Advanced Calculus: MATH 410
Real Numbers

Professor David Levermore

6 August 2020

1. Real Number System

1.1. **Introduction.** Numbers are at the heart of mathematics. By now you must be fairly familiar with them. Some basic sets of numbers are:

$$\text{natural numbers, } \mathbb{N} = \{0, 1, 2, \cdots\};$$

$$\text{integers (die Zahlen), } \mathbb{Z} = \{\cdots, -2, -1, 0, 1, 2, \cdots\};$$

$$\text{rational numbers (quotients), } \mathbb{Q} = \left\{\frac{p}{q} \,:\, p, q \in \mathbb{Z},\, q \neq 0\right\};$$

$$\text{real numbers, } \mathbb{R} = (-\infty, \infty);$$

$$\text{complex numbers, } \mathbb{C} = \{x + y\,\mathrm{i} \,:\, x, y \in \mathbb{R}\}.$$

Each of these sets is endowed with arithmetic operations (like 'addition' and 'multiplication') by which their elements are manipulated. It is fairly clear how $\mathbb{N}$, $\mathbb{Z}$, and $\mathbb{Q}$ are related through an increasingly richer algebraic structure. It is also fairly clear that $\mathbb{R}$ and $\mathbb{C}$ bear a similar relationship. What is less clear is the relationship between $\mathbb{Q}$ and $\mathbb{R}$. In particular, what are the properties that allow $\mathbb{R}$ and not $\mathbb{Q}$ to be identified with a 'line'? The answer will have something to do with an order relation (like 'less than') by which their elements are compared. In this chapter we address some of these issues.

We start by explaining why the rational numbers are inadequate for mathematical analysis. Simply put, the rationals do not allow us to solve equations that we would like to solve. This was also the reason behind the introduction of the negative integers and the rational numbers. The negative integers allow us to solve equations like $x+m = n$, where m, $n \in \mathbb{N}$. The rationals allow us to solve equations like $mx = n$, where m, $n \in \mathbb{Z}$ with $m \neq 0$. However, the rationals do not allow us to solve the rather simple equation $x^2 = 2$.

Proposition 1.1. *There esists no $x \in \mathbb{Q}$ such that $x^2 = 2$.*

Proof. We argue by assuming the contrary, and showing that it leads to a contradiction. Suppose there is such an $x \in \mathbb{Q}$. Then we can write it as $x = p/q$, where p and q are nonzero integers with no common factors. Because $x^2 = 2$, we see that $p^2 = 2q^2$. Hence, 2 is a factor of p^2, which implies that 2 must also be a factor of p. Therefore we can set $p = 2r$ for some nonzero integer r. Because $p^2 = 2q^2$, we see that $2^2r^2 = 2q^2$, which is the same as $q^2 = 2r^2$. Hence, 2 is a factor of q^2, which implies that 2 must also be a factor of q. It follows that 2 is a factor of both p and q, which contradicts our assumption that p and q have no common factors. Therefore no such $x \in \mathbb{Q}$ exists. □

There is nothing special about 2 in our argument. The same result is obtained for equations like $x^2 = n$ where n is any positive integer that is not a perfect square. More generally, the same result is obtained for equations like $x^m = n$ where m and n are positive integers such that $n \neq k^m$ for some integer k. The problem is that there are too many "holes" like this in $\mathbb{Q}$. In this chapter we will see how $\mathbb{R}$ fills these holes so as to allow the solution of such equations.

1.2. **Fields.** The sets $\mathbb{Q}$, $\mathbb{R}$, and $\mathbb{C}$ endowed with their natural arithmetic operations are each an example of a general algebraic structure known as a *field*.

Definition 1.1. *A* field *is a set $\mathbb{F}$ equipped with two distinguished binary operations, called* addition *and* multiplication*, that satisfy the addition, multiplication, and distributive axioms presented below. Taken together, these axioms constitute the so-called* field axioms*.*

Addition axioms. Addition maps any two $x, y \in \mathbb{F}$ to their *sum* $x + y \in \mathbb{F}$ such that:

A1: $\forall x, y \in \mathbb{F}$	$x + y = y + x,$	— commutativity;
A2: $\forall x, y, z \in \mathbb{F}$	$(x + y) + z = x + (y + z),$	— associativity;
A3: $\exists 0 \in \mathbb{F}$ such that $\forall x \in \mathbb{F}$	$x + 0 = x,$	— identity;
A4: $\forall x \in \mathbb{F}\ \exists - x \in \mathbb{F}$, such that	$x + (-x) = 0,$	— invertibility.

Multiplication axioms. Multiplication maps any two $x, y \in \mathbb{F}$ to their *product* $xy \in \mathbb{F}$ such that:

M1: $\forall x, y \in \mathbb{F}$	$xy = yx,$	— commutativity;
M2: $\forall x, y, z \in \mathbb{F}$	$(xy)z = x(yz),$	— associativity;
M3: $\exists 1 \in \mathbb{F}$ such that $1 \neq 0$ and $\forall x \in \mathbb{F}$	$x1 = x,$	— identity;
M4: $\forall x \in \mathbb{F}$ such that $x \neq 0\ \exists x^{-1} \in \mathbb{F}$ such that	$xx^{-1} = 1,$	— invertibility.

Distributive axiom. Addition and multiplication are related by:

D: $\forall x, y, z \in \mathbb{F}$	$x(y + z) = xy + xz,$	— distributivity.

Examples. When addition and multiplication have their usual meaning, the field axioms clearly hold in $\mathbb{Q}$, $\mathbb{R}$, and $\mathbb{C}$, but not in $\mathbb{N}$ or $\mathbb{Z}$. They also hold in $\mathbb{Z}_n \equiv \mathbb{Z}/(n\mathbb{Z})$ (the integers mod n) when n is prime. If you do not know this last example, do not worry. It is not critical in this course, but is covered in basic algebra courses.

All the usual rules for algebraic manipulations involving addition, subtraction, multiplication, and division can be developed from the field axioms. This is not as easy as it sounds!

1.2.1. *Consequences of the Addition Axioms.* We begin by isolating the addition axioms.

Definition 1.2. *Any set $\mathbb{G}$ equipped with a distinguished binary operation that satisfies the addition axioms is called an* Abelian group *or a* commutative group*.*

Examples. When addition has its usual meaning, the axioms for an Abelian group clearly hold in $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$, and $\mathbb{C}$, but not in $\mathbb{N}$. (As defined here, $\mathbb{N}$ satisfies all these axioms but A4.) They also hold in $\mathbb{Z}_n$ for every positive integer n.

The addition axioms immediately imply the following.

Proposition 1.2. *Let $\mathbb{G}$ be an Abelian group.*

(a) *If $x, y, z \in \mathbb{G}$ and $x + y = x + z$ then $y = z$.*
(b) *If $x, y \in \mathbb{G}$ and $x + y = x$ then $y = 0$.*
(c) *If $x, y \in \mathbb{G}$ and $x + y = 0$ then $y = -x$.*
(d) *If $x, y \in \mathbb{G}$ then $-(x + y) = (-x) + (-y)$.*
(e) *If $x \in \mathbb{G}$ then $-(-x) = x$.*

Proof. Exercise.

Assertion (a) states that addition enjoys a so-called cancellation law. Assertion (b) states that there is a unique additive identity of the type assumed in A3. This unique additive identity

is called *zero*. All other elements of $\mathbb{G}$ are said to be *nonzero*. Assertion (c) states that for every $x \in \mathbb{G}$ there is a unique additive inverse of the type assumed in A4. This unique additive inverse is called the *negative* of x. The map defined for every $x \in \mathbb{G}$ by $x \mapsto -x$ is called negation. Assertion (d) states that the negative of a sum is the sum of the negatives. Assertion (e) states that every $x \in \mathbb{G}$ is the negative of its negative.

When working with Abelian groups, it is both convenient and common to write

$$x - y\,, \qquad x + y + z\,, \qquad 2x\,, \qquad 3x\,, \qquad \cdots\,,$$

rather than

$$x + (-y)\,, \quad x + (y + z)\,, \quad x + x\,, \quad x + x + x\,, \quad \cdots\,.$$

More precisely, the symbol nx is can be defined for every group element x and every natural number n by induction. We set $0x = 0$, where the second 0 is the additive identity, and define $(n + 1)x = nx + x$ for every $n \in \mathbb{N}$. This notation satisfies the following properties.

Proposition 1.3. *Let $\mathbb{G}$ be an Abelian group.*

(a) *If $x \in \mathbb{G}$ and $m, n \in \mathbb{N}$ then $(m + n)x = mx + nx$ and $(mn)x = n(mx)$.*
(b) *If $x, y \in \mathbb{G}$ and $n \in \mathbb{N}$ then $n(x + y) = nx + ny$.*
(c) *If $x \in \mathbb{G}$ and $n \in \mathbb{N}$ then $n(-x) = -(nx)$.*

Proof. Exercise.

Motivated by these facts, for every group element x the definition of the symbol nx can be extended to every integer n by setting $nx = (-n)(-x)$ when n is negative.

1.2.2. *Consequences of the Multiplcation Axioms.* The only connection of the multiplication axioms to addition is through the references to zero in M3 and M4. An immediate consequence of M3 is that every field has at least two elements — 0 and 1. It is also clear that the nonzero elements of a field considered with the operation of multiplication form an Abelian group.

Examples. When addition and multiplication have their usual meaning, the addition and multiplication axioms clearly hold in $\mathbb{Q}$, $\mathbb{R}$, and $\mathbb{C}$, but not in $\mathbb{N}$ or $\mathbb{Z}$. They also hold in $\mathbb{Z}_n$ when n is prime.

The multiplication axioms immediately imply the following.

Proposition 1.4. *Let $\mathbb{F}$ be a field.*

(a) *If $x, y, z \in \mathbb{F}$, $x \neq 0$, and $xy = xz$ then $y = z$.*
(b) *If $x, y \in \mathbb{F}$, $x \neq 0$, and $xy = x$ then $y = 1$.*
(c) *If $x, y \in \mathbb{F}$, $x \neq 0$, and $xy = 1$ then $y = x^{-1}$.*
(d) *If $x, y \in \mathbb{F}$, $x \neq 0$ and $y \neq 0$ then $xy \neq 0$ and $(xy)^{-1} = x^{-1}y^{-1}$.*
(e) *If $x \in \mathbb{F}$ and $x \neq 0$ then $(x^{-1})^{-1} = x$.*

Proof. Exercise.

Assertion (a) states that multiplication enjoys a so-called cancellation law. Assertion (b) states that there is a unique multiplicative identity of the type assumed in M3. This unique multiplicative identity is called *one*. Assertion (c) states that for every nonzero $x \in \mathbb{F}$ there is a unique multiplicative inverse of the type assumed in M4. This unique multiplicative inverse is called the *reciprocal* of x. The map defined for every nonzero $x \in \mathbb{F}$ by $x \mapsto x^{-1}$ is called reciprocation. Assertion (d) states that the reciprocal of a product is the product of the reciprocals. Assertion (e) states that every nonzero $x \in \mathbb{F}$ is the reciprocal of its reciprocal.

When working with fields, it is both convenient and common to write

$$x/y\,, \qquad xyz\,, \qquad x^2\,, \qquad x^3\,, \quad \cdots\,,$$

rather than

$$xy^{-1}\,, \quad x(yz)\,, \quad xx\,, \quad xxx\,, \quad \cdots\,.$$

More precisely, the symbol x^n is can be defined for every field element x and every positive interger n by induction. We set $x^1 = x$ and define $x^{n+1} = x^n x$ for every $n \in \mathbb{Z}_+$, where $\mathbb{Z}_+$ denotes the positive integers. This notation satisfies the following properties.

