Chapter 2. Real Numbers [*]

1 Algebraic Properties of Real Numbers

Number systems

$$\mathbb{N} \subset \mathbb{Z} \subset \mathbb{Q} \subset \mathbb{R} \subset \mathbb{C}.$$

The set of integers $\mathbb{Z}$ is a (Abelian) group with addition. The set of rational numbers $\mathbb{Q}$ is a *field* with the addition and multiplication.

(I) $\mathbb{Q}$ is a group under addition.
There is a binary operation $+ : \mathbb{Q} \times \mathbb{Q} \to \mathbb{Q}$ with the following properties,

(i) there is an identity $x + 0 = 0 + x = x, \quad \forall x \in \mathbb{Q};$

(ii) inverse $x + (-x) = (-x) + x = 0, \quad \forall x \in \mathbb{Q};$

(iii) associativity $(x + y) + z = x + (y + z), \quad \forall x, y, z \in \mathbb{Q};$

(iv) commutativity $x + y = y + x, \quad \forall x, y \in \mathbb{Q}.$

Note. We call $\mathbb{Q}$ is a *group* under the *group law* '+', since the properties (a), (b), (c) hold; and more specially, we say it is an *Abelian* (or, *commutative*) *group*, because of (d).

(II) $\mathbb{Q}_0 := \mathbb{Q} \setminus \{0\}$ is a group under multiplication.
There is a binary operation $\times : \mathbb{Q} \times \mathbb{Q} \to \mathbb{Q}$ with the following properties

(v) there is an identity $x \times 1 = 1 \times x = x, \quad \forall x \in \mathbb{Q}_0;$

(vi) inverse $x \times x^{-1} = x^{-1} \times x = 1 \quad \forall x \in \mathbb{Q}_0;$

(vii) associativity $(x \times y) \times z = x \times (y \times z), \quad \forall x, y, z \in \mathbb{Q}_0;$

(viii) commutativity $x \times y = y \times x, \quad \forall x, y \in \mathbb{Q}_0.$

[*] Lecture notes for CUHKSZ course MAT2006: Elementary Real Analysis.

(I,II) Connection between addition and multiplication.

$$\text{(iv)} \quad \text{distribution law} \quad x \times (y+z) = x \times y + x \times z, \qquad \forall z, y, z \in \mathbb{Q}.$$

(III) Order: There is a binary relation '$\leq$' between any two numbers $x, y \in \mathbb{Q}$, one can determine whether $x \leq y$ or not.

$$\begin{array}{rllll} \text{(x)} & \text{reflexivity} & x \leq x & \forall x \in \mathbb{Q}; & \\ \text{(xi)} & \text{antisymmetry} & x \leq y \quad \text{and} \quad y \leq x & \Longrightarrow x = y; \\ \text{(xii)} & \text{transitivity} & x \leq y \quad \text{and} \quad y \leq z & \Longrightarrow x \leq z; \\ \text{(xiii)} & \text{connexity} & x \leq y \quad \text{or} \quad y \leq x & \forall x, y \in \mathbb{Q}. \end{array}$$

Note. In general, we say a binary relation defined over a set is called an *partial order* if the properties (x)–(xii) are fulfilled. In addition, if (xiii) is also satisfies, we call it a *total order* (or, *linear order*). A typical example of partial order but not total order relation is the inclusion (i.e. subset) relation of sets.

(I,III) and (II,III) Connections between addition (multiplication) and order.

$$\begin{array}{rl} \text{(xiv)} & x \leq y \quad \Longrightarrow \quad x + z \leq y + z \qquad \forall x, y, z \in \mathbb{Q}; \\ \text{(xv)} & 0 \leq x \quad \text{and} \quad 0 \leq y \quad \Longrightarrow \quad 0 \leq x \times y \qquad x, y \in \mathbb{Q}; \end{array}$$

In algebra, the set $\mathbb{Q}$ equipped with the two operations '+' and '×' and with the above properties (i)–(iv), is called a *field*. Moreover, because of the order properties (x)–(xv) are also valid, we say $\mathbb{Q}$ is an *ordered field*. Hence, from the algebraic (or arithmetic) point of view, the rational number set $\mathbb{Q}$ is perfect. However, it is not suitable for the analysis purpose, because it is lack of another important property, namely, the *completeness*.

The real number set $\mathbb{R}$ is an extension of the rational numbers $\mathbb{Q}$: $\mathbb{R}$ is also an ordered field, and it is complete in a sense that every length along the number line – such as $\sqrt{2}$ – to correspond to a real number and vice versa.

2 The Axiom of Completeness

To introduce the Axiom of Completeness, we shall make use of the term least upper bound of a bounded above set A.

2.1 Least upper bound and greatest lower bound

Definition 1 (bounded set)**.** A set $A \subset \mathbb{R}$ is *bounded above* if there exists a number $b \in \mathbb{R}$ such that $a \leq b$ for all $a \in A$. The number b is called an *upper bound* for A.

