An Introduction to Elementary Set Theory

Guram Bezhanishvili and Eachan Landreth*

1 Introduction

In this project we will learn elementary set theory from the original historical sources by two key figures in the development of set theory, Georg Cantor (1845–1918) and Richard Dedekind (1831–1916). We will learn the basic properties of sets, how to define the size of a set, and how to compare different sizes of sets. This will enable us to give precise definitions of finite and infinite sets. We will conclude the project by exploring a rather unusual world of infinite sets.

Georg Cantor, the founder of set theory, considered by many as one of the most original minds in the history of mathematics, was born in St. Petersburg, Russia in 1845. His parents moved the family to Frankfurt, Germany in 1856. Cantor entered the Wiesbaden Gymnasium at the age of 15, and two years later began his university career in Zürich, Switzerland. In 1863 he moved to the University of Berlin, which during Cantor's time was considered the world's leading center of mathematical research. Four years later Cantor received his doctorate under the supervision of the great Karl Weierstrass (1815–1897). In 1869 Cantor obtained an unpaid lecturing post at the University of Halle. Ten years later he was promoted to a full professor. However, Cantor never achieved his dream of holding a Chair of Mathematics at Berlin. It is believed that one of the main reasons was the nonacceptance of his theories of infinite sets by the leading mathematicians of that time, most noticeably by Leopold Kronecker (1823–1891), a professor at the University of Berlin and a very influential figure in German mathematics, both mathematically and politically.

Cantor married in 1874 and had two sons and four daughters. Ten years later Cantor suffered the first of several mental breakdowns that were to plague him for the rest of his life. Cantor died in 1918 in a mental hospital at Halle. By that time his revolutionary ideas were becoming accepted by some of the leading figures of the new century. One of the greatest mathematicians of the twentieth century, David Hilbert (1862–1943), described Cantor's new mathematics as "the most astonishing product of mathematical thought" [17, p. 359], and claimed that "no one shall ever expel us from the paradise which Cantor has created for us" [17, p. 353]. More on Georg Cantor can be found in [8, 11, 12, 15, 17, 19] and in the literature cited therein.

Richard Dedekind was an important German mathematician, who was also a friend to, and an ally of, Cantor. He was born in Braunschweig, Germany in 1831. In 1848 Dedekind entered the Collegium Carolinum in Braunschweig, and in 1850 he entered the University of Göttingen—an important German center of mathematics and the home of the great Carl Friedrich Gauss (1777–1855). Dedekind became the last student of Gauss. In 1852 Dedekind received his doctorate, and spent the next two years at the University of Berlin—the mecca of mathematics of the second half of the nineteenth century. At the University of Berlin, Dedekind became friends with Bernhard Riemann (1826–1866). They both were awarded the Habilitation in 1854, upon which Dedekind

*Department of Mathematical Sciences; New Mexico State University; Las Cruces, NM 88003; gbezhani@nmsu.edu, eachan@nmsu.edu.

returned to Göttingen to teach as Privatdozent.[1] In Göttingen, Dedekind became friends with Lejeune Dirichlet (1805–1859). After Dirichlet's death, Dedekind edited Dirichlet's lectures on number theory, which were published in 1863. He also edited the works of Gauss and Riemann. From 1858 to 1862 Dedekind taught at the Polytechnic Institute in Zürich. In 1862 his alma mater the Collegium Carolinum was upgraded to a Technische Hochschule (Institute of Technology), and Dedekind returned to his native Braunschweig to teach at the Institute. He spent the rest of his life there. Dedekind retired in 1894, but continued active mathematical research until his death.

Dedekind is mostly known for his research in algebra and set theory. He was the first to define real numbers by means of cuts of rational numbers. To this day many schools around the globe teach the theory of real numbers based on Dedekind's cuts. Dedekind was the first to introduce the concept of an ideal—a key concept in modern algebra—generalizing the ideal numbers of Ernst Kummer (1810–1893). His contributions to set theory as well as to the study of natural numbers and modular lattices are equally important. In fact, his 1900 paper on modular lattices is considered the first publication in a relatively new branch of mathematics called lattice theory. Dedekind was a well-respected mathematician during his lifetime. He was elected to the Academies of Berlin and Rome as well as to the French Academy of Sciences, and also received honorary doctorates from the universities of Oslo, Zürich, and Braunschweig. More on Richard Dedekind can be found in [15, 17, 22, 24] and in the literature cited therein.

The beginning of Dedekind's friendship with Cantor dates back to 1874, when they first met each other while on holidays at Interlaken, Switzerland. Their friendship and mutual respect lasted until the end of their lives. Dedekind was one of the first who recognized the importance of Cantor's ideas, and became his important ally in promoting set theory.

It is only fitting to study set theory from the writings of Cantor and Dedekind. In this project we will be working with the original historical source by Cantor "Beiträge zur Begründung der transfiniten Mengenlehre" ("Contributions to the founding of the theory of transfinite numbers. I") [5] which appeared in 1895, and the original historical source by Dedekind "Was sind und was sollen die Zahlen?" ("The nature and meaning of numbers") [9] which appeared in 1888. An English translation of Cantor's source is available in [6], and an English translation of Dedekind's source is available in [10].

2 Sets

In the first half of the project our main subject of study will be *sets*. This is how Cantor defined a set:

By an "aggregate" we are to understand any collection into a whole M of definite and separate objects m of our intuition or our thought. These objects are called the "elements" of M. [6, p. 85]

The German word for a set is *Menge*, which is the reason Cantor denotes a set by M and its elements by m. In [6] Menge is translated as an aggregate, but it has since become common to use the word set instead.

[1] In Germany, as well as in some other European and Asian countries, Habilitation is the highest academic qualification a scholar can achieve. It is earned after obtaining a Ph.D., and requires the candidate to write a second thesis, known as a Habilitation thesis. The level of a Habilitation thesis has to be considerably higher than that of a Ph.D. thesis, and must be accomplished independently. The Habilitation qualifies the holder to independently supervise doctoral candidates. In Germany such a post is known as Privatdozent. After serving as a Privatdozent, one is eligible for full professorship [16].

Examples of sets are to be found everywhere around us. For example, we can speak of the set of all living human beings, the set of all cities in the US, the set of all sentences of some language, the set of all prime numbers, and so on. Each living human being is an element of the set of all living human beings. Similarly, each prime number is an element of the set of all prime numbers, and so on.

If S is a set and s is an element of S, then we write $s \in S$. If it so happens that s is not an element of S, then we write $s \notin S$. If S is the set whose elements are s, t, and u, then we write $S = \{s, t, u\}$. The left brace and right brace visually indicate the "bounds" of the set, while what is written within the bounds indicates the elements of the set. For example, if $S = \{1, 2, 3\}$, then $2 \in S$, but $4 \notin S$.

Sets are determined by their elements. The order in which the elements of a given set are listed does not matter. For example, $\{1, 2, 3\}$ and $\{3, 1, 2\}$ are the same set. It also does not matter whether some elements of a given set are listed more than once. For instance, $\{1, 2, 2, 2, 3, 3\}$ is still the set $\{1, 2, 3\}$.