Proposition 1.5. *Let $\mathbb{F}$ be a field.*

(a) *If $x \in \mathbb{F}$ and $m, n \in \mathbb{Z}_+$ then $x^{m+n} = x^m x^n$ and $x^{mn} = (x^m)^n$.*
(b) *If $x, y \in \mathbb{F}$ and $n \in \mathbb{Z}_+$ then $(xy)^n = x^n y^n$.*
(c) *If $x \in \mathbb{F}$, $x \neq 0$, and $n \in \mathbb{Z}_+$ then $x^n \neq 0$ and $(x^n)^{-1} = (x^{-1})^n$.*

Proof. Exercise.

Motivated by these facts, for every nonzero field element x the definition of the symbol x^n can be extended to every integer n by setting $x^0 = 1$, where the 1 is the multiplicative identity, and $x^n = (x^{-1})^{-n}$ when n is negative. The symbol 0^n remains undefined when n is not positive.

Exercise. Let $\mathbb{F}$ be a field. Extend Proposition 1.5 to $\mathbb{Z}$ by proving the following.

(a) If $x \in \mathbb{F}$, $x \neq 0$, and $m, n \in \mathbb{Z}$ then $x^{m+n} = x^m x^n$ and $x^{mn} = (x^m)^n$.
(b) If $x, y \in \mathbb{F}$, $x \neq 0$, $y \neq 0$, and $n \in \mathbb{Z}$ then $(xy)^n = x^n y^n$.
(c) If $x \in X$, $x \neq 0$, and $n \in \mathbb{Z}$ then $x^n \neq 0$ and $(x^n)^{-1} = (x^{-1})^n$.

1.2.3. *Consequences of the Distributive Axiom.* The distributive axiom gives the key connection between addition and multiplication. Taken together, the field axioms imply the following.

Proposition 1.6. *Let $\mathbb{F}$ be a field.*

(a) *If $x \in \mathbb{F}$ then $x0 = 0$.*
(b) *If $x, y \in \mathbb{F}$ and $xy = 0$ then $x = 0$ or $y = 0$.*
(c) *If $x, y \in \mathbb{F}$ then $(-x)y = -(xy) = x(-y)$.*
(d) *If $x \in \mathbb{F}$ and $x \neq 0$ then $(-x)^{-1} = -x^{-1}$.*

Proof. Exercise.

Assertion (a) states that the product of anything with zero is zero. In particular, it shows that zero cannot have a multiplicative inverse. Hence, an element has a multiplicative inverse if and only if it is nonzero. Assertion (b) states that if a product is zero, at least one of its factors must be zero. This should be compared with (d) of Proposition 1.4. Assertions (c) and (d) state how negation, multiplication, and reciprocation relate.

The field axioms allow us to extend to any field many of the formulas that you have known for years in the context of $\mathbb{R}$ or $\mathbb{C}$. For example, we can establish the following formulas.

Proposition 1.7. *Let $\mathbb{F}$ be a field. Then for every $x, y \in \mathbb{F}$ and every $n \in \mathbb{Z}_+$ we have the* difference of powers *and* binomial *formulas*

$$x^{n+1} - y^{n+1} = (x - y)\left(x^n + x^{n-1}y + \cdots + x^{n-k}y^k + \cdots + x\,y^{n-1} + y^n\right)\,, \tag{1.1}$$

$$(x + y)^n = x^n + n\,x^{n-1}y + \cdots + \frac{n!}{(n-k)!\,k!}\,x^{n-k}y^k + \cdots + n\,x\,y^{n-1} + y^n\,. \tag{1.2}$$

Proof. Exercise.

1.3. **Ordered Sets.** The sets $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{Q}$, and $\mathbb{R}$ endowed with their natural order relation are each an example of a general structure known as an ordered set.

Definition 1.3. *An* ordered set $(X, <)$ *is a set X equipped with a distinguished binary relation "$<$", called an* order*, that satisfies the order axioms presented below.*

Order axioms. A binary relation "$<$" on a set X is called an *order* whenever:

O1: if $x, y, z \in X$ then $x < y$ and $y < z$ implies $x < z$, — transitivity;
O2: if $x, y \in X$ then exactly one of $x < y$, $x = y$, or $y < x$ is true, — trichotomy.

Examples. When "$<$" has its usual meaning of "less than", the order axioms clearly hold in $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{Q}$, and $\mathbb{R}$. When "$<$" has the unusual meaning of "greater than", the order axioms also clearly hold in $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{Q}$, and $\mathbb{R}$. We will stick with the usual meaning of "$<$" in what follows.

When working with ordered sets, it is both convenient and common to use the notation

$$x > y\,, \qquad x \leq y\,, \qquad x \geq y\,,$$
$$x < y < z\,, \qquad x < y \leq z\,, \qquad \cdots\,,$$

to respectively mean

$$y < x\,, \qquad x < y \text{ or } x = y\,, \qquad y < x \text{ or } x = y\,,$$
$$x < y \text{ and } y < z\,, \qquad x < y \text{ and } y \leq z\,, \qquad \cdots\,.$$

1.3.1. *Bounds.* Ordered sets have associated notions of bounds and boundedness.

Definition 1.4. *Let $(X, <)$ be an ordered set.*

- *A point $x \in X$ is an* upper bound *(a* lower bound*) of a set $S \subset X$ whenever $y \leq x$ ($x \leq y$) for every $y \in S$.*
- *If $S \subset X$ has an upper bound (a lower bound) then S is said to be* bounded above *(*bounded below*).*
- *If $S \subset X$ is both bounded above and bounded below then it is said to be* bounded*.*

Next, ordered sets have associated notions of supremum and infimum.

Definition 1.5. *Let $(X, <)$ be an ordered set, and let $S \subset X$ be bounded above. A point $x \in X$ is a* least upper bound *or* supremum *of S whenever:*

(i) *x is an upper bound of S;*
(ii) *if $y \in X$ is also an upper bound of S then $x \leq y$.*

We similarly define a greatest lower bound *or* infimum *of S.*

If a supremum or infimum of S exists then it must be unique. The supremum of S is denoted $\sup\{S\}$ or $\sup\{z : z \in S\}$, while the infimum is denoted $\inf\{S\}$ or $\inf\{z : z \in S\}$.

Finally, the notions of supremum and infimum should not be confused with those of maximum and minimum.

Definition 1.6. *Let $(X, <)$ be an ordered set, and let $S \subset X$. A point $x \in S$ is a* maximum *(*minimum*) of S whenever x is an upper (lower) bound of S.*

If a maximum or minimum of S exists then it must be unique. The maximum of S is denoted $\max\{S\}$ or $\max\{z : z \in S\}$, while the minimum is denoted $\min\{S\}$ or $\min\{z : z \in S\}$. Moreover, if a maximum (minimum) of S exists then

$$\sup\{S\} = \max\{S\} \qquad \big(\inf\{S\} = \min\{S\}\big)\,.$$

Examples. Any bounded open interval (a, b) in $\mathbb{R}$ has neither a maximum nor a minimum, yet $\sup\{(a, b)\} = b$ and $\inf\{(a, b)\} = a$. For any bounded closed interval $[a, b]$ in $\mathbb{R}$ we have $\max\{[a, b]\} = b$ and $\min\{[a, b]\} = a$. The same is true if these intervals and their endpoints are restricted to elements of $\mathbb{Q}$.

1.3.2. *Least Upper Bound Property.* What distinguishes $\mathbb{R}$ from $\mathbb{Q}$ is the following property.

Definition 1.7. *Let $(X, <)$ be an ordered set. Then X is said to have the* least upper bound property *whenever every nonempty subset of X with an upper bound has a least upper bound.*

Remark. It may seem we should also define a "greatest lower bound property", but the next proposition shows that this is unnecessary because it is exactly the same property.

Proposition 1.8. *Let $(X, <)$ be an ordered set. Let X have the* least upper bound property*. Then every nonempty subset of X with a lower bound has a greatest lower bound.*

Proof. Let $S \subset X$ be a nonempty set with a lower bound. Let $L \subset X$ be the set of all lower bounds of S. It is nonempty and bounded above by any element of S. Therefore $\sup\{L\}$ exists. It is easy to check that $\sup\{L\} = \inf\{S\}$. □

Examples. When "$<$" has its usual meaning of "less than", the sets $\mathbb{N}$ and $\mathbb{Z}$ have the least upper bound property. However, as we will show in the next proposition, the set $\mathbb{Q}$ does not.

Proposition 1.9. *The set $\mathbb{Q}$ does not have the least upper bound property.*

Proof. Consider the sets

$$S = \big\{r \in \mathbb{Q} : r > 0\,,\, r^2 < 2\big\}\,, \qquad \tilde{S} = \big\{r \in \mathbb{Q} : r > 0\,,\, r^2 > 2\big\}\,.$$

These sets are clearly nonempty because $1 \in S$ and $2 \in \tilde{S}$. We can show that every point in $\tilde{S}$ is an upper bound for S. In order to show that S has no least upper bound, we first shows that there is no $r \in \mathbb{Q}$ such that $r^2 = 2$. It follows (by trichotomy) that if p is a least upper bound of S then either $p \in S$ or $p \in \tilde{S}$. We will show that neither can be the case. More specifically, we will show that if $p \in S$ then p is not an upper bound of S, and that if $p \in \tilde{S}$ then p is not a least upper bound of S.

Let $p \in S$. We will construct a $q \in S$ such that $p < q$, thereby showing that p is not an upper bound of S. There are many ways to construct such a q. We are seeking a rational approximation of $\sqrt{2}$ from below that is better than p. This can be done by taking one iteration of Newton's method applied to $f(x) = 1 - 2/x^2 = 0$. Set

$$q = p - \frac{f(p)}{f'(p)} = p - \frac{p^2 - 2}{4}\,p = \frac{6 - p^2}{4}\,p\,.$$

Because $x \mapsto f(x)$ is increasing and concave over $x > 0$, a picture alone should convince you this is a suitable q. Indeed, it is clear from the above formula that $0 < p < q$. A skeptic only needs to check that $q^2 < 2$. We confirm this fact by the calculation

$$2 - q^2 = 2 - \frac{36 - 12p^2 + p^4}{16}\,p^2 = \frac{(2 - p^2)^2(8 - p^2)}{16} > 0\,.$$

Now let $p \in \tilde{S}$. We will construct a $q \in \tilde{S}$ such that $q < p$, thereby showing that p is not a least upper bound of S. Once again, there are many ways to construct such a q. This time we are seeking a rational approximation of $\sqrt{2}$ from above that is better than p. This can be done by taking one iteration of Newton's method and applied to $f(x) = x^2 - 2 = 0$. Set

$$q = p - \frac{f(p)}{f'(p)} = p - \frac{p^2-2}{2p} = \frac{p^2+2}{2p}\,.$$

Because $x \mapsto f(x)$ is increasing and convex over $x > 0$, a picture alone should convince you this is a suitable p. Indeed, it is clear from the above formula that $q < p$ and that $q > 0$. A skeptic only needs to check that $q^2 > 2$. We confirm this fact by the calculation

$$q^2 - 2 = \frac{p^4+4p^2+4}{4p^2} - 2 = \frac{p^4-4p^2+4}{4p^2} = \left(\frac{p^2-2}{2p}\right)^2 > 0\,.$$

□

Remark. An alternative construction that can be used for both cases in the above proof is

$$q = p - \frac{p^2-2}{p+2} = \frac{2p+2}{p+2}\,.$$

Then

$$q^2 - 2 = \frac{4p^2+8p+4}{p^2+4p+4} - 2 = 2\frac{p^2-2}{(p+2)^2}\,.$$

While this construction yields a slicker proof, the underlying geometric picture seems less clear.