Similarly, the set A is *bounded below* if there exists a lower bound $l \in \mathbb{R}$ satisfying $l \leq a$ for every $a \in A$. The number l is called a *lower bound* for A.

Definition 2 (least upper bound, or, supremum)**.** A real number s is the *least upper bound* for a set $A \subset \mathbb{R}$ if it meets the following two criteria:

(i) s is an upper bound for A;

(ii) if b is any upper bound for A, then $s \leq b$.

The least upper bound is also frequently called the *supremum* of the set A, and we write $s = \sup A$.

Remark. The *greatest lower bound* or *infimum* for A, denoted by $\inf A$, is defined in a similar way.

Although a set can have a host of upper bounds, it can have only one least upper bound. If s_1 and s_2 are both least upper bounds for a set A, then by property (ii) in Definition 2 we can assert $s_1 \leq s_2$ and $s_2 \leq s_1$. The conclusion is that $s_1 = s_2$ and **least upper bounds are unique**.

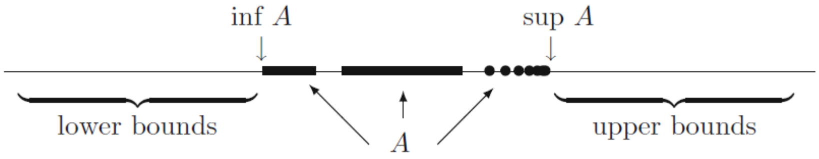


Figure 1: Definition of $\sup A$ and $\inf A$.

Example 2.1. Let

$$A = \left\{ \frac{1}{n} \,|\, n \in \mathbb{N} \right\} = \left\{ 1, \frac{1}{2}, \frac{1}{3}, \dots \right\}.$$

The set A is bounded above and below. We have

$$\sup A = 1 \quad \text{and} \quad \inf A = 0.$$

An important lesson to take from the above Example is that $\sup A$ and $\inf A$ may or may not be elements of the set A. This issue is tied to understanding the crucial difference between the maximum and the supremum (or the minimum and the infimum) of a given set.

Definition 3 (maximum and minimum)**.** A real number a_0 is a *maximum* of the set A if a_0 is an element of A and $a_0 \geq a$ for all $a \in A$. Similarly, a number a_1 is a *minimum* of A if $a_1 \in A$ and $a_1 \leq a$ for every $a \in A$.

Example 2.2. Consider the open, closed and half-open-half-closed intervals

$$\begin{aligned} A &:= (0,1) = \{x \in \mathbb{R} \,|\, 0 < x < 1\}, \\ B &:= [0,1] = \{x \in \mathbb{R} \,|\, 0 \leq x \leq 1\}, \\ C &:= (0,1] = \{x \in \mathbb{R} \,|\, 0 < x \leq 1\}, \\ D &:= [0,1) = \{x \in \mathbb{R} \,|\, 0 \leq x < 1\}. \end{aligned}$$

We have

$$\sup A = \sup B = \sup C = \sup D = 1, \qquad \inf A = \inf B = \inf C = \inf D = 0,$$

and

$$\min B = \min D = 0, \qquad \max B = \max C = 1,$$

$$\max A, \quad \min A, \quad \min A, \quad \max D \qquad \text{do not exist.}$$

Thus, the supremum can exist and not be a maximum, but when a maximum exists, then it is also the supremum.

2.2 Axiom of Completeness (AoC)

The *Axiom of completeness* is the defining difference between $\mathbb{R}$ and $\mathbb{Q}$.

Axiom of Completeness. *Every nonempty set of real numbers that is bounded above has a least upper bound.*

Although we can see now that not every nonempty bounded set contains a maximum, the Axiom of Completeness asserts that every such set does have a least upper bound. We are not going to prove this. An *axiom* in mathematics is an accepted assumption, to be used without proof. Preferably, an axiom should be an elementary statement about the system in question that is so fundamental that it seems to need no justification.

Let's see why it is not a valid statement about $\mathbb{Q}$.

Example 2.3. Consider again the set

$$S = \{x \in \mathbb{Q} \,|\, x^2 < 2\},$$

and pretend for the moment that our world consists only of rational numbers. The set S is certainly bounded above, for instance 2 is an upper bound. For any $x_0 \in \mathbb{Q}$ being an upper bound of S, then $x_0^2 > 2$, since there is no rational number x such that $x^2 = 2$, according to Theorem 1 of Chapter 1. One can verify that

$$x_1 = \frac{x_0}{2} + \frac{1}{x_0} \in \mathbb{Q}, \qquad x_1 < x_0, \qquad x_1^2 > 2,$$

and hence x_1 is also an upper bound which is smaller than x_0. There is no least upper bound of S in $\mathbb{Q}$ (proof by contradiction.)

