2 Fundamentals of Real Analysis

In this Chapter, we will discuss the essential properties of the real number system $\mathbb{R}$. Although it is possible to give a formal construction of this system on the basis of a more primitive set (such as the set $\mathbb{N}$ of natural numbers, the set $\mathbb{Z}$ of integers, or the set $\mathbb{Q}$ of rational numbers)[1], we have chosen not to do so. Instead, we exhibit a list of fundamental properties associated with the real numbers and show how further properties can be deduced from them. This kind of activity is much more useful in learning the tools of analysis than examining the logical difficulties of constructing a model for $\mathbb{R}$. The natural numbers $\mathbb{N}$ are constructed using five axioms, known as the *Peano's Axioms*. One could then recursively define addition and multiplication, and verify that they obeyed the usual laws of algebra. Next, the integers $\mathbb{Z}$ are constructed by taking differences of the natural numbers, $a - b$. Then, the rationals $\mathbb{Q}$ are constructed by taking quotients of the integers, $\frac{a}{b}$, although we need to exclude division by zero in order to keep the laws of algebra reasonable. Readers interested in learning about the explicit constructions of $\mathbb{N}$, $\mathbb{Z}$, and $\mathbb{Q}$ may go through **Chapters 1, 3, 4** of the book by **T. Tao**.

There are three methods that are often used to construct the real numbers. Each method has its advantages and disadvantages. Each method leads to a model for the real numbers, that is, a set with addition, multiplication, and ordering that satisfy the axioms for *complete ordered field*. The three models are respectively referred to as the *Weierstrass-Stolz model* (decimal expansions, the most intuitive model), the *Dedekind model* (Dedekind cuts, the slickest model), and the *Meray-Cantor model* (completion of a metric space, the most far-reaching model). W. Rudin, in his renowned book, has stated that *"it is pedagogically unsound (though logically correct) to start off with the construction of the real numbers from the rational numbers. At the beginning, most students simply fail to appreciate the need for doing this."* Following his words, we introduce the real number system simply as an ordered field with the least upper bound (supremum) property, and then go on to learn a few interesting applications of this property.

2.1 Algebraic and Order Properties of $\mathbb{R}$

The real numbers $\mathbb{R}$ have some rather unexpected properties. In fact, there are many things that are difficult to prove rigorously. For example, how do we know that $\sqrt{2}$ exists? In other words, how can we be sure that there is

[1]The symbols $\mathbb{N}$, $\mathbb{Q}$, and $\mathbb{R}$ stand for "natural", "quotient", and "real" respectively. $\mathbb{Z}$ stands for "Zahlen", the German word for number. There is also the complex numbers $\mathbb{C}$, which obviously stands for "complex".

some real number whose square is 2? Also, it is easy to convince yourself that $2+3=3+2$. Can you be so sure about $\sqrt{2}+\sqrt{3}=\sqrt{3}+\sqrt{2}$ or $e+\pi=\pi+e$, if you can really write down what those numbers are? In fact, our intuition works pretty well about what should be true for $\mathbb{N}$ or $\mathbb{Z}$ or even for $\mathbb{Q}$. Things don't get hard until we are forced to admit the existence of irrationals. There are constructive methods for making the full set $\mathbb{R}$ from $\mathbb{Q}$. The first rigorous construction was given by Richard Dedekind in 1872. He developed the idea first in 1858 though he did not publish it until 1872. This is what he wrote at the beginning of the article: *"As professor in the Polytechnic School in Zürich I found myself for the first time obliged to lecture upon the ideas of the differential calculus and felt more keenly than ever before the lack of a really scientific foundation for arithmetic. In discussing the notion of the approach of a variable magnitude to a fixed limiting value, and especially in proving the theorem that every magnitude which grows continuously but not beyond all limits, must certainly approach a limiting value, I had recourse to geometric evidences. . . . This feeling of dissatisfaction was so overpowering that I made the fixed resolve to keep mediating on the question till I should find a purely arithmetic and perfectly rigorous foundation for the principles of infinitesimal analysis."* Dedekind was one of the last research students of Gauss. His arithmetization of analysis was his most important contribution to Mathematics, although it was not enthusiastically received by leading mathematicians of his day. Readers interested in learning briefly about Dedekind cuts and related concepts may go through **Section 8.4** of the book by **S. Abbot**. Readers interested in learning briefly about the decimal representation of real numbers and related concepts may go through **Section 2.5** of the book by **Bertle & Sherbert**.

In **Section 2.1.1**, we introduce the algebraic properties of $\mathbb{R}$ that are based on the two binary operations of addition $(+)$ and multiplication $(\cdot)$. Next, we introduce the order properties of $\mathbb{R}$ in **Section 2.1.2** which are based on the notion of positivity and then we derive some consequences of these properties and illustrate their use in working with inequalities.

2.1.1 Algebraic Properties of $\mathbb{R}$

We begin with a brief discussion of the *algebraic (or field) properties* of $\mathbb{R}$ with respect to the binary operations of addition $(+)$ and multiplication $(\cdot)$. All other algebraic properties can be derived from these basic properties. In the terminology of abstract algebra, the system of real numbers is a *field* with respect to addition and multiplication. The basic properties listed below are known as the *field axioms*:
(A1) $a+b=b+a$ for all a,b in $\mathbb{R}$ (*commutative property of addition*);
(A2) $(a+b)+c=a+(b+c)$ for all a,b,c in $\mathbb{R}$ (*associative property of*

addition);
(A3) there exists an element 0 in $\mathbb{R}$ such that $0+a=a$ and $a+0=a$ for all a in $\mathbb{R}$ (*existence of a zero element*);
(A4) for each a in $\mathbb{R}$, there exists an element $-a$ in $\mathbb{R}$ such that $a+(-a)=0$ and $(-a)+a=0$ (*existence of negative elements*);
(M1) $a \cdot b = b \cdot a$ for all a, b in $\mathbb{R}$ (*commutative property of multiplication*);
(M2) $(a \cdot b) \cdot c = a \cdot (b \cdot c)$ for all a, b, c in $\mathbb{R}$ (*associative property of multiplication*);
(M3) there exists an element 1 in $\mathbb{R}$ such that $1 \cdot a = a$ and $a \cdot 1 = a$ for all a in $\mathbb{R}$ (*existence of a unit element*);
(M4) for each a in $\mathbb{R}$, there exists an element $\frac{1}{a}$ in $\mathbb{R}$ such that $a \cdot \frac{1}{a} = 1$ and $\frac{1}{a} \cdot a = 1$ (*existence of reciprocals*);
(D) $a \cdot (b+c) = (a \cdot b) + (a \cdot c)$ and $(b+c) \cdot a = (b \cdot a) + (c \cdot a)$ for all a, b, c in $\mathbb{R}$ (*distributive property of multiplication over addition*).

These properties should be familiar to the reader. The first four are concerned with addition, the next four with multiplication, and the last one connects the two operations. We now present a few simple results. All other properties can be deduced from the nine properties listed above.

Theorem 2.1: (a) If z and a are elements in $\mathbb{R}$ with $z+a=a$, then $z=0$ (0 *is unique*).
(b) If u and $b \neq 0$ are elements in $\mathbb{R}$ with $u \cdot b = b$, then $u=1$ (1 *is unique*).
(c) If $a \in \mathbb{R}$, then $a \cdot 0 = 0$ (*multiplication by* 0 *always results in* 0).

Proof: Exercise!

Theorem 2.2: (a) If $a \neq 0$ and b in $\mathbb{R}$ are such that $a \cdot b = 1$, then $b = \frac{1}{a}$ (*uniqueness of reciprocals*).
(b) If $a \cdot b = 0$, then either $a=0$ or $b=0$.

Proof: Exercise!