1.4. **Ordered Fields.** The sets $\mathbb{Q}$ and $\mathbb{R}$ endowed with their natural algebraic operations and order relation are each an example of a general algebraic structure known as an ordered field.

Definition 1.8. *An ordered set $(\mathbb{F}, <)$ where $\mathbb{F}$ is a field is called an* ordered field *whenever*

OF1: *if $x, y, z \in \mathbb{F}$ then $x < y$ implies $x + z < y + z$;*
OF2: *if $x, y \in \mathbb{F}$ then $0 < x$ and $0 < y$ implies $0 < xy$.*

If $x > 0$ ($x < 0$, $x \geq 0$, $x \leq 0$) then we say x is positive (negative, nonnegative, nonpositive*). The set of all positive (negative) elements of $\mathbb{F}$ is denoted $\mathbb{F}_+$ ($\mathbb{F}_-$).*

Examples. When addition, multiplication, and "<" have their usual meanings, the fields $\mathbb{Q}$ and $\mathbb{R}$ are ordered fields. In an algebra course you can learn that many other ordered fields arise in Galois theory. One such example is

$$\mathbb{F} = \left\{x + y\sqrt{2} \,:\, x, y \in \mathbb{Q}\right\}.$$

Exercise. Show that $\mathbb{F}$ in the above example is an ordered field.

Exercise. Show that $\mathbb{F}$ in the above example does not have the least upper bound property.

Examples. When n is prime and addition and multiplication have their usual meanings for the field $\mathbb{Z}_n$ and the binary relation "<" has its usual meeaning on $\{0, 1, \cdots, n-1\}$ then the ordered set $(\mathbb{Z}_n, <)$ is not an ordered field.

Exercise. Consider the complex numbers $\mathbb{C}$. For every $x, y \in \mathbb{C}$ define $x < y$ to mean

$$\mathrm{Re}(x) < \mathrm{Re}(y) \qquad \text{or} \qquad \mathrm{Re}(x) = \mathrm{Re}(y) \quad \& \quad \mathrm{Im}(x) < \mathrm{Im}(y)\,.$$

Show that $(\mathbb{C}, <)$ is an ordered set, but is not an ordered field.

1.4.1. *Consequences of the Ordered Field Axioms.* The ordered field axioms allow us to extend to any ordered field many of the rules for working with inequalities that you have known for years in the context of $\mathbb{R}$. For example, the rule that multiplying both sides of an inequality by a positive (negative) quantity will preserve (reverse) the inequality. Some of these rules are given in the following proposition.

Proposition 1.10. *Let $\mathbb{F}$ be an ordered field.*

(a) *If $x > 0$ then $-x < 0$, and vice versa.*
(b) *If $x > 0$ and $y < z$ then $y < x + z$ and $xy < xz$.*
(c) *If $x < 0$ and $y < z$ then $x + y < z$ and $xy > xz$.*
(d) *If $x \neq 0$ then $x^2 > 0$.*
(e) *If $0 < x < y$ and $n \in \mathbb{Z}_+$ then $0 < x^n < y^n$ and $0 < y^{-n} < x^{-n}$.*

Proof. Exercise.

The above proposition shows that $\mathbb{F}_+$ satisfies the following.

P1: If $x, y \in \mathbb{F}_+$ then $x + y \in \mathbb{F}_+$ and $xy \in \mathbb{F}_+$.
P2: For every $x \in \mathbb{F}$ exactly one of $x \in \mathbb{F}_+$, $-x \in \mathbb{F}_+$, or $x = 0$ is true.

These so-called positivity properties alone characterize the order relation on the field $\mathbb{F}$.

Proposition 1.11. *Let $\mathbb{F}$ be a field. Let $\mathbb{F}_+ \subset \mathbb{F}$ satisfy the positivity properties P1 and P2. Define the binary relation $<$ on $\mathbb{F}$ by*

$$x < y \qquad \textit{means} \qquad y - x \in \mathbb{F}_+ \,. \tag{1.3}$$

Then $(\mathbb{F}, <)$ is an ordered field.

Proof. Exercise.

Remark. Proposition 1.11 implies that we could have defined an ordered field as a field $\mathbb{F}$ that has a subset $\mathbb{F}_+$ satisfying the positivity properties, P1 and P2. In that case the positivity properties become the positivity axioms and, upon defining the order on $\mathbb{F}$ by (1.3), the order axioms O1, O2, OF1, and OF2 become order properties. This is the approach taken in Fitzpatrick's book.

1.4.2. *Absolute Value Function.* There is a natural absolute value function on any ordered field.

Definition 1.9. *Let $\mathbb{F}$ be an ordered field. The* absolute value *function on $\mathbb{F}$ is defined by*

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

Some of its properites are given in the following proposition. They should look familiar to you. But here the goal is to see how they follow from Definition 1.9 and the ordered field axioms.

Proposition 1.12. *Let $\mathbb{F}$ be an ordered field. Then for every $x, y \in \mathbb{F}$*

(a) $|x| \geq 0$, — *nonnegativity;*
(b) $|x| = 0$ *if and only if* $x = 0$, — *definiteness;*
(c) $|x + y| \leq |x| + |y|$, — *triangle inequality;*
(d) $|xy| = |x|\,|y|$, — *multiplicativity;*
(e) $\big||x| - |y|\big| \leq |x - y|$, — *difference inequality.*

Proof. Exercise.

With the absolute value function we can define the *distance* between points $x, y \in \mathbb{F}$ by $d(x, y) = |x - y|$. This distance function satisfies the following proposition.

Proposition 1.13. *Let $\mathbb{F}$ be an ordered field. Let $d(x, y) = |x - y|$ where $|\cdot|$ is given by Definition 1.9. Then for every $x, y, z \in \mathbb{F}$*

(a) $d(x, y) \geq 0$, — *nonnegativity;*
(b) $d(x, y) = 0$ *if and only if* $x = y$, — *definiteness;*
(c) $d(x, y) = d(y, x)$, — *symmetry;*
(d) $d(x, z) \leq d(x, y) + d(y, z)$, — *triangle inequality.*

Proof. Exercise.

We can also characterize bounded sets with the absolute value function.

Proposition 1.14. *Let $\mathbb{F}$ be an ordered field. Then $S \subset \mathbb{F}$ is bounded if and only if there exists an $m \in \mathbb{F}_+$ such that*

$$x \in S \quad \Longrightarrow \quad |x| \leq m\,.$$

Proof. Exercise.

1.5. **Real Numbers.** We now state without proof the main theorem of this chapter.

Theorem 1.1. *There exists a unique (up to an isomorphism) ordered field with the least upper bound property that contains $\mathbb{Q}$ (up to an isomorphism) as a subfield.*

Proof. Proofs of this theorem are quite long and technical. A proof of all but the uniqueness is given in the book by W. Rudin, *Principles of Mathematical Analysis*, McGraw-Hill, New York, 1976, and another is given in the book by T. Tao, *Analysis I*, Hindustan Book Agency, 2006 (available through the American Mathematical Society). The proof in Rudin is based upon a construction due to Richard Dedekind in which the real numbers are built from subsets of the rationals now called Dedekind cuts. The proof in Tao is based upon a construction due to Georg Cantor in which the real numbers are built from equivalence classes of Cauchy sequences within the rationals. Both Dedekind and Cantor published their constructions in 1872. □

Definition 1.10. *The* real numbers *are defined to be the unique ordered field with the least upper bound property whose existence is guaranteed by Theorem 1.1. This field is denoted $\mathbb{R}$.*

Remark. The least upper bound property that sets $\mathbb{R}$ apart from $\mathbb{Q}$. As we will see, it is why $\mathbb{R}$ can be identified with a line.

Remark. The real numbers became widely used in Europe only after the 1585 publication of the booklets *De Theinde*, and *L'arithmétique*, by Simon Stevin. The first advocated for the use of decimal fractions and decimal arithmetic. The latter advocated for the inclusion of zero and negatives as numbers. It showed how this view of numbers simplified the solution of quadratic equations. This view of real numbers persisted until the construction of Weierstrass in 1860, and those of Dedekind and Cantor in 1872. Of these, that of Cantor is the most universal. It generalizes to the completion of metric or uniform spaces. Roughly speaking, it views an irrational number as the class of all rational sequences that approximate it.

1.5.1. *Powers.* Recall that we showed that $y^2 = 2$ has no solution within $\mathbb{Q}$. One of the most important consequences of the fact $\mathbb{R}$ has least upper bound property is the existence of solutions to such equations.

Proposition 1.15. *For every $x \in \mathbb{R}_+$ and every $n \in \mathbb{Z}_+$ there exists a unique $y \in \mathbb{R}_+$ such that $y^n = x$.*

Proof. The uniqueness of such a y is clear because if $y < z$ then $y^n < z^n$. So we only have to show such a y exists. Consider the sets

$$S = \left\{ r \in \mathbb{R}_+ \,:\, r^n < x \right\}, \qquad \tilde{S} = \left\{ r \in \mathbb{R}_+ \,:\, r^n > x \right\}.$$

The set S is nonempty because $s = x/(1+x) < 1$ implies $s^n < s < x$, whereby $s \in S$. The set $\tilde{S}$ is nonempty because $1 + x \in \tilde{S}$. We can show that every point in $\tilde{S}$ is an upper bound for S. Let $y = \sup\{S\}$. Then (by trichotomy) either $y \in S$, $y \in \tilde{S}$, or $y^n = x$. We will show that the first two cases cannot occur, which will thereby prove the theorem. More specifically, we will show that no point in S is an upper bound of S, and that no point in $\tilde{S}$ is the least upper bound of S.

Let $p \in S$. We can construct a $q \in S$ such that $p < q$, thereby showing that p is not an upper bound of S. This can be done by taking one iteration of Newton's method applied to $f(r) = 1 - x/r^n = 0$. The details are left as an exercise.

Now suppose $p \in \tilde{S}$. We can construct a $q \in \tilde{S}$ such that $q < p$, thereby showing that p is not a least upper bound of S. This can be done by taking one iteration of Newton's method and applied to $f(r) = r^n - x = 0$. The details are left as an exercise. □

We designate the number y asserted in Proposition 1.15 by $x^{\frac{1}{n}}$.

Proposition 1.16. *Let $x \in \mathbb{R}_+$. Then*

$$(x^{\frac{1}{n}})^m = (x^m)^{\frac{1}{n}} \qquad \text{for every } m \in \mathbb{Z} \text{ and } n \in \mathbb{Z}_+ .$$

Proof. Just check that $(x^{\frac{1}{n}})^m$ and $(x^m)^{\frac{1}{n}}$ each solve the equation $y^n = x^m$. Equality then follows by Proposition 1.15. □

By setting $x^{\frac{m}{n}} = (x^{\frac{1}{n}})^m = (x^m)^{\frac{1}{n}}$, we can define x^p for every $x \in \mathbb{R}_+$ and $p \in \mathbb{Q}$. We then define x^r for every $x \in \mathbb{R}_+$ and $r \in \mathbb{R}$ by

$$x^r = \begin{cases} \sup\{x^p \,:\, p \in \mathbb{Q}\,,\, p < r\} & \text{for } x \geq 1\,, \\ \sup\{x^p \,:\, p \in \mathbb{Q}\,,\, p > r\} & \text{for } 0 < x < 1\,. \end{cases}$$

We can then show the following.

Proposition 1.17. *Let $x, y \in \mathbb{R}_+$ and $r, s \in \mathbb{R}$. Then*

(a) $x^{r+s} = x^r x^s$,
(b) $(xy)^r = x^r y^r$,
(c) $x^{rs} = (x^r)^s$, *and* $(x^r)^{-1} = (x^{-1})^r$.