In the real numbers, there is a least upper bound of S. Back in $\mathbb{R}$, the Axiom of Completeness states that we may set $\alpha = \sup S$ and be confident that such a number exists. In the next section, we will prove that $\alpha^2 = 2$. But according to Theorem 1 of Chapter 1, this implies α is not a rational number. If we are restricting our attention to only rational numbers, then α is not an allowable option for $\sup S$, and the search for a least upper bound goes on indefinitely. Whatever rational upper bound is discovered, it is always possible to find one smaller.

2.3 Some properties of supremum and infimum.

Example 2.4. Let $A \subset \mathbb{R}$ be nonempty and bounded above, and let $c \in \mathbb{R}$. Define the set $c + A$ by

$$c + A = \{c + a \,|\, a \in A\}.$$

Then $\sup(c + A) = c + \sup A$.

To properly verify this we focus separately on each part of Definition 2. Setting $s = \sup A$, we see that $a \leq s$ for all $a \in A$, which implies $c + a \leq c + s$ for all $a \in A$. Thus, $c + s$ is an upper bound for $c + A$ and condition (i) is verified.

For (ii), let b be an arbitrary upper bound for $c + A$; i.e., $c + a \leq b$ for all $a \in A$. This is equivalent to $a \leq b - c$ for all $a \in A$, from which we conclude that $b - c$ is an upper bound for A. Because s is the least upper bound of A, $s \leq b - c$, which can be rewritten as $s + c \leq b$. This verifies part (ii) of Definition 2, and we conclude $\sup(c + A) = c + \sup A$.

Exercise 1. Let A and B be two nonempty subsets of $\mathbb{R}$, both bounded above. Show that

$$\sup(A \cup B) = \max\{\sup A, \sup B\}.$$

Lemma 1. *Assume $s \in \mathbb{R}$ is an upper bound for a set $A \subset \mathbb{R}$. Then, $s = \sup A$ if and only if, for every choice of $\epsilon > 0$, there exists an element $a \in A$ satisfying $s - \epsilon < a$.*

Proof. There are two directions to be verified.

($\Rightarrow$) For the forward direction, we assume $s = \sup A$ and consider $s - \epsilon$, where $\epsilon > 0$ has been arbitrarily chosen. Because $s - \epsilon < s$, part (ii) of Definition 2 implies that $s - \epsilon$ is not an upper bound for A. If this is the case, then there must be some element $a \in A$ for which $s - \epsilon < a$ (because otherwise $s - \epsilon$ would be an upper bound). This proves the lemma in one direction.

($\Leftarrow$) Conversely, assume s is an upper bound with the property that no matter how $\epsilon > 0$ is chosen, $s - \epsilon$ is no longer an upper bound for A. Notice that what this implies is that if b is any number less than s, then b is not an upper bound. (Just let $\epsilon = s - b$.) To prove that $s = \sup A$, we must verify part (ii) of Definition 2. Because we have just argued that any number smaller than s cannot be an upper bound, it follows that if b is some other upper bound for A, then $s \leq b$. □

2.4 Dedekind's Cut Property*

The *Dedekind*[1] *Cut Property* of the real numbers is the following:

If A and B are nonempty, disjoint sets with $A \cup B = \mathbb{R}$ and $a < b$ for all $a \in A$ and $b \in B$, then there exists $c \in \mathbb{R}$ such that $x \leq c$ whenever $x \in A$ and $x \geq c$ whenever $x \in B$.

Theorem 2. *The Axiom of Completeness is equivalent to Dedekind's Cut Property.*

Proof. ($\Rightarrow$) A is bounded above since any $b \in B$ is an upper bound. By the AoC, there is a least upper bound of A and we set $c = \sup A$. Then $x \leq c$ for every $x \in A$ since c is an upper bound of A. Moreover, since every $b \in B$ is an upper bound of A, we have $c \leq b$ for all $b \in B$.

($\Leftarrow$) Assuming Cut Property is true. Let $E \subset \mathbb{R}$ which is bounded above, we are going to show that $\sup E$ exists. If $\max E$ exists, it must be equal to $\sup E$ and nothing to proof. Hence, we shall assume $\max E$ does not exist. Denote B to be the set of upper bounds of E, that is

$$B = \{b \in \mathbb{R} \,|\, x \leq b, \quad \forall x \in E\},$$

and set

$$A = \mathbb{R} \setminus B.$$

Since $\max E$ does not exist, every element in E is not an upper bound of E, we have $E \subset A$. Note that A and B are disjoint and $A \cup B = \mathbb{R}$. To see they form a cut of $\mathbb{R}$, we must show that $a < b$ for all $a \in A$ and $b \in B$. For every $a \in A$, then $a \notin B$, which means a is not

[1] Richard Dedekind (1831–1916) was a German mathematician who made important contributions to abstract algebra, axiomatic foundation for the natural numbers, algebraic number theory and the definition of the real numbers.

an upper bound of E. Thus there exists $x \in E$ such that $a < x$. (For otherwise, a is an upper bound by definition). Since every $b \in B$ is an upper bound of E, we have $x \leq b$, and consequently $a < b$. Therefore A and B form a cut of $\mathbb{R}$. According to the Cut Property, there exists $c \in \mathbb{R}$ such that

$$a \leq c \leq b, \qquad \forall a \in A, \quad \forall b \in B.$$

We assert that $c = \sup E$.