Many sets are given a shorthand notation in mathematics because they are used so frequently. A few elementary examples are the set of natural numbers,

$$\{0, 1, 2, \dots\},$$

denoted by the symbol $\mathbb{N}$, the set of integers,

$$\{\dots, -2, -1, 0, 1, 2, \dots\},$$

denoted by the symbol $\mathbb{Z}$, the set of rational numbers, denoted by the symbol $\mathbb{Q}$, and the set of real numbers, denoted by the symbol $\mathbb{R}$.

A set may be defined by a property. For instance, the set of all planets in the solar system, the set of all even integers, the set of all polynomials with real coefficients, and so on. For a property P and an element s of a set S, we write $P(s)$ to indicate that s has the property P. Then the notation $A = \{s \in S : P(s)\}$ indicates that the set A consists of all elements s of S having the property P. The colon : is commonly read as "such that," and is also written as "$|$." So $\{s \in S \mid P(s)\}$ is an alternative notation for $\{s \in S : P(s)\}$. For a concrete example, consider $A = \{x \in \mathbb{R} : x^2 = 1\}$. Here the property P is "$x^2 = 1$." Thus, A is the set of all real numbers whose square is one.

Exercise 2.1. In the following sentences, identify the property, and translate the sentence to set notation.

1. The set of all even integers.
2. The set of all odd prime numbers.
3. The set of all cities with population more than one million people.

Exercise 2.2. Give an alternative description of the sets specified below.

1. $\{x \in \mathbb{R} : x^2 = 1\}$.
2. $\{x \in \mathbb{Z} : x > -2 \text{ and } x \leq 3\}$.
3. $\{x \in \mathbb{N} : x = 2y \text{ for some } y \in \mathbb{N}\}$.

2.1 Subset relation

For two sets, we may speak of whether or not one set is contained in the other. Here is how Dedekind defines this relation between sets. Note that Dedekind calls sets *systems*.

A system A is said to be *part* of a system S when every element of A is also an element of S. Since this relation between a system A and a system S will occur continually in what follows, we shall express it briefly by the symbol $A \prec S$. [10, p. 46]

Modern notation for $A \prec S$ is $A \subseteq S$, and we say that A is a *subset* of S. Thus,

$$A \subseteq S \text{ if, and only if, for all } x, \text{ if } x \in A, \text{ then } x \in S.$$

When A is not a subset of S, we write $A \not\subseteq S$.

Exercise 2.3. Describe what it means for $A \not\subseteq S$ that is similar to the description of $A \subseteq S$ given above.

Dedekind goes on to show that the subset relation satisfies the following properties.

Exercise 2.4.

1. Show that $A \subseteq A$.
2. Show that if $A \subseteq B$ and $B \subseteq C$, then $A \subseteq C$.

The first property is usually referred to as *reflexivity* and the second as *transitivity*. Thus, Exercise 2.4 establishes that the subset relation between sets is both reflexive and transitive. Dedekind also defines what it means for A to be a proper part of S.

A system A is said to be a *proper* part of S, when A is part of S, but...S is not a part of A, i.e., there is in S an element which is not an element of A. [10, p. 46]

Nowadays we say that A is a *proper subset* of S, and write $A \subset S$. If A is not a proper subset of S, then we write $A \not\subset S$.

Exercise 2.5.

1. Describe what it means for A to be a proper subset of S.
2. Describe what it means for A not to be a proper subset of S.
3. Show that if $A \subset S$, then $A \subseteq S$.
4. Does the converse hold? Justify your answer.
5. Show that $A \not\subset A$ for each set A.
6. Prove that if $A \subset B$ and $B \subset C$, then $A \subset C$.

The fifth property is usually referred to as *irreflexivity*. Thus, it follows from Exercise 2.5 that being a proper subset is an irreflexive and transitive relation.

As we have already seen, the subset relation $\subseteq$ is defined by means of the membership relation $\in$. However, the two behave quite differently.

Exercise 2.6.

1. Give an example of a set A such that there is a set B with $B \in A$ but $B \not\subseteq A$.

2. Give an example of a set A such that there is a set B with $B \subseteq A$ but $B \notin A$.

2.2 Set equality

We already discussed the membership and subset relations between sets. But when are two sets equal? Dedekind addresses this issue as follows.

...a system S...is completely determined when with respect to every thing it is determined whether it is an element of S or not.[2] The system S is hence the same as the system T, in symbols $S = T$, when every element of S is also element of T, and every element of T is also element of S. [10, p. 45]

Thus, two sets A and B are equal, in notation $A = B$, when they consist of the same elements; that is,

$$A = B \text{ if, and only if, for all } x,\ x \in A \text{ if, and only if, } x \in B.$$

Exercise 2.7. Prove that $A = B$ if and only if $A \subseteq B$ and $B \subseteq A$.

If two sets A and B are not equal, we write $A \neq B$.

Exercise 2.8. Let P be the property "is a prime number" and O be the property "is an odd integer." Consider the sets $A = \{x \in \mathbb{N} : P(x)\}$ and $B = \{x \in \mathbb{N} : O(x)\}$.

1. Examine A and B with respect to the subset relation. What can you conclude? Justify your answer.

2. Are A and B equal? Justify your answer.

Exercise 2.9. Consider the sets

$$A = \{x \in \mathbb{Z} : x = 2(y - 2) \text{ for some } y \in \mathbb{Z}\}$$

and

$$B = \{x \in \mathbb{Z} : x = 2z \text{ for some } z \in \mathbb{Z}\}.$$

Are A and B equal? Justify your answer.

---

[2] We give Dedekind's footnote in full, where he opposes Kronecker's point of view and sides with Cantor in his mathematical battles with Kronecker. "In what manner this determination is brought about, and whether we know a way of deciding upon it, is a matter of indifference for all that follows; the general laws to be developed in no way depend upon it; they hold under all circumstances. I mention this expressly because Kronecker not long ago (*Crelle's Journal*, Vol. 99, pp. 334–336) has endeavored to impose certain limitations upon the free formation of concepts in mathematics which I do not believe to be justified; but there seems to be no call to enter upon this matter with more detail until the distinguished mathematician shall have published his reasons for the necessity or merely the expediency of these limitations."

2.3 Set operations

So far we have studied the membership, subset, and equality relations between sets. But we can also define operations on sets that are somewhat similar to the operations of addition, multiplication, and subtraction of numbers that you are familiar with.

The sum of a collection of sets is obtained by combining the elements of the sets. Nowadays we call this operation *union*. This is how Dedekind defines it.

By the system *compounded* out of any systems $A, B, C, \ldots$ to be denoted $\mathfrak{M}(A, B, C, \ldots)$ we mean that system whose elements are determined by the following prescription: a thing is considered as element of $\mathfrak{M}(A, B, C, \ldots)$ when and only when it is element of some one of the systems $A, B, C, \ldots$, i.e., when it is element of A, *or* B, *or* $C, \ldots$ [10, pp. 46–47]

In the particular case of two sets A and B, the union of A and B is the set consisting of the elements that belong to either A or B. Modern notation for $\mathfrak{M}(A, B)$ is $A \cup B$. Thus,

$$A \cup B = \{x : x \in A \text{ or } x \in B\}.$$

Here the meaning of "or" is inclusive; that is, if it so happens that an element x belongs to both A and B, then x belongs to the union $A \cup B$.