Proof. Exercise.

Finally, positive powers preserve order while negative powers reverse order.

Proposition 1.18. *Let $x, y \in \mathbb{R}_+$ and $r \in \mathbb{R}_+$. Then $x < y$ implies $x^r < y^r$ and $x^{-r} > y^{-r}$.*

Proof. Exercise.

1.5.2. *Intervals.* Intervals are special subsets of $\mathbb{R}$ that will play a leading role in our study. They are denoted with the so-called interval notation. The empty set $\emptyset$ is considered to be an interval. For every $a \in \mathbb{R}$ we define $[a, a] = \{a\}$. For every $a, b \in \mathbb{R}$ such that $a < b$ we define

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

We call a and b respectively the *left endpoint* and the *right endpoint* of these intervals. Unless it is stated explicitly otherwise, when we write (a, b), $[a, b)$, or $(a, b]$ it is implied that $a < b$, while when we write $[a, b]$ it is implied that $a \leq b$.

Similarly, for every $a, b \in \mathbb{R}$ we define

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

We call a the *left endpoint* of (a, ∞) and $[a, \infty)$ and b the *right endpoint* of $(-\infty, b)$ and $(-\infty, b]$. Finally, we define $(-\infty, \infty) = \mathbb{R}$. The symbols ∞ and $-\infty$ have no meaning in $\mathbb{R}$ outside of the context of an unbounded interval. They are not considered endpoints of intervals in $\mathbb{R}$.

An endpoint of an interval is said to be *closed* if it is contained in the interval, and is said to be *open* otherwise. More specifically, $a \in \mathbb{R}$ is a closed endpoint of the intervals

$$[a, b], \qquad [a, b), \qquad [a, \infty),$$

while it is an open endpoint of the intervals

$$(a, b], \qquad (a, b), \qquad (a, \infty).$$

Similarly, $b \in \mathbb{R}$ is a closed endpoint of the intervals

$$[a, b], \qquad (a, b], \qquad (-\infty, b],$$

while it is an open endpoint of the intervals

$$[a, b), \qquad (a, b), \qquad (-\infty, b).$$

An interval is said to be *closed* if all its endpoints are closed and is said to be *open* if all its endpoints are open. More specifically, for every $a, b \in \mathbb{R}$ with $a < b$ the intervals

$$[a, b], \quad [a, \infty), \quad \text{and} \quad (-\infty, b] \quad \text{are closed},$$

while the intervals

$$(a, b), \quad [a, \infty), \quad \text{and} \quad (-\infty, b] \quad \text{are open}.$$

The intervals $[a, b)$ and $(a, b]$ are neither closed nor open. The intervals $\emptyset$ and $(-\infty, \infty)$ are both closed and open because they have no endpoints.

Exercise. Prove that for every $a \in \mathbb{R}$ and $r \in \mathbb{R}_+$ we have

$$\{x \in \mathbb{R} : |x - a| < r\} = (a - r, a + r), \qquad \{x \in \mathbb{R} : |x - a| \leq r\} = [a - r, a + r].$$

1.5.3. *Some Properties of the Real Numbers.* The following properties relate the reals $\mathbb{R}$ with the positive integers $\mathbb{Z}_+$, the integers $\mathbb{Z}$, and the rationals $\mathbb{Q}$.

Proposition 1.19. *The following hold.*

- *If $x, y \in \mathbb{R}$ and $x > 0$ then there exists $n \in \mathbb{Z}_+$ such that*
$$nx > y\,.$$
- *If $x \in \mathbb{R}$ then there exists a unique $m \in \mathbb{Z}$ such that*
$$m \in (x-1, x] \quad \big(\textit{or equivalently } x \in [m, m+1)\big)\,.$$
- *If $x, y \in \mathbb{R}$ and $x < y$ then there exists a $q \in \mathbb{Q}$ such that*
$$x < q < y\,.$$

Remark. The first assertion above says that real numbers are comparable and is called the *Archimedean property* of $\mathbb{R}$, the second is a statement about the uniform distribution of the integers, while the third asserts that $\mathbb{Q}$ is *dense* in $\mathbb{R}$ — i.e. that between any two reals lies a rational. The least upper bound property will be used to prove them, but it is not required. Indeed, the proposition remains true if $\mathbb{R}$ is replaced by $\mathbb{Q}$, but would need a different proof.

Proof. Suppose the first assertion is false. Then the set $S = \{nx : n \in \mathbb{N}\}$ is bounded above by y. By the least upper bound property S has a supremum. Let $z = \sup\{S\}$. Because $x > 0$ we have that $z - x < z$. Hence, $z - x$ is not an upper bound for S because $z = \sup\{S\}$. This implies there exists some $n \in \mathbb{N}$ such that $z - x < nx$. But then $z < (n+1)x$, which contradicts the fact z is an upper bound of S. Therefore the first assertion holds.

To prove the second assertion, by the first assertion there exists $k, l \in \mathbb{Z}_+$ such that $-x < k$ and $x < l$. It follows that $-k < x < l$. Because
$$x \in (-k, l) \subset \bigcup_{m=-k}^{l-1} [m.m+1)\,,$$
there exists some $m \in \mathbb{Z}$ such that
$$-k \le m < l \qquad \text{and} \qquad m \le x < m+1\,.$$
It then follows that $m \in (x-1, x]$. To prove uniqueness, suppose that $m, n \in \mathbb{Z}$ and $m, n \in (x-1, x]$. Without loss of generality we may suppose that $m \le n$ — otherwise simply exchange m and n. Because $x - 1 < m \le n \le x$, we see that $n - m \in \mathbb{Z}$ satisfies
$$0 \le n - m < x - (x-1) = 1\,.$$
It follows that $n - m = 0$, thereby establishing the uniqueness in the second assertion.

To prove the third assertion, because $y - x > 0$, by the first assertion there exists $n \in \mathbb{Z}_+$ such that $n(y-x) > 1$. Then by the second assertion there exists a unique $m \in (nx, nx+1]$. Combining these facts yields
$$nx < m \le nx + 1 < nx + n(y-x) = ny\,.$$
Because n is positive, we conclude that
$$x < \frac{m}{n} < y\,.$$
Therefore the third assertion holds with $q = m/n$. □

1.6. **Extended Real Numbers.** It is convenient to extend the real numbers $\mathbb{R}$ by appending two new elements designated $-\infty$ and ∞. This enlarged set is called the extended real numbers and is denoted by $\mathbb{R}_{\text{ex}}$.

1.6.1. *Order.* The order $<$ on $\mathbb{R}$ is extended to $\mathbb{R}_{\text{ex}}$ by defining

$$-\infty < x < \infty \quad \text{for every } x \in \mathbb{R}\,.$$

The ordered set $(\mathbb{R}_{\text{ex}}, <)$ has the property that ∞ ($-\infty$) is an upper (lower) bound for every subset of $\mathbb{R}_{\text{ex}}$. It also has the least upper bound property. Indeed, the supremum of any $S \subset \mathbb{R}_{\text{ex}}$ is given by

$$\sup\{S\} = \begin{cases} \infty & \text{if } S \cap \mathbb{R} \text{ has no upper bound in } \mathbb{R} \text{ or } \infty \in S\,, \\ -\infty & \text{if } S = \{-\infty\} \text{ or } S = \emptyset\,, \\ \sup\{S \cap \mathbb{R}\} & \text{otherwise}\,, \end{cases}$$

where $\emptyset$ denotes the empty set. In particular, every $S \subset \mathbb{R}$ that has no upper bound in $\mathbb{R}$ (and therefore no supremum in $\mathbb{R}$) has $\sup\{S\} = \infty$ in $\mathbb{R}_{\text{ex}}$. Similar statements hold for lower bounds and infimums. This implies that every $S \subset \mathbb{R}_{\text{ex}}$ is bounded in the sense of having both an upper bound and a lower bound in $\mathbb{R}_{\text{ex}}$.

Remark. Because some $S \subset \mathbb{R}$ can be both bounded in the ordered set $\mathbb{R}_{\text{ex}}$ and unbounded in the ordered set $\mathbb{R}$, when using terms like "bounded" and "unbounded" we must make clear which notion of bounded is being employed. Unless we specify otherwise, the terms "bounded" and "unbounded" will be used in the sense of the ordered field $\mathbb{R}$.

1.6.2. *Algebraic Operations.* The operations of addition and multiplication on $\mathbb{R}$ cannot be extended so as to make $\mathbb{R}_{\text{ex}}$ into a field. However, it is natural to extend addition by defining for every $x \in \mathbb{R}$

$$x + \infty = \infty + x = \infty\,, \qquad x - \infty = -\infty + x = -\infty\,,$$

and by defining

$$\infty + \infty = \infty\,, \qquad -\infty - \infty = -\infty\,,$$

while leaving $\infty - \infty$ and $-\infty + \infty$ undefined. In particular, $-\infty$ and ∞ do not have additive inverses.

Similarly, it is natural to extend multiplication by defining for every nonzero $x \in \mathbb{R}$

$$x\,\infty = \infty\,x = \begin{cases} \infty & \text{if } x > 0\,, \\ -\infty & \text{if } x < 0\,, \end{cases} \qquad x\,(-\infty) = (-\infty)\,x = \begin{cases} -\infty & \text{if } x > 0\,, \\ \infty & \text{if } x < 0\,, \end{cases}$$

and by defining

$$\infty\,\infty = (-\infty)\,(-\infty) = \infty\,, \qquad \infty\,(-\infty) = (-\infty)\,\infty = -\infty\,,$$

while leaving $0\,\infty$, $\infty\,0$, $0\,(-\infty)$, and $(-\infty)\,0$ undefined. In particular, 0, $-\infty$ and ∞ do not have multiplicative inverses.

1.6.3. *Intervals.* Interval notation extends naturally to $\mathbb{R}_{\text{ex}}$. Given any $a \in \mathbb{R}_{\text{ex}}$ we define $[a, a] = \{a\}$. Given any a, $b \in \mathbb{R}_{\text{ex}}$ such that $a < b$ we define the sets

$$\begin{aligned} (a,b) &= \left\{x \in \mathbb{R}_{\text{ex}} : a < x < b\right\}, & [a,b) &= \left\{x \in \mathbb{R}_{\text{ex}} : a \le x < b\right\}, \\ (a,b] &= \left\{x \in \mathbb{R}_{\text{ex}} : a < x \le b\right\}, & [a,b] &= \left\{x \in \mathbb{R}_{\text{ex}} : a \le x \le b\right\}. \end{aligned} \tag{1.4}$$

When these sets are contained within $\mathbb{R}$ the notation coincides with the interval notation we introduced earlier. Therefore we call these sets *intervals* too. The new intervals are the ones that contain either $-\infty$ or ∞ — namely, ones that have the form $[-\infty, b)$, $[-\infty, b]$, $(a, \infty]$, or $[a, \infty]$ for some a, $b \in \mathbb{R}_{\text{ex}}$. In particular, we have $\mathbb{R}_{\text{ex}} = [-\infty, \infty]$.

Here the symbols ∞ and $-\infty$ denote points in $\mathbb{R}_{\text{ex}}$. For every $a \in \mathbb{R}_{\text{ex}}$ the point ∞ is the closed right endpoint of the intervals $[a, \infty]$ and $(a, \infty]$, and is the open right endpoint of the intervals $[a, \infty)$ and (a, ∞). Similarly, for every $b \in \mathbb{R}_{\text{ex}}$ the point $-\infty$ is the closed left endpoint of the intervals $[-\infty, b]$ and $[-\infty, b)$, and is the open left endpoint of the intervals $(-\infty, b]$ and $(-\infty, b)$. Therefore every interval in $\mathbb{R}_{\text{ex}}$ that contains more than one point has two endpoints.

