem*. The problem initially arose in connection with its most familiar instance: Is $\mathbb{R}$ the next size infinite set after $\mathbb{N}$? Or are there sets of intermediate size? To put the question quite concretely: Is there a subset of $\mathbb{R}$ that is larger in size than $\mathbb{N}$ but smaller than $\mathbb{R}$? This question seems unambiguous and is easy to understand. So what's the answer?

It turns out that presently the answer is both yes and no. There are two obvious approaches to take to solving this problem. Since $\mathbb{N} \subseteq \mathbb{R}$, we could try to squeeze an intermediate size set S in between them: find S so that $\mathbb{N} \subset S \subset \mathbb{R}$ and $\mathbb{N} \prec S \prec \mathbb{R}$. Failing that, we could try instead to demonstrate that this is impossible (which, of course, is different from merely failing to find one). Mathematicians achieved no success following either approach. No sets were found of an intermediate size; every set tried was either countably infinite or else was just as big as $\mathbb{R}$. On the other hand, no one was able to construct a proof showing there couldn't be any such sets, either.

Cantor's intuition was that $2^{\aleph_0}$ is the immediate successor of $\aleph_0$. His conjecture to this effect became known as the *Continuum Hypothesis* (*CH*). His formulation of this hypothesis took different forms at different times but goes back to 1874, and more definitively to 1883.

CONJECTURE: *Continuum Hypothesis*
There are no sets with cardinality between $|\mathbb{N}| = \aleph_0$ *and* $|\mathbb{R}| = 2^{\aleph_0}$.

Cantor expended much time and effort trying to prove this result throughout his life. There were times when he was sure he had succeeded, only to discover a flaw in his reasoning. In the end, though, he failed. We now know that there was good reason for his failure: it *cannot be proved*, given the usual axioms of *Set Theory*. On the other hand, it also cannot be disproved.


Kurt Gödel

As mentioned above, Hilbert featured Cantor's *Continuum Problem* as the very first in his list of difficult problems for twentieth-century mathematicians to work on. This gave a great impetus to attempts to solve it. The first significant achievement regarding its solvability was due to the Austrian Kurt Gödel, already then one of the leading mathematical logicians of the twentieth century. In 1940 he proved that Cantor's *Continuum Hypothesis* was consistent with the rest of *Set Theory*.

THEOREM*: *Consistency of the Continuum Hypothesis (***Gödel 1940***)
The Continuum Hypothesis is consistent with the axioms of Set Theory.

Of course, being consistent with the axioms of *Set Theory* is a far cry from being a logical consequence of the axioms, so the *Continuum Hypothesis* was still awaiting proof. In 1963, however, the American mathematician Paul Cohen, using a new technique of mathematical logic developed by him for the occasion, showed that *CH* could not be proved as a theorem of *Set Theory* because its negation was also consistent with the axioms of *Set Theory*.

THEOREM*: *Independence of the Continuum Hypothesis (***Cohen 1963***)
The negation of the Continuum Hypothesis is consistent with the axioms of Set Theory.

The upshot of these results is that we are free to assume either *CH* or some form of its opposite without contradicting the rest of *Set Theory*. In other words, the truth of the *Continuum Hypothesis* cannot be decided on the basis of the usual axioms of *Set Theory*; it is an *undecidable* result, whose proof or disproof requires a new idea. Exploring various axioms strong enough to decide this issue has been an active area of set-theoretical research since the 1960s. Many set theorists believe the *Continuum Hypothesis* is probably false, but no one has proposed a widely accepted axiom to decide the matter. Some mathematicians take a more formalistic attitude, saying the hypothesis is neither true nor false in any absolute sense; like the paralllel postulate in geometry, its truth value depends on what system of mathematics you are developing. Given the familiar nature of both $\mathbb{N}$ and $\mathbb{R}$, however, this outlook is unsatisfying. Surely there either is or is not a set S between $\mathbb{N}$ and $\mathbb{R}$ of strictly intermediate size!

Delving into the *Continuum Hypothesis* and related matters are usually reserved for graduate level courses in *Set Theory* or *Mathematical Logic*; they are also a favorite topic of contemporary philosophers of mathematics. We have gone as deeply into this matter as we can in an introductory course. Our last pass through *Set Theory* in Section 5.3 will talk about its standard axiomatization and some related matters, drawing upon what we've already learned about the rather bizarre world of infinite sets.

EXERCISE SET 5.2

Some of the following exercises are non-trivial. Give them your best shot, but don't get discouraged if you run stuck. As we noted above, the material in this section moves us toward the outer limits of elementary Set Theory.

Problems 1 - 4: True or False
Are the following statements true or false? Explain your answer.

*1. Two-dimensional space, considered as an infinite set of points, is more numerous than one-dimensional space but less numerous than three-dimensional space.

*2. If S and T are uncountably infinite sets, then $S \sim T$.

3. $\mathcal{P}(\mathbb{N}) \prec \mathbb{R}$.

4. For a countable set S, $S \sim \mathcal{P}(S)$.

***Problems 5 - 8: Intervals Equinumerous with* (0,1)**
The following problems deal with finite intervals of real numbers.

*5. *Example* 1
   *a. Find a functional matching to show geometrically that $(0, 1) \sim (-1, 3)$ (see Example 1). Then explain why your matching is a one-to-one correspondence.
   b. Find a functional matching to show geometrically that $(0, 1) \sim (-1/2, 2/3)$ (see Example 1). Then explain why your matching is a one-to-one correspondence.

*6. *Equinumerous Unit Intervals*
Prove that $(0, 1] \sim (0, 1)$ in the following two ways.

a. Graph some (discontinuous) function that gives a one-to-one correspondence from $(0, 1]$ to $\mathbb{R}$ and find a formula that works for it. Then use the transitivity of equinumerosity to draw your final conclusion.

*b. Show that adjoining a single element to $(0, 1)$ doesn't change its cardinality. Either use *Corollary* 2 to *Theorem* 5.1-10 (but only if you've worked Problem 5.1-33b, proving that corollary), or else prove *Corollary* 2 for this particular case by using an argument similar to the proof of *Theorem* 5.1-10.

*7. What is wrong with the following proof, which purports to show that $(0, 1)$ is uncountable.
Start with the whole interval and choose the midpoint 1/2. Now take each of the two subintervals created by 1/2 as a cut point, and choose their midpoints as the next numbers to list: 1/4, 3/4. Continue this process: at stage k choose 2^{k-1} midpoints as new cut points to list in order. This procedure eventually lists all numbers of the form $m/2^n$. However, since 1/3, for example, is omitted from this list, the set $(0, 1)$ is uncountable.

*8. *Proposition* 2
The following outline two proofs of *Proposition* 2: *$(a, b) \sim (c, d)$ for any real numbers $a < b$, $c < d$.*

*a. First prove this result pictorially. Given two intervals of different finite lengths, place the shorter one above and parallel to the other one and then show how the two intervals can be matched up point for point to yield a one-to-one correspondence. Describe your process in words.

EC b. Then prove this result more formally by determining a formula that does such a matching. Hint: match x in (a, b) to the y in the same relative position in the interval (c, d).

Problems 9 - 14: Sets Equinumerous with $\mathbb{R}$

The following problems deal with sets equinumerous with $\mathbb{R}$.

9. *Proposition* 3
Show in the following way that the formula $y = \dfrac{x - .5}{x(x - 1)}$ given in the proof of *Proposition* 3 is a one-to-one correspondence between $(0, 1)$ and $\mathbb{R}$. Carefully graph the formula and explain how points in the unit interval are related to real numbers. Then tell why the graph exhibits a one-to-one correspondence between the two sets.

*10. *Intervals Equinumerous with $\mathbb{R}$*

*a. Show that $\mathbb{R} \sim (-1, 1)$ in the following way. Construct a unit circle centered about the point $(0, 1)$ and then consider just the lower semicircle minus its top "endpoints". Show how to project the points on this open semicircle onto the points on the x-axis. Then explain why this induces a one-to-one correspondence between $(-1, 1)$ and $\mathbb{R}$. Why does this mean in turn that $(0, 1) \sim \mathbb{R}$?

b. Modify the given formula of Problem 9 to show directly that $(-1, 1) \sim \mathbb{R}$. Explain why your formula is a one-to-one correspondence.

11. *Equinumerous Finite Intervals*
Show that any two intervals of finite length, whether open, closed, or half-open/half-closed are equinumerous with one another and with $\mathbb{R}$. Use whatever results are already known about these cases, including earlier problems.

12. *Equinumerous Infinite Intervals*
Show that any two intervals of infinite length, whether open or half-closed are equinumerous with one another and with $\mathbb{R}$. Use whatever results are already known about these cases.