(i) The facts that c is an upper bound of A and $E \subset A$ imply that c is also an upper bound of E.

(ii) The fact that $c \leq b$ for every $b \in B$ means c is the least upper bound of E.

Hence, the least upper bound of E exists and $\sup E = c$.

□

Remark. Some authors call the property that least upper bound for a bounded-above nonempty set exists as the Least Upper Bound Principle. Since the Least Upper Bound Principle and the Cut Property are equivalent, one may take the Cut Property as the Axiom of Completeness alternatively; see for example [Zorich].

3 Consequences of Axiom of Completeness

3.1 Cantor's nested interval property.

Theorem 3 (Nested Interval Property)**.** *For each $n \in \mathbb{N}$, assume we are given a closed interval $I_n = [a_n, b_n] = \{x \in \mathbb{R}, \,|\, a_n \leq x \leq b_n\}$. Assume also that each I_n contains I_{n+1}. Then, the resulting nested sequence of closed intervals*

$$I_1 \supset I_2 \supset I_3 \supset \cdots$$

has a nonempty intersection; that is $\bigcap_{n=1}^{\infty} I_n \neq \emptyset$.

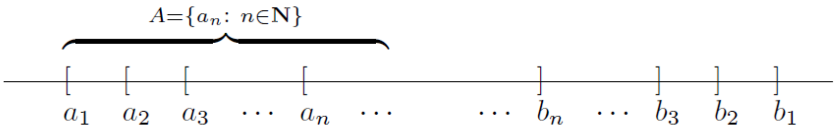
Proof. In order to show that $\bigcap_{n=1}^{\infty} I_n$ is not empty, we are going to use the Axiom of Completeness (AoC) to produce a single real number x satisfying $x \in I_n$ for every $n \in \mathbb{N}$. Now, AoC is a statement about bounded sets, and the one we want to consider is the set

$$A = \{a_n \,|\, n \in \mathbb{N}\}$$

of left-hand endpoints of the intervals.

Because the intervals are nested, we see that every b_n serves as an upper bound for A. Thus, we are justified in setting

$$x = \sup A$$



Now, consider a particular $I_n = [a_n, b_n]$. Because x is an upper bound for A, we have $a_n \leq x$. The fact that each b_n is an upper bound for A and that x is the least upper bound implies $x \leq b_n$.

Altogether then, we have $a_n \leq x \leq b_n$, which means $x \in I_n$ for every choice of $n \in \mathbb{N}$. Hence, $x \in \bigcap_{n=1}^{\infty} I_n$, and the intersection is not empty. □

3.2 The density of $\mathbb{Q}$ in $\mathbb{R}$.

The set $\mathbb{Q}$ is an extension of $\mathbb{N}$, and $\mathbb{R}$ in turn is an extension of $\mathbb{Q}$. The next few results indicate how $\mathbb{N}$ and $\mathbb{Q}$ sit inside of $\mathbb{R}$.

Theorem 4 (Archimedean Property)**.** (i) *Given any number $x \in \mathbb{R}$, there exists an $n \in \mathbb{N}$ satisfying $n > x$.*

(ii) *Given any real number $y > 0$, there exists an $n \in \mathbb{N}$ satisfying $1/n < y$.*

Proof. Assume, for contradiction, that $\mathbb{N}$ is bounded above. By the Axiom of Completeness (AoC), $\mathbb{N}$ should then have a least upper bound, and we can set $\alpha = \sup \mathbb{N}$. If we consider $\alpha - 1$, then we no longer have an upper bound, and therefore there exists an $n \in \mathbb{N}$ satisfying $\alpha - 1 < n$. But this is equivalent to $\alpha < n + 1$. Because $n + 1 \in \mathbb{N}$, we have a contradiction to the fact that α is supposed to be an upper bound for $\mathbb{N}$. (Notice that the contradiction here depends only on AoC and the fact that $\mathbb{N}$ is closed under addition.)

Part (ii) follows from (i) by letting $x = 1/y$.

□

This familiar property of $\mathbb{N}$ is the key to an extremely important fact about how $\mathbb{Q}$ fits inside of $\mathbb{R}$.