The operation of subtraction is defined by $a-b = a+(-b)$ for a, b in $\mathbb{R}$. Similarly, division is defined for a, b in $\mathbb{R}$ with $b \neq 0$ by $\frac{a}{b} = a \cdot \left(\frac{1}{b}\right)$. In the following, we will use this customary notation for subtraction and division, and we will use all the familiar properties of these operations. We will ordinarily drop the use of the "dot" to indicate multiplication and write ab for $a \cdot b$. Similarly, we will use the usual notation for exponents and write a^2 for aa, a^3 for $(a^2)a$; and, in general, we define $a^{n+1} = (a^n)a$ for $n \in \mathbb{N}$. We agree to adopt the convention that $a^1 = a$. Further, if $a \neq 0$, we write $a^0 = 1$ and a^{-1} for $\frac{1}{a}$, and if $n \in \mathbb{N}$, we will write a^{-n} for $\left(\frac{1}{a}\right)^n$, when it is convenient to do so. In general, we will freely apply all the usual techniques of algebra without further elaboration.

2.1.2 Order Properties of $\mathbb{R}$

The *order properties* of $\mathbb{R}$ refer to the notions of positivity and inequalities between real numbers. As with the algebraic structure of $\mathbb{R}$, we proceed by isolating three basic properties from which all other order properties and calculations with inequalities can be deduced. The simplest way to do this is to identify a special subset of $\mathbb{R}$ by using the notion of positivity.

There is a non-empty subset P of $\mathbb{R}$, called the set of positive real numbers, that satisfies the following properties:
(a) If a, b belong to P, then $a + b$ belongs to P.
(b) If a, b belong to P, then ab belongs to P.
(c) If a belongs to $\mathbb{R}$, then exactly one of the following holds:

$$a \in P, \qquad a = 0, \qquad -a \in P.$$

The first two properties ensure the compatibility of order with the operations of addition and multiplication, respectively. Property (c) is usually called the *Trichotomy Property*, since it divides $\mathbb{R}$ into three distinct types of elements. It states that the set $\{-a : a \in P\}$ of negative real numbers has no elements in common with the set P of positive real numbers, and, moreover, the set $\mathbb{R}$ is the union of three disjoint sets.

If $a \in P$, we write $a > 0$ and say that a is a *positive* (or a *strictly positive*) real number. If $a \in P \cup \{0\}$, we write $a \geq 0$ and say that a is a *nonnegative* real number. Similarly, if $-a \in P$, we write $a < 0$ and say that a is a *negative* (or a *strictly negative*) real number. If $-a \in P \cup \{0\}$, we write $a \leq 0$ and say that a is a *nonpositive* real number.

The notion of inequality between two real numbers will now be defined in terms of the set P of positive elements.

Definition 2.1: Let a, b be elements of $\mathbb{R}$.
(a) If $a - b \in P$, then we write $a > b$ or $b < a$.
(b) If $a - b \in P \cup \{0\}$, then we write $a \geq b$ or $b \leq a$.

The Trichotomy Property implies that for $a, b \in \mathbb{R}$, exactly one of the following will hold:

$$a > b, \qquad a = b, \qquad a < b.$$

Therefore, if both $a \leq b$ and $b \leq a$, then $a = b$.

For notational convenience, we will write $a < b < c$ to mean that both $a < b$ and $b < c$ are satisfied. The other similar inequalities $a \leq b < c$, $a \leq b \leq c$, and $a < b \leq c$ are defined in a similar manner.

To illustrate how the basic order properties are used to derive the rules of inequalities, we will now establish several results which you may have used in earlier mathematics courses.

Theorem 2.3: Let a, b, c be any elements of $\mathbb{R}$.
(a) If $a > b$ and $b > c$, then $a > c$.
(b) If $a > b$, then $a + c > b + c$.
(c) If $a > b$ and $c > 0$, then $ca > cb$.
(d) If $a > b$ and $c < 0$, then $ca < cb$.

Proof: Exercise!

Theorem 2.4: (a) If $a \in \mathbb{R}$ and $a \neq 0$, then $a^2 > 0$.
(b) $1 > 0$.
(c) If $n \in \mathbb{N}$, then $n > 0$.

Proof: Exercise!

It is worth noting that *no smallest positive real number can exist.* This follows by observing that if $a > 0$, then since $\frac{1}{2} > 0$, we have that $0 < \frac{1}{2}a < a$. Thus, if it is claimed that a is the smallest positive real number, we can exhibit a smaller positive number $\frac{1}{2}a$. This observation leads us to the next result, which will be used frequently as a method of proof. For instance, to prove that a number $a \geq 0$ is actually equal to zero, we see that it suffices to show that a is smaller than an arbitrary positive number.

Theorem 2.5: (a) If $a \in \mathbb{R}$ is such that $0 \leq a < \epsilon$ for every $\epsilon > 0$, then $a = 0$.
(b) If $a \in \mathbb{R}$ is such that $0 \leq a \leq \epsilon$ for every $\epsilon > 0$, then $a = 0$.

Proof: (a) Suppose to the contrary that $a > 0$. Then if we take $\epsilon_0 = \frac{1}{2}$, we have $0 < \epsilon_0 < a$. Therefore, it is false that $a < \epsilon$ for every $\epsilon > 0$ and we conclude that $a = 0$.

(b) Exercise!

The product of two positive numbers is positive. However, the positivity of a product of two numbers does not imply that each factor is positive. The correct conclusion is given in the next theorem. It is an important tool in working with inequalities.

Theorem 2.6: If $ab > 0$, then either
(a) $a > 0$ and $b > 0$, or
(b) $a < 0$ and $b < 0$.

Proof: Exercise!

Corollary 2.6: If $ab < 0$, then either
(a) $a < 0$ and $b > 0$, or
(b) $a > 0$ and $b < 0$.

Proof: Exercise!

2.2 Absolute Value and the Real Line

From the Trichotomy Property, we are assured that if $a \in \mathbb{R}$ and $a \neq 0$, then exactly one of the numbers a and $-a$ is positive. The absolute value of $a \neq 0$ is defined to be the positive one of these two numbers. The absolute value of 0 is defined to be 0.

Definition 2.2: The absolute value of a real number a, denoted by $|a|$, is defined by

$$|a| = \begin{cases} a & \text{if } a > 0, \\ -a & \text{if } a < 0. \end{cases}$$

We see from the definition that $|a| \geq 0$ for all $a \in \mathbb{R}$, and that $|a| = 0$ if and only if $a = 0$. Also, $|-a| = |a|$ for all $a \in \mathbb{R}$. Some additional properties are as follows.

Theorem 2.7: (a) $|ab| = |a||b|$ for all $a, b \in \mathbb{R}$.
(b) $|a|^2 = a^2$ for all $a \in \mathbb{R}$.
(c) If $c \geq 0$, then $|a| \leq c$ if and only if $-c \leq a \leq c$.
(d) $-|a| \leq a \leq |a|$ for all $a \in \mathbb{R}$.

Proof: Exercise!

Theorem 2.8 (Triangle Inequality): If $a, b \in \mathbb{R}$, then $|a + b| \leq |a| + |b|$.

Proof: Exercise!

Corollary 2.8.1: If $a, b \in \mathbb{R}$, then

(a) $||a| - |b|| \leq |a - b|$.
(b) $|a - b| \leq |a| + |b|$.

Proof: Exercise!

A straightforward application of Mathematical Induction extends the Triangle Inequality to any finite number of elements of $\mathbb{R}$.

Corollary 2.8.2: If $a_1, a_2, \ldots, a_n$ are any real numbers, then

$$|a_1 + a_2 + \cdots + a_n| \leq |a_1| + |a_2| + \cdots + |a_n|.$$

Proof: Exercise!

A convenient and familiar geometric interpretation of the real number system is the *real line*. In this interpretation, the absolute value $|a|$ of an element a in $\mathbb{R}$ is regarded as the distance from a to the origin 0. More generally, the distance between elements a and b in $\mathbb{R}$ is $|a - b|$.

2.3 Boundedness in $\mathbb{R}$

Thus far, we have discussed the algebraic and the order properties of $\mathbb{R}$. In this section, we shall present one more property of $\mathbb{R}$ that is often called the *Completeness Property*. The system $\mathbb{Q}$ of rational numbers also has the algebraic and order properties but we know that $\sqrt{2}$ cannot be represented as a rational number; therefore $\sqrt{2}$ does not belong to $\mathbb{Q}$. This observation shows the necessity of an additional property to characterize the real number system. This additional property, the Completeness Property, is an essential property of $\mathbb{R}$, and with this final assumption on $\mathbb{R}$, we say that $\mathbb{R}$ is a *complete ordered field*. It is this special property that permits us to define and develop the various limiting procedures that will be discussed in the courses that follow. There are several different ways to describe the Completeness Property. We choose to give what is probably the most efficient approach by assuming that each non-empty bounded subset of $\mathbb{R}$ has a supremum.