13. *Theorem* 2

a. Show that the modified correspondence set up between $(0, 1] \times (0, 1]$ and $(0, 1]$ in the proof of *Theorem* 5.2-2 is a genuine one-to-one correspondence; i.e., show that each ordered pair of positive fractional numbers has a unique positive real number it is associated with, that no positive real numbers are left without a mate, and that given any positive fractional real number z there is a unique ordered pair (x, y) that it is related to.

b. Prove *Corollary* 2 to *Theorem* 2: $\mathbb{R}^n \sim \mathbb{R}$.

*14. *Irrational Numbers*
   a. Prove that the set of irrational numbers is uncountably infinite and is equinumerous with $\mathbb{R}$.
   b. Knowing how rational and irrational numbers are represented by certain sorts of infinite decimal expansions, show that between every two rational numbers there is an irrational number, and between every two irrational numbers there is a rational number. Do you find this paradoxical? Explain.

*15. Prove that if $S \sim \mathbb{R}$ and $T \sim \mathbb{R}$, then $S \cup T \sim \mathbb{R}$. Hint: first prove this result for the case of disjoint sets, using what you know about the size relationship between $\mathbb{R}$ and its intervals.

Problems 15 - 20: Uncountable Sets

*16. *Proposition 1*
   a. Prove *Proposition* 1a: *If S is uncountably infinite and $S \sim T$, then T is uncountably infinite.*
   b. Prove *Proposition* 1b: *If S is uncountably infinite and $S \subseteq T$, then T is uncountably infinite.*
   *c. Prove *Proposition* 1c: *If S is countably infinite and T is uncountably infinite, then $S \prec T$.*

*17. Is the full counterpart to *Proposition* 5.1-2 (namely, *If T is uncountably infinite, then S is uncountably infinite iff $S \sim T$*) true or false? If it is true, prove it. If it is false, give a counterexample.

18. Prove that the collection of all *co-infinite sets* in $\mathbb{N}$ (i.e., all subsets of $\mathbb{N}$ whose complement in $\mathbb{N}$ is infinite) is uncountably infinite. What size is it? Hint: first calculate the cardinality of all co-finite sets.

*19. Using a diagonalization argument, prove that the countably infinite Cartesian product of countably infinite sets $\prod_{i=1}^{\infty} S_i = \{(x_1, x_2, \dots) : x_i \in S_i\}$ is uncountable.

20. Show that if T is any infinite set and $S \prec T$, then $S \cup T \sim T$.

EC 21. *Cantor's Theorem and Diagonalization*
Analyze the proof of *Cantor's Theorem* for the special case when $S = \mathbb{N}$. Represent each subset M of $\mathbb{N}$ by an infinite sequence of 0s and 1s, putting a 1 in the n^{th} place if $n \in M$ and otherwise putting in a 0. Explain what choosing the subset N amounts to in this case. Then explain why the proof of *Cantor's Theorem* is considered a (generalized) diagonalization argument.

HINTS TO STARRED EXERCISES 5.2

1. [No hint.]

2. [No hint.]

5. a. Draw a graph of the function and use it to help you establish equinumerosity.

6. b. This result can be argued using the proof of Theorem 5.1-10, taking $T = (0, 1]$, $x_0 = 1$, and S to be a sequence $\{x_n\}$ of your choice in $(0, 1]$. Work through the proof for the specifics of this case and show that $(0, 1] \sim (0, 1)$

7. Review what it means to be uncountably infinite. Also note that if this sort of proof were valid, it would show that the rationals in the interval $(0, 1]$ are uncountably infinite (why is this?).

8. a. Begin by connecting the corresponding interval endpoints by two line segments. Continue these segments until they meet. Use this intersection point to match up the intervals.
   b. This is EC, but it can be done by anyone familiar with the coordinate geometry of straight lines.

10. The hints are already given in the problem itself.

14. a. See the end of Section 5.1 for ideas.
    b. Rationals are repeating decimals.

15. Use the fact that all sorts of finite intervals are equinumerous with $\mathbb{R}$.

16. c. Use Theorem 5.1-9 along with the relevant definitions of the properties and relations involved.

17. Cantor's Theorem is relevant here.

19. Review how the uncountability of $\mathbb{R}$ was proved.

5.3 Formal Set Theory & the Halting Problem

Cautious mathematicians and philosophers have always thought that if you mess with infinity long enough you'll get into some sort of trouble. Cantor, however, judged earlier thinkers' scruples regarding the actual infinite misplaced and so confidently developed a theory of transfinite *Set Theory*. The linchpin of his theory was the notion of a one-to-one correspondence, used for comparing cardinalities of sets.

His theory seemed on pretty solid ground, but Cantor came to realize that besides transfinite quantities, which were still extendible (shades of Aristotle's potential infinity?), there was an absolute infinity, which could not be further extended: it was as large as it could get. In the mid-1880s he speculated that such an infinity must be connected with God, but in addition to his theological stance on this, he thought it could not be treated mathematically in the same way as the other infinities or difficulties would arise.

Mathematicians largely ignored Cantor's theological ruminations on infinity, but they too would soon discover that it was important to distinguish between the infinity of everything and the infinities associated with familiar mathematical objects and constructions. Right around the turn of the twentieth century, mathematicians and logicians learned publicly that informal *Set Theory* harbored within it certain paradoxes (a polite term for unexpected logical contradictions) if such an extreme notion was permitted. These paradoxes could be pragmatically avoided by steering clear of absolute infinity (just forbid using it as if it were an ordinary infinity), but a more intellectually honest approach would be to systematically and consciously fence out the paradoxes without using such *ad hoc* stratagems. This required first axiomatizing *Set Theory* and then demonstrating that the sets which give rise to the paradoxes no longer exist according to the theory and so cannot create theoretical havoc.

Paradoxes and the Need for Axiomatic Set Theory

This concern for logical consistency is often cited as the key motivation for formalizing *Set Theory*. This is certainly crucial from our later vantage point, but historically speaking, it was only one impetus leading into the formulation of axiomatic *Set Theory*, and maybe not the most important one at the time.* Any mathematical theory with a rich enough body of results simply begs to be deductively organized. Once sufficiently many results are known, mathematicians naturally try to put them into a single coherent system. Complex propositions are demonstrated from simpler ones, starting from a limited number of fundamental principles chosen as the deductive basis of the whole theory. Naturally, if a well-developed theory also has problems that need straightening out, that will give an even stronger impulse toward theoretical organization.

Evidently this is what happened in *Set Theory*. In 1908, Ernst Zermelo published the first axiomatization of *Set Theory*, in two papers. Zermelo's most immediate goal was to substantiate a rather surprising result he had proved four years earlier, namely, that sets of any size whatsoever can be well-ordered.** His initial proof and the result itself had been vigorously challenged by various mathematicians in the interim, so Zermelo offered a second proof, identifying in detail the various assumptions that were needed for its execution. In the process, he began to axiomatize *Set Theory* as a whole, which he considered foundational to mathematics. This was done in a way that turned out also

Ernst Zermelo

* This issue is discussed in detail in chapter 3 of Gregory H. Moore's book, ***Zermelo's Axiom of Choice: Its Origins, Development, and Influence***.

** A set is *well-ordered* by an order relation iff every non-empty subset has a first element according to that order. A countable set is ordered by the sequence that lists it, but could an uncountable set be well-ordered?

to be fruitful for handling the paradoxes, which was an explicit concern of his second paper. Zermelo was aware of the paradoxes and showed that they disappeared under his axiomatization. At that time the real cause of the paradoxes of infinity was still being debated; some thought the contradictions might be avoided by reforming logic rather than mathematics, but Zermelo thought it should be done by restricting the way sets were postulated.

We will follow Zermelo's approach here. Logic has now been developed in a rather definitive and consistent way, yet, as we will show next, informal *Set Theory* readily generates contradictions. We are in need of an axiomatization to systematize *Set Theory*, but we need it just as much to exclude the contradictions. Before we begin looking at the broad contours of such an axiomatization, then, we will show just what goes wrong with informal *Set Theory*.

Russell's Paradox and Inconsistent Sets

In 1901 Bertrand Russell was pondering *Cantor's Theorem* (*Theorem* 5.2-3), which uses power sets to establish the existence of sets of ever increasing size. Thinking about the result, he began to wonder what would happen if the universe $\mathcal{U}$ of all possible sets were used as the base set. Would $\mathcal{P}(\mathcal{U})$ be larger than $\mathcal{U}$? The theorem says yes, but if $\mathcal{U}$ really is the collection of *all* sets, wouldn't *it* have to be the largest set? This was baffling, so Russell explored the matter further.