Theorem 5 (Density of $\mathbb{Q}$ in $\mathbb{R}$)**.** *For every two real numbers a and b with $a < b$, there exists a rational number r satisfying $a < r < b$.*

Proof. We want to show that there exists $m, n \in \mathbb{Z}$ such that

$$a < \frac{m}{n} < b. \tag{3.1}$$

Using the Archimedean Property (Theorem 4), we may pick $n \in \mathbb{N}$ large enough so that

$$\frac{1}{n} < b - a. \tag{3.2}$$

Inequality (3.1) (which we are trying to prove) is equivalent to $na < m < nb$. With n already chosen, the idea now is to choose m to be the smallest integer greater than na. In other words, pick $m \in \mathbb{Z}$ so that

$$(m-1) \leq na < m \tag{3.3}$$

Now, the second inequality in (3.3) immediately yields $a < m/n$, which leaves us to show $m/n < b$. Noting that inequality (3.2) is equivalent to $a < b - 1/n$. Then the first inequality in (3.3) yields

$$\begin{aligned} m \leq& na + 1 \\ <& n\left(b - \frac{1}{n}\right) + 1 \\ =& nb, \end{aligned}$$

which implies $m/n < b$, and we have $a < m/n < b$, as desired. □

Theorem 5 is paraphrased by saying that $\mathbb{Q}$ is *dense* in $\mathbb{R}$. We can use this result to show that the irrational numbers are dense in $\mathbb{R}$ as well.

Corollary 6. *Given any two real numbers $a < b$, there exists an irrational number t satisfying $a < t < b$.*

Hint. Applying Theorem 5 to the two real numbers $a - \sqrt{2}$ and $b - \sqrt{2}$.

3.3 The existence of square roots

Theorem 7. *There exists a real number $\alpha \in \mathbb{R}$ satisfying $\alpha^2 = 2$.*

Proof. Consider the set

$$T = \{t \in \mathbb{R}, \mid t^2 < 2\}$$

and set $\alpha = \sup T$. We are going to prove $\alpha^2 = 2$ by ruling out the possibilities $\alpha^2 < 2$ and $\alpha^2 > 2$. Keep in mind that there are two parts to the definition of $\sup T$, and they will both be important. (This always happens when a supremum is used in an argument.) The strategy is to demonstrate that $\alpha^2 < 2$ violates the fact that α is an upper bound for T, and $\alpha^2 > 2$ violates the fact that it is the least upper bound.

If we assume $\alpha^2 < 2$. In search of an element of T that is larger than α, write

$$\left(\alpha + \frac{1}{n}\right)^2 = \alpha^2 + \frac{2\alpha}{n} + \frac{1}{n^2} < \alpha^2 + \frac{2\alpha + 1}{n}.$$

By the Archimedean Property, we can choose $n_0 \in \mathbb{N}$ large enough so that

$$\frac{1}{n_0} < \frac{2 - \alpha^2}{2\alpha + 1}.$$

This implies $\frac{2\alpha+1}{n_0} < 2 - \alpha^2$, and consequently that

$$\left(\alpha + \frac{1}{n_0}\right)^2 < \alpha^2 + \frac{2\alpha + 1}{n_0} < 2.$$

Thus, $\alpha + \frac{1}{n_0} \in T$, contradicting the fact that α is an upper bound for T. We conclude that $\alpha^2 < 2$ cannot happen.

If we assume $\alpha^2 > 2$. We shall find an upper bound of T that is smaller than α. We write

$$\left(\alpha - \frac{1}{n}\right)^2 = \alpha^2 - \frac{2\alpha}{n} + \frac{1}{n^2} > \alpha^2 - \frac{2\alpha}{n}.$$

By the Archimedean Property again, we can choose $n_1 \in \mathbb{N}$ large enough so that

$$\frac{1}{n_1} < \frac{\alpha^2 - 2}{2\alpha}.$$

This implies $\frac{2\alpha}{n_1} < \alpha^2 - 2$, and consequently that

$$\left(\alpha - \frac{1}{n_1}\right)^2 > \alpha^2 - (2 - \alpha^2) = 2.$$

Thus $\alpha - \frac{1}{n_1}$ is an upper bound of T that is smaller than α, contradicting the fact that α is the least upper bound for T. We conclude that $\alpha^2 < 2$ cannot neither happen. Hence, we must have $\alpha^2 = 2$. □

A small modification of this proof can be made to show that $\sqrt{x}$ exists for any $x \geq 0$. The binomial formula for $(a + b)^m$ can be used to show that $\sqrt[m]{x}$ exists for arbitrary values of $m \in \mathbb{N}$ and $x \geq 0$.

4 Cardinality

At the moment, we have an image of $\mathbb{R}$ as consisting of rational and irrational numbers, continuously packed together along the real line. We have seen that both $\mathbb{Q}$ and $\mathbb{I}$ (the set of irrationals) are dense in $\mathbb{R}$, meaning that in every interval (a, b) there exist rational and irrational numbers alike. Mentally, there is a temptation to think of $\mathbb{R}$ and ii as being intricately mixed together in equal proportions, but this turns out not to be the case. In a way that Cantor made precise, the irrational numbers far outnumber the rational numbers in making up the real line.

4.1 1–1 Correspondence

The term cardinality is used in mathematics to refer to the size of a set. The cardinalities of finite sets can be compared simply by attaching a natural number to each set. But how might we draw this same conclusion without referring to any numbers? Cantor's idea was to attempt to put the sets into a 1–1 correspondence with each other. The advantage of this method of comparing the sizes of sets is that it works equally well on sets that are infinite.