We now introduce the notions of upper bound and lower bound for a set of real numbers. These ideas will be of utmost importance in later sections.

Definition 2.3: Let S be a non-empty subset of $\mathbb{R}$.
(a) The set S is said to be *bounded above* if there exists a number $u \in \mathbb{R}$ such that $s \leq u$ for all $s \in S$. Each such number u is called an *upper bound* of S.
(b) The set S is said to be *bounded below* if there exists a number $w \in \mathbb{R}$

such that $w \leq s$ for all $s \in S$. Each such number w is called an *lower bound* of S.
(c) A set is said to be *bounded* if it is both bounded above and bounded below. A set is said to be *unbounded* if it is not bounded.

Example 2.1: The set $S = \{x \in \mathbb{R} : x < 2\}$ is bounded above; the number 2 and any number larger than 2 is an upper bound of S. This set has no lower bounds, so that the set is not bounded below. Thus it is unbounded (even though it is bounded above).

2.4 Supremum and Infimum

Note that if a set has one upper bound, then it has infinitely many upper bounds, because if u is an upper bound of S, then the numbers $u+1, u+2, \ldots$ are also upper bounds of S (a similar observation is valid for lower bounds). In the set of upper bounds of S and the set of lower bounds of S, we single out their least and greatest elements, respectively, for special attention in the following definition.

Definition 2.4: Let S be a non-empty subset of $\mathbb{R}$.
(a) If S is bounded above, then a number u is said to be a *supremum* (or a *least upper bound*) of S, denoted sup S, if it satisfies the conditions:
(1) u is an upper bound of S, and
(2) if v is any upper bound of S, then $u \leq v$.
(b) If S is bounded below, then a number w is said to be a *infimum* (or a *greatest lower bound*) of S, denoted inf S, if it satisfies the conditions:
(1) w is a lower bound of S, and
(2) if t is any upper bound of S, then $t \leq w$.

It is not difficult to see that *there can be only one supremum of a given subset S of* $\mathbb{R}$. Then we can refer to *the* supremum of a set instead of *a* supremum. For, suppose that u_1 and u_2 are both suprema of S. If $u_1 < u_2$, then the hypothesis that u_2 is a supremum implies that u_1 cannot be an upper bound of S. Similarly, we see that $u_2 < u_1$ is not possible. Therefore, we must have $u_1 = u_2$. A similar argument can be given to show that the infimum of a set is uniquely determined. Note that *the empty set is bounded above by every real number, so it has no supremum.*

It needs to be emphasized that in order for a non-empty set S in $\mathbb{R}$ to have a supremum, it must have an upper bound. Thus, not every subset of $\mathbb{R}$ has a supremum; similarly, not every subset of $\mathbb{R}$ has an infimum. Indeed, there are four possibilities for a non-empty subset S of $\mathbb{R}$; it can have
(a) both a supremum and an infimum,

(b) a supremum but no infimum,
(c) a infimum but no supremum,
(d) neither a supremum nor an infimum.

It is also important to note that that in order to show that $u = \sup S$ for some non-empty subset S of $\mathbb{R}$, we need to show that both (1) and (2) of **Definition 2.4 (a)** hold. It will be instructive to reformulate these statements. First, the reader should see that the following two statements about a number u and a set S are equivalent:
(1a) u is an upper bound of S,
(1b) $s \leq u$ for all $s \in S$.

Also, the following statements about an upper bound u of a set S are equivalent:
(2a) if v is any upper bound of S, then $u \leq v$,
(2b) if $z < u$, then z is not an upper bound of S,
(2c) if $z < u$, then there exists $s_1 \in S$ such that $z < s_1$,
(2d) if $\epsilon > 0$, then there exists $s_2 \in S$ such that $u - \epsilon < s_2$.

Therefore, we can state two alternate formulations for the supremum. We now state two important lemmas.

Lemma 2.1: A number u is the supremum of a non-empty subset S of $\mathbb{R}$ if and only if u satisfies the conditions:
(1) $s \leq u$ for all $s \in S$,
(2) if $v < u$, then there exists $s' \in S$ such that $v < s'$.

Proof: Exercise!

Lemma 2.2: An upper bound u of a non-empty set S in $\mathbb{R}$ is the supremum of S if and only if for every $\epsilon > 0$, there exists an $s_0 \in S$ such that $u - \epsilon < s_0$.

Proof: If u is an upper bound of S that satisfies the stated condition and if $v < u$, then we put $\epsilon = u - v$. Then $s > 0$, so there exists $s_0 \in S$ such that $v = u - \epsilon < s_0$. Therefore, v is not an upper bound of S, and we conclude that $u = \sup S$.

Conversely, suppose that $u = \sup S$ and let $\epsilon > 0$. Since $u - \epsilon < u$, then $u - \epsilon$ is not an upper bound of S. Therefore, some element $s_0 \in S$ must be greater than $u - \epsilon$; that is, $u - \epsilon < s_0$.

It is important to realize that the supremum of a set may or may not be an element of the set. Sometimes it is and sometimes it is not, depending on the particular set. We consider a few examples.

Example 2.2: If a non-empty set S_1 has a finite number of elements, then it can be shown that S_1 has a largest element u and a least element w. Then $u = \sup S_1$ and $w = \inf S_1$, and they are both members of S_1 (this is clear if S_1 has only one element, and it can be proved by induction on the number of elements in S_1).

Example 2.3: The set $S_2 = \{x : 0 \leq x \leq 1\}$ clearly has 1 for an upper bound. We prove that 1 is its supremum as follows. If $v < 1$, there exists an element $s' = \frac{v+1}{2} \in S_2$ such that $v < s'$. Therefore v is not an upper bound of S_2 and, since v is an arbitrary number $v < 1$, we conclude that $\sup S_2 = 1$. It is similarly shown that $\inf S_2 = 0$. Note that both the supremum and the infimum of S_2 are contained in S_2.

Example 2.4: The set $S_3 = \{x : 0 < x < 1\}$ clearly has 1 for an upper bound. Using the same argument as given in **Example 2.3**, we see that $\sup S_3 = 1$. In this case, the set S_3 does not contain its supremum. Similarly, $\inf S_3 = 0$ is not contained in S_3.

2.5 Completeness Property of $\mathbb{R}$ and its Applications

It is not possible to prove on the basis of the algebraic and order properties of $\mathbb{R}$ (discussed in **Sections 2.1.1 & 2.1.2** respectively) that every non-empty subset of $\mathbb{R}$ that is bounded above has a supremum in $\mathbb{R}$. However, it is a deep and fundamental property of the real number system that this is indeed the case. We will make frequent and essential use of this property, especially in our discussion of limiting processes. The property can be formally stated as follows.

The Completeness Property of $\mathbb{R}$: Every non-empty set of real numbers that has an upper bound also has a supremum in $\mathbb{R}$.

This property is also called the *Supremum Property* of $\mathbb{R}$ or sometimes the *Least Upper Bound Property* of $\mathbb{R}$. The analogous property for infimum, known as the *Infimum Property* of $\mathbb{R}$ or sometimes the *Greatest Lower Bound Property* of $\mathbb{R}$, can be deduced from the Completeness Property as follows. Suppose that S is a non-empty subset of $\mathbb{R}$ that is bounded below. Then the non-empty set $T = \{-s : s \in S\}$ is bounded above, and the Supremum Property implies that $u = \sup T$ exists in $\mathbb{R}$. It can be easily verified that $-u$ is the infimum of S (**Proof!**).