Carefully pouring over the proof to discover the flaw either in his or Cantor's reasoning, Russell was led to consider the class of all sets that are not elements of themselves. (Such sets are, as a matter of fact, the usual case: a set does not normally contain itself as an *element*, though certainly the collection $\mathcal{U}$ of all sets would have to contain itself.) We'll denote this class of normal sets by $\mathcal{N} = \{X : X \notin X\}$. Using this notation, Russell's next question can be formulated as follows: Is $\mathcal{N}$ itself a normal set? That is, does $\mathcal{N}$ contain itself as an element or not? He was led to conclude, paradoxically enough, that if it does, then it doesn't; while if it doesn't, then it does. This is now known as *Russell's Paradox*.*

The argument goes as follows. First suppose $\mathcal{N} \in \mathcal{N}$. This makes $\mathcal{N}$ a normal set, since only normal sets belong to $\mathcal{N}$. But then $\mathcal{N} \notin \mathcal{N}$: that's what it means to be normal. On the other hand, suppose that $\mathcal{N} \notin \mathcal{N}$. Then $\mathcal{N}$ is a normal set and so must belong to the set $\mathcal{N}$, which consists of all such sets. Thus $\mathcal{N} \in \mathcal{N}$. Combining these conclusions, we end up with $\mathcal{N} \in \mathcal{N} \leftrightarrow \mathcal{N} \notin \mathcal{N}$, a contradiction of the form $\mathbf{P} \leftrightarrow \neg\mathbf{P}$.


Bertrand Russell

Try as he might, Russell could not explain the paradox away. The above reasoning is absolutely impeccable, as he was finally forced to concede; the contradictory conclusion follows. The problem, then, seems to lie not with the logic of the argument but with the mathematics.** *Set Theory* itself must be the source of the contradiction. To purge *Set Theory* of this blight, set formation will somehow need to be restricted so that such "inconsistent sets," as Cantor had called them, cannot be generated. This was done by Zermelo in axiomatizing *Set Theory*. In particular, the main axiom that rescues us from the paradoxes, as we will see shortly, is the *Axiom of Separation*. Before considering this axiom and its consequences, we will pause momentarily to take up a few preliminary matters.

---

* The more popular *Barber Paradox* is similar in structure to *Russell's Paradox*. See Exercise 1.

** Russell did not draw this conclusion. He still believed that logic, which for him included a theory of classes, was at fault. He therefore developed a theory of logical types to handle the problem. This follows a different route than the one we do here.

Syntax and Semantics of Set Theory

The formal syntax and semantics underlying *Set Theory* is merely that of *Predicate Logic*. The only thing peculiar to the grammar of *Set Theory* is its specialized symbols. We need symbols for particular sets, such as $\emptyset$, symbols for operations, such as $\cap$, and symbols for relations, such as $\subseteq$. Many of these symbols can be introduced by means of definitions, as we have already done. However, we need to take the membership relation "$\in$" as primitive. The sentence "$x \in S$" asserts a relation whose intended meaning is intuitively clear but which cannot be defined in simpler terms.

Informal *Set Theory* often distinguishes between individuals and sets by using lower-case and upper-case letters. We could continue this practice on a formal level if we wanted. Some axiomatizations of *Set Theory* do permit both individuals and sets in their universe of discourse.* However, most axiomatizations today assume that sets are all there are (for the purpose of developing *Set Theory*, of course). We will follow this latter course. Letters, whether upper or lower case, will thus represent one type of object only: sets. We can continue to use upper case letters to emphasize a set's role as containing elements and lower case letters when we want to emphasize its role as an element of another set, but considered individually, all variables will denote sets. Their elements, if they have any, will themselves be sets, without exception. This may strike you as highly counterintuitive (weird), but it simplifies the theory by having objects of only one sort to deal with. Such an approach may also seem overly restrictive, but in fact it is adequate. It can be used, for example, to develop/model familiar theories, such as that of *Peano Arithmetic*, inside *Set Theory*.

The Axiom of Extensionality

Predicate Logic already has a fixed, standard notion of identity. Given the laws governing identity, it is clear that $S = T$ logically implies $(\forall x)(x \in S \leftrightarrow x \in T)$: merely apply *Sub* to the last clause of the tautology $x \in S \leftrightarrow x \in S$ and universally generalize. Thus, identical sets must contain exactly the same elements. However, we are unable to turn this claim around and say that sets having the same elements must be identical; that is, *set equality for co-extensive sets doesn't follow from logic alone*. To consider sets as being determined purely by their extensions (set membership), we need to postulate such a result as an axiom.

AXIOM* 5.3 - 1: *Axiom of Extensionality
$\forall x(x \in S \leftrightarrow x \in T) \rightarrow S = T.$

This axiom, combined with the other direction from logic, easily proves the criteria for equal sets that we earlier stated as *Definition* 4.1-1. We will state it here as a proposition.

PROPOSITION* 5.3 - 1: *Equal Sets
$S = T \leftrightarrow \forall x(x \in S \leftrightarrow x \in T)$

***Proof*:**
This follows immediately, in the way noted above. ■

The Axiom of Separation and Some Consequences

We are now ready to extricate ourselves from the predicament we landed in with the discovery of paradoxical sets. Let's take another look at *Russell's Paradox*. We concluded that the argument there was valid and that *Set Theory* needed to change to avoid the contradiction.

* This was actually Zermelo's 1908 approach. It is also the approach of P. Suppes in ***Axiomatic Set Theory***.

But what is there to change? Not very much of *Set Theory* is even involved in the offending argument! About all we assumed in constructing $\mathcal{N}$ is what we might call *Cantor's Comprehension Principle*, which says that all elements having a given property compose a set. This is a principle that lies at the very heart of informal *Set Theory*. Is it really illegitimate??

Cantor had conceived of sets as arising any time a collection of definite objects is present in thought or actuality. Thus, any well-defined membership requirement $P(x)$ generates a set containing precisely those elements satisfying it; every set $\{x : P(x)\}$, defined by means of set descriptor notation, is a bona fide set from an informal point of view. Is this false? Regrettably, it seems that it must be, for such a naive attitude leads to a contradiction like *Russell's Paradox*.

A natural way to get around *Russell's Paradox* is to reject *Cantor's Comprehension Principle*. Thus, we can no longer be absolutely certain in using set descriptor notation that what we've constructed is really a *set*. In particular, we want to reject having to accept things such as $\mathcal{N}$, the collection of all normal sets, and the full universe of sets $\mathcal{U}$ as being genuine sets.

So then, which collections defined by means of a membership description really are sets, and which ones are not? Must we give up presenting sets by means of property descriptions of their elements? Are we left only with sets whose elements can be listed? That would completely gut *Set Theory*, so that can't be the right solution.

It seems reasonable to suspect that *Russell's Paradox* arises not because a property is used to pick out the members of a set, but because the size of the universe in which the property is allowed to operate is too large.* A property ought not to function in the unrestricted context of the entire universe of possible objects, but within a more circumscribed domain. Given a set *U already known to exist*, a property can then be legitimately used to *separate off within this set* all those elements having the given property. Surely the result of such an operation is itself a set; surely each subclass of an existing set U is still a set. Here U functions as a restricted universe of discourse for S.

Zermelo's *Axiom of Separation* codifies this approach. His axiom schema provides the formal justification for using restricted set-descriptor notation. Given a set U, known to exist, the collection $S = \{x \in U : P(x)\}$ is also a set; in fact, it is well-defined. Uniqueness can be proved here and elsewhere using the *Axiom of Extensionality*, but we will assume this without further ado, instead of belaboring the fact by proving each time that the sets being considered are unique.

AXIOM* 5.3 - 2: *Axiom of Separation
$\forall U \exists S(S = \{x \in U : P(x)\})$ **

A set S will exist, according to the *Axiom of Separation*, whenever there is an existing superset U of S and a defining property P holding for all and only those members inside U that belong to S. You can think of S as being *separated out* from the rest of the universe of discourse U by means of a criterion $P(x)$.

A defining property $P(x)$ will therefore no longer be allowed to operate in an unrestricted fashion. In fact, we can (re)prove the following constructive version of *Russell's Paradox*, which demonstrates that the class of all normal sets is *not* a set/does not exist.

THEOREM* 5.3 - 1: *Russell's Paradox (Constructive Version)
$\neg \exists S \forall x(x \in S \leftrightarrow x \notin x)$

***Proof*:**
This is proved by *Contradiction* (*NI*).

---

* This is Zermelo's intuition, but it also lies behind an alternative axiomatization of *Set Theory* due to von Neumann, Bernays, and Gödel, who distinguish between large classes and sets.

** We will formulate this axiom and later axioms somewhat informally using set descriptor notation, though this notation abbreviates a still more basic *Set Theory* statement (see the first footnote in Section 4.1).