Definition 4. A function $f : A \to B$ is *one-to-one* (or *injective*) if $a_1 \neq a_2$ in A implies that $f(a_1) \neq f(a_2)$ in B. The function f is *onto* (or *surjective*) if, given any $b \in B$, it is possible to find an element $a \in A$ for which $f(a) = b$.

If f is both 1–1 and onto, it is a *1–1 correspondence* (or *bijection*) between the two sets A and B.

Definition 5. The set A has the same *cardinality* as B if there exists $f : A \to B$ that is a 1–1 correspondence. In this case, we write $A \sim B$.

Example 4.1. (i) If we let $E = \{2, 4, 6, \dots\}$ be the set of even natural numbers, then we can show $\mathbb{N} \sim E$. To see why, let $f : \mathbb{N} \to E$ be given by $f(n) = 2n$.

It is certainly true that E is a proper subset of $\mathbb{N}$, and for this reason it may seem logical to say that E is a "smaller" set than $\mathbb{N}$. This is one way to look at it, but it represents a point of view that is heavily biased from an overexposure to finite sets. The definition of cardinality is quite specific, and from this point of view E and $\mathbb{N}$ are equivalent.

(ii) To make this point again, note that although $\mathbb{N}$ is contained in $\mathbb{Z}$ as a proper subset, we can show $\mathbb{N} \sim \mathbb{Z}$. Let

$$f(n) = \begin{cases} (n-1)/2, & \text{if } n \text{ is odd} \\ -n/2, & \text{if } n \text{ is even.} \end{cases}$$

(iii) Show that $(-1, 1) \sim \mathbb{R}$ by considering $f(x) = \tan(\frac{\pi}{2}x)$, another choice is $f(x) = \frac{x}{1-x^2}$.

5 Countable Sets

Definition 6. A set A is *countable* if $\mathbb{N} \sim A$. An infinite set that is not countable is called an *uncountable* set.

Theorem 8. (i) *The set* $\mathbb{Q}$ *is countable;* (ii) *The set* $\mathbb{R}$ *is uncountable.*

Proof. Set $A_1 = \{0\}$ and for each $n \geq 2$, let A_n be the set given by

$$A_n = \left\{ \pm\frac{p}{q} \quad | \quad p, q, \in \mathbb{N}, \text{ having no common factor}, \quad p + q = n \right\}.$$

Note that each A_n is finite and every rational number appears in exactly one of these sets. Our 1–1 correspondence with $\mathbb{N}$ is then achieved by consecutively listing the elements in each A_n. Every rational number appears in the correspondence exactly once. Given, say, $22/7$, we have that $22/7 \in A_{29}$. Because the set of elements in $A_1, \ldots, A_{28}$ is finite, it is clear that $22/7$ eventually gets included in the sequence. The fact that this line of reasoning applies to any rational number p/q is our proof that the correspondence is onto. To verify that it is 1–1, we observe that the sets A_n were constructed to be disjoint so that no rational number appears twice. This completes the proof of (i).

(ii) This part is an unexpected one, and its proof is done by contradiction.

Assume that there does exist a 1–1, onto function $f : \mathbb{N} \to \mathbb{R}$, which suggests that it is possible to enumerate the elements of $\mathbb{R}$. If we let $x_1 = f(1)$, $x_2 = f(2)$, and so on, then our assumption that f is onto means that we can write

$$\mathbb{R} = \{x_1, x_2, x_3, x_4, \ldots\} \tag{5.1}$$

and every real number appears somewhere on the list. We will now use the Nested Interval Property to produce a real number that is not there.

Let I_1 be a closed interval that does not contain x_1. Next, let I_2 be a closed interval, contained in I_1, which does not contain x_2. The existence of such an I_2 is easy to verify. Certainly I_1 contains two smaller disjoint closed intervals, and x_2 can only be in one of these. In general, given an interval I_n, construct I_{n+1} to satisfy

(a) $I_{n+1} \subset I_n$, and

(b) $x_{n+1} \notin I_{n+1}$.

We now consider the intersection $\bigcap_{n=1}^{\infty} I_n$. If x_{n_0} is some real number from the list in (5.1), then we have $x_{n_0} \notin I_{n_0}$, and it follows that

$$x_{n_0} \notin \bigcap_{n=1}^{\infty} I_n.$$

Now, we are assuming that the list in (5.1) contains every real number, and this leads to the conclusion that

$$\bigcap_{n=1}^{\infty} I_n = \emptyset.$$

However, the Nested Interval Property (NIP) asserts that $\bigcap_{n=1}^{\infty} I_n \neq \emptyset$. By NIP, there is at least one $x \in \bigcap_{n=1}^{\infty} I_n$ that, consequently, cannot be on the list in (5.1). This contradiction means that such an enumeration of $\mathbb{R}$ is impossible, and we conclude that $\mathbb{R}$ is an uncountable set. □

What exactly should we make of this discovery? It is an important exercise to show that any subset of a countable set must be either countable or finite. This should not be too surprising. If a set can be arranged into a single list, then deleting some elements from this list results in another (shorter, and potentially terminating) list. This means that countable sets are the smallest type of infinite set. Anything smaller is either still countable or finite.