An interval is said to be *closed in* $\mathbb{R}_{\text{ex}}$ if is contains all its endpoints. An interval is said to be *open in* $\mathbb{R}_{\text{ex}}$ if is contains none of its endpoints.

Remark. Intervals in $\mathbb{R}$ like $[a, \infty)$ and $(-\infty, b]$ are closed in $\mathbb{R}$ but are not closed in $\mathbb{R}_{\text{ex}}$. When using the term "closed" we must make clear which notion is being employed. Unless we specify otherwise, the term "closed" will be used in the sense of being closed in $\mathbb{R}$.

We will often use the following characterization of intervals.

Proposition 1.20. (Interval Characterization Theorem.) *A set $S \subset \mathbb{R}_{\text{ex}}$ is an interval if and only if it has the property that*

$$\forall x, y \in \mathbb{R} \qquad x, y \in S \quad \textit{and} \quad x < y \qquad \Longrightarrow \qquad (x, y) \subset S\,. \tag{1.5}$$

Proof. ($\Longrightarrow$) It is clear from (1.4) that if S is an interval then it has property (1.5). In particular, the empty set and every singleton set (a set with only a single point in it) have property (1.5).

($\Longleftarrow$) The empty set and every singleton set is an interval. So we only have to consider sets that contain at least two points.

Let $S \subset \mathbb{R}_{\text{ex}}$ contain at least two points and have property (1.5). Because $\mathbb{R}_{\text{ex}}$ has the least upper bound property, while every subset of $\mathbb{R}_{\text{ex}}$ is bounded, we can set $a = \inf\{S\}$ and $b = \sup\{S\}$. Clearly $a < b$ because S has at least two points in it.

First, we show that $(a, b) \subset S$. Let $z \in (a, b)$. We claim that there exists $x \in (a, z)$ and $y \in (z, b)$ such that x, $y \in S$. (Otherwise z is either a lower or upper bound for S, which contradicts either $a = \inf\{S\}$ or $b = \sup\{S\}$.) Hence, property (1.5) implies that $z \in (x, y) \subset S$. Therefore, $(a, b) \subset S$.

Next, we claim that if $x < a$ or $x > b$ then $x \notin S$ because that would contradict either $a = \inf\{S\}$ or $b = \sup\{S\}$. When this fact is combined with the fact that $(a, b) \subset S$ it follows that S is an interval with

$$S = \begin{cases} (a,b) & \text{if } a \notin S \text{ and } b \notin S\,, \\ [a,b) & \text{if } a \in S \text{ and } b \notin S\,, \\ (a,b] & \text{if } a \notin S \text{ and } b \in S\,, \\ [a,b] & \text{if } a \in S \text{ and } b \in S\,, \end{cases}$$

where the sets (a, b), $[a, b)$, $(a, b]$, and $[a, b]$ are defined by (1.4). □

1.7. **Algebraic Extensions of the Real Numbers*.** The set $\mathbb{R}_{\text{ex}}$ of extended real numbers that was presented in the previous section behaves well with respect to the least upper bound property, but does not have a nice algebraic structure. Algebraic extensions of the real numbers have binary operations of addition and multiplication that are extensions of those operations on $\mathbb{R}$. Here we present a few. This material is optional and is not used in any later chapter.

The complex numbers $\mathbb{C}$ is one such extension of $\mathbb{R}$. It is given by

$$\mathbb{C} = \left\{ x + y\,\mathrm{i} \,:\, x, y \in \mathbb{R} \right\},$$

where i satisfies the relation $\mathrm{i}^2 = -1$. Recall that the sum and product of any two complex numbers $z_1 = x_1 + y_1\mathrm{i}$ and $z_2 = x_2 + y_2\mathrm{i}$ are respectively defined by

$$\begin{aligned} z_1 + z_2 &= (x_1 + x_2) + (y_1 + y_2)\,\mathrm{i}\,, \\ z_1 z_2 &= (x_1 x_2 - y_1 y_2) + (x_1 y_2 + x_2 y_1)\,\mathrm{i}\,. \end{aligned}$$

The set $\mathbb{C}$ endowed with these operations is a field. If we identify the real numbers with those complex numbers with no imaginary part then we see that these operations reduce to their real counterparts by setting $y_1 = y_2 = 0$.

Remark. The complex numbers were introduced in the 1500's in order to make sense of the formula for the roots of a cubic polynomial. However, they were not accepted by most mathematicians because at the time only positive numbers were considered "real". Zero and the negative numbers became "real" after works by Girard and Descartes in the 1600's. It was only after Euler made extensive use of complex numbers in the 1700's that they worked their way into the mainstream of mathematics during the early 1800's. Euler introduced the notation i for $\sqrt{-1}$, which was subsequently adopted by Gauss and others in the 1800's. Gauss also introduced the name *complex numbers*.

In 1843 William R. Hamilton discovered a four (real) dimensional algebraic extension of $\mathbb{R}$ that he dubbed the *quaternions*. It is given by

$$\mathbb{H} = \left\{ x_0 + x_1\,\mathrm{i} + x_2\,\mathrm{j} + x_3\,\mathrm{k} \,:\, x_0, x_1, x_2, x_3 \in \mathbb{R} \right\},$$

where the symbols i, j, and k satisfy the relations

$$\mathrm{i}^2 = \mathrm{j}^2 = \mathrm{k}^2 = -1\,, \quad \mathrm{i} = \mathrm{jk} = -\mathrm{kj}\,, \quad \mathrm{j} = \mathrm{ki} = -\mathrm{ik}\,, \quad \mathrm{k} = \mathrm{ij} = -\mathrm{ji}\,. \tag{1.6}$$

He called x_0 the *scalar part* and $x_1\mathrm{i}+x_2\mathrm{j}+x_3\mathrm{k}$ the *vector part* of the quaternion $x_0+x_1\mathrm{i}+x_2\mathrm{j}+x_3\mathrm{k}$. The sum and product of any two quaternions $x = x_0+x_1\mathrm{i}+x_2\mathrm{j}+x_3\mathrm{k}$ and $y = y_0+y_1\mathrm{i}+y_2\mathrm{j}+y_3\mathrm{k}$ are respectively defined by

$$\begin{aligned} x + y &= (x_0 + y_0) + (x_1 + y_1)\,\mathrm{i} + (x_2 + y_2)\,\mathrm{j} + (x_3 + y_3)\,\mathrm{k}\,, \\ xy &= (x_0y_0 - x_1y_1 - x_2y_2 - x_3y_3) + (x_0y_1 + x_1y_0 + x_2y_3 - x_3y_2)\,\mathrm{i} \\ &\quad + (x_0y_2 - x_1y_3 + x_2y_0 + x_3y_1)\,\mathrm{j} + (x_0y_3 + x_1y_2 - x_2y_1 + x_3y_0)\,\mathrm{k}\,. \end{aligned}$$

The set $\mathbb{H}$ endowed with these operations is *not* a field because this multiplication is *not* commutative. However, all the other field axioms are satisfied. The quaternions form an *associative normed division algebra*. If we identify the real numbers with those quaternions that have no vector part then we see that these operations reduce to their real counterparts by setting

$$x_1 = x_2 = x_3 = y_1 = y_2 = y_3 = 0\,.$$

Because the words scalar and vector now play a central role in linear algebra, the scalar part of a quaternion is now called the *real part*, while the *vector part* is called the imaginary part.

Exercise. Show that relations (1.6) follow from $\mathrm{i}^2 = \mathrm{j}^2 = \mathrm{k}^2 = \mathrm{ijk} = -1$.

Inspired by Hamilton's discovery of the quaternions in 1843, John Graves and Arthur Cayley independently discovered an eight (real) dimensional algebraic extension of the reals. They both published this extension in 1845 and it became known as the *octonions*. It is given by

$$\mathbb{O} = \left\{ x_0 + x_1\mathrm{e}_1 + x_2\mathrm{e}_2 + x_3\mathrm{e}_3 + x_4\mathrm{e}_4 + x_5\mathrm{e}_5 + x_6\mathrm{e}_6 + x_7\mathrm{e}_7 \,:\, x_0, x_1, x_2, x_3, x_4, x_5, x_6, x_7 \in \mathbb{R} \right\},$$

where the symbols $\mathrm{e}_1, \cdots, \mathrm{e}_7$ satisfy the relations

$$\begin{aligned}
\mathrm{e}_1^2 = \mathrm{e}_2^2 = \mathrm{e}_3^2 = \mathrm{e}_4^2 = \mathrm{e}_5^2 = \mathrm{e}_6^2 = \mathrm{e}_7^2 &= -1\,, \\
\mathrm{e}_1 = \mathrm{e}_2\mathrm{e}_3 = -\mathrm{e}_3\mathrm{e}_2 = \mathrm{e}_4\mathrm{e}_5 = -\mathrm{e}_5\mathrm{e}_4 = \mathrm{e}_7\mathrm{e}_6 = -\mathrm{e}_6\mathrm{e}_7\,, \\
\mathrm{e}_2 = \mathrm{e}_3\mathrm{e}_1 = -\mathrm{e}_1\mathrm{e}_3 = \mathrm{e}_4\mathrm{e}_6 = -\mathrm{e}_6\mathrm{e}_4 = \mathrm{e}_5\mathrm{e}_7 = -\mathrm{e}_7\mathrm{e}_5\,, \\
\mathrm{e}_3 = \mathrm{e}_1\mathrm{e}_2 = -\mathrm{e}_2\mathrm{e}_1 = \mathrm{e}_4\mathrm{e}_7 = -\mathrm{e}_7\mathrm{e}_4 = \mathrm{e}_6\mathrm{e}_5 = -\mathrm{e}_5\mathrm{e}_6\,, \\
\mathrm{e}_4 = \mathrm{e}_5\mathrm{e}_1 = -\mathrm{e}_1\mathrm{e}_5 = \mathrm{e}_6\mathrm{e}_2 = -\mathrm{e}_2\mathrm{e}_6 = \mathrm{e}_7\mathrm{e}_3 = -\mathrm{e}_3\mathrm{e}_7\,, \\
\mathrm{e}_5 = \mathrm{e}_1\mathrm{e}_4 = -\mathrm{e}_4\mathrm{e}_1 = \mathrm{e}_7\mathrm{e}_2 = -\mathrm{e}_2\mathrm{e}_7 = \mathrm{e}_3\mathrm{e}_6 = -\mathrm{e}_6\mathrm{e}_3\,, \\
\mathrm{e}_6 = \mathrm{e}_1\mathrm{e}_7 = -\mathrm{e}_7\mathrm{e}_1 = \mathrm{e}_2\mathrm{e}_4 = -\mathrm{e}_4\mathrm{e}_2 = \mathrm{e}_5\mathrm{e}_3 = -\mathrm{e}_3\mathrm{e}_5\,, \\
\mathrm{e}_7 = \mathrm{e}_6\mathrm{e}_1 = -\mathrm{e}_1\mathrm{e}_6 = \mathrm{e}_2\mathrm{e}_5 = -\mathrm{e}_5\mathrm{e}_2 = \mathrm{e}_3\mathrm{e}_4 = -\mathrm{e}_4\mathrm{e}_3\,.
\end{aligned} \tag{1.7}$$

The associated multiplication is *not* commutative and *not* associative. However, all the other field axioms are satisfied. The octonions form a *nonassociative normed division algebra*.