Another useful operation on sets is taking their common part. Nowadays this operation is known as *intersection*. This is how Dedekind defines it.

A thing g is said to be *common* element of the systems $A, B, C, \ldots$, if it is contained in each of these systems (that is in A *and* in B *and* in $C \ldots$). Likewise a system T is said to be a *common part* of $A, B, C, \ldots$ when T is part of each of these systems; and by the *community* of the systems $A, B, C, \ldots$ we understand the perfectly determinate system $\mathfrak{G}(A, B, C, \ldots)$ which consists of all the common elements g of $A, B, C, \ldots$ and hence is likewise a common part of those systems. [10, pp. 48–49]

In the particular case of two sets A and B, the intersection of A and B is the set consisting of the elements of both A and B. Modern notation for $\mathfrak{G}(A, B)$ is $A \cap B$. Thus,

$$A \cap B = \{x : x \in A \text{ and } x \in B\}.$$

We may also define the difference of two sets A and B as the set consisting of those elements of A that do not belong to B. This operation is called *set complement* and is denoted by $-$. Thus,

$$A - B = \{x : x \in A \text{ and } x \notin B\}.$$

The notations for the set operations $\cup, \cap, -$, for the membership relation $\in$, and for the subset relation $\subseteq$ that we use today were first introduced by the famous Italian mathematician Giuseppe Peano (1858–1932).[3]

[3]More on the life and work of Giuseppe Peano can be found in [13, 15, 18]. Also, our webpage `http://www.cs.nmsu.edu/historical-projects/` offers a variety of historical projects, including an historical project on Peano's work on natural numbers (see [3]).

Exercise 2.10. Let $A = \{2, 3, 5, 7, 11, 13\}$ and $B = \{A, 2, 11, 18\}$.

1. Find $A \cup B$.
2. Find $A \cap B$.
3. Find $A - B$.

Usually the sets that we work with are subsets of some ambient set. For instance, even numbers, odd numbers, and prime numbers are all subsets of the set of integers $\mathbb{Z}$. Such an ambient set is referred to as a *universal set* (or a *set of discourse*) and is denoted by U. In other words, a universal set is the underlying set that all the sets under examination are subsets of. We may thus speak of the set difference $U - A$, which is the set of those elements of U that do not belong to A. The set difference $U - A$ is usually denoted by A^c. Thus,

$$A^c = U - A = \{x \in U : x \notin A\}.$$

Exercise 2.11. Let $A = \{x \in \mathbb{R} : x^2 = 2\}$ and $B = \{x \in \mathbb{R} : x \geq 0\}$.

1. Find $A \cap B$.
2. Find $A \cup B$.
3. Find $A - B$.
4. For $U = \mathbb{R}$, find A^c and B^c.
5. Find $\mathbb{N} - B$.

2.4 Empty set

As we saw in Exercise 2.11, the set operations may yield a set containing no elements.

Exercise 2.12.

1. Let A be any set and let E be a set containing no elements. Prove that $E \subseteq A$.
2. Conclude that there is a unique set containing no elements.

We call the set containing no elements the *empty set* (or *null set*) and denote it by $\emptyset$.

Exercise 2.13. Give a definition of the empty set.

Exercise 2.14. Consider the following sets:

1. $A = \{x \in \mathbb{Q} : x^2 = 2\}$,
2. $B = \{x \in \mathbb{R} : x^2 + 1 = 0\}$,
3. $C = \{x \in \mathbb{N} : x^2 + 1 < 1\}$.

Can you give an alternative description of each of these sets? Justify your answer.

2.5 Set identities

There are a number of set identities that the set operations of union, intersection, and set difference satisfy. They are very useful in calculations with sets. Below we give a table of such set identities, where U is a universal set and A, B, and C are subsets of U.

Commutative Laws:	$A \cup B = B \cup A$	$A \cap B = B \cap A$
Associative Laws:	$(A \cup B) \cup C = A \cup (B \cup C)$	$(A \cap B) \cap C = A \cap (B \cap C)$
Distributive Laws:	$A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$	$A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$
Idempotent Laws:	$A \cup A = A$	$A \cap A = A$
Absorption Laws:	$A \cap (A \cup B) = A$	$A \cup (A \cap B) = A$
Identity Laws:	$A \cup \emptyset = A$	$A \cap U = A$
Universal Bound Laws:	$A \cup U = U$	$A \cap \emptyset = \emptyset$
DeMorgan's Laws:	$(A \cup B)^c = A^c \cap B^c$	$(A \cap B)^c = A^c \cup B^c$
Complement Laws:	$A \cup A^c = U$	$A \cap A^c = \emptyset$
Complements of U and $\emptyset$:	$U^c = \emptyset$	$\emptyset^c = U$
Double Complement Law:	$(A^c)^c = A$	
Set Difference Law:	$A - B = A \cap B^c$	

Each of these laws asserts that the set on the right-hand side is equal to the set on the left-hand side. As we now know, this means that the two sets consist of the same elements. For example, to verify the de Morgan law $(A \cup B)^c = A^c \cap B^c$, we need to show that for each x, we have $x \in (A \cup B)^c$ if, and only if, $x \in A^c \cap B^c$. But $x \in (A \cup B)^c$ is equivalent to $x \notin A \cup B$. This is equivalent to $x \notin A$ and $x \notin B$, which is clearly equivalent to $x \in A^c$ and $x \in B^c$. Therefore, $x \in (A \cup B)^c$ is equivalent to $x \in A^c \cap B^c$. Thus, we have verified that $(A \cup B)^c$ and $A^c \cap B^c$ consist of the same elements, which means that $(A \cup B)^c = A^c \cap B^c$. Other set identities in the table can be verified by a similar argument. The next three exercises invite you to verify the remaining set identities in the table. The laws are grouped in these exercises according to the level of difficulty, from very simple to more difficult.

Exercise 2.15.

1. Prove the commutative laws.
2. Prove the associative laws.
3. Prove the idempotent laws.
4. Prove the identity laws.
5. Prove the universal bound laws.

Exercise 2.16.

1. Prove the complement laws.
2. Prove the complement of U and $\emptyset$ laws.
3. Prove the double complement law.
4. Prove the difference law.

Exercise 2.17.

1. Prove the absorbtion laws.
2. Prove the second DeMorgan law.
3. Prove the distributive laws.

Exercise 2.18. Prove the following using only set identities:

1. $(A \cup B) - C = (A - C) \cup (B - C)$.
2. $(A \cup B) - (C - A) = A \cup (B - C)$.
3. $A \cap (((B \cup C^c) \cup (D \cap E^c)) \cap ((B \cup B^c) \cap A^c)) = \emptyset$.

2.6 Cartesian products and powersets

Next we introduce two more operations on sets. Both will play an important role in the second part of the project when we start developing the theory of finite and infinite sets. The first one plays an important role in defining the concept of function between sets, which is one of the key concepts in mathematics. The second one is of great importance in building sets of bigger and bigger sizes.