Suppose $\exists S \forall x(x \in S \leftrightarrow x \notin x)$, and let $\mathcal{N}$ denote such a set.
Thus, $\forall x(x \in \mathcal{N} \leftrightarrow x \notin x)$.
Since $\mathcal{N}$ is (assumed to be) a set, we can instantiate this universal sentence to $\mathcal{N}$:
$\mathcal{N} \in \mathcal{N} \leftrightarrow \mathcal{N} \notin \mathcal{N}$.
But this is a contradiction.
Thus, we conclude that $\neg\exists S \forall x(x \in S \leftrightarrow x \notin x)$. ■

The *Axiom of Separation* plays a central role in the axiomatic development of *Set Theory*. It or one of its consequences comes into play every time we want to demonstrate the existence of a set that is not automatically guaranteed by an axiom of *Set Theory*. We'll illustrate this by proving that the intersection of two sets exists (is a set). Showing the existence of set differences will be left as an exercise. As mentioned above, these sets are uniquely defined.

PROPOSITION* 5.3 - 2: *Existence of Intersections
$\exists I(I = \{x \in S : x \in T\})$

***Proof*:**
In the *Axiom of Separation*, take S in place of U and '$x \in T$' in place of '$P(x)$'. ■

Given this result, the intersection $S \cap T$ of two sets S and T can be defined as follows (cf. *Definition* 4.1-7): $S \cap T = \{x \in S : x \in T\}$. Other set notation can be similarly defined; we will not explicitly do this for each new case, since we introduced them all earlier.

PROPOSITION* 5.3 - 3: *Existence of Set Differences
$\exists D(D = \{x \in S : x \notin T\})$; i.e., $S - T$ is a set.

***Proof*:**
See Exercise 10. ■

The Empty Set

As crucial as the *Axiom of Separation* is, it only helps us generate new sets, such as intersections or set differences, given the *pre-existence* of larger old ones. In this respect it functions much like the induction step does in *Peano Arithmetic*: it provides a procedure that can be applied to what you already have, but it requires an appropriate initialization to get the whole process started. In order for this axiom to actually give us something, we need to begin with an existing set that can function as a superset. This isn't furnished by the *Axiom of Separation* or the *Axiom of Extensionality*; they are both universal, not existential, statements. The existence of a set or collection of sets must be postulated by some existential axiom in order to get going.

Once we postulate the existence of some universe of discourse U, we can then prove that subsets of U exist via the *Axiom of Separation*. In particular, we can prove that the *empty set* exists, as follows. From the *Axiom of Separation*, we could immediately conclude that $\{x \in U : x \neq x\}$ is a set, and we could name this set $\emptyset$. This very nearly agrees with our definition of $\emptyset$ in *Definition* 4.1-4, except there we used the now-outlawed unrestricted set-descriptor notation. Even with our new approach, however, the *empty set* membership criterion follows: $(\forall x)(x \notin \emptyset)$ (see Exercise 7). This approach generates empty sets inside each set U, but it is easy to show that all such sets must be the same: the *empty set* is unique (see Exercise 8).

Unfortunately, we have no way to say "there exists a universal set U" in the language of *Set Theory*. Given the semantics of *Set Theory*, all of our variables refer to sets, but we can't just say $\exists U$; without a defining property, we have a syntactically incomplete sentence. And we certainly can't postulate $(\exists U)(\forall x)(x \in U)$: such a set leads us right back into *Russell's Paradox* (see Exercise 3). What we need to say, therefore, is that a *particular set* U exists.

Let's be really stingy about this so we won't get large problematic sets; let's only claim that the *empty set* exists. We already know that if anything exists, then $\emptyset$ must, so this choice is the bare minimum. Of course, whether we can do anything important if we're this thrifty is another issue, but let's see what happens.

We will take the following, then, as our third axiom, which is the membership criterion for the *empty set*, stated positively. Uniqueness follows, as noted above, which we will prove here to illustrate the process and because it involves a slightly unusual argument.

AXIOM* 5.3 - 3: *Empty Set Axiom
$\exists E \forall x (x \notin E)$

PROPOSITION* 5.3 - 4: *Unique Existence of the Empty Set
$\exists! E \forall x (x \notin E)$

***Proof*:**
Existence of such a set follows from the *Empty Set Axiom*.
To show that such a set is unique, let E_1 and E_2 denote two such sets and consider any x.
Then $x \notin E_1$ and $x \notin E_2$.
But this weakens to $x \notin E_1 \leftrightarrow x \notin E_2$ (consult an extended truth table, if necessary).
By the biconditional counterpart to *Conpsn*, $x \in E_1 \leftrightarrow x \in E_2$.
Generalizing and applying *Proposition* 1, $E_1 = E_2$.
So the *empty set* E postulated by *Axiom* 3 is unique. ■

Since the *empty set* exists and is unique, we can introduce the standard symbol $\emptyset$ for it. This was done earlier as *Definition* 4.1-4, which assumed an unrestricted universe; we now have (via *Axiom* 3) that this is a set and is unique.

Finite Unions

We are now in a position to assert that sets exist. Mind you, we don't have much of an arsenal to back this claim up (only $\emptyset$), but it's a start. Given the existence of $\emptyset$, we could take it as our reference set U and generate other sets by means of the *Axiom of Separation* or by taking intersections or set differences. It should be pretty obvious, though, that this won't generate anything new. The sole subset of $\emptyset$ is $\emptyset$ itself, so nothing new can be separated out of $\emptyset$. Taking the intersection or set difference of $\emptyset$ with itself also yields $\emptyset$.

It seems, then, that postulating the existence of $\emptyset$ wasn't very fruitful. We've run into a dead end. Or have we? We might reasonably blame the ineffectiveness of $\emptyset$ on its size; it's as meager a set as you can find. But an alternative reason for its unproductiveness can also be assigned; namely, the equipment available for *working on* $\emptyset$. Our construction tools to this point – separating subsets, intersecting sets, and taking set differences – always move us *downward into a given set* instead of *upward and out from that set* into the rest of the universe of sets, whatever that might be.

Here we really see the restrictive nature of the *Axiom of Separation*. It nicely helps us avoid contradictions, such as *Russell's Paradox*, but it also puts severe limitations on what can be certified as a bona fide set. In order to proceed further in *Set Theory*, we will have to adopt one or more axioms that give us a way to move outward into a bigger universe. This is needed regardless of the size of the *empty set*. So we'll first stop to address this problem before reassessing the possible need for a bigger initial set.

The operation of taking the union of two sets S and T goes beyond the given sets and thus requires a new axiom. It can't simply be separated out of an existing set like intersection and set difference. We will assume the following as our axiom.

AXIOM* 5.3 - 4: *Finite Union Axiom
$\exists U(U = \{x : x \in S \vee x \in T\})$; i.e., $U = S \cup T$ is a set.

From this axiom, we can obtain the existence of the union of a finite collection of sets via recursion (hence the axiom's name). To get unions of arbitrary collections, however, a stronger assertion is needed. We'll turn to this next.

Total Intersections and Unions

The axioms adopted to this point justify the operations of intersection, set difference, and union. Since finite intersections and unions are gotten by repeating ordinary intersection and union so many times, they are also legitimated by what we have done. Furthermore, the total intersection of a given non-empty collection of sets can be shown to exist: it's inside each member of the collection and so exists by the *Axiom of Separation* (see Exercise 12).

Total unions of arbitrary collections of sets cannot be shown to exist given the machinery at our disposal so far. We must therefore adopt another axiom asserting their existence in order to make them available.

AXIOM* 5.3 - 5: *Total Union Axiom
$\forall \mathcal{C}\, \exists U(U = \{x : \exists S[S \in \mathcal{C} \wedge x \in S]\})$; i.e., $U = \bigcup_{S \in \mathcal{C}} S$ is a set.

Given this axiom, the definition and notation given earlier for this notion (*Definition* 4.2-2) is properly supported. Also, once this axiom is accepted, along with some others, the *Finite Union Axiom* can be proved as a theorem (see Exercise 38).

Existence of Sets of Sets: Pairs

Let's pause for a brief progress report on the existence of particular sets: the *Finite Union Axiom* and the *Total Union Axiom* help us not one whit if all we have is the *empty set* to start with, for all they give back is $\emptyset$. Our inventory thus remains the same: just $\emptyset$. Stay tuned, though; we may get a breakthrough from another quarter soon.

The above axioms justify the formation of subsets, intersections, set differences, and unions of sets, but they do not legitimate sets whose *elements* are the given sets. Nothing so far lets us generate *collections of sets*. You might be ready to say that such a collection of sets ought to exist if all of its elements already do. However, in the wake of *Russell's Paradox*, we have to be careful here: the existence of all the elements of a class is no guarantee that the entire class is a set. The class of all sets cannot exist (isn't a bona fide set), for instance. So the existence of sets of sets must be carefully postulated by means of separate axioms. We have no a priori guarantee that these sets will not lead to contradictory results, but we can at least try to make sure that the known contradictions are avoided.

In Section 4.2 we argued intuitively for the legitimacy of familiar sets of collections; here we will introduce axioms that stipulate the existence of certain sorts of conservative collections. The first axiom of this type enables us to pair up any two sets to form a new one. How much trouble could this cause?