Remark. Both the quaternions and the octonions have not been elevated to the status of numbers. Yet they share many important properties with the complex numbers. For example, they have notions of conjugate and absolute value that satisfy

$$|x| = \sqrt{\bar{x}x}\,, \qquad |xy| = |x|\,|y|\,. \tag{1.8}$$

In each case the absolute value is the associated Euclidean norm. While not as central to mathematics as the complex numbers, both the quaternions and the octonions have proven to be useful. For example, the quaternions effeciently encode rotations in some computer graphics algorithms. With hindsight we can see that their algebraic structure implicitly appears in the 1840 work of Olinde Rodrigues on the displacement of solids and in the 1748 work of Euler on the four squares identiity. The octonions have fewer applications, but currently play a role in particle physics. Their algebraic structure implicitly appears in the 1818 work of Carl Ferdinand Degen on the eight squares identity.

Remark. It was shown by Adolph Hurwitz in 1898 that $\mathbb{R}$, $\mathbb{C}$, $\mathbb{H}$, and $\mathbb{O}$ are the only normed division alegbras over the reals. There is a sixteen (real) dimensional algebraic extension of $\mathbb{R}$ called the *sedenions* that has much less structure. For example, relation (1.8) is not satisfied. Consequently, it has found far fewer applications.

Remark. The discovery of the quaternions inspired the discovery of many other so-called *hypercomplex systems* in addition to the octonions and sedenions. The list is too long to give here. While of intrinsic mathematical interest, these have found few applications.

Remark. The discovery of the quaternions inspired developments by Hermann Grassmann, Arthur Cayley, James Sylvester, Willard Gibbs, Oliver Heaviside, and others later during the 1800s in what would become linear algebra, multilinear algebra, and vector analysis. Linear and multilinear algebra became central to mathematics because linear spaces provided a more flexible setting than the quaternions. For example, they played a major role in developments in topology and geometry. Vector analysis became the setting for much of science and engineering.

2. Sequences of Real Numbers

2.1. **Sequences and Subsequences.** Sequences play a central role in analysis. We introduce them here in the context of an arbitrary set X before specializing to sets of real numbers.

Definition 2.1. *A* sequence *in a set X is a map from $\mathbb{N}$ into X, often denoted $\{x_k\}$ or $\{x_k\}_{k\in\mathbb{N}}$, where $k \mapsto x_k$ maps the* index $k \in \mathbb{N}$ *to the point $x_k \in X$.*

Remark. Any countable ordered set may be used as the index set instead of $\mathbb{N}$.

2.1.1. *Monotonic Sequences.* Whenever X is an ordered set, sequences that either preserve or reverse order are special.

Definition 2.2. *Let $(X, <)$ be an ordered set. A sequence $\{x_k\}_{k\in\mathbb{N}}$ in X is called*

$$\begin{aligned}
\text{increasing} \quad & \textit{whenever } x_l > x_k \textit{ for every } k, l \in \mathbb{N} \textit{ with } l > k\,,\\
\text{nondecreasing} \quad & \textit{whenever } x_l \geq x_k \textit{ for every } k, l \in \mathbb{N} \textit{ with } l > k\,,\\
\text{decreasing} \quad & \textit{whenever } x_l < x_k \textit{ for every } k, l \in \mathbb{N} \textit{ with } l > k\,,\\
\text{nonincreasing} \quad & \textit{whenever } x_l \leq x_k \textit{ for every } k, l \in \mathbb{N} \textit{ with } l > k\,.
\end{aligned}$$

It is called monotonic *whenever it is either nondecreasing or nonincreasing. It is called* strictly monotonic *whenever it is either increasing or decreasing.*

The following characterizations are often used to establish the monotonicity of a sequence.

Proposition 2.1. *Let $(X, <)$ be an ordered set and $\{x_k\}_{k\in\mathbb{N}}$ be a sequence in X. The following characterizations hold.*

$$\begin{aligned}
&\{x_k\}_{k\in\mathbb{N}} \textit{ is increasing} && \textit{if and only if} && x_{k+1} > x_k \textit{ for every } k \in \mathbb{N}\,;\\
&\{x_k\}_{k\in\mathbb{N}} \textit{ is nondecreasing} && \textit{if and only if} && x_{k+1} \geq x_k \textit{ for every } k \in \mathbb{N}\,;\\
&\{x_k\}_{k\in\mathbb{N}} \textit{ is decreasing} && \textit{if and only if} && x_{k+1} < x_k \textit{ for every } k \in \mathbb{N}\,;\\
&\{x_k\}_{k\in\mathbb{N}} \textit{ is nonincreasing} && \textit{if and only if} && x_{k+1} \leq x_k \textit{ for every } k \in \mathbb{N}\,.
\end{aligned}$$

Proof. Exercise. □

2.1.2. *Eventually and Frequently.* It is convenient to introduce the concepts of *eventually* and *frequently* in the context of sequences.

Definition 2.3. *Let $\mathcal{A}(x)$ be any assertion about any $x \in X$. (For example, $\mathcal{A}(x)$ could be the assertion "$x \in S$" for a given $S \subset X$.) Let $\{x_k\}$ be a sequence in X. Then we say:*

$$\begin{aligned}
&\text{"}\mathcal{A}(x_k) \text{ eventually } \textit{as } k \to \infty\text{"} \textit{ when } \exists m \in \mathbb{N} \textit{ such that } \forall k \geq m\ \mathcal{A}(x_k)\,;\\
&\text{"}\mathcal{A}(x_k) \text{ frequently } \textit{as } k \to \infty\text{"} \textit{ when } \forall m \in \mathbb{N}\ \exists k \geq m \textit{ such that } \mathcal{A}(x_k)\,.
\end{aligned}$$

When there is no possible confusion as to the index set, we say simply "$\mathcal{A}(x_k)$ eventually" or "$\mathcal{A}(x_k)$ frequently", dropping the "as $k \to \infty$".

Exercise. Show that $2^{-k} < .001$ eventually.

Exercise. Let $\{x_k\}$ be a sequence in X. Let $\mathcal{A}(x)$ be any assertion about any $x \in X$ and let $\sim \mathcal{A}(x)$ be its negation. Show that the negation of "$\mathcal{A}(x_k)$ eventually" is "$\sim \mathcal{A}(x_k)$ frequently".

Exercise. Show that $\cos(k) > .5$ frequently, but not eventually.

2.1.3. *Subsequences.* Another useful concept is that of a *subsequence.*

Definition 2.4. *A* subsequence $\{x_{n_k}\}_{k\in\mathbb{N}}$ *of a sequence* $\{x_k\}_{k\in\mathbb{N}}$ *in a set* X *is a map from* $\mathbb{N}$ *into* X *given by* $k \mapsto x_{n_k}$*, where* $\{n_k\}_{k\in\mathbb{N}}$ *is an increasing sequence in* $\mathbb{N}$*.*

Example. If $\{x_k\}$ is a sequence in a set X, then $\{x_{2k}\}$ is the subsequence with indices that are even, while $\{x_{k^3}\}$ is the subsequence with indices that are cubes.

Exercise. Consider the sequence $\{2^k\}$. Write out the first three terms (i.e. $k = 0, 1, 2$) in the subsequences $\{2^{3k}\}$, $\{2^{2k+1}\}$, and $\{2^{k^3}\}$.

In an ordered set, subsequences of monotonic sequences are again monotonic.

Proposition 2.2. *Let* $(X, <)$ *be an ordered set. Let* $\{x_k\}_{k\in\mathbb{N}}$ *be a sequence in* X *that is increasing (nondecreasing, decreasing, nonincreasing). Then every subsequence of* $\{x_k\}_{k\in\mathbb{N}}$ *is also increasing (nondecreasing, decreasing, nonincreasing).*

Proof. Exercise.

You should test your understanding of the concepts in this section by proving the following.

Proposition 2.3. *Let* $\{x_k\}_{k\in\mathbb{N}}$ *be a sequence in a set* X *and let* $\mathcal{A}(x)$ *be an assertion about any* $x \in X$*. We have the following characterizations.*

i) $\mathcal{A}(x_k)$ *eventually as* $k \to \infty$ *if and only if for every subsequence* $\{x_{n_k}\}_{k\in\mathbb{N}}$ *we have* $\mathcal{A}(x_{n_k})$ *eventually as* $k \to \infty$*.*

ii) $\mathcal{A}(x_k)$ *frequently as* $k \to \infty$ *if and only if there exists a subsequence* $\{x_{n_k}\}_{k\in\mathbb{N}}$ *such that* $\mathcal{A}(x_{n_k})$ *eventually as* $k \to \infty$*.*

Proof. Exercise.

Remark. The implications ($\Longrightarrow$) in Proposition 2.3 are more useful than the ($\Longleftarrow$) ones.

Exercise. Let $\{x_k\}_{k\in\mathbb{N}}$ be a sequence in a set X and let $\{x_{n_k}\}_{k\in\mathbb{N}}$ be a subsequence of it. Show that if $\mathcal{A}(x)$ be an assertion about any $x \in X$ and $\mathcal{A}(x_k)$ frequently as $k \to \infty$ then we cannot generally conclude that $\mathcal{A}(x_{n_k})$ frequently as $k \to \infty$.

2.2. **Convergence and Divergence.** The most important concept related to sequences is that of convergence. Here we see it in the context of real sequences.

Definition 2.5. *A sequence* $\{a_k\}_{k\in\mathbb{N}} \subset \mathbb{R}$ *is said to* converge *or is said to be* convergent *whenever there exists a point* $a \in \mathbb{R}$ *such that*

$$\text{for every } \epsilon > 0 \qquad |a_k - a| < \epsilon \text{ eventually as } k \to \infty\,. \tag{2.1}$$

We then say that the sequence converges to a*. This is denoted as*

$$a_k \to a \text{ as } k \to \infty \qquad \text{or as} \qquad \lim_{k\to\infty} a_k = a\,.$$

A sequence that does not converge is said to diverge *or is said to be* divergent*.*

An immediate consequence of this definition is the following geometric characterization.

Proposition 2.4. *A sequence* $\{a_k\}_{k\in\mathbb{N}} \subset \mathbb{R}$ *converges to* $a \in \mathbb{R}$ *if and only if its associated sequence of distances* $\{|a_k - a|\}_{k\in\mathbb{N}}$ *converges to* 0*.*

Proof. Exercise.

An immediate consequence of Definition 2.5 and some ideas from the previous section is the following proposition.

Proposition 2.5. *A sequence $\{a_k\}_{k\in\mathbb{N}} \subset \mathbb{R}$ diverges if and only if for every $a \in \mathbb{R}$ there exists an $\epsilon_a > 0$ such that*

$$|a_k - a| \geq \epsilon_a \text{ frequently as } k \to \infty\,.$$

Proof. Exercise.

Definition 2.5 does not assert that there is a unique number a that satisfies (2.1). The following proposition establishes this and more.

Proposition 2.6. *If a sequence $\{a_k\}_{k\in\mathbb{N}} \subset \mathbb{R}$ converges, there is a unique point in $\mathbb{R}$ to which it converges. Moreover, the set $\{a_k\} \subset \mathbb{R}$ is bounded.*

Proof. Here we prove only the boundedness assertion. The proof of the uniqueness assertion is left as an exercise.

Let $\{a_k\}_{k\in\mathbb{N}}$ converge to $a \in \mathbb{R}$. Then $\exists m \in \mathbb{N}$ such that $\forall k \geq m\ |a_k - a| < 1$. In particular, for every $k \geq m$ we have that $a - 1 < a_k < a + 1$. Then for every $k \in \mathbb{N}$ we have

$$|a_k| < 1 + \max\left\{|a_0|, |a_1|, \ldots, |a_{m-1}|, |a|\right\}.$$

Therefore the sequence $\{a_k\}_{k\in\mathbb{N}}$ is bounded. □

Definition 2.6. *The unique point to which a convergent sequence in $\mathbb{R}$ converges is called the* limit *of the sequence.*

An important characterization of the limit of a convergent sequence is given by the following.