For two sets A and B, we define the *Cartesian product* of A and B to be the set of all ordered pairs (a, b), where $a \in A$ and $b \in B$. This operation on sets is somewhat similar to the product of two numbers. We denote the Cartesian product of A and B by $A \times B$. Thus,

$$A \times B = \{(a, b) : a \in A \text{ and } b \in B\}.$$

Exercise 2.19. Let $A = \{1, 2, 3\}$ and $B = \{a, b\}$.

1. Determine $A \times B$ and $B \times A$.
2. Are $A \times B$ and $B \times A$ equal? Justify your answer.

Exercise 2.20.

1. Let A consist of 4 elements and B consist of 5 elements. How many elements are in $A \times B$? Justify your answer.
2. More generally, let A consist of n elements and B consist of m elements. How many elements are in $A \times B$? Justify your answer.

Given a set A, we may speak of the set of all subsets of A. This is yet another operation on sets which, as we will see, is of great importance. We call the set of all subsets of A the *powerset* of A and denote it by $P(A)$. Thus,

$$P(A) = \{B : B \subseteq A\}.$$

For example, if $A = \{1, 2\}$, then the subsets of A are $\emptyset, \{1\}, \{2\}$, and A. Therefore, $P(A) = \{\emptyset, \{1\}, \{2\}, A\}$.

Exercise 2.21.

1. Determine $P(\emptyset)$.

2. Determine $P(\{1\})$.
3. Determine $P(\{1, 2, 3\})$.

Exercise 2.22.

1. Calculate $P(\{\emptyset\})$.
2. Calculate $P(\{\emptyset, \{\emptyset\}\})$.
3. Calculate $P(\{\{\emptyset\}\})$.
4. Calculate $P(P(\emptyset))$.
5. Calculate $P(P(\{\emptyset\}))$.

2.7 Russell's paradox

We conclude the first half of the project by the celebrated Russell's paradox. As we saw earlier in the project, different properties give rise to different sets. If every set was determined by some property, then the whole of set theory would be derivable from the general principles of logic. Since all of mathematics is based on set theory, it would follow that the whole of mathematics is derivable from the general principles of logic. This was the grand plan, known as *logicism*, of the great German mathematician, philosopher, and one of the founders of modern logic, Gottlob Frege (1848–1925).[4] Unfortunately, soon after Frege published his program, the famous British philosopher, mathematician, and antiwar activist Bertrand Russell (1872–1970) found a fatal flaw in Frege's arguments. This became known as *Russell's paradox*.[5] For the history of Russell's paradox, including the excerpt from his 1902 letter to Frege, we refer to [23], where different versions of the paradox, as well as paradoxes of a similar nature can also be found. Below we give one of the most popular versions of Russell's paradox, which is perfectly suited for our purposes. It is taken from [21, pp. 1–2].

By a set, we mean any collection of objects — for example, the set of all even integers or the set of all saxophone players in Brooklyn. The objects that make up a set are called its members or elements. Sets may themselves be members of sets; for example, the set of all sets of integers has sets as its members. Most sets are not members of themselves; the set of cats, for example, is not a member of itself because the set of cats is not a cat. However, there may be sets that do belong to themselves — for example, the set of all sets. Now, consider the set A of all those sets X such that X is not a member of X. Clearly, by definition, A is a member of A if any only if A is not a member of A. So, if A is a member of A, then A is also not a member A; and if A is not a member of A, then A is a member of A. In any case, A is a member of A and A is not a member of A.

Let A be the set of all those sets that are not members of themselves.

Exercise 2.23. Single out the property that defines the set A.

[4] More on the life and work of Frege can be found in [14, 15]. Also, another article in this series [20] offers an historical project on Frege's development of propositional logic.

[5] More on the life and work of Russell can be found in [2, 15]. Also, another article in this series [4] offers an historical project on Russell's work on logic.

The question we will examine is whether A is a member of itself.

Exercise 2.24.

1. First assume that $A \in A$ and conclude that $A \notin A$. Justify your argument.
2. Next assume that $A \notin A$ and conclude that $A \in A$. Justify your argument.
3. What can you conclude from (1) and (2)? Explain.
4. Discuss why Russell's paradox contradicts Frege's program.
5. How would you resolve the situation? Explain.

3 Functions, one-to-one correspondences, and cardinal numbers

So far in this project we have studied such basic relations between sets as membership, subset, and equality relations. We have also studied basic operations on sets such as union, intersection, set difference, Cartesian product, and powerset. In the second half of the project we will discuss the "size" of a set. We have already encountered sets of large and small sizes. Some sets that we have encountered were finite and some were infinite. Our next goal is to formalize the concept of the size of a set. As we will see, this can be done by means of *functions*—one of the key concepts in mathematics.

We will learn about functions, one-to-one and onto functions, one-to-one correspondences, and how they allow us to formalize the concept of the size of a set. A formalization of the size of a set is known as the *cardinality* of a set. We will discuss what it means for two sets to have the same size, and study how to compare the sizes of different sets. We will introduce countable sets and show that many sets are countable, including the set of integers and the set of rational numbers. We will also discuss Cantor's diagonalization method which allows us to show that not every infinite set is countable. This will yield infinite sets of different sizes. In particular, we will show that the set of real numbers is not countable. We will also examine the cardinal number $\aleph_0$, the first in the hierarchy of infinite cardinal numbers, and obtain a method that allows us to create infinitely many infinite cardinal numbers.

3.1 Functions

You have probably already encountered functions from real numbers to real numbers in the first course of calculus. More generally, given two sets A and B, a *function* from A to be B is a rule associating with each element of A *one and only one* element of B. This is how Dedekind defines a function. Note that he refers to functions as *transformations* (and to sets as *systems*).

By a *transformation* ϕ of a system S we understand a law according to which to every determinate element s of S there *belongs* a determinate thing which is called the *transform* of s and denoted by $\phi(s)$; we say also that $\phi(s)$ *corresponds* to the element s, that $\phi(s)$ *results* or is *produced* from s by the transformation ϕ, that s is *transformed* into $\phi(s)$ by the transformation ϕ. [10, p. 50]

When there is a function f from A to B, we write $f : A \to B$. The set A is referred to as the *domain* of f, and the set B is referred to as the *codomain* of f. Since the function f associates with each $a \in A$ a unique $b \in B$, we say that f *maps* a *to* b and write $f(a) = b$. A convenient way to think about a function $f : A \to B$ is as the set of ordered pairs (a, b), where $a \in A$, $b \in B$, and f maps a to b. It follows from the definition of a function that we cannot have two ordered pairs (a, b) and (a, c) with $b \neq c$. Thus, we can think of functions from A to B as subsets F of the Cartesian product $A \times B$ which satisfy the following property: For each $a \in A$ there exists a unique $b \in B$ such that $(a, b) \in F$.

Exercise 3.1. Are the following functions? Justify your answer.