The force of Theorem 8 is that the cardinality of $\mathbb{R}$ is, informally speaking, a larger type of infinity. The real numbers so outnumber the natural numbers that there is no way to map $\mathbb{N}$ onto $\mathbb{R}$. No matter how we attempt this, there are always real numbers to spare. The set $\mathbb{Q}$, on the other hand, is countable. As far as infinite sets are concerned, this is as small as it gets. What does this imply about the set $\mathbb{I}$ of irrational numbers? By imitating the demonstration that $\mathbb{N} \sim \mathbb{Z}$, we can prove that the union of two countable sets must be countable. Because $\mathbb{R} = \mathbb{Q} \cup \mathbb{I}$, it follows that $\mathbb{I}$ cannot be countable because otherwise $\mathbb{R}$ would be. The inescapable conclusion is that, despite the fact that we have encountered so few of them, the irrational numbers form a far greater subset of $\mathbb{R}$ than $\mathbb{Q}$.

The properties of countable sets described in this discussion are useful, we state them as some final propositions.

Theorem 9. *If $A \subset B$ and B is countable, then A is either countable or finite.*

Theorem 10. (i) *If $A_1, A_2, \ldots, A_m$ are each countable sets, then the union $A_1 \cup A_2 \cup \cdots \cup A_m$ is countable.*

(ii) *If A_n is a countable set for each $n \in \mathbb{N}$, then $\bigcup_{n=1}^{\infty} A_n$ is countable.*

Theorem 11 (Schröder–Bernstein Theorem)**.** *Assume there exists a 1–1 function $f : X \to Y$ and another 1–1 function $g : Y \to X$. Then there exists a 1–1, onto function $h : X \to Y$ and hence $X \sim Y$.*

Sketch of proof. The strategy is to partition X and Y into components

$$X = A \cup A' \quad \text{and} \quad Y = B \cup B'$$

with $A \cap A' = \emptyset$ and $B \cap B' = \emptyset$, in such a way that f maps A onto B, and g maps B' onto A'.

(a) Set $A_1 = X \setminus g(Y) = \{x \in X \,|\, x \notin g(Y)\}$ (If $A_1 = \emptyset$, proof done), and inductively define a sequence of sets by letting $A_{n+1} = g(f(A_n))$. Show that $\{A_n \,|\, n \in \mathbb{N}\}$ is a pairwise disjoint collection of subsets of X, while $\{f(A_n) \,|\, n \in \mathbb{N}\}$ is a similar collection in Y.

(b) Let $A = \bigcup_{n=1}^{\infty} A_n$ and $B = \bigcup_{n=1}^{\infty} f(A_n)$. Then f maps A onto B.

(c) Let $A' = X \setminus A$ and $B' = Y \setminus B$. Then g maps B' onto A'.

(d) A 1–1 correspondence between X and Y is then given as

$$h(x) = \begin{cases} f(x) & \text{when } x \in A \\ g^{-1}(x) & \text{when } x \in A'. \end{cases}$$

□

6 Cantor's Theorem

Cantor's work into the theory of infinite sets extends far beyond the conclusions of Theorem 8. Although initially resisted, his creative and relentless assault in this area eventually produced a revolution in set theory and a paradigm shift in the way mathematicians came to understand the infinite.

6.1 Cantor's Diagonalization Method

Cantor published his discovery that $\mathbb{R}$ is uncountable in 1874. Although it has some modern polish on it, the argument presented in Theorem 8 (ii) is actually quite similar to the one Cantor originally found. In 1891, Cantor offered another proof of this same fact that is startling in its simplicity. It relies on decimal representations for real numbers, which we will accept and use without any formal definitions.

Theorem 12. *The open interval* $(0, 1) = \{x \in \mathbb{R} \,|\, 0 < x < 1\}$ *is uncountable.*

Exercise 2. Show that $(0, 1)$ is uncountable if and only if $\mathbb{R}$ is uncountable. This shows that Theorem 12 is equivalent to Theorem 8.

Proof. As with Theorem 8, we proceed by contradiction and assume that there does exist a function $f : \mathbb{N} \to (0, 1)$ that is 1–1 and onto. For each $m \in \mathbb{N}$, $f(m)$ is a real number between 0 and 1, and we represent it using the decimal notation

$$f(m) = 0.a_{m1}a_{m2}a_{m3}a_{m4}\ldots.$$

What is meant here is that for each $m, n \in \mathbb{N}$, a_{mn} is the digit from the set $\{0, 1, 2, \ldots, 9\}$ that represents the nth digit in the decimal expansion of $f(m)$. The 1–1 correspondence between $\mathbb{N}$ and $(0, 1)$ can be summarized in the doubly indexed array (a matrix)