It can be shown that the real number system is essentially the only complete ordered field (an ordered field in which the Completeness Property holds); that is, if an alien from another planet were to construct a mathematical system with the algebraic, order, and the completeness properties, the alien's system would differ from the real number system only in that the alien might use different symbols for the real numbers and $+$, $\cdot$, $<$, etc.. Note that $\mathbb{Q}$ does not have the Supremum Property. For example, the subset $S = \{x \in \mathbb{Q} : x > 0 \text{ and } x^2 < 2\}$ of $\mathbb{Q}$ is a non-empty subset of $\mathbb{Q}$ bounded above but sup S does not belong to $\mathbb{Q}$. Another interesting example is the subset $S = \{1, 1 + \frac{1}{1!}, 1 + \frac{1}{1!} + \frac{1}{2!}, \cdots\}$ of $\mathbb{Q}$ which is bounded above because each element of the set is less than 3, but there is no rational number which is the supremum of S. In fact, the supremum of the set is e, an irrational number.

We will now discuss how to work with suprema and infima. We will also give some very important applications of these concepts to derive fundamental properties of $\mathbb{R}$. We begin with examples that illustrate useful techniques in applying the ideas of supremum and infimum.

It is an important fact that taking suprema and infima of sets is compatible with the algebraic properties of $\mathbb{R}$. The following example establishes the compatibility of taking suprema and addition.

Example 2.5: Let S be a non-empty subset of $\mathbb{R}$ that is bounded above, and let a be any number in R. Define the set $a+S = \{a+s : s \in S\}$. Prove that

$$\sup(a + S) = a + \sup S.$$

Solution: If we let $u = sup\ S$, then $x \leq u$ for all $x \in S$, so that $a+x \leq a+u$. Therefore, $a+u$ is an upper bound for the set $a+S$; consequently, we have $\sup(a + S) \leq a + u$. Now if v is any upper bound of the set $a + S$, then $a + x \leq v$ for all $x \in S$. Consequently $x \leq v - a$ for all $x \in S$, so that $v - a$ is an upper bound of S. Therefore, $u = \sup S \leq v - a$, which gives us $a + u \leq v$. Since v is any upper bound of $a + S$, we can replace v by $\sup(a + S)$ to get $a + u \leq \sup(a + S)$. Combining these inequalities, we conclude that $\sup(a + S) = a + u = a + \sup S$.

Example 2.6: If A, B are subsets of $\mathbb{R}$ such that $A \subseteq B$ and $c \in \mathbb{R}$, we define

$$cA = \{d : d = cx \text{ for some } x \in A\}.$$

Prove that
(a) $\sup cA = c \sup A$ and $\inf cA = c \inf A$, if $c \geq 0$,
(b) $\sup cA = c \inf A$ and $\inf cA = c \sup A$, if $c < 0$.

Solution: The result is obvious if $c = 0$. If $c > 0$, then $cx \le M$ if and only if $x \le \frac{M}{c}$, which shows that M is an upper bound of cA if and only if $\frac{M}{c}$ is an upper bound of A, so $\sup cA = c \sup A$. If $c < 0$, then $cx \le M$ if and only if $x \ge \frac{M}{c}$, so M is an upper bound of cA if and only if $\frac{M}{c}$ is a lower bound of A, so $\sup cA = c \inf A$. The remaining results follow similarly.

Example 2.7: Suppose that A, B are subsets of $\mathbb{R}$ and $A \subseteq B$. Prove that if $\sup A$, $\sup B$ exist, then $\sup A \le \sup B$, and if $\inf A$, $\inf B$ exist, then $\inf A \ge \inf B$.

Solution: Since $\sup B$ is an upper bound of B and $A \subseteq B$, it follows that $\sup B$ is an upper bound of A, so $\sup A \le \sup B$. The proof for the infimum is similar, or we may apply the result for the supremum to $-A \subseteq -B$.

The above example suggests that for subsets A, B of $\mathbb{R}$ such that $A \subseteq B$, we have $\inf B \le \inf A \le \sup A \le \sup B$.

Example 2.8: Suppose that A and B are non-empty subsets of $\mathbb{R}$ that satisfy the property: $a \le b$ for all $a \in A$ and all $b \in B$. Prove that

$$\sup A \le \inf B.$$

Solution: Given $b \in B$, we have $a \le b$ for all $a \in A$. This means that b is an upper bound of A, so that $\sup A \le b$. Next, since the last inequality holds for all $b \in B$, we see that the number $\sup A$ is a lower bound for the set B. Therefore, we conclude that $\sup A \le \inf B$.

Example 2.9: If A, B be non-empty subsets of $\mathbb{R}$, we define

$$A + B = \{d : d = x + y \text{ for some } x \in A, y \in B\}$$

and

$$A - B = \{d : d = x - y \text{ for some } x \in A, y \in B\}.$$

Prove that
(a) $\sup(A + B) = \sup A + \sup B$,
(b) $\sup(A - B) = \sup A - \inf B$,
(c) $\inf(A + B) = \inf A + \inf B$,
(d) $\inf(A - B) = \inf A - \sup B$.

Solution: (a) The set $A + B$ is bounded from above if and only if A and B are both bounded from above, $\sup(A + B)$ exists if and only if both $\sup A$ and $\sup B$ exist. In that case, if $x \in A$ and $y \in B$, then

$$x + y \le \sup A + \sup B,$$

so $\sup A + \sup B$ is an upper bound of $A + B$ and therefore

$$\sup(A + B) \leq \sup A + \sup B.$$

To get the inequality in the opposite direction, suppose that $\epsilon > 0$. Then there exists $x \in A$ and $y \in B$ such that

$$x > \sup A - \frac{\epsilon}{2} \quad \text{and} \quad y > \sup B - \frac{\epsilon}{2}.$$

It follows that

$$x + y > \sup A + \sup B - \epsilon$$

for every $\epsilon > 0$, which implies that

$$\sup(A + B) \geq \sup A + \sup B.$$

Thus,

$$\sup(A + B) = \sup A + \sup B.$$

(b) It follows from **(a)** and **Example 2.6 (b)** that

$$\sup(A - B) = \sup A + \sup(-B) = \sup A - \inf B.$$

(c) & (d) Exercise!

Example 2.10: Let S be a non-empty bounded subset of $\mathbb{R}$ with $\sup S = M$ and $\inf S = m$. Prove that the set $T = \{|x - y| : x \in S, y \in S\}$ is bounded above and $\sup T = M - m$.

Solution: For $x, y \in S$, we have $m \leq x \leq M$ and $m \leq y \leq M$. Therefore

$$m - M \leq x - y \leq M - m, \text{ i.e., } |x - y| \leq M - m.$$

This shows that the set T is bounded above, $M - m$ being an upper bound.

Let $a \in S$. Then $|a - a| \in T$ showing that T is non-empty. By the Supremum Property of $\mathbb{R}$, $\sup T$ exists. We now prove that no real number less than $M - m$ is an upper bound of T. If possible, let $p < M - m$ be an upper bound of T. Let $(M - m) - p = 2\epsilon$. Then $\epsilon > 0$ and $p + \epsilon = M - m - \epsilon$. Since $\sup S = M$, there exists an element $x \in S$ such that

$$M - \frac{\epsilon}{2} < x \leq M.$$

Again, since $\inf S = m$, there exists an element $y \in S$ such that

$$m \leq y < m + \frac{\epsilon}{2}.$$

Now, $x - y > M - m - \epsilon$, i.e., $x - y > p + \epsilon$. This shows that p is not an upper bound of T. Therefore, no real number less than $M - m$ is an upper bound of T, i.e., $\sup T = M - m$.

The idea of upper bound and lower bound is applied to functions by considering the range of a function. Given a function $f : D \to \mathbb{R}$, we say that f is bounded above if the set $f(D) = \{f(x) : x \in D\}$ is bounded above in $\mathbb{R}$; that is, there exists $B \in R$ such that $f(x) \leq B$ for all $x \in D$. Similarly, the function f is bounded below if the set $f(D)$ is bounded below. We say that f is bounded if it is bounded above and below; this is equivalent to saying that there exists $B \in \mathbb{R}$ such that $|f(x)| \leq B$ for all $x \in D$. The following examples illustrate how to work with suprema and infima of functions.

Example 2.11: Suppose $D \subseteq \mathbb{R}$ and $f, g : D \to \mathbb{R}$ and $f \leq g$. Prove that if g is bounded from above, then $\sup\limits_D f \leq \sup\limits_D g$ and if f is bounded from below, then $\inf\limits_D f \leq \inf\limits_D g$.