1. $f(x) = x^2$ with domain and codomain $\mathbb{R}$.
2. $g(x) = 2x + 1$ with domain and codomain $\mathbb{Q}$.
3. $h(x) = \pm x$ with domain and codomain $\mathbb{Z}$.
4. $u(x) = \sqrt{x}$ with domain and codomain $\mathbb{N}$.

Exercise 3.2. Write each of the following functions as a set of ordered pairs.

1. $f : \mathbb{R} \to [-1, 1]$ defined by $f(x) = \cos(x)$.
2. $g : \mathbb{R} \to \mathbb{R}$ defined by $g(x) = 300x$.
3. $h : \mathbb{R}^+ \to \mathbb{R}$ defined by $h(x) = \ln(x)$. Here and below $\mathbb{R}^+ = \{x \in \mathbb{R} : x > 0\}$.

For a function $f : A \to B$, we call the set of all values of f the *range* or *image* of f. Thus, the image of f is the set

$$\mathrm{Im}(f) = \{b \in B : b = f(a) \text{ for some } a \in A\}.$$

Exercise 3.3. Consider the function $f = \{(1, 2), (2, 3), (3, 3), (4, 5), (5, -1), (6, 2)\}$. Identify the domain and image of f.

3.2 Images and inverse images

Let $f : A \to B$ be a function, $S \subseteq A$, and $T \subseteq B$. The image of S with respect to f is the set of those elements of B which the elements of S are mapped to. Therefore, the *image* of S with respect to f is the set

$$f(S) = \{f(s) : s \in S\}.$$

On the other hand, the inverse image of T with respect to f is the set of those elements of A that are mapped to some element of T. Thus, the *inverse image* of T with respect to f is the set

$$f^{-1}(T) = \{a \in A : f(a) \in T\}.$$

Exercise 3.4. For each of the following functions determine the image of $S = \{x \in \mathbb{R} : 9 \leq x^2\}$.

1. $f : \mathbb{R} \to \mathbb{R}$ defined by $f(x) = |x|$.
2. $g : \mathbb{R} \to \mathbb{R}^+$ defined by $g(x) = e^x$.
3. $h : \mathbb{R} \to \mathbb{R}$ defined by $h(x) = x - 9$.

Exercise 3.5. For each of the following functions determine the inverse image of $T = \{x \in \mathbb{R} : 0 \leq x^2 - 25\}$.

1. $f : \mathbb{R} \to \mathbb{R}$ defined by $f(x) = 3x^3$.
2. $g : \mathbb{R}^+ \to \mathbb{R}$ defined by $g(x) = \ln(x)$.
3. $h : \mathbb{R} \to \mathbb{R}$ defined by $h(x) = x - 9$.

3.3 When are two functions equal?

Let $f, g : A \to B$ be two functions. We say that f *equals* g and write $f = g$ if $f(a) = g(a)$ for each $a \in A$.

Exercise 3.6. Let $f : \mathbb{Z} \to \mathbb{Z}$ be defined by $f(a) = 2a^2 - a$ and $g : \mathbb{Z} \to \mathbb{Z}$ be defined by $g(x) = x(2x - 1)$. Determine whether f is equal to g. Justify your answer.

3.4 Composition

Given two functions $f : A \to B$ and $g : B \to C$, we can produce a new function $h : A \to C$ by *composing* f and g. This is how Dedekind defines the composition of two functions.

If ϕ is a transformation of a system S, and ψ a transformation of the transform $S' = \phi(S)$, there always results a transformation θ of S, *compounded* out of ϕ and ψ, which consists of this that to every element s of S there corresponds the transform

$$\theta(s) = \psi(s') = \psi(\phi(s)),$$

where again we have put $\phi(s) = s'$. This transformation θ can be denoted briefly by the symbol $\psi.\phi$ or $\psi\phi$, the transform $\theta(s)$ by $\psi\phi(s)$ where the order of the symbols ϕ, ψ is to be considered... [10, p. 52]

Thus, if $f : A \to B$ and $g : B \to C$ are two functions, then their *composition* is defined as the function $h : A \to C$ such that $h(a) = g(f(a))$ for each $a \in A$. We denote the composition of f and g by $g \circ f$. Dedekind goes on to show that for any functions $f : A \to B$, $g : B \to C$, and $h : C \to D$, we have $h \circ (g \circ f) = (h \circ g) \circ f$.

If now χ signifies a transformation of the system $\psi(s') = \psi\phi(s)$ and η the transformation $\chi\psi$ of the system S' compounded out of ψ and χ, then is $\chi\theta(s) = \chi\psi(s') = \eta(s') = \eta\phi(s)$; therefore the compound transformations $\chi\theta$ and $\eta\phi$ coincide for every element s of S, i.e., $\chi\theta = \eta\phi$. In accordance with the meaning of θ and η this theorem can finally be expressed in the form

$$\chi.\psi\phi = \chi\psi.\phi,$$

and this transformation compounded out of ϕ, ψ, χ can be denoted briefly by $\chi\psi\phi$. [10, pp. 52–53]

In the next exercise we examine Dedekind's proof.

Exercise 3.7. Let $f : A \to B$, $g : B \to C$, and $h : C \to D$ be functions.

1. State what you need to show to conclude that $h \circ (g \circ f) = (h \circ g) \circ f$.

2. Consider now some $a \in A$. Calculate $h((g \circ f)(a))$ and $(h \circ g)(f(a))$. Are they equal?

3. Use your solutions to (1)–(2) to conclude that $h \circ (g \circ f) = (h \circ g) \circ f$.

Let f and g be functions with the same domain and codomain. Then we can form $g \circ f$ and $f \circ g$, and both of these functions have the same domain and codomain as f and g. In the next exercise we examine whether the functions $g \circ f$ and $f \circ g$ have to be equal.

Exercise 3.8. Let $f, g : A \to A$ be functions with the same domain and codomain A.

1. Give a brief justification of why $g \circ f, f \circ g : A \to A$ both have the same domain and codomain A.

2. Either give a proof that $g \circ f$ and $f \circ g$ are equal or show that they are not equal by constructing a counterexample.

One of the simplest functions is the so-called identity function. This is how Dedekind defines it.

The simplest transformation of a system is that by which each of its elements is transformed into itself; it will be called the *identical* transformation of the system. [10, p. 50]

In other words, the *identity function* on a set A is the function $i_A : A \to A$ defined by $i_A(a) = a$ for each $a \in A$.

Exercise 3.9. Let $f : A \to B$ be a function.

1. Show that for the identity function i_A on A we have $f \circ i_A = f$.

2. Show that for the identity function i_B on B we have $i_B \circ f = f$.

3.5 One-to-one and onto functions, one-to-one correspondences

We already encountered functions $f : A \to B$ that map different elements of A to the same element of B. For example, the absolute value function of Exercise 3.4 has this property. We say that f is a *one-to-one function* (or an *injective function*) if f maps different elements of A to different elements of B. Thus, f is one-to-one if for each $a_1, a_2 \in A$, from $a_1 \neq a_2$ it follows that $f(a_1) \neq f(a_2)$.

Exercise 3.10. Let $f : A \to B$ be a function. Show that the following two conditions are equivalent:

1. f is one-to-one.