AXIOM* 5.3 - 6: *Pairing Axiom
$\forall x \forall y \exists P(P = \{z : z = x \vee z = y\})$; i.e., $P = \{x, y\}$ is a set.

Uniqueness of pairs follows in the usual way (see Exercise 14). With this axiom, we can obtain doubletons; singletons are then defined as special doubletons.

DEFINITION* 5.3 - 1: *Doubletons
$\{x, y\} = \{z : z = x \vee z = y\}$

DEFINITION* 5.3 - 2: *Singletons
$\{x\} = \{x, x\}$

Singletons are special doubletons, so results shown for doubletons also apply to them. This takes some getting used to, but going from doubletons to singletons is the standard approach in formal set theory. For a different approach to singletons and doubletons, see Exercises 22 – 27.

Sets with three, four, or any finite number of set-members can now be defined in a standard fashion (see Exercises 17 – 21) and given appropriate names. Whether or not collections of these sizes exist, of course, depends on how many distinct sets are available to collect together in the first place.

So let's take stock of our supply of sets once again. We have $\emptyset$ more or less by fiat, and so far nothing else. Using the *Pairing Axiom*, though, we can now generate $\{\emptyset\} = \{\emptyset, \emptyset\}$. This is something new! It's a set with one element, the set $\emptyset$, while the *empty set* has no elements.* So now we have two sets: $\emptyset$ and $\{\emptyset\}$. With these two sets, we can form the new singleton $\{\{\emptyset\}\}$, and we can also pair them up to get $\{\emptyset, \{\emptyset\}\}$, which is also new. Continuing in this way, we can generate sets of increasing size and complexity. It's not clear yet whether the universe of sets so obtained is rich enough to do anything worthwhile from a mathematical point of view (it nearly is), but we can already see that the *empty set* coupled with the *Pairing Axiom* is far more fertile than we might have expected.

The Power Set Axiom

The *Axiom of Separation* allows us to form subsets, but it does not justify collecting all the subsets of a given set S together to form the power set $\mathcal{P}(S)$. Pairs of subsets can be formed by means of the *Pairing Axiom*, but this does not always generate the entire power set (only the finite subsets). To guarantee the existence of the full power set in general, a new axiom is required.

AXIOM* 5.3 - 7: *Power Set Axiom
$\forall S \exists P(P = \{R : R \subseteq S\})$; i.e., $P = \mathcal{P}(S)$ is a set.

Based on this axiom, we can define the power set $\mathcal{P}(S)$ as we did above (*Definition* 4.2-5). The *Power Set Axiom* also provides a rigorous foundation for the results we stated earlier about power sets (see Section 4.2)

Given any set S, $\mathcal{P}(S)$ contains all the subsets of S as elements. As we already know from *Cantor's Theorem*, the power set is larger than the original set. Does this increase the collection of sets we can now prove to exist? We already have $\emptyset$, $\{\emptyset\}$, $\{\emptyset, \{\emptyset\}\}$, and so on, obtained from the *empty set* by means of the pairing operator. The power set operator applied to $\emptyset$ also gives us $\mathcal{P}(\emptyset) = \{\emptyset\}$ and then $\mathcal{P}(\mathcal{P}(\emptyset)) = \{\emptyset, \{\emptyset\}\}$. Nothing new arises here, though, because we have begun with $\emptyset$. The pairing set operator (in conjunction with taking unions) works slower than the power set operator to generate successive power sets, but it can eventually construct them (see Exercise 28). This is definitely not true in general, however, if infinite sets exist and are used as a base set; the power set operator is a very strong operator, yielding sets that cannot be gotten in any other way. We cannot construct $\mathcal{P}(\mathbb{N})$ without the *Power Set Axiom*, for example; using pairing and then unions on the subsets gotten by the *Axiom of Separation* will only generate a finite portion of $\mathcal{P}(\mathbb{N})$ (see Exercises 29-30).

* In general, a singleton $\{x\}$ is distinct from its element x, but this cannot be proved without an additional more technical axiom, known as the *Axiom of Foundation*.

The Axiom of Foundation (Optional)

We noted above that sets formed from the *empty set* $\emptyset$ by means of the pairing operation are new sets: $\emptyset \neq \{\emptyset\} \neq \{\{\emptyset\}\}$ and so on. Intuitively, this would seem to be true for all sets, that $S \neq \{S\}$. And, in fact, if S has more than one element, this result can be proved (see Exercise 36). However, our present list of axioms do not universally rule out $S = \{S\}$; nothing prohibits a single element set S from containing just itself. That is, we cannot rule out the possibility that $S \in S$ for singletons.

A similar thing holds true for any size non-empty set: we have no way to rule out $S \in S$. Such sets seem abnormal and may even be thought to be partly responsible for *Russell's Paradox*, but they have not been banned from our universe of sets at this point. If we want to exclude them, we need to do it by means of an additional axiom.

The occurrence of a set S containing itself is certainly bizarre from a constructive point of view. For if we consider sets as being formed by the process of gathering their elements together in some way, whether through listing them or through separating them from others inside a given superset, then sets can't be considered present prior to their formation. Yet this is just what is required if they are to contain themselves as an element.

Taking the point of view that the elements of a set must at least be present prior to the formation of the set itself (in order to avoid such self-reference), we can see that other situations must also be guarded against. For elements of a set that are themselves sets, *their* elements must be present prior to their formation and hence also to the given set's formation. That is, we must not allow the following situation to occur, either: $S \in T \wedge T \in S$, abbreviated by $S \in T \in S$. For then S must be present prior to its formation. Nor can we allow $S \in T \in R \in S$, etc. It should be clear that we must avoid any such closed cycle of whatever length: S cannot be an interior member at any level inside itself.

If membership cycles $S \in T_n \in T_{n-1} \in \cdots \in T_1 \in S$ were permitted, the following irregularity would occur. Passing ever further inward along this chain of sets, proceeding from each set to one located inside it, we could continue without stopping. The set S contains the set T_1, which contains T_2, which contains $\cdots T_n$, which contains S, which contains T_1, and so on. Allowing such a cycle means we can never get to the "bottom" of this descending membership chain. Such a cycle would have no "foundation;" it would involve us in an infinite regress.

A descending chain of sets might also occur without repetition or circularity, generated not by a cycle but by a sequence of distinct sets. This is no less irregular from our constructive perspective than the former case. An infinite chain such as $\cdots T_n \in T_{n-1} \in \cdots \in T_1 \in S$, in which all T_i are distinct, is as strange as one generated by a finite cycle. It, too, has no ultimate foundation.

The *Axiom of Foundation* rules out all of these bizarre possibilities. Infinite descending membership chains of whatever sort are forbidden.* Starting with a given set, it ought to be possible to move inside it via the membership relation to reach rock bottom after only a finite number of moves. If this cannot be done, then the set construction cannot be started or completed.

Saying in the language of *Set Theory* that each set must have a foundation is rather technical and required some ingenuity to think up in the first place. Since we won't be making any use of it later in the text, we will omit the precise formulation. The *Axiom of Foundation*, also called the *Axiom of Regularity*, was first proposed by John von Neumann in 1929 and put into its present form the following year by Ernst Zermelo.

---

* This is naturally different from saying that there cannot be an infinite *ascending* membership chain $T_1 \in T_2 \in T_3 \in \cdots \in T_n \cdots$ in S, where each $T_n \in S$. See Exercise 37.

The Axiom of Infinity (Optional)

None of *Set Theory's* axioms thus far guarantee the existence of infinite sets. Starting with $\emptyset$ and applying the power set operator, we can get sets of arbitrarily large finite size, but unless we have more than the *empty set* to begin with, we can't generate an infinite set.

It is possible to model all of the natural numbers inside set theory based on the above axioms, using the definitions $0 = \emptyset$ and $\mathcal{S}(n) = n \cup \{n\}$ (see Exercises 41-48), but without adopting another axiom, there is no way to collect all of these sets together into an infinite set $\mathbb{N}$. What we have so far, then, is insufficient for mathematical purposes. To get sets like $\mathbb{N}$, we need an *Axiom of Infinity*. That such a result cannot be proved and must be postulated was first recognized by Zermelo.

Asserting the existence of an infinite set can be done in a way that relates to the way natural numbers are modeled inside *Set Theory*. This axiom basically affirms the existence of inductive sets that contain all individual natural numbers. Note that this axiom, like the *Empty Set Axiom*, is a simple existential sentence, and thus gives us the unconditional existence of another set, a starting point for generating sets that are infinite in size by means of set operations. Since the axiom's formulation involves successor sets, we will first define this notion for all sets whatsoever, not just for numbers. The unique existence of such sets is immediate given what we already have.