Solution: If $f \leq g$ and g is bounded from above, then for every $x \in D$, we have $f(x) \leq g(x) \leq \sup\limits_D g$. Thus, f is bounded from above by $\inf\limits_D g$, so $\sup\limits_D$ $f \leq \sup\limits_D g$. Similarly, g is bounded from below by $\inf\limits_D g$, so $\inf\limits_D g \leq \inf\limits_D f$.

Note: The hypothesis $f(x) \leq g(x)$ for all $x \in D$ in **Example 2.10** does not imply any relation between $\sup\limits_D f$ and $\inf\limits_D g$. For example, if $f(x) = x^2$ and $g(x) = x$ with $D = \{x : 0 \leq x \leq 1\}$, then $f(x) \leq g(x)$ for all $x \in D$. However, we see that $\sup f(D) = 1$ and $\inf g(D) = 0$. Since $\sup g(D) = 1$, the conclusion of **(a)** holds.

However[2], if $f(x) \leq g(y)$ for all $x, y \in D$, then we may conclude that $\sup f(D) \leq \inf g(D)$, which we may written as $\sup\limits_{x\,\in\,D} f(x) \leq \inf\limits_{y\,\in\,D} g(y)$ (note that the functions in the above example do not satisfy this hypothesis). The proof proceeds in two stages as in **Example 2.8**.

Like limits, the supremum and infimum do not preserve strict inequalities in general. For example, if we define $f : [0, 1] \to \mathbb{R}$ by

$$f(x) = \begin{cases} x & \text{if } 0 \leq x < 1, \\ 0 & \text{if } x = 1, \end{cases}$$

then $f < 1$ on $[0, 1]$ but $\sup\limits_{[0,\,1]} f = 1$.

[2]Note that $\sup\limits_D f$, $\sup\limits_{x\,\in\,D} f(x)$, and $\sup f(D)$ refer to the same quantity although they are notationally different. The same holds for inf also.

Next, we consider the supremum and infimum of linear combinations of functions. Scalar multiplication by a positive constant multiplies the inf or sup, while multiplication by a negative constant switches the inf and sup.

Example 2.12: Let $f : D \to \mathbb{R}$ is a bounded function and $c \in \mathbb{R}$. Prove that
(a) $\sup cf = c \sup f$ and $\inf cf = c \inf f$, if $c \geq 0$,
(b) $\sup cf = c \inf f$ and $\inf cf = c \sup f$, if $c < 0$.

Solution: Exercise! (**Hint:** Apply the results of **Example 2.6** to the set $\{cf(x) : x \in D\} = c\{f(x) : x \in D\}$).

Example 2.13: Let $f, g : D \to \mathbb{R}$ be bounded functions. Prove that
(a) $\sup(f + g) \leq \sup f + \sup g$,
(b) $\inf(f + g) \geq \inf f + \inf g$.

Solution: Since $f(x) \leq \sup f$ and $g(x) \leq \sup g$ for every $x \in [a, b]$, we have

$$f(x) + g(x) \leq \sup f + \sup g.$$

Thus, $f + g$ is bounded from above by $\sup f + \sup g$, so $\sup\,(f + g) \leq \sup f + \sup g$. The proof for the infimum is analogous (or apply the result for the supremum to the functions $-f, -g$).

Thus, for sums of functions, we get an inequality. We may have strict inequality because f and g may take values close to their suprema (or infima) at different points. Let us consider the following example: Define $f, g : [0, 1] \to \mathbb{R}$ by $f(x) = x$, $g(x) = 1 - x$. Then

$$\sup f = \sup g = \sup\,(f + g) = 1,$$

so $\sup\,(f + g) < \sup f + \sup g$.

Example 2.14: Let $f, g : D \to \mathbb{R}$ be bounded functions. Prove that
(a) $|\sup f - \sup g| \leq \sup|f - g|$,
(b) $|\inf f - \inf g| \leq \sup|f - g|$.

Solution: Since $f = f - g + g$ and $f - g \leq |f - g|$, we get from **Examples 2.11 and 2.13** that

$$\sup f \leq \sup\,(f - g) + \sup g \leq \sup\,|f - g| + \sup g,$$

so

$$\sup f - \sup g \leq \sup\,|f - g|.$$

Exchanging f and g in this inequality, we get

$$\sup\, g - \sup\, f \leq \sup\, |f - g|$$

which implies that

$$|\sup\, f - \sup\, g| \leq \sup|f - g|.$$

Replacing f by $-f$ and g by $-g$ in this inequality and using the identity $\sup(-f) = -\inf\, f$, we get

$$|\inf\, f - \inf\, g| \leq \sup|f - g|.$$

Example 2.15: Let $f, g : D \to \mathbb{R}$ be bounded functions such that

$$|f(x) - f(y)| \leq |g(x) - g(y)| \quad \text{ for all } x, y \in D.$$

Prove that $\sup\, f - \inf\, f \leq \sup\, g - \sup\, g$.

Solution: The condition implies that for all $x, y \in D$, we have

$$f(x)-f(y) \leq |g(x)-g(y)| = \max\{g(x), g(y)\} - \min\{g(x), g(y)\} \leq \sup g - \inf g$$

which implies that

$$\sup\{f(x) - f(y) : x, y \in D\} \leq \sup\, g - \inf\, g.$$

From **Example 2.9**,

$$\sup\{f(x) - f(y) : x, y \in D\} \leq \sup\, f - \inf\, f,$$

so the result follows.

2.6 The Archimedean Property

Because of your familiarity with the set $\mathbb{R}$ and the customary picture of the real line, it may seem obvious that the set $\mathbb{N}$ of natural numbers is not bounded in $\mathbb{R}$. How can we prove this obvious fact? In fact, we cannot do so by using only the Algebraic and Order Properties given in **Sections 2.1.1 & 2.1.2**. Indeed, we must use the Completeness Property of $\mathbb{R}$ as well as the Inductive Property of $\mathbb{N}$ (that is, if $n \in \mathbb{N}$, then $n + 1 \in \mathbb{N}$). The absence of upper bounds for $\mathbb{N}$ means that given any real number x there exists a natural number n (depending on x) such that $x < n$.

Theorem 2.9 If $x \in \mathbb{R}$, then there exists $n_x \in \mathbb{N}$ such that $x < n_x$.

Proof: If the assertion is false, then $n \leq x$ for all $n \in \mathbb{N}$. Therefore, x

is an upper bound of $\mathbb{N}$. Therefore, by the Completeness Property, the non-empty set $\mathbb{N}$ has a supremum $u \in \mathbb{R}$. Subtracting 1 from u gives a number $u - 1$ which is smaller than the supremum u of $\mathbb{N}$. Therefore $u - 1$ is not an upper bound of $\mathbb{N}$, so there exists $m \in \mathbb{N}$ with $u - 1 < m$. Adding 1 gives $u < m + 1$, and since $m + 1 \in \mathbb{N}$, this inequality contradicts the fact that u is an upper bound of $\mathbb{N}$.

Corollary 2.9.1: If $S = \{\frac{1}{n} : n \in \mathbb{N}\}$, then $\inf S = 0$.

Proof: Since $S \neq \phi$ is bounded below by 0, it has an infimum and we let $w = \inf S$. It is clear that $w \geq 0$. For any $\epsilon > 0$, **Theorem 2.9** implies that there exists $n \in \mathbb{N}$ such that $\frac{1}{\epsilon} < n$, which implies $\frac{1}{n} < \epsilon$. Therefore we have

$$0 \leq w \leq \frac{1}{n} < \epsilon.$$

But since $\epsilon > 0$ is arbitrary, it follows from **Theorem 2.5(a)** that $w = 0$.

Corollary 2.9.2: If $y > 0$, there exists $n_y \in \mathbb{N}$ such that $0 < \frac{1}{n_y} < y$.

Proof: Since $\inf \{\frac{1}{n} : n \in \mathbb{N}\} = 0$ and $y > 0$, then y is not a lower bound for the set $\{\frac{1}{n} : n \in \mathbb{N}\}$. Thus there exists $n_y \in \mathbb{N}$ such that $0 < \frac{1}{n_y} < y$.