2. For each $a_1, a_2 \in A$, whenever $f(a_1) = f(a_2)$, then $a_1 = a_2$.

In fact, both of these conditions are equivalent to a third condition stating that $S = f^{-1}(f(S))$ for each $S \subseteq A$. But this is a little more challenging to prove. (Try!)

We also encountered functions $f : A \to B$ such that the image of f is a proper subset of the codomain of f. Again, the absolute value function of Exercise 3.4 has this property. We say that f is an *onto function* (or a *surjective function*) if the image of f equals the codomain of f. Thus, f is onto if for each $b \in B$ there exists at least one $a \in A$ such that $f(a) = b$. One can show that f is onto if and only if $T = f(f^{-1}(T))$ for each $T \subseteq B$. This is a little more challenging to prove. (Give it a try!)

Let $f : A \to B$ be a function. If it happens that f is both one-to-one and onto, then we say that f is a *one-to-one correspondence* (or a *bijection*) between A and B.

Exercise 3.11. Consider the following two functions:

1. $f : \mathbb{R} \to \mathbb{R}$ defined by $f(x) = 4x - 15$.
2. $g : \mathbb{R} \to \mathbb{R}$ defined by $f(x) = 15x^3$.

Prove that both f and g are one-to-one correspondences.

Let $f : A \to B$ be a one-to-one correspondence. Then to each $b \in B$ there corresponds a unique $a \in A$ such that $f(a) = b$. We define $f^{-1} : B \to A$ by

$$f^{-1}(b) = \text{the unique } a \text{ such that } f(a) = b.$$

Exercise 3.12. Let $f : A \to B$ be a one-to-one correspondence.

1. Prove that f^{-1} is a function.
2. Prove that f^{-1} is one-to-one.
3. Prove that f^{-1} is onto.
4. Conclude that $f^{-1} : B \to A$ is a one-to-one correspondence.

Exercise 3.13. Let $f : A \to B$ be a one-to-one correspondence. By Exercise 3.12, $f^{-1} : B \to A$ is also a one-to-one correspondence.

1. Prove that $f^{-1} \circ f = i_A$.
2. Prove that $f \circ f^{-1} = i_B$.

3.6 Set equivalence

We are finally in a position to give a formal definition of the size of a set and to compare different sizes of sets. Informally speaking, if $f : A \to B$ is a one-to-one function, then since different elements of A are mapped to different elements of B, the size of B is at least as large as the size of A. On the other hand, if f is onto, then since each element in B has at least one element in A that is mapped to it, the size of B is no greater than the size of A. Thus, one-to-one correspondences provide us with a means to compare the sizes of sets. This key observation of Cantor led him to the notion of two sets being equivalent. Let us read how Cantor defines that two sets are equivalent.

We say that two aggregates M and N are "equivalent," in signs

$$M \sim N \quad \text{or} \quad N \sim M,$$

if it is possible to put them, by some law, in such a relation to one another that to every element of each one of them corresponds one and only one element of the other. [6, p. 86]

Next Cantor states that each set is equivalent to itself, and that if a set is equivalent to two other sets, then the two sets are also equivalent.

Every aggregate is equivalent to itself:

$$M \sim M.$$

If two aggregates are equivalent to a third, they are equivalent to one another; that is to say:

from $M \sim P$ and $N \sim P$ follows $M \sim N$. [6, p. 87]

Exercise 3.14. Prove the above two claims of Cantor.

3.7 Cardinality of a set, cardinal numbers

As we saw in the previous section, two sets A and B having the same size can be formalized by saying that the sets A and B are equivalent. All equivalent sets have the same size. One of the key breakthroughs of Cantor was to introduce new numbers, which he called *cardinal numbers*, measuring the size of sets. Let us read how Cantor defined the cardinality of a set.

Every aggregate M has a definite "power," which we will also call its "cardinal number."

We will call by the name "power" or "cardinal number" of M the general concept which, by means of our active faculty of thought, arises from the aggregate M when we make abstraction of the nature of its various elements m and of the order in which they are given.

We denote the result of this double act of abstraction, the cardinal number or power of M, by

$\overline{\overline{M}}$. [6, p. 86]

Nowadays it is more customary to denote the cardinal number of a set S by $|S|$.

Exercise 3.15. Describe in your own words Cantor's definition of a cardinal number. Given a set S consisting of ten round marbles, each of a different color, what is $|S|$?

In the next excerpt, Cantor connects the two key notions, that of cardinality and that of set equivalence.

Of fundamental importance is the theorem that two aggregates M and N have the same cardinal number if, and only if, they are equivalent: thus,

from $M \sim N$ we get $\overline{\overline{M}} = \overline{\overline{N}}$,

and

from $\overline{\overline{M}} = \overline{\overline{N}}$ we get $M \sim N$.

Thus the equivalence of aggregates forms the necessary and sufficient condition for the equality of their cardinal numbers. [6, pp. 87–88]

Exercise 3.16. Let S be the set of all perfect squares

$$\{0, 1,\ 4,\ 9,\ 16,\ 25,\ \ldots\}.$$

From Cantor's statement above, do S and $\mathbb{N}$ have the same cardinality? Justify your answer.

Exercise 3.17. Do $\mathbb{N}$ and $\mathbb{Z}$ have the same cardinality? Justify your answer.

Exercise 3.18. Do $\mathbb{N}$ and $\mathbb{N} \times \mathbb{N}$ have the same cardinality? Justify your answer. (Hint: Draw a picture of $\mathbb{N} \times \mathbb{N}$. Can you label each element of $\mathbb{N} \times \mathbb{N}$ by a unique natural number?)

Exercise 3.19. Does $\mathbb{Q}$ have the same cardinality as $\mathbb{N}$? Justify your answer. (Hint: Establish a one-to-one correspondence between $\mathbb{Q}$ and a subset of $\mathbb{Z} \times (\mathbb{N} - \{0\})$ and modify your solution to Exercise 3.18.)

3.8 Ordering of cardinal numbers

Some sets have larger size than others. Since cardinal numbers measure the size of sets, it is natural to speak about one cardinal number being less than the other. This is exactly what Cantor does in the next excerpt. We assume that Cantor's definition of "part" is the same as that of Dedekind given in Section 2.1.

If for two aggregates M and N with the cardinal numbers $\mathfrak{a} = \overline{\overline{M}}$ and $\mathfrak{b} = \overline{\overline{N}}$, both the conditions:

(a) There is no part of M which is equivalent to N,

(b) There is a part N_1 of N, such that $N_1 \sim M$,

are fulfilled, it is obvious that these conditions still hold if in them M and N are replaced by two equivalent aggregates M' and N'. Thus they express a definite relation of the cardinal numbers $\mathfrak{a}$ and $\mathfrak{b}$ to one another.

Further, the equivalence of M and N, and thus the equality of $\mathfrak{a}$ and $\mathfrak{b}$, is excluded...

Thirdly, the relation of $\mathfrak{a}$ to $\mathfrak{b}$ is such that it makes impossible the same relation of $\mathfrak{b}$ to $\mathfrak{a}$; for if in (a) and (b) the parts played by M and N are interchanged, two conditions arise which are contradictory to the former ones.