DEFINITION* 5.3 - 3: *Successors
$\mathcal{S}(X) = X \cup \{X\}$

AXIOM* 5.3 - 8: *Axiom of Infinity: Existence of Inductive Sets
$\exists I(\emptyset \in I \wedge \forall X(X \in I \rightarrow \mathcal{S}(X) \in I))$

Such an inductive set need not be unique (many sets I may contain $\emptyset$ and be closed under the successor operator), but by using this axiom it is possible to define $\mathbb{N}$ as the intersection of all such sets. Again, since we will not be developing set-theoretic *Peano Arithmetic* any further, we will omit the details (see Exercise 41).

Other Axioms for Set Theory

We now have nine axioms for *Set Theory*. Not all of them are independent of one another, however. For example, the *Finite Union Axiom* is redundant, given the *Total Union Axiom* and the *Pairing Axiom* (see Exercise 38). Also, the existence of the *empty set* $\emptyset$ follows if any set whatsoever exists: separate it out of any such set via the property $x \neq x$ (see Exercise 40). Thus, by adopting the *Axiom of Infinity*, we can drop the *Empty Set Axiom* without losing a thing. In addition, the *Pairing Axiom* is superfluous given the *Power Set Axiom* and a few others (see Exercise 39). So it is clear that we can do with several fewer axioms than we have taken.

On the other hand, there are still some results in *Set Theory* that cannot be proved with the axiomatic basis we have thus far. The proposition that got Zermelo started on axiomatizing *Set Theory* in the first place, the *Well-Ordering Theorem*, is a deeper result that requires the *Axiom of Choice* for its proof (see the last part of Section 5.1). This axiom and its many striking consequences are treated in more advanced texts on *Set Theory*.

The *Axiom of Choice* is sometimes avoided because of the controversial nature of some of its consequences, but there is one final axiom whose legitimacy is taken to be on a par with that of the other axioms. This is the *Axiom of Replacement*, first introduced by Abraham Fraenkel in 1922. The basic idea of the *Axiom of Replacement* is that given a set U and a partial function f defined on U (i.e., $f(x)$ is either undefined or uniquely defined for each input $x \in U$), the range $R = f(U) = \{f(x) : x \in U\}$ is also a set. Because functions can be used to compare the sizes of sets (one-to-one correspondences are functions of a specialized sort), this

axiom helps to distinguish between classes that are too big to be sets and those that are bona fide sets. Given this axiom, the *Axiom of Separation* can be proved, and so it, too, may be omitted from our list of axioms without any loss of deductive power.

Taken altogether, these axioms form the deductive basis for what is known as *Zermelo-Fraenkel Set Theory* (ZF); if the *Axiom of Choice* is included, the theory is known as ZFC. Other axioms have also been proposed in advanced set theoretical research. These mainly have to do with the existence of large cardinals. We will not go into this topic here. It arises in the context of developing axiomatic *Set Theory's* metatheory, which considers the nature of possible models for *Set Theory* along with the logical relationships holding between various axioms and propositions (consistency and independence results). Consequences of these axioms for some other areas of mathematics have also been explored, but this goes far beyond the scope of an elementary discrete mathematics text.

Our discussion in Chapters 4 and 5 provides a good introduction to *Set Theory*, but there is still a vast expanse remaining to be explored. There are several excellent books on the topic, some of them a bit older. Paul R. Halmos' book *Naive Set Theory* is a good informal introduction to the field. Patrick Suppes' *Axiomatic Set Theory* presents a formal development on a level that can be understood by someone with a background in introductory logic. The approach of Robert Stoll's *Set Theory and Logic* and of Herbert Enderton's *Elements of Set Theory* lies somewhere between that of Halmos and Suppes. Azriel Levy's *Basic Set Theory* is a careful, formal development on a more advanced level. Abraham Fraenkel's *Set Theory and Logic* is a dated but very interesting and well written treatment of certain foundational aspects of *Set Theory* by one of its creators. Joseph W. Dauben's *Georg Cantor: His Mathematics and Philosophy of the Infinite* is both an intellectual biography and a history of the genesis of transfinite *Set Theory*. Ivor Grattan-Guinness's recent book *The Search for Mathematical Roots: 1870 – 1940* deals with *Set Theory* in the context of the history of mathematical logic and the foundations of mathematics from Cantor to Gödel. *Zermelo's Axiom of Choice* by Gregory Moore treats certain later developments connected with the axiomatization of *Set Theory*. Cantor's original articles on transfinite numbers are well worth reading and are available in English translation in a Dover paperback, titled *Contributions to the Founding of the Theory of Transfinite Numbers*.

Set Theory and Computer Science: The Halting Problem

We've explored *Set Theory* now both informally and formally in order to develop some mathematical intuition about sets and to catch a glimpse of how *Set Theory* is deductively organized using the formalism and tools of *Predicate Logic*. *Set Theory* is foundational for many (most?) modern mathematical theories. It is also used in various contexts in computer science. Some applications use only the basic ideas considered in Chapter 4. But there are also areas of theoretical computer science, such as the theory of computation, that draw upon the more advanced parts of *Set Theory*.* One advanced proof technique that has found application here is the technique of diagonalization. We looked at this in a couple of guises, first in connection with demonstrating the uncountability of the real numbers but then also in proving *Cantor's Theorem*. There it was less clear how the argument had anything to do with diagonalization (cf. Exercise 5.2-19); self-referentiality seems to be the essence of the technique. It is in this latter form that diagonalization has found fruitful application in foundations of mathematics and computer science; it is also a standard ploy of modern analytical philosophy.

In this concluding subsection, we will briefly look at how diagonalization enters into computer science by arguing for a famous result that "solves" the *Halting Problem*. Our presen-

* It is interesting to note in this connection that several twentieth-century pioneers in computer science, such as Alan Turing and John von Neuman, were also active researchers in mathematical logic and *Set Theory*. Others, such as Kurt Gödel and Alonzo Church, focused on computatibility issues in logic, though without contributing to developments in computer science.

tation is necessarily informal, since we haven't rigorously defined what a computer program is (an effective procedure or algorithm) nor how it gets numerically encoded so a computer can implement it. This approach has its advantages, however, simplicity being one of them, since someone with only a basic familiarity with the way computers work should be able to follow it. Our argument can be formalized and made perfectly rigorous, however. This and similar results, all originating in work done since the mid-1930s, are proved in more advanced courses in mathematical logic, set theory, and computer science.

A rudimentary concern of computer programmers is not only to design a program that seems to do what it is intended to do but to make sure it behaves as it should for all possible inputs. There is always the potential of overlooking something so that a program works well for most values but fails spectacularly for outliers. In particular, a program might contain an unsuspected infinite loop, so that the computer will never halt, producing no output.

It would be great, therefore, if a super-program could be designed that, given any program together with any input, would be able to determine whether or not the program halts/produces an output. Can such a universal program-checker be designed or not? This is known as the *Halting Problem*. It was first tackled and answered by the British computer scientist Alan Turing in 1936. His answer is: no such super-program exists. You no doubt suspected this (else someone would be super-rich-and-famous in the computer world and lots of people would be out of work), but Turing actually *proved* that no such program *can* exist, using a proof-by-contradiction diagonalization argument.

Alan Turing

THEOREM*: *Insolubility of the Halting Problem

No universal program-checker exists that can determine for all possible programs and inputs whether a given program will halt and produce an output for a given input.

***Proof*:**

Suppose to the contrary that such a super-program exists. Call it H since it solves the *Halting Problem*. Suppose H returns a 1 when a program halts and a 0 when it doesn't: we'll write $H(P, x) = 1$ when program P halts upon being fed input x; else $H(P, x) = 0$.

Since computer programs are formulated as character strings, which get encoded as sequences of bits, just like inputs do, programs may themselves be treated as input data. In fact, that's needed in order for H to operate on a program P.

Thus, given any procedure P, we can consider what happens when H is given program P with input data P: H will produce a 1 iff P halts upon being fed itself as input.

We're now ready to design a procedure – call it C – which acts on programs in the following contrary way: if $H(P, P) = 0$, $C(P)$ will halt; while if $H(P, P) = 1$, C will go into an infinite loop and fail to halt. In other words, using H to do its checking, C halts given program P acting on input data P iff P itself does not halt with that input data.

You can maybe guess what's coming next; it's a self-focused maneuver inspired by Cantor's diagonalization argument/*Russell's Paradox*.

Question: What does C do when given itself as the program to act upon?

Answer: C applied to itself halts iff C acting on itself does not halt.

This is a blatant contradiction of the form $Q \leftrightarrow \neg Q$.

Thus we have no choice but to reject the original supposition: no universal program-checker H exists. The *Halting Problem* is intrinsically insoluble. ■

This is just one problem in theoretical computer science known to be insoluble. There are others as well. There are also problems known to be difficult but solvable given enough time. This moves us into the realm of algorithmic complexity, which is another area of elementary discrete mathematics. We will not explore this topic here due to time constraints, but it is

certainly accessible to anyone with a background in computer science and the knowledge of discrete mathematics provided by this text.