Corollary 2.9.3: If $z > 0$, there exists $n_z \in \mathbb{N}$ such that $n_z - 1 \leq z < n_z$.

Proof: **Theorem 2.9** ensures that the subset $E_z = \{m \in \mathbb{N} : z < m\}$ of $\mathbb{N}$ is not empty. By the Well-Ordering Property, E_z has a least element, which we denote by n_z. Then $n_z - 1$ does not belong to E_z, and hence we have $n_z - 1 \leq z < n_z$.

Theorem 2.10 (The Archimedean Property of $\mathbb{R}$)[3]: If $x, y \in \mathbb{R}$ and $x > 0, y > 0$, then there exists a natural number n such that $ny > x$.

Proof: If possible, let there exist no natural number n for which $ny > x$. Then for every natural number k, $ky \leq x$. Therefore, the set $S = \{ky : k \in \mathbb{N}\}$ is bounded above, x being an upper bound. S is non-empty because $y \in S$. By the Completeness Property of $\mathbb{R}$, $\sup S$ exists. Let $\sup S = b$. Then $ky \leq b$ for all $k \in \mathbb{N}$. Now, $b - y < b$ since $y > 0$. This shows that $b - y$ is not an upper bound of S and therefore there exists a natural number p such that $b - y < py \leq b$. This implies $(p + 1)y > b$. This shows that b is not the supremum of S since $p \in \mathbb{N} \implies p + 1 \in \mathbb{N}$ and there-

[3] Collectively, the **Corollaries 2.9.1–2.9.3** are sometimes referred to as the Archimedean Property of $\mathbb{R}$.

fore $(p+1)y \in S$. This leads to a contradiction. Therefore, our assumption is wrong and the existence of a natural number n satisfying $ny > x$ is proved.

The importance of the Supremum Property lies in the fact that it guarantees the existence of real numbers under certain hypotheses. We shall make use of it in this way many times. In fact, it can be used to prove the existence of a positive real number x such that $x^2 = 2$; that is, the positive square root of 2. You already know that x cannot be a rational number (**Proof!**); thus, you will be deriving the existence of at least one irrational[4] number.

Theorem 2.11: For every real number $a > 0$ and every integer $n > 0$, there exists a unique positive real number[5] y such that $y^n = a$.

Proof: Let $S = \{s \in \mathbb{R} : s > 0 \text{ and } s^n < a\}$. Let $t = \frac{a}{1+a}$. Then, $0 < t < 1$ and also $0 < t < a$. This implies $t^n < t < a$. Now, $t > 0$ and $t^n < a \implies t \in S$, provided that S is non-empty. Let $u = 1 + a$. Then, $u > 1$ and $u > a$. This implies $u^n > u > a$. Since $u^n > a$ and $u > 0$, u is an upper bound of S. Thus, S is a non-empty subset of $\mathbb{R}$, bounded above and hence sup S exists. Let $y = \sup S$. Clearly, $y > 0$. We prove that $y^n = a$. Suppose, if possible, either $y^n > a$ or $y^n < a$ (by the law of trichotomy).

Case I: Let $y^n > a$. Then $\frac{y^n - a}{(1+y)^n - y^n} > 0$. By the Archimedean Property of $\mathbb{R}$, there exists a natural number m such that

$$0 < \frac{1}{m} < \frac{y^n - a}{(1+y)^n - y^n}$$

or,

$$y^n - a > \frac{1}{m}\left[\binom{n}{1}y^{n-1} + \binom{n}{2}y^{n-2} + \cdots + \binom{n}{n}\right].$$

Now,

$$\begin{aligned}\left(y - \frac{1}{m}\right)^n &= y^n - \binom{n}{1}y^{n-1}\cdot\frac{1}{m} + \cdots + (-1)^n\binom{n}{n}\cdot\frac{1}{m^n} \\ &> y^n - \frac{1}{m}\left[\binom{n}{1}y^{n-1} + \binom{n}{2}y^{n-2} + \cdots + \binom{n}{n}\right] \\ &> y^n - (y^n - a) = a.\end{aligned}$$

[4]The ancient Greeks were aware of the existence of irrational numbers as early as 500 B.C. However, a satisfactory theory of such numbers was not developed until late in the nineteenth century.

[5]This real number y is written $\sqrt[n]{a}$ or $a^{\frac{1}{n}}$.

This shows that $y - \frac{1}{m}$ is an upper bound of S and this contradicts that $y = \sup S$.

Case II: Let $y^n < a$. Then $\frac{a-y^n}{(1+y)^n - y^n} > 0$. By the Archimedian Property of $\mathbb{R}$, there exists a natural number m such that

$$0 < \frac{1}{k} < \frac{a - y^n}{(1+y)^n - y^n}$$

or,

$$a - y^n > \frac{1}{k}\left[\binom{n}{1}y^{n-1} \cdot \frac{1}{k} + \binom{n}{2}y^{n-2} + \cdots + \binom{n}{n}\right].$$

Now,

$$\begin{aligned}\left(y + \frac{1}{k}\right)^n &= y^n + \binom{n}{1}y^{n-1} \cdot \frac{1}{k} + \cdots + \binom{n}{n} \cdot \frac{1}{k^n} \\ &< y^n + \frac{1}{k}\left[\binom{n}{1}y^{n-1} + \binom{n}{2}y^{n-2} + \cdots + \binom{n}{n}\right] \\ &< y^n - (a - y^n) = a.\end{aligned}$$

This shows that $y + \frac{1}{k} \in S$ and this contradicts that $x = \sup S$.

In view of the Cases I and II, we have $y^n = a$.

We now prove that y is unique. If possible, let $x \neq y$ and $x^n = a$. Then, $x > 0$, $y > 0$, and $x \neq y \implies y^n \neq x^n$. Therefore, $x^n \neq a$. So, y is unique.

Corollary 2.11: There exists a unique positive real number y such that $y^2 = 2$.

Proof: Exercise!

An ordered field is called an *Archimedian ordered field* if the Archimedian Property holds in it. Thus $\mathbb{R}$ is an Archimedian ordered field. $\mathbb{Q}$ is also an Archimedian ordered field. But $\mathbb{Q}$ is not a complete Archimedian ordered field, while $\mathbb{R}$ is so.

We end this section by discussing the geometric representation of the Archimedian Property of $\mathbb{R}$. Geometrically, it implies that of two unequal curves, surfaces or bodies, the larger of the two can become smaller than the quantity obtained by a suitable number of repition of the smaller. Let A_1 be any point on a straight line between two arbitrarily chosen points A & B. Take the points $A_2, A_3, A_4, \ldots$ so that A_1 lies between A and A_2, A_2 between

A_1 and A_3, A_3 between A_2 and A_4, and so on. Moreover, let the segments $AA_1, A_1A_2, A_2A_3, \ldots$ be equal to one another. Then, among this series of points, there always exists a certain point A_n such that B lies between A and A_n.

2.7 Density of Rational Numbers in $\mathbb{R}$

We next show that the set of rational numbers is "dense" in $\mathbb{R}$ in the sense that given any two real numbers there is a rational number between them (in fact, there are infinitely many such rational numbers).

Theorem 2.12: If x and y are any real numbers with $x < y$, then there exists a rational number $r \in \mathbb{Q}$ such that $x < r < y$.

Proof: It is no loss of generality to assume that $x > 0$. Since $y - x > 0$, it follows that (**why?**) there exists $n \in \mathbb{N}$ such that $\frac{1}{n} < y - x$. Therefore, we have $nx + 1 < ny$. Then, for $nx > 0$, we can obtain (**why?**) $m \in \mathbb{N}$ with $m-1 \leq nx < m$. Therefore, $m \leq nx+1 < ny$, whence $nx < m < ny$. Thus, the rational number $r = \frac{m}{n}$ satisfies $x < r < y$.

To round out the discussion of the interlacing of rational and irrational numbers, we have the same "betweenness property" for the set of irrational numbers.

Corollary 2.12: If x and y are real numbers with $x < y$, then there exists an irrational number z such that $x < z < y$.