We express the relation of $\mathfrak{a}$ to $\mathfrak{b}$ characterized by (a) and (b) by saying: $\mathfrak{a}$ is "less" than $\mathfrak{b}$ or $\mathfrak{b}$ is "greater" than $\mathfrak{a}$; in signs

$$\mathfrak{a} < \mathfrak{b} \text{ or } \mathfrak{b} > \mathfrak{a}.$$ [6, pp. 89–90]

Exercise 3.20. Describe in your own words what it means for two cardinals $\mathfrak{a} = |A|$ and $\mathfrak{b} = |B|$ to be in the relation $\mathfrak{a} < \mathfrak{b}$.

Cantor goes on to state the following:

We can easily prove that,

if $\mathfrak{a} < \mathfrak{b}$ and $\mathfrak{b} < \mathfrak{c}$, then we always have $\mathfrak{a} < \mathfrak{c}$. [6, pp. 90]

Exercise 3.21. Prove the above claim of Cantor.

Exercise 3.22.

1. Let $\mathfrak{a}$ and $\mathfrak{b}$ be two cardinal numbers. Modify Cantor's definition of $\mathfrak{a} < \mathfrak{b}$ to define $\mathfrak{a} \leq \mathfrak{b}$. (Hint: Examine what happens if you drop condition (a) from Cantor's definition of $\mathfrak{a} < \mathfrak{b}$.)

2. Prove that $\mathfrak{a} \leq \mathfrak{a}$.

3. Prove that if $\mathfrak{a} \leq \mathfrak{b}$ and $\mathfrak{b} \leq \mathfrak{c}$, then $\mathfrak{a} \leq \mathfrak{c}$.

4. Do you think that $\mathfrak{a} \leq \mathfrak{b}$ and $\mathfrak{b} \leq \mathfrak{a}$ imply $\mathfrak{a} = \mathfrak{b}$? Explain your reasoning. (Hint: This is not as trivial as it might look.[6])

A fundamental property of cardinal numbers is that they are *comparable*. Namely, if $\mathfrak{a}$ and $\mathfrak{b}$ are cardinal numbers, then it is the case that $\mathfrak{a} < \mathfrak{b}$, $\mathfrak{b} < \mathfrak{a}$, or $\mathfrak{a} = \mathfrak{b}$. This property is known as the *law of trichotomy*. Its proof is based on an important principle in mathematics known as the *Axiom of Choice*, and is beyond the scope of this project. We will encounter the Axiom of Choice again in Exercise 3.27.

3.9 Finite and infinite sets

Now that we have a good understanding of cardinal numbers and how they compare to each other, we are ready to define finite and infinite sets. Intuitively, a set is finite if it consists of finitely many elements and it is infinite otherwise. This intuitive idea can be formalized by saying that a set A is *finite* or has *finite cardinality* if there is $n \in \mathbb{N}$ such that A is equivalent to the set $\{0, \ldots, n-1\} \subset \mathbb{N}$. On the other hand, A is *infinite* or has *infinite cardinality* if A is not equivalent to $\{0, \ldots, n-1\}$ for any $n \in \mathbb{N}$. If A is a finite set, then we call its cardinal number $|A|$ *finite*.

Exercise 3.23.

1. Prove that A is an infinite set if, and only if, A is not equivalent to any finite subset of $\mathbb{N}$.

2. Describe finite cardinal numbers.

The definition of finite and infinite sets given above is actually how Cantor defined them. Dedekind gave a different definition of finite and infinite sets. Note that Dedekind calls equivalent sets *similar*.

[6] In fact, this is known as the Cantor-Bernstein-Schröder theorem, named after Georg Cantor, Felix Bernstein (1878–1956), and Ernst Schröder (1841–1902). The history and different proofs of the theorem can be found in [7].

A system S is said to be *infinite* when it is similar to a proper part of itself; in the contrary case S is said to be a *finite* system. [10, p. 63]

Exercise 3.24.

1. According to Dedekind's definition, are $\mathbb{N}$, $\mathbb{Z}$, and $\mathbb{Q}$ infinite sets? Explain.
2. Compare Cantor's and Dedekind's definitions.

We will say more about the comparison of Cantor's and Dedekind's definitions of finite and infinite sets later in the project.

3.10 Countable sets

Our next goal is to identify a special infinite cardinal number that Cantor calls aleph-zero. Note that Cantor calls infinite sets and infinite cardinal numbers *transfinite*.

The first example of a transfinite aggregate is given by the totality of finite cardinal numbers ν; we call its cardinal number "Aleph-zero," and denote it by $\aleph_0$; [6, pp. 103–104]

Note that $\aleph_0$ is the first letter of the Hebrew alphabet.

Exercise 3.25. Show that $\aleph_0$ is the cardinal number of $\mathbb{N}$; that is, show that $|\mathbb{N}| = \aleph_0$.

Cantor's next claim is that $\aleph_0$ is greater than any finite cardinal number:

The number $\aleph_0$ is greater than any finite number μ:

$$\aleph_0 > \mu.$$ [6, p. 104]

Exercise 3.26. Prove the above claim of Cantor.

In modern terminology, a set whose cardinal number is $\aleph_0$ is called *countably infinite*, and sets that are either finite or countably infinite are called *countable*. There are many examples of countable sets. For instance, finite sets, $\mathbb{N}$, $\mathbb{Z}$, and $\mathbb{Q}$ are all examples of countable sets. The next natural question is whether there exist *uncountable* sets, and if so, how to construct them.

3.11 Uncountable sets and higher levels of infinity

Before discussing the issue of existence of uncountable sets, we address the following claim of Cantor that $\aleph_0$ is the smallest among infinite cardinal numbers.

...$\aleph_0$ is the least transfinite cardinal number. If $\mathfrak{a}$ is any transfinite cardinal number different from $\aleph_0$, then

$$\aleph_0 < \mathfrak{a}.$$ [6, p. 104]

Exercise 3.27. Prove the above claim of Cantor. (Hint: Let A be an infinite set and let $\mathfrak{a} = |A|$. Can you define a one-to-one function from $\mathbb{N}$ to A? What can you conclude about the relationship between $\aleph_0$ and $\mathfrak{a}$? To define such a function you will need to make countably-many choices of some elements of A. The ability to make such choices depends on what is known as the *Axiom of Choice* — an important but rather controversial principle in mathematics. A detailed discussion of this axiom and its history can be found in [1] and in the references therein.)

Exercise 3.28. Reexamine Exercise 3.24.2. In particular, can you use your answer to Exercise 3.27 to convince yourself that Cantor's and Dedekind's definitions of finite and infinite sets are equivalent? Explain.

We are finally in a position to show that there exist uncountable sets. In fact, we will show that the set of real numbers is uncountable. We will also examine a method that allows us to build sets of larger and larger cardinality.