Proposition 2.7. *Let $\{b_k\}_{k\in\mathbb{N}}$ be a convergent sequence in $\mathbb{R}$. Let $b \in \mathbb{R}$. Then*

$$\lim_{k\to\infty} b_k = b\,,$$

if and only if for every $(a, c) \subset \mathbb{R}$

$$b \in (a, c) \qquad \Longrightarrow \qquad b_k \in (a, c) \quad \textit{eventually}\,.$$

Proof. Exercise.

Subsequences of a convergent sequence are also convergent, and have the same limit.

Proposition 2.8. *If a sequence $\{a_k\}_{k\in\mathbb{N}} \subset \mathbb{R}$ converges to a limit $a \in \mathbb{R}$ then every subsequence of $\{a_k\}_{k\in\mathbb{N}}$ also converges to a.*

Proof. Exercise.

The main theorem regarding basic operations, order, and limits is the following.

Proposition 2.9. *Let $\{a_k\}$ and $\{b_k\}$ be convergent sequences in $\mathbb{R}$ with $a_k \to a$ and $b_k \to b$ as $k \to \infty$. Then*

(i) $(a_k + b_k) \to (a + b)\,,$

(ii) $-a_k \to -a\,,$

(iii) $a_k b_k \to ab\,,$

(iv) $1/a_k \to 1/a$ *provided no division by zero occurs*,

(v) $|a_k| \to |a|\,.$

Moreover, if $a_k \leq b_k$ frequently then $a \leq b$. (Equivalently, if $a < b$ then $a_k < b_k$ eventually.)

Proof. Exercise.

When working with real sequences, it is useful to distinguish two of the many ways in which a sequence might diverge — namely, those when the sequence "approaches" either ∞ or $-\infty$.

Definition 2.7. *A sequence* $\{a_k\}_{k\in\mathbb{N}} \subset \mathbb{R}$ *is said to* diverge to ∞ (to $-\infty$) *if for every* $b \in \mathbb{R}$ *we have that*

$$a_k > b \text{ eventually } \ (a_k < b \text{ eventually}) \ \text{ as } k \to \infty.$$

This is denoted as

$$a_k \to \infty \ \ (a_k \to -\infty) \ \text{ as } k \to \infty\,,$$

or as

$$\lim_{k\to\infty} a_k = \infty \quad \left(\lim_{k\to\infty} a_k = -\infty\right).$$

Remark. Some sources say that such a sequence *approaches* ∞ (*approaches* $-\infty$). Our choice of language is evident from the following exercise.

Exercise. Show that if a sequence diverges to either ∞ or $-\infty$ then it is divergent.

2.3. **Monotonic Sequences.** For monotonic sequences the least upper bound property can be employed to show the existence of limits.

Proposition 2.10. (Monotonic Sequence Theorem) *Let* $\{a_k\}_{k\in\mathbb{N}}$ *be a sequence in* $\mathbb{R}$ *that is nondecreasing (nonincreasing). Then it converges if and only if it is bounded above (bounded below). Moreover, if it converges then*

$$\lim_{k\to\infty} a_k = \sup\{a_k\} \quad \left(\lim_{k\to\infty} a_k = \inf\{a_k\}\right),$$

while if it diverges then

$$\lim_{k\to\infty} a_k = \infty \quad \left(\lim_{k\to\infty} a_k = -\infty\right).$$

Proof. We give the proof for the nondecreasing case only; the nonincreasing case goes similarly. The proof of the assertion in the case of divergence in is left as an exercise.

($\Rightarrow$) This follows from Proposition 2.6, which states that every convergent sequence is bounded.

($\Leftarrow$) Because $\{a_k\}$ is bounded above, we can set $a = \sup\{a_k\,;\, k \in \mathbb{N}\}$ by the least upper bound property. We claim that $a_k \to a$ as $k \to \infty$. Let $\epsilon > 0$ be arbitrary. There exists some $m_\epsilon \in \mathbb{N}$ such that $0 \le a - a_{m_\epsilon} < \epsilon$. (For if not, it would mean that $a_k \le a - \epsilon$ for every $k \in \mathbb{N}$, which would contradict the definition of a.) But then for every $k > m_\epsilon$ we have that $0 \le a - a_k \le a - a_{m_\epsilon} < \epsilon$, which establishes the claim. □

For monotonic sequences it is enough to know what happens to a single subsequence.

Proposition 2.11. *Let* $\{a_k\}$ *be a monotonic sequence in* $\mathbb{R}$. *Then* $\{a_k\}$ *is convergent if and only if it has a convergent subsequence.*

Proof. Exercise.

The Monotonic Sequence Theorem has the following consequence, often attributed to Cantor.

Proposition 2.12. (Nested Interval Theorem.) *Let $\{[a_k, b_k]\}_{k\in\mathbb{N}}$ be a sequence of closed, bounded intervals in $\mathbb{R}$ that is nested in the sense that*

$$[a_{k+1}, b_{k+1}] \subset [a_k, b_k] \quad \textit{for every } k \in \mathbb{N}\,.$$

Then the sequences $\{a_k\}$ and $\{b_k\}$ converge and

$$\bigcap_{k\in\mathbb{N}} [a_k, b_k] = [a, b]\,, \quad \textit{where} \quad a = \lim_{k\to\infty} a_k\,, \quad b = \lim_{k\to\infty} b_k\,, \quad \textit{with } a \le b\,.$$

Proof. Because the sequence of intervals $\{[a_k, b_k]\}$ is nested, it follows that the sequence $\{a_k\}$ is nondecreasing while the sequence $\{b_k\}$ is nonincreasing. Hence, for every $m, n \in \mathbb{N}$ we have $a_m \le a_n \le b_n$ when $m \le n$ and $a_m \le b_m \le b_n$ when $m \ge n$. This implies the sequence $\{a_k\}$ is bounded above by every b_n, while the sequence $\{b_k\}$ is bounded below by every a_n. The Monotone Sequence Theorem then implies that the sequences $\{a_k\}$ and $\{b_k\}$ converge with

$$a = \lim_{k\to\infty} a_k = \sup_k\{a_k\} \le b_n\,, \qquad b = \lim_{k\to\infty} b_k = \inf_k\{b_k\} \ge a_n\,.$$

Because $a_k \le b_k$ for every $k \in \mathbb{N}$ it follows that $a \le b$.

If $x \in [a, b]$ then because $[a, b] \subset [a_k, b_k]$ for every $k \in \mathbb{N}$, it follows that $x \in \cap_k [a_k, b_k]$. If $x < a$ then because $a_k \to a$, it follows that $x < a_k$ eventually, whereby $x \notin \cap_k [a_k, b_k]$. Similarly, if $b < x$ then because $b_k \to b$, it follows that $b_k < x$ eventually, whereby $x \notin \cap_k [a_k, b_k]$. □

Remark. This theorem shows the above intersection of nested intervals is always nonempty. In particular, when $a = b$ this intersection consists of a single point. To better appreciate significance of this result, you should do the following exercise.

Exercise. Show that a nested sequence of closed, bounded intervals in $\mathbb{Q}$ can have an empty intersection.

2.4. **Limits and *e*.** While e emerged in the 1600's, it was named by Leonhard Euler by 1731. Euler published the first detailed study of e in 1748. One thing that he showed was

$$\lim_{k\to\infty} \left(1 + \frac{1}{k}\right)^k = \lim_{k\to\infty} \left(1 - \frac{1}{k}\right)^{-k} = e\,. \tag{2.2}$$

The first limit had already appeared in the 1683 work of Jacob Bernoulli on compound interest. Here we use the Monotonic Sequence Theorem to prove that these limits exist and are equal. We will then define e by these limits. In fact, we prove that more general limits exist.

Proposition 2.13. *For every $x \in \mathbb{R}$ we have the convergent limits*

$$\lim_{k\to\infty} \left(1 + \frac{x}{k}\right)^k = \lim_{k\to\infty} \left(1 - \frac{x}{k}\right)^{-k}. \tag{2.3}$$

Moreover, when $x \neq 0$ the sequence in the first limit is increasing for $k > -x$ while that in the second limit is decreasing for $k > x$.

Remark. Later we will see these limits converge to e^x. Here we only show they converge.

Proof. Let $x \in \mathbb{R}$. When $x = 0$ the result is easy because both sequences reduce to the constant sequence $\{1\}$, so we only need to consider the case $x \neq 0$.

We will use two inequalities derived from the difference of powers formula (1.1), which for every $n \in \mathbb{Z}_+$ and every $y, z \in \mathbb{R}$ gives

$$z^{n+1} - y^{n+1} = (z - y)\big(z^n + z^{n-1}y + \cdots + zy^{n-1} + y^n\big)\,.$$

Because there are $n+1$ terms in the sum that appears inside the last factor, we see that for every $n \in \mathbb{Z}_+$ and every $z > y > 0$ we have the inequalities

$$(n+1)(z-y)y^n < z^{n+1} - y^{n+1} < (n+1)(z-y)z^n \,. \tag{2.4}$$

We will now use the inequalities (2.4) to prove the monotonicity of the sequences that appear in (2.3). First, if $x > 0$ and $k \in \mathbb{Z}_+$ then setting $z = 1 + \frac{x}{k}$, $y = 1 + \frac{x}{k+1}$, and $n = k$ into the second inequality of (2.4) yields

$$\left(1 + \frac{x}{k}\right)^{k+1} - \left(1 + \frac{x}{k+1}\right)^{k+1} < \frac{x}{k}\left(1 + \frac{x}{k}\right)^{k} .$$

Similarly, if $x < 0$ and $k \in \mathbb{Z}_+$ with $k > -x$ then setting $z = 1 + \frac{x}{k+1}$, $y = 1 + \frac{x}{k}$, and $n = k$ into the first inequality of (2.4) yields

$$-\frac{x}{k}\left(1 + \frac{x}{k}\right)^{k} < \left(1 + \frac{x}{k+1}\right)^{k+1} - \left(1 + \frac{x}{k}\right)^{k+1} .$$

By combining these two inequalities we see that

$$a_k = \left(1 + \frac{x}{k}\right)^{k} < \left(1 + \frac{x}{k+1}\right)^{k+1} = a_{k+1} \qquad \text{for every } k > -x \,.$$

Upon replacing x with $-x$ and taking reciprocals we obtain

$$b_{k+1} = \left(1 - \frac{x}{k+1}\right)^{-(k+1)} < \left(1 - \frac{x}{k}\right)^{-k} = b_k \qquad \text{for every } k > x \,.$$

Therefore $\{a_k\}$ is increasing for $k > -x$ while $\{b_k\}$ is decreasing for $k > x$.