Proof: If we apply **Theorem 2.12** to the real numbers $\frac{x}{\sqrt{2}}$ and $\frac{y}{\sqrt{2}}$, we obtain a rational number $r \neq 0$ (**why?**) such that

$$\frac{x}{\sqrt{2}} < r < \frac{y}{\sqrt{2}}.$$

Thus, $z = r\sqrt{2}$ is irrational (**why?**) and satisfies $x < z < y$.

2.8 Geometrical Representation of Real Numbers

The real numbers can be represented by points on a straight line. Let $X'X$ be a directed line. We take a point O on the line. O divides the line into two parts. The part to the right of O is called the positive side, the part to the left of O is called the negative side. Let us take a point A to the right of O. Let O represent the real number *zero* and A represent the real number *one*. Taking the distance OA as the unit distance on some chosen scale, each real number can be represented by a unique point on the line; a positive real number by a point lying to the right of O and a negative real number by a point lying to the left of O. A point that represents a rational number is called a rational point and a point that represents an irrational number is called an irrational point. By the density property of $\mathbb{R}$, between any two points on the line, there lie infinitely many rational points as well as infinitely many irrational points. Having a complete representation of the set $\mathbb{R}$ as points on the line, the question arises — "Does there exist any other point on the line that does not correspond to a real number?" The answer to the question is provided by *Cantor-Dedekind axiom* which states that there is a one-to-one correspondence between the set of all points on a line and the set of all real numbers. Therefore, each point on the line corresponds to only one real number and conversely, each real number is represented by only one point on the line.

Note: It will be convenient for us to suppose that a straight line is composed of points which correspond to all the numbers in the set $\mathbb{R}$. The points on the line can be looked upon as images of the numbers in $\mathbb{R}$. In view of the one-to-one correspondence between the two sets (the set of points on the line and the set of numbers in $\mathbb{R}$), we shall use the word "a point" for "a real number" and vice versa.

2.9 Intervals

The order relation on $\mathbb{R}$ determines a natural collection of subsets called *intervals*. The notations and terminology for these special sets will be familiar from earlier courses. If $a, b \in \mathbb{R}$ satisfy $a < b$, then the *open inteval* determined by a and b is the set

$$(a, b) = \{x \in \mathbb{R} : a < x < b\}.$$

The points a and b are called the endpoints of the interval; however, the endpoints are not included in an open interval. If both endpoints are adjoined to this open interval, then we obtain the *closed inteval* determined by a and b; namely, the set

$$[a, b] = \{x \in \mathbb{R} : a \leq x \leq b\}.$$

The two *half-open* (or *half-closed*) *intervals* determined by a and b are $[a, b)$, which includes the endpoint a, and $(a, b]$, which includes the endpoint b.

Each of these four intervals is bounded and has length defined by $b - a$. If $a = b$, the corresponding open interval is the empty set $(a, a) = \phi$, whereas the corresponding closed interval is the singleton set $[a, a] = \{a\}$.

There are five types of unbounded intervals for which the symbols ∞ (or

$+\infty$) and $-\infty$ are used as notational convenience in place of the endpoints. The *infinite open intervals* are the sets of the form

$$(a, \infty) = \{x \in \mathbb{R} : x > a\} \qquad \text{and} \qquad (-\infty, b) = \{x \in \mathbb{R} : x < b\}.$$

The first set has no upper bounds and the second one has no lower bounds. Adjoining endpoints gives us the *infinite closed intervals*:

$$[a, \infty) = \{x \in \mathbb{R} : x \leq a\} \qquad \text{and} \qquad (-\infty, b] = \{x \in \mathbb{R} : x \leq b\}.$$

It is often convenient to think of the entire set $\mathbb{R}$ as an infinite interval; in this case, we write $(-\infty, +\infty) = \mathbb{R}$. No point is an endpoint of $(-\infty, +\infty)$. We call $-\infty$ and $+\infty$ *points at infinity*. If S is a non-empty set of reals, we write $\sup S = +\infty$ to indicate that S is unbounded above, and $\inf S = -\infty$ to indicate that S is unbounded below. The real number system with $-\infty$ and $+\infty$ adjoined is called the *extended real number system*, or simply the *extended reals*. It must be emphasized that ∞ and $-\infty$ are **NOT** elements of $\mathbb{R}$, but only convenient symbols. The arithmetic relationships among $-\infty$, $+\infty$, and the real numbers are defined as follows.
(a) If a is any real number, then

$$\begin{aligned} a + \infty &= \infty + a = \infty, \\ a - \infty &= -\infty + a = -\infty, \\ \frac{a}{\infty} &= \frac{a}{-\infty} = 0. \end{aligned}$$

(b) If $a > 0$, then

$$\begin{aligned} a \cdot \infty &= \infty \cdot a = \infty, \\ a \cdot (-\infty) &= (-\infty) \cdot a = -\infty. \end{aligned}$$

(c) If $a < 0$, then

$$\begin{aligned} a \cdot \infty &= \infty \cdot a = -\infty, \\ a \cdot (-\infty) &= (-\infty) \cdot a = \infty. \end{aligned}$$

We also define

$$\infty + \infty = \infty \cdot \infty = (-\infty) \cdot (-\infty) = \infty$$

and

$$-\infty - \infty = \infty \cdot (-\infty) = (-\infty) \cdot \infty = -\infty.$$

Finally, we define

$$|\infty| = |-\infty| = \infty.$$

It is not useful to define $\infty - \infty$, $0 \cdot \infty$, $\frac{\infty}{\infty}$, and $\frac{0}{0}$. They are called *indeterminate forms*, and left undefined. You have studied indeterminate forms in calculus; we will look at them more carefully later.

An obvious property of intervals is that if two points x, y with $x < y$ belong to an interval I, then any point lying between them also belongs to I. That is, if $x < t < y$, then the point t belongs to the same interval as x and y. In other words, if x and y belong to an interval I, then the interval $[x, y]$ is contained in I. We now state the characterization theorem for intervals.

Theorem 2.13: If S is a subset of $\mathbb{R}$ that contains at least two points and has the property

$$\text{if } \ x, y \in S \ \text{ and } \ x < y, \ \text{ then } [x, y] \subseteq S$$

then S is an interval.

Definition 2.5: A sequence of intervals I_n, $n \in \mathbb{N}$, is said to be *nested* if the following chain of inclusions holds:

$$I_1 \supseteq I_2 \supseteq \cdots \supseteq I_n \supseteq I_{n+1} \supseteq \cdots$$

For example, if $I_n = [0, \frac{1}{n}]$ for $n \in \mathbb{N}$, then $I_n \supseteq I_{n+1}$ for each $n \in \mathbb{N}$ so that this sequence of intervals is nested. In this case, the element 0 belongs to all I_n and the Archimedean Property can be used to show that 0 is the only such common point (**Proof!**). We denote this by writing $\bigcap_{n=1}^{\infty} I_n = \{0\}$.

It is important to realize that, in general, a nested sequence of intervals need not have a common point. For example, if $J_n = (0, \frac{1}{n})$ for $n \in \mathbb{N}$, then this sequence of intervals is nested, but there is no common point, since for every $x > 0$, there exists $m \in \mathbb{N}$ such that $\frac{1}{m} < x$ so that $x \neq J_m$. Similarly, the sequence of intervals $K_n = (n, \infty)$, $n \in \mathbb{N}$, is nested but has no common point (**why?**).

However, it is an important property of $\mathbb{R}$ that every nested sequence of closed, bounded sequence of intervals does have a common point, as we will now prove. Notice that the Completeness Property of $\mathbb{R}$ plays an essential role in establishing this property.

Theorem 2.14 (Nested Intervals Property): If $I_n = [a_n, b_n]$, $n \in \mathbb{N}$, is a nested sequence of closed, bounded intervals, then there exists a number $\xi \in \mathbb{R}$ such that $\xi \in I_n$ for all $n \in \mathbb{N}$.