Let $[0, 1]$ be the set of all real numbers between 0 and 1. Our goal is to show that $[0, 1]$ is uncountable. For this we will employ what is known as *Cantor's diagonalization method*. We start by representing elements of $[0, 1]$ as infinite decimals which do not end in infinitely repeating 9's. Our proof is by contradiction. Suppose that $[0, 1]$ is countable. Then, since $[0, 1]$ is infinite, $\mathbb{N} \sim [0, 1]$. Therefore, to each infinite decimal one can assign a unique natural number, so the infinite decimals can be enumerated as follows:

$$\begin{array}{c} .a_{11}a_{12}\ldots a_{1n}\ldots \\ .a_{21}a_{22}\ldots a_{2n}\ldots \\ \vdots \\ .a_{n1}a_{n2}\ldots a_{nn}\ldots \\ \vdots \end{array}$$

Exercise 3.29.

1. Construct an infinite decimal $.b_1b_2\ldots b_n\ldots$ such that $a_{nn} \neq b_n$ for each positive n.

2. Does this contradict the above assumption that $[0, 1]$ was countable? Explain.

3. What can you conclude from the above contradiction? Explain.

4. Is the cardinal number of $[0, 1]$ strictly greater than $\aleph_0$? Justify your answer.

5. Is $|\mathbb{R}|$ strictly greater than $\aleph_0$? Justify your answer.

As a result, we see that not every infinite set is countable. The next natural question is whether we can create larger and larger infinite sets. The answer to this question is again *yes*. The proof of this important fact is based on the *generalized version of Cantor's diagonalization method*.

Exercise 3.30.

1. Let A be a set and let $P(A)$ be the powerset of A. Prove the following claim of Cantor:

$$|P(A)| > |A|.$$

   Hint: Employ Cantor's generalized diagonalization method. Assume that $A \sim P(A)$. Then there is a one-one correspondence $f : A \to P(A)$. Consider the set $B = \{a \in A : a \notin f(a)\}$. Can you deduce that $B \in P(A)$ is not in the range of f? Does this imply a contradiction?[7]

2. Describe an infinite increasing sequence of infinite cardinal numbers.

References

[1] "Axiom of choice," `http://en.wikipedia.org/wiki/Axiom_of_choice`

[2] "Bertrand Russell," `http://en.wikipedia.org/wiki/Bertrand_Russell`

[3] Bezhanishvili, G., *Peano Arithmetic*, `http://www.cs.nmsu.edu/historical-projects/`

[4] Bezhanishvili, G. and Fussner, W., *An Introduction to Symbolic Logic,* Loci: Convergence, published on-line in this series. Also available from `http://www.cs.nmsu.edu/historical-projects/`

[5] Cantor, G., *Beiträge zur Begründung der transfiniten Mengenlehre. I*, Mathematische Annalen **46** (1895), 481–512.

[6] Cantor, G., *Contributions to the Founding of the Theory of Transfinite Numbers*, Philip Jourdain (translator), The Open Court Publishing Company, La Salle, Illinois, 1941.

[7] "Cantor-Bernstein-Schroeder theorem," `http://en.wikipedia.org/wiki/Cantor-Bernstein-Schroeder_theorem`

[8] Dauben, J., *Georg Cantor: His Mathematics and Philosophy of the Infinite*, Princeton University Press, 1990.

[9] Dedekind, R., *Was sind und was sollen die Zahlen?*, Vieweg, Braunschweig, 1888.

[10] Dedekind, R., *Essays on the Theory of Numbers*, W.W. Beman (translator), 3rd Printing, The Open Court Publishing Company, Chicago, 1924.

[11] Dunham, W., *Journey Through Genius: The Great Theorems of Mathematics*, John Wiley & Sons Inc., New York, 1990.

[12] "Georg Cantor," `http://en.wikipedia.org/wiki/Georg_Cantor`

[13] "Giuseppe Peano," `http://en.wikipedia.org/wiki/Giuseppe_Peano`

[14] "Gottlob Frege," `http://en.wikipedia.org/wiki/Frege`

[7] Note that there are similarities between Russell's paradox and Cantor's generalized diagonalization method!

[15] Grattan-Guinness, I., *The Search for Mathematical Roots 1870-1940*. Princeton University Press, 2000.

[16] "Habilitation," http://en.wikipedia.org/wiki/Habilitation

[17] Hollingdale, S., *Makers of Mathematics*, Penguin Books, New York, 1994.

[18] Kennedy, H. C., *Peano. Life and works of Giuseppe Peano*, Studies in the History of Modern Science, Vol. 4, D. Reidel Publishing Co., Dordrecht, 1980.

[19] Laubenbacher, R., Pengelley, D., *Mathematical Expeditions: Chronicles by the Explorers*, Springer Verlag, New York, 1999.

[20] Lodder, J., *Deduction through the Ages: A History of Truth,* Loci: Convergence, published on-line in this series. Also available from http://www.cs.nmsu.edu/historical-projects/

[21] Mendelson, E., *Introduction to Mathematical Logic*, Fourth Edition, Chapman & Hall, 1997.

[22] "Richard Dedekind," http://en.wikipedia.org/wiki/Richard_Dedekind

[23] "Russell's paradox," http://en.wikipedia.org/wiki/Russell's_paradox

[24] Sonar, Th., *Brunswick's second mathematical star: Richard Dedekind (1831–1916)*. Math. Intelligencer 34 (2012), no. 2, 63-67.

Notes to the instructor

This project is a self-contained treatment of the topics from elementary set theory typically covered in a first course in discrete mathematics. It may be used as a text for the set theory unit of a standard one semester course at the freshman or sophomore level, and should require approximately three/four weeks of class time to complete. A first course in discrete mathematics typically covers logic, set theory, and number theory units. This project may be covered immediately after the logic unit or at the end of the course when both the logic and number theory units are already covered. No specific prerequisites are assumed other than a basic familiarity with pre-calculus and/or college algebra.

The first part of the project covers standard material about basic properties of sets; membership, subset, and equality relations between sets; and set operations. It ends with the discussion of Russell's paradox, which typically stimulates interesting in-class discussion. In particular, the purpose of Exercise 2.24.5 is to stimulate such a discussion. A rigorous treatment of it is, of course, beyond the scope of this project.

The second part of the project covers set equivalence, cardinal numbers, and countable and uncountable sets. Several exercises, especially in the second part of the project, are slightly open-ended. In our opinion, this stimulates independent thinking, as well as provides an opportunity for further in-class discussion. In our experience, such discussions enhance students' understanding of the material.

Typically, students have little to no difficulty understanding the material presented in the first part of the project. However, the material in the second part of the project requires instructor guidance. It is advisable to have a detailed class discussion on some of the excerpts from Cantor and Dedekind, as well as on set equivalence, the cardinality of a set, and countable and uncountable sets. In particular, the open-ended exercises about the Cantor-Bernstein-Schröder theorem, the Cantor and Dedekind definitions of finite and infinite sets, and the Axiom of Choice provide the instructor with an opportunity to have a more detailed class discussion on these topics.

The project may be covered in its entirety or only parts of it may be selected for class discussions. Alternatively the instructor may wish to assign it as outside class reading. Each part of the project may be assigned independently, and each requires approximately one/two weeks of class time. The project has a large variety of exercises. The instructor may wish to pick and choose which exercises to assign, depending on what parts of the project will be covered.