EXERCISE SET 5.3

Problems 1 - 4: *Russell's Paradox*

The following problems deal with various aspects of Russell's Paradox.

*1. *Russell's Paradox and the Barber Paradox*

   *a. Legend has it that a barber in a small town (call it *Russellville*) shaves all and only those people who do not shave themselves. If this is so, who shaves the barber? Explain your answer.

   *b. Compare the *Barber Paradox* of part *a* with *Russell's Paradox*. How are they similar?

2. *Russell's Paradox vs. Theorem 1*
   Explain why the set $\mathcal{N}$ formed in connection with *Russell's Paradox* leads us to a contradiction, forcing us to conclude that *Set Theory* is contradictory, while the same (sort of) set formed in proving *Theorem* 1 is used merely to claim that such a set doesn't exist, not that *Set Theory* is contradictory.

3. *No Universal Set Exists*
   Show that adopting $\exists U \forall x(x \in U)$ as an axiom of *Set Theory* leads to *Russell's Paradox*, thus contradicting *Theorem* 1 and making *Set Theory* inconsistent.

4. *Russell's Paradox and Cantor's Theorem*
   Review the proof of *Cantor's Theorem*, and explain what happens if the set used as a base set there is allowed to be the "set" of all sets $\mathcal{S}$. See whether you can discover how Russell might have been led to consider his paradoxical set $\mathcal{N}$.

Problems 5 - 6: True or False

Are the following statements true or false? Explain your answer.

5. The existence of all the elements belonging to a given class guarantees the existence of that class as a set.

*6. *Set Theory* was axiomatized primarily in order to deal with *Russell's Paradox* and like results.

Problems 7 - 8: Properties of Interior Empty Sets

Let $\emptyset_U = \{x \in U : x \neq x\}$ for any set U. Then prove the following, without using the Empty Set Axiom.

7. *Empty Set Membership Criterion*
   $\forall x(x \notin \emptyset_U)$

8. *Uniqueness of the Empty Set*
   If U and V are any sets, then $\emptyset_U = \emptyset_V$.

Problems 9 - 15: Existence and Uniqueness of Sets

Use the Axiom of Separation and the Axiom of Extensionality to prove the following existence and uniqueness results.

9. *Sets Formed by Separation are Unique/Well-Defined*
   a. Prove that sets formed by means of the *Axiom of Separation* are unique: $\forall U \exists! S(S = \{x \in U : P(x)\})$.
   b. Prove that intersections (see *Proposition* 2) are unique: $\forall S \forall T \exists! I(I = \{x \in S : x \in T\})$.

*10. *Set Differences*
   Prove *Proposition* 3, and show that set differences are unique: $\forall S \forall T \exists! D(D = \{x \in S : x \notin T\})$.

*11. *Finite Union Axiom*

   *a. Prove the unique existence of the union of two sets: $\forall S \forall T \exists! U(U = \{x : x \in S \vee x \in T\})$.

   b. Prove that if $S_1, S_2, \ldots, S_n$ are sets, then $S_1 \cup S_2 \cup \cdots \cup S_n$ is a unique set, for any n.

12. *Total Intersections*
Prove that if $\mathcal{C}$ is any non-empty collection of sets, then its total intersection $\bigcap_{S\in\mathcal{C}} S$ exists and is unique.

13. *Total Unions*
Prove that if $\mathcal{C}$ is any non-empty collection of sets, then its total union $\bigcup_{S\in\mathcal{C}} S$ exists and is unique.

14. *Pairs*
Prove the unique existence of pairs of sets: $\forall x \forall y \exists P(P = \{z : z = x \vee z = y\})$.

15. *Power Sets*
Prove the unique existence of power sets: $\forall S \exists P(P = \{R : R \subseteq S\})$.

EC 16. *Set Equality from Above*
Prove that two sets are identical iff they are members of exactly the same sets; i.e., prove $x = y \leftrightarrow \forall S(x \in S \leftrightarrow y \in S)$. Hint: use what you know about singletons to help you prove one direction.

Problems 17 - 21: Defining Tripletons, Quadrupletons, and N-pletons

The following problems explore ways to define tripletons and other size sets.

17. Give two distinctly different ways to define the notion of a *tripleton*, using an appropriate union of doubletons or singletons or both.

18. Taking your favorite definition from Problem 17 as the official definition, formulate and prove the obvious proposition giving the set membership criterion for tripletons: $x \in \{a, b, c\}$ iff _____ .

19. Take the definition from Problem 17 that you did *not* use in Problem 18 and show that the set so defined is equal to the one officially defined as the tripleton. (You may use the result of Problem 18 to prove this identity.) Thus, you have demonstrated that the two definitions you gave in Problem 17 are equivalent.

20. What do you think would be a natural definition for a quadrupleton (a set with four elements)? Would you restrict yourself to using only doubletons? Or should you make use of tripletons in your definition?

21. Give a recursive definition of an n-pleton for any positive integer $n \geq 2$. (Reflect on what you did in Problems 17 – 20 and then standardize the process if necessary.)

Problems 22 - 27: Defining Singletons and Doubletons

The following exercises are intended as an alternative approach to defining singletons and doubletons, so base all your work on Problem X. Do not assume any material about them from the text unless you can demonstrate its legitimacy here.

22. Formulate an axiom that affirms the existence of singletons directly.

23. State and prove a uniqueness proposition for singletons. Then define singletons, using the standard notation.

24. State and prove the set membership criterion for singletons.

25. Given your definition of singletons from Problem 23, define doubletons in terms of them and whatever else you need. What guarantees that doubletons are unique?

26. State and prove the set membership criterion for doubletons.

27. Compare the approach taken in the text with what you did in Problems 22-26. What are the relative merits of each approach? Why do you think that the standard approach is the one given in the text?

Problems 28 - 30: Generating Power Sets Using Pairs and Unions

The following problems explore how much of a power set can be obtained using pairs and unions.

*28. *Power by Pairing*

*a. Indicate how the pairing operation (yielding doubletons) and union can be used on the set $\emptyset$ to generate both the set $\{\emptyset, \{\emptyset\}\}$ and its power set, $\{\emptyset, \{\emptyset\}, \{\{\emptyset\}\}, \{\emptyset, \{\emptyset\}\}\}$.

b. Does the pairing operation alone on $\emptyset$ suffice to obtain the results of part a?

29. Let $A_1, A_2, \ldots, A_n$ be any n subsets of $\mathbb{N}$. Explain how $\{A_1, A_2, \ldots, A_n\}$ can be obtained using the pairing and union operators. Thus, finite portions of $\mathcal{P}(\mathbb{N})$ can be obtained without the power set operator.

EC 30. What is wrong with the following argument, showing that a countably infinite portion of $\mathcal{P}(\mathbb{N})$ can be obtained using pairing and unions. Let P_k denote the set $\{A_1, A_2, \ldots, A_k\}$, which we know exists using pairing and union. Then the infinite collection of subsets $\{A_1, A_2, \ldots A_n, \ldots\} = \bigcup_{k=1}^{\infty} P_k$ can be obtained by taking the total union of these sets.

Problems 31 - 35: Defining Ordered Pairs and Ordered Triples

The following explore potential definitions for ordered pairs and triples.

31. Show that $(x, y) = \{x, \{y\}\}$ would not be a good definition for ordered pairs. Hint: show that two different ordered pairs can be made to yield the same unordered pair. Use $\emptyset$, $\{\emptyset\}$, and $\{\{\emptyset\}\}$ in some combination.

32. Show that $(x, y) = \{x, \{x, y\}\}$ could be taken as a definition for ordered pairs; i.e., that it, too, leads to the fundamental result about equal ordered pairs: $(a, b) = (c, d)$ iff $a = c \wedge b = d$. What axiom plays a key role in your proof? Explain why you think this definition is not the one settled upon by Wiener or Kuratowski.

33. Wiener's original definition for an ordered pair was $(x, y) = \{\{\{x\}, \emptyset\}, \{\{y\}\}\}$. Show with this definition that the fundamental result about equal ordered pairs still holds: $(a, b) = (c, d)$ iff $a = c \wedge b = d$. Compare Wiener's definition with the official one, due to Kuratowski, stated in Exercise Set 4.3 just prior to Problem 17.

34. You might think that ordered triples can be defined directly, in a way that is similar to the definition for ordered pairs, by $(x, y, z) = \{\{x\}, \{x, y\}, \{x, y, z\}\}$. Determine what is wrong with this definition.

EC 35. Given the official definition of ordered pairs (see Exercise Set 4.3, just prior to Problem 17), show how to generate the Cartesian product $S \times T$ of two sets S and T using the power set operator and finite unions. Hint: you must first find an appropriate universal set to use for separating $S \times T$ out.

Problems 36 - 37: Axiom of Foundation

The following problems relate to the Axiom of Foundation.