$$\begin{array}{ccccccccccc}
\mathbb{N} & & (0,1) & & & & & & & & \\
\hline
1 & \leftrightarrow & f(1) & = & 0.\boldsymbol{a_{11}} & a_{12} & a_{13} & a_{14} & a_{15} & a_{16} & \cdots \\
2 & \leftrightarrow & f(2) & = & 0.a_{21} & \boldsymbol{a_{22}} & a_{23} & a_{24} & a_{25} & a_{26} & \cdots \\
3 & \leftrightarrow & f(3) & = & 0.a_{31} & a_{32} & \boldsymbol{a_{33}} & a_{34} & a_{35} & a_{36} & \cdots \\
4 & \leftrightarrow & f(4) & = & 0.a_{41} & a_{42} & a_{43} & \boldsymbol{a_{44}} & a_{45} & a_{46} & \cdots \\
5 & \leftrightarrow & f(5) & = & 0.a_{51} & a_{52} & a_{53} & a_{54} & \boldsymbol{a_{55}} & a_{56} & \cdots \\
6 & \leftrightarrow & f(6) & = & 0.a_{61} & a_{62} & a_{63} & a_{64} & a_{65} & \boldsymbol{a_{66}} & \cdots \\
\vdots & & \vdots & & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots
\end{array}$$

The key assumption about this correspondence is that every real number in $(0, 1)$ is assumed to appear somewhere on the list. Now for the pearl of the argument. Define a real number $x \in (0, 1)$ with the decimal expansion $x = 0.b_1b_2b_3b_4\ldots$ using the rule

$$b_n = \begin{cases} 2 & \text{if} \quad a_{nn} \neq 2 \\ 3 & \text{if} \quad a_{nn} = 2. \end{cases}$$

Then, the real number $x = 0.b_1b_2b_3b_4\ldots$ can not be $f(1)$, or $f(2)$, and in general, can not be $f(n)$ for any $n \in \mathbb{N}$, and hence it is not in the list $\{f(n)\}_{n=1}^{\infty}$, we arrive at a contradiction. Therefore, $(0, 1)$ is uncountable.

□

Exercise 3. Why the argument in the above proof does not apply to show that $\mathbb{Q} \cap (0, 1)$ is uncountable?

Having distinguished between the countable infinity of $\mathbb{N}$ and the uncountable infinity of $\mathbb{R}$, a new question that occupied Cantor was whether or not there existed an infinity "above" that of $\mathbb{R}$. This is logically treacherous territory. The same care we gave to defining the relationship "has the same cardinality as" needs to be given to defining relationships such as "has cardinality greater than" or "has cardinality less than or equal to." Nevertheless, without getting too weighed down with formal definitions, one gets a very clear sense from the next result that there is a hierarchy of infinite sets that continues well beyond the continuum of $\mathbb{R}$.

6.2 Power Sets and Cantor's Theorem

Definition 7. Given a set A, the *power set* $P(A)$ refers to the collection of all subsets of A.

Exercise 4. (a) Let $A = \{a, b, c\}$. List the eight elements of $P(A)$. (Do not forget that $\emptyset$ is considered to be a subset of every set.)

(b) If A is finite with n elements, show that $P(A)$ has 2^n elements. Because of this, the power set of A is sometimes denoted by 2^A.

Theorem 13 (Cantor's theorem)**.** *Given any set A, there does not exist a function $f : A \to P(A)$ that is onto.*

Proof. Assume, for contradiction, that $f : A \to P(A)$ is onto. Unlike the usual situation in which we have sets of numbers for the domain and range, f is a correspondence between a set and its power set. For each element $a \in A$, $f(a)$ is a particular subset of A. The assumption that f is onto means that every subset of A appears as $f(a)$ for some $a \in A$. To arrive at a contradiction, we will produce a subset $B \subset A$ that is not equal to $f(a)$ for any $a \in A$.

Construct B using the following rule. For each element $a \in A$, consider the subset $f(a)$. This subset of A may contain the element a or it may not. This depends on the function f. If $f(a)$ does not contain a, then we include a in our set B. More precisely, let

$$B = \{a \in A \,|\, a \notin f(a)\}.$$

Because we have assumed that our function $f : A \to P(A)$ is onto, it must be that $B = f(a')$ for some $a' \in A$. The contradiction arises when we consider whether or not a' is an element of B.

(i) If $a' \in B$. Then according to the definition of B, we have $a' \notin f(a') = B$, contradiction.

(ii) If $a' \notin B = f(a')$. Then according to the definition of B, we have $a' \in B$, contradiction.

Therefore, there does not exist function $f : A \to P(A)$ that is onto.

□

To get an initial sense of its broad significance, let's apply this result to the set of natural numbers. Cantor's Theorem states that there is no onto function from $\mathbb{N}$ to $P(\mathbb{N})$; in other words, the power set of the natural numbers is uncountable.

Exercise 5. Show that $P(\mathbb{N}) \sim \mathbb{R}$. Hint: using the Schröder–Bernstein theorem.