Proof: Since the intervals are nested, we have $I_n \subseteq I_1$ for all $n \in \mathbb{N}$, so that $a_n \leq b_1$ for all $n \in \mathbb{N}$. Hence, the non-empty set $\{a_n : n \in \mathbb{N}\}$ is bounded above, and we let ξ be its supremum. Clearly, $a_n \leq \xi$ for all $n \in \mathbb{N}$. We claim also that $\xi \in b_n$ for all n. This is established by showing that for any particular n, the number b_n is an upper bound for the set $\{a_k : k \in \mathbb{N}\}$. We consider two cases. (i) If $n \leq k$, then since $I_n \supseteq I_k$, we have $a_k \leq b_k \leq b_n$. (ii) If $k < n$, then since $I_k \supseteq I_n$, we have $a_k \leq a_n \leq b_n$. Thus, we conclude that $a_k \leq b_n$ for all k, so that b_n is an upper bound of the set $\{a_k : k \in \mathbb{N}\}$. Hence, $\xi \leq b_n$ for each $n \in \mathbb{N}$. Since $a_n \leq \xi \leq b_n$ for all n, we have $\xi \in I_n$ for all $n \in \mathbb{N}$.

Theorem 2.15: If $I_n = [a_n, b_n]$, $n \in \mathbb{N}$, is a nested sequence of closed, bounded intervals such that the lengths $b_n - a_n$ of I_n satisfy

$$\inf\{b_n - a_n : n \in \mathbb{N}\} = 0,$$

then the number ξ contained in I_n for all $n \in \mathbb{N}$ is unique.

Proof: If $\eta = \inf\{b_n : n \in \mathbb{N}\}$, then an argument similar to the proof of **Theorem 2.14** can be used to show that $a_n \leq \eta$ for all n, and hence that $\xi \leq \eta$. In fact, $x \in I_n$ for all $n \in \mathbb{N}$ if and only if $\xi \leq x \leq \eta$. If we have $\inf\{b_n - a_n : n \in \mathbb{N}\} = 0$, then for any $\epsilon > 0$, there exists $m \in \mathbb{N}$ such that $0 \leq \eta - \xi \leq b_m - a_m < \epsilon$. Since this holds for all $\epsilon > 0$, it follows that $\eta - \xi = 0$. Therefore, we conclude that $\xi = \eta$ is the only point that belongs to I_n for every $n \in \mathbb{N}$.

2.10 Countable & Uncountable Sets

The notions of "finite" and "infinite" are extremely primitive, and it is very likely that the reader has never examined these notions very carefully. We define these terms precisely and establish a few basic results and state some other important results that seem obvious but whose proofs are a bit tricky.

Definition 2.6: (a) A set is called an *empty set* (denoted by ϕ) if it has 0 elements.
(b) If $n \in \mathbb{N}$, a set S is said to have n elements if there exists a bijection from the set $\mathbb{N}_n = \{1, 2, \ldots, n\}$ onto S.
(c) A set S is said to be *finite* if it is either empty or it has n elements for some $n \in \mathbb{N}$.
(d) A set S is said to be *infinite* if it is not finite.

Since the inverse of a bijection is a bijection, it is easy to see that a set S has n elements if and only if there is a bijection from S onto the set $\{1, 2, \ldots, n\}$. Also, since the composition of two bijections is a bijection, we see that a set S_1 has n elements if and only there is a bijection from S_1 onto another set S_2 that has n elements. Further, a set T_1 is finite if and only if there is a bijection from T_1 onto another set T_2 that is finite. It is now necessary to establish some basic properties of finite sets to be sure that the definitions do not lead to conclusions that conflict with our experience of counting.

Theorem 2.16 (Uniqueness Theorem): If S is a finite set, then the number of elements in S is a unique number in $\mathbb{N}$.

Theorem 2.17: The set $\mathbb{N}$ of natural numbers is an infinite set.

Theorem 2.18: (a) If A is a set with m elements and B is a set with n elements and if $A \cap B = \phi$, then $A \cup B$ has $m + n$ elements.
(b) If A is a set with $m \in N$ elements and $C \subseteq A$ is a set with 1 element, then $A \setminus C$ is a set with $m - 1$ elements.
(c) If C is an infinite set and B is a finite set, then $C \setminus B$ is an infinite set.

Theorem 2.19: Suppose that S and T are sets and that $T \subseteq S$.
(a) If S is a finite set, then T is a finite set.
(b) If T is an infinite set, then S is an infinite set.

We now introduce an important type of infinite set.

Definition 2.7: (a) A set S is said to be *denumerable* (or *enumerable* or *countably infinite*) if there exists a bijection of $\mathbb{N}$ onto S.
(b) A set S is said to be *countable* if it is either finite or denumerable.
(c) A set S is said to be *uncountable* if it is not countable.

From the properties of bijections, it is clear that S is denumerable if and only if there exists a bijection of S onto $\mathbb{N}$. Also a set S_1 is denumerable if and only if there exists a bijection from S_1 onto a set S_2 that is denumerable. Further, a set T_1 is countable if and only if there exists a bijection from T_1 onto a set T_2 that is countable. Finally, an infinite countable set is denumerable.

Example 2.16: (a) The set $E = \{2n : n \in \mathbb{N}\}$ of even natural numbers is denumerable, since the mapping $f : \mathbb{N} \to E$ defined by $f(n) = 2n$ for $n \in \mathbb{N}$, is a bijection of $\mathbb{N}$ onto E. Similarly, the set $O = \{2n - 1 : n \in \mathbb{N}\}$ of odd natural numbers is denumerable.

(b) The set $\mathbb{Z}$ of all integers is denumerable. To construct a bijection of

$\mathbb{N}$ onto $\mathbb{Z}$, we map 1 onto 0, we map the set of even natural numbers onto the set $\mathbb{N}$ of positive integers, and we map the set of odd natural numbers onto the negative integers. This mapping can be displayed by the enumeration:

$$\mathbb{Z} = \{0, 1, -1, 2, -2, 3, -3, \ldots\}.$$

(c) The union of two disjoint denumerable sets is denumerable. Indeed, if $A = \{a_1, a_2, a_3, \ldots\}$ and $B = \{b_1, b_2, b_3, \ldots\}$, we can enumerate the elements of $A \cup B$ as:

$$a_1, b_1, a_2, b_2, a_3, b_3, \ldots$$

Theorem 2.20: The set $\mathbb{N} \times \mathbb{N}$ is denumerable.

2.11 Uncountability of $\mathbb{R}$

We will now use the Nested Interval Property to prove that the set $\mathbb{R}$ is an uncountable set. The proof was given by Georg Cantor in 1874 in the first of his papers on infinite sets. He later published a proof, known as the *Cantor's second proof*, which is the elegant "diagonal" argument based on decimal representations of real numbers. Interested readers may go through the detailed proof in **Section 2.5** of the book by **Bertle & Sherbert**.

Theorem 2.21: The set $\mathbb{R}$ of real numbers is not countable.

Proof: We will prove that the unit interval $I = [0, 1]$ is an uncountable set. This implies that the set $\mathbb{R}$ is an uncountable set, for if $\mathbb{R}$ were countable, then the subset I would also be countable (**why?**). The proof is by contradiction. If we assume that I is countable, then we can enumerate the set as $I = \{x_1, x_2, \ldots, x_n, \ldots\}$. We first select a closed subinterval I_1 of I such that $x_1 \notin I_1$, then select a closed subinterval I_2 of I such that $x_2 \notin I_2$, and so on. In this way, we obtain non-empty closed intervals

$$I_1 \supseteq I_2 \supseteq \cdots \supseteq I_n \supseteq \cdots$$

such that $I_n \subseteq I$ and $x_n \notin I_n$ for all n. Therefore $\xi \neq x_n$ for all $n \in \mathbb{N}$, so the enumeration of I is not a complete listing of the elements of I, as claimed. Hence, I is an uncountable set.

The fact that the set $\mathbb{R}$ of real numbers is uncountable can be combined with the fact that the set $\mathbb{Q}$ of rational numbers is countable to conclude that the set $\mathbb{R} \setminus \mathbb{Q}$ of irrational numbers is uncountable. Indeed, since the union of two countable sets is countable, if $\mathbb{R} \setminus \mathbb{Q}$ is countable, then since

$\mathbb{R} = \mathbb{Q} \cup (\mathbb{R} \setminus \mathbb{Q})$, we conclude that $\mathbb{R}$ is also a countable set, which is a contradiction. Therefore, the set of irrational numbers $\mathbb{R} \setminus \mathbb{Q}$ is an uncountable set.