*36. Suppose that $x \in X$ and $y \in X$, where $x \neq y$. Show that $X \neq \{X\}$, without appealing to the *Axiom of Foundation*.

37. The *Axiom of Foundation* rules out infinite descending membership chains T_i such that $\cdots T_n \in T_{n-1} \in \cdots \in T_2 \in T_1 \in S$. Does it also rule out infinite ascending membership chains $T_1 \in T_2 \in \cdots \in T_n \cdots$ inside a set S, where all $T_i \in S$? Consider in this connection the set S whose elements are $\emptyset, \{\emptyset\}, \{\{\emptyset\}\}, \cdots$. Does S have either infinite descending or infinite ascending membership chains? You may assume that collecting all these sets together yields a bona fide set.

Problems 38 - 40: Redundancy of Axioms

The following problems look at deductive relations among the axioms of Set Theory.

*38. *Proving the Finite Union Axiom*
Using the *Pairing Axiom* and the *Total Union Axiom*, prove the *Finite Union Axiom*. In other words, using the fact that you can form pairs and take total unions, show that you can generate $S \cup T$ from S and T without using the *Finite Union Axiom*.

39. *Proving the Pairing Axiom*
Using the *Finite Union Axiom* and the *Power Set Axiom*, along with the *Axiom of Extensionality* and the *Axiom of Separation*, prove the *Pairing Axiom*.

40. *Proving the Empty Set Axiom*
Using the *Axiom of Infinity* along with the *Axiom of Separation* and the *Axiom of Extensionality*, prove the *Empty Set Axiom*.

Problems 41 - 48: Peano Arithmetic Inside Set Theory

The following problems look at how natural numbers can be introduced inside Set Theory. Prove the following results in order, which show among other things that the Peano Postulates can be proved in the context of Set Theory.

The set $\mathbb{N}$ *is defined to be the intersection of all possible inductive sets (sets satisfying the Axiom of Infinity). In addition,* 0 *and successor are defined by the equations* $0 = \emptyset$ *and* $\mathcal{S}(n) = n \cup \{n\}$.

41. $\mathbb{N}$ is the smallest inductive set; i.e., it satisfies the *Axiom of Infinity* and is a subset of all inductive sets.
42. $0 \in \mathbb{N}$
43. $\forall n(n \in \mathbb{N} \rightarrow S(n) \in \mathbb{N})$

*44. $\forall n(0 \neq S(n))$

45. $\forall n \forall m(S(n) = S(m) \rightarrow n = m)$
46. $(\forall P \subseteq \mathbb{N})(0 \in P \wedge \forall n(n \in P \rightarrow \mathcal{S}(n) \in P) \rightarrow P = \mathbb{N})$
47. Write the natural numbers 1, 2, 3, and 4 in the following two ways.
    a. Using set-braces and the numbers 0, 1, 2, and 3.
    b. Using only set-braces and the *empty set* $\emptyset$.

EC 48. Given the pattern emerging in Problem 47, state a proposition that gives $\mathcal{S}(n)$ in terms of earlier natural numbers. Then prove your result using the recursive definition for $\mathcal{S}(n)$ and mathematical induction.

Problems 49 - 57: Properties of the Membership Relation Restricted to $\mathbb{N}$

Assuming the definition of 0 *and* $\mathcal{S}(n)$ *preceding Problem 41, prove the following interconnected properties holding for the membership relation on* $\mathbb{N}$. *Some results can be proved via UG (and may be true of more sets than just natural numbers); others will require mathematical induction (use the successor form for these). Note that some parts involve the exclusive-or connective* $\veebar$. *The key result here is the trichotomy property of Problem 55. You may wish to try to prove it from scratch without the benefit of first proving Problems 49 – 54; that will help you see just why the earlier exercises were included and where they enter the argument.*

*49. $(\forall n \in \mathbb{N})(n \in \mathcal{S}(n))$

*50. $(\forall m \in \mathbb{N})(\forall n \in \mathbb{N})(m \in n \rightarrow m \in \mathcal{S}(n))$

EC 51. $(\forall m \in \mathbb{N})(\forall n \in \mathbb{N})(\forall p \in \mathbb{N})(m \in n \in p \rightarrow m \in p)$ [Thus the membership relation $\in$ is *transitive* on $\mathbb{N}$].

52. $(\forall m \in \mathbb{N})(\forall n \in \mathbb{N})(m \in n \rightarrow \mathcal{S}(m) \in \mathcal{S}(n))$
53. $(\forall m \in \mathbb{N})(\forall n \in \mathbb{N})(m \in n \rightarrow n = \mathcal{S}(m) \veebar \mathcal{S}(m) \in n)$
54. $(\forall m \in \mathbb{N})(m = 0 \veebar 0 \in m)$
55. $(\forall m \in \mathbb{N})(\forall n \in \mathbb{N})(m \in n \veebar m = n \veebar n \in m)$ [Thus a trichotomy property holds for the relation $\in$].
56. $(\forall m \in \mathbb{N})(\forall n \in \mathbb{N})(m \in n \leftrightarrow m \subset n)$ [Note that *proper inclusion* is involved here.]
57. $(\forall n \in \mathbb{N})(\forall m)(m \in n \rightarrow m \in \mathbb{N})$
    [Note: nothing is initially required of m except that it is a set belonging to n. This result shows that $\mathbb{N}$ is an extension of the sequence of natural numbers, a "number" of sorts following all the natural numbers, the first "transfinite ordinal number."]

Problems 58 - 61: Set–Theoretical Definition of the $<$ ***Relation***

The order relation $<$ *for natural numbers is defined by the biconditional* $m < n \leftrightarrow m \in n$.
The order relation $\leq$ *on* $\mathbb{N}$ *is then defined in the obvious way by* $m \leq n \leftrightarrow m < n \vee m = n$.

58. Taking the definition of $<$, translate the results of Problems 49 – 55 into the language of numerical inequality.
59. Given the definitions of $<$ and $\leq$, show by means of the results in Problems 55 – 56 that $m \leq n \leftrightarrow m \subseteq n$.
60. Using the above definitions and the results of Problems 49 – 56, prove that $m \leq n \wedge n \leq m \rightarrow m = n$ for any natural numbers m and n.

61. Show that $\mathbb{N}$ is *well–ordered*; i.e., show that any non–empty set S of natural numbers has a least element m, in the sense that $m \leq x$ for every element $x \in S$.
Hint: note that the intersection of two natural numbers is always the smaller one, and use the *Axiom of Foundation* and the relevant results of Problems 49 – 56 and the above problems. *Proof by Mathematical Induction* is not directly needed here, though it is still active and present in the results being used.

EC 62. Discuss the set theoretic approach to *Peano Arithmetic* outlined in Problems 41 – 61. How do you feel about the way in which natural numbers and $\mathbb{N}$ are defined there? What aspects of these definitions or the subsequent theoretical developments, if any, bother you? What value might such an approach have?

Problems 63 - 64: The Halting Problem

The following problems deal with the Halting Problem.

63. Speedy Gonzalez claims that given today's advanced technology any computer program that will produce an output for a given input will do so in less than 100 hours. Comment on his claim based on what you know about the *Halting Problem*.

*64. Ellen Touring solved the *Halting Problem* by means of the following procedure S: *Given a program P and an input x, let P run with input x. If it produces an output, let $S(P, x) = 1$; otherwise let $S(P, x) = 0$.* Explain the problem with her solution.

HINTS TO STARRED EXERCISES 5.3

1. a. Try two possible answers: the barber does; the barber doesn't. What follows from each of these?
   b. The comparison should be clear.

6. [No hint.]

10. Use the *Axiom of Separation* for existence; the *Axiom of Extensionality* is needed for uniqueness.

11. a. Suppose there are two such sets and show they are identical. Existence follows from *Axiom* 4.

28. a. Keep making the elements and sets you need in stages via *Pairing*, and then take the union of appropriate sets to get the full power set.

36. Use the *Axiom of Extensionality* to show why these two sets cannot be identical.

38. Recall that *Total Unions* takes unions of collections of sets and that small collections can be formed via *Pairing*. What collection of sets do you need to form to arrive at $S \cup T$?

44. Use the definitions of 0 and $\mathcal{S}(n)$ given just before Problem 41. Show why the two sets are different via Proposition 1.

49. Use the definition of $\mathcal{S}(n)$ given just before Problem 41. Mathematical induction is not needed.

50. Use the definition of $\mathcal{S}(n)$ given just before Problem 41. Mathematical induction is not needed.

51. This one is EC, but it is a more interesting result than the last two; try it if you have time. (It gets translated into the transitivity of $<$ via the definition given prior to Problem 58.) Here induction is required: use PMI on p.

64. This contradicts the theorem on the unsolvability of the *Halting Problem*, but what specifically is wrong with the procedure offered here?