Next, we show that the sequence $\{a_k\}$ is bounded above and the sequence $\{b_k\}$ is bounded below. For $k > |x|$ we have

$$\frac{a_k}{b_k} = \left(1 + \frac{x}{k}\right)^{k}\left(1 - \frac{x}{k}\right)^{k} = \left(1 - \frac{x^2}{k^2}\right)^{k} < 1 \,. \tag{2.5}$$

When this fact is combined with the monotonicity facts we conclude that

$$a_k \le a_{\max\{k,l\}} < b_{\max\{k,l\}} \le b_l \qquad \text{for all integers } k, l \ge |x| \,.$$

Hence, $\{a_k\}$ is bounded above by every b_l while $\{b_l\}$ is bounded below by every a_k. By the Monotonic Sequence Theorem (Proposition 2.10) these sequences are convergent with

$$\lim_{k\to\infty} a_k \le \lim_{l\to\infty} b_l \,.$$

Finally, we show that the sequences $\{a_k\}$ and $\{b_k\}$ have the same limit. If $k > 1$ then setting $z = 1$, $y = 1 - \frac{x^2}{k^2}$, and $n = k - 1$ into the second inequality of (2.4) yields

$$1 - \left(1 - \frac{x^2}{k^2}\right)^{k} < \frac{x^2}{k} \,.$$

Therefore we see from the equalities in (2.5) that

$$\left(1 - \frac{x^2}{k}\right) b_k < a_k < b_k \qquad \text{for every } k > 1 + |x| \,.$$

By passing to the limit in these inequalities we see that $\{a_k\}$ and $\{b_k\}$ have the same limit. $\square$

2.5. **Wallis Product Formula*.** In 1656, before Newton devoloped calculus, John Wallis computed the area of a semidisk and showed that $\frac{\pi}{2}$ could be expressed as the following limit.

Proposition 2.14.

$$\lim_{k\to\infty} \frac{2^2\cdot 4^2\cdot 6^2\cdots(2k-2)^2\cdot(2k)^2}{1\cdot 3^2\cdot 5^2\cdots(2k-1)^2\cdot(2k+1)} = \frac{\pi}{2}\,. \tag{2.6}$$

Proof. We will prove this limit after we borrow some facts from calculus that we have not covered yet in this course, but that you already know. For every $n\in\mathbb{N}$ we define

$$c_n = \int_0^{\frac{\pi}{2}} \cos(x)^n\,\mathrm{d}x\,.$$

Because $0\le\cos(x)\le 1$ for every $x\in[0,\frac{\pi}{2}]$ we see that $\{c_n\}_{n\in\mathbb{N}}$ is a nonincreasing sequence. We obtain $c_0=\frac{\pi}{2}$ and $c_1=1$ by direct integration. We can use integration by parts to derive the recursion formula

$$c_n = \frac{n-1}{n}\,c_{n-2} \qquad \text{for every } n\ge 2\,.$$

With these facts in hand, we can prove the Wallis product formula (2.6).

First, we can use induction to show that for every $k\in\mathbb{N}$

$$c_{2k} = \frac{\pi}{2}\cdot\frac{1}{2}\cdot\frac{3}{4}\cdot\frac{5}{6}\cdots\frac{2k-3}{2k-2}\cdot\frac{2k-1}{2k}\,, \qquad c_{2k+1} = 1\cdot\frac{2}{3}\cdot\frac{4}{5}\cdot\frac{6}{7}\cdots\frac{2k-2}{2k-1}\cdot\frac{2k}{2k+1}\,.$$

Because $\{c_n\}_{n\in\mathbb{N}}$ is a nonincreasing sequence, we have $c_{2k+2}\le c_{2k+1}\le c_{2k}$ for every $k\in\mathbb{N}$. Upon dividing this by c_{2k} and multiplying the result by $\frac{\pi}{2}$, we obtain the inequalities

$$\frac{2k+1}{2k+2}\cdot\frac{\pi}{2} \le \frac{2^2\cdot 4^2\cdot 6^2\cdots(2k-2)^2\cdot(2k)^2}{1\cdot 3^2\cdot 5^2\cdots(2k-1)^2\cdot(2k+1)} \le \frac{\pi}{2}\,.$$

By passing to the limit in these inequalities we obtain the Wallis product formula (2.6). □

2.6. **De Moivre-Stirling Formula*.** Another famous limit is the de Moivre-Stirling formula.

Proposition 2.15. (de Moivre-Stirling)

$$\lim_{n\to\infty} \frac{n!}{\sqrt{2\pi n}\,n^n e^{-n}} = 1\,. \tag{2.7}$$

Remark. This shows that $n!$ behaves like $\sqrt{2\pi n}\,(n/e)^n$ for large n. The convergence of the limit was established in 1730 by Abraham de Moivre in his work on probability. Later that same year James Stirling used the Wallis product formula to evaluate the limit. Formula (2.7) is commonly called the Stirling Formula or the Stirling Approximation, which are names that give Stirling all the credit for a result that was due largely to de Moivre.

We first use the Monotonic Sequence Theorem to show that the limit exists.

Proposition 2.16. (De Moivre Convergence Theorem)

$$\lim_{n\to\infty} \frac{n!}{\sqrt{n}\,n^n e^{-n}} \quad \textit{converges to a positive limit}\,. \tag{2.8}$$

Moreover, the sequence that appears in the limit is decreasing.

Remark. The following proof uses facts from calculus that we have not covered yet. For example, it uses the functions $\log(x)$, e^x, and their basic properties. Strictly speaking, our proof will be incomplete until all these logical gaps are filled. However, we give it now to illustrate a substantial application of the Monotonic Sequence Theorem.

Proof. We will bound $\ell_n = \log(n!)$. By properties of logarithms

$$\ell_n = \log(n!) = \log\bigl(1 \cdot 2 \cdots (n-1) \cdot n\bigr) = \sum_{k=1}^{n} \log(k)\,. \tag{2.9}$$

This looks like a numerical approximation to $\int_1^n \log(x)\,\mathrm{d}x$. We will used the facts that:

- $\log(x)$ is strictly concave (down) over $(0,\infty)$;
- if $f(x)$ is a differentiable strictly concave function then the trapezoidal approximation $\mathcal{T}_h[f]$ with uniform subintervals of length h underestimates the definite integral $\int_a^b f(x)\,\mathrm{d}x$ with a correction that satisfies the bounds

$$0 < \int_a^b f(x)\,\mathrm{d}x - \mathcal{T}_h[f] < \frac{h^2}{8}\bigl(f'(a) - f'(b)\bigr)\,.$$

In particular, for $a = 1$, $b = n$, $f(x) = \log(x)$, and $h = 1$ we have $\mathcal{T}_1[\log] = \ell_n - \frac{1}{2}\log(n)$ and $f'(x) = 1/x$, whereby

$$0 < \int_1^n \log(x)\,\mathrm{d}x + \tfrac{1}{2}\log(n) - \ell_n < \frac{1}{8}\left(1 - \frac{1}{n}\right) < \frac{1}{8}\,. \tag{2.10}$$

Now define the sequence $\{c_n\}_{n=2}^{\infty}$ by

$$c_n = \int_1^n \log(x)\,\mathrm{d}x + \tfrac{1}{2}\log(n) - \ell_n = \sum_{k=1}^{n-1}\left(\int_k^{k+1} \log(x)\,\mathrm{d}x - \frac{\log(k) + \log(k+1)}{2}\right)\,. \tag{2.11}$$

The terms in this sum are the corrections of the trapezoidal approximation to $\int_k^{k+1}\log(x)\,\mathrm{d}x$ with $h = 1$, so they are each positive. Therefore the sequence $\{c_n\}_{n=2}^{\infty}$ is increasing. It is also bounded with $0 < c_n < \frac{1}{8}$ by (2.10).

Because $x\log(x) - x$ is a primitive (antiderivative) of $\log(x)$, we have

$$c_n = \int_1^n \log(x)\,\mathrm{d}x + \tfrac{1}{2}\log(n) - \ell_n = (n + \tfrac{1}{2})\log(n) - n + 1 - \ell_n\,.$$

Therefore, by properties of exponentials and logarithms we have

$$r_n = \frac{n!}{\sqrt{n}\,n^n e^{-n}} = \exp\bigl(\ell_n - (n+\tfrac{1}{2})\log(n) + n\bigr) = e^{1-c_n}\,. \tag{2.12}$$

Because $\{c_n\}_{n=2}^{\infty}$ is an increasing sequence that satifies the bounds $0 < c_n < \frac{1}{8}$ while e^x is an increasing function of x, we see that $\{r_n\}_{n=2}^{\infty}$ is a decreasing sequence that satifies the bounds

$$e^{\frac{7}{8}} < r_n = \frac{n!}{\sqrt{n}\,n^n e^{-n}} < e \qquad \text{for every } n \ge 2\,. \tag{2.13}$$

Therefore the Monotonic Sequence Theorem (Proposition 2.10) implies that $\{r_n\}_{n=2}^{\infty}$ converges. Moreover, its limit lies within $[e^{\frac{7}{8}}, e)$ and is thereby positive. $\square$

We now prove the de Moivre-Stirling formula (2.7) by using the Wallis product formula (2.6) to evaluate the limit (2.8), thereby establishing Proposition 2.15.

Proof. Let r_n be defined by (2.12). By (2.13) and the de Moivre Convergence Theorem

$$\lim_{n\to\infty} r_n = r \qquad \text{for some } r \in [e^{\frac{7}{8}}, e)\,.$$

To establish (2.7) we need to show that $r = \sqrt{2\pi}$. The Wallis product formula (2.6) states

$$\lim_{n\to\infty} w_n = \frac{\pi}{2} \qquad \text{where} \quad w_n = \frac{2^2\cdot 4^2\cdot 6^2\cdots(2n-2)^2\cdot(2n)^2}{1\cdot 3^2\cdot 5^2\cdots(2n-1)^2\cdot(2n+1)}\,.$$

It is easily checked that

$$\begin{aligned}\frac{r_n^2}{r_{2n}} &= \frac{(n!)^2}{n\,n^{2n}e^{-2n}}\cdot\frac{\sqrt{2n}\,(2n)^{2n}e^{-2n}}{(2n)!} = \sqrt{\frac{2}{n}}\cdot\frac{(2^n\,n!)^2}{(2n)!}\\ &= \sqrt{\frac{2}{n}}\cdot\frac{2\cdot 4\cdot 6\cdots(2n-2)\cdot 2n}{1\cdot 3\cdot 5\cdots(2n-3)\cdot(2n-1)} = \sqrt{\frac{2}{n}\,(2n+1)\,w_n}\,.\end{aligned}$$

The de Moivre Convergence Theorem (Proposition 2.16) and the properties of limits from Propositions 2.8 and 2.9 then imply that

$$r = \lim_{n\to\infty}\frac{r_n^2}{r_{2n}} = \lim_{n\to\infty}\sqrt{\frac{2}{n}\,(2n+1)\,w_n} = \sqrt{4\,\frac{\pi}{2}} = \sqrt{2\pi}\,.$$

Therefore the de Moivre-Stirling formula (2.7) holds. □

Exercise. Prove the factorial-root limit

$$\lim_{n\to\infty}\frac{(n!)^{\frac{1}{n}}}{n} = \frac{1}{e}\,. \tag{2.14}$$

Exercise. Show that the bounds (2.13) can be sharpened to

$$\sqrt{2\pi} < r_n = \frac{n!}{\sqrt{n}\,n^n e^{-n}} < e \qquad \text{for every } n \geq 2\,.$$

Remark. $\sqrt{2\pi}$ is a bit more than 2.5 while $e^{\frac{7}{8}}$ is a bit less than 2.4.

Exercise. Prove that

$$\lim_{n\to\infty}\left(1+\frac{1}{n}\right)^{n+\frac{1}{2}} = e\,,$$

and that the sequence that appears in this limit is decreasing.

Exercise. Define the sequence $\{d_n\}_{n=2}^{\infty}$ by

$$d_n = \int_1^{n+\frac{1}{2}} \log(x)\,\mathrm{d}x - \ell_n\,,$$

where ℓ_n is given by (2.9). Show that $\{d_n\}_{n=2}^{\infty}$ is a decreasing sequence such that $c_n < d_n$ for every $n \geq 2$, where c_n is given by (2.11). Show moreover that $\{d_n\}_{n=2}^{\infty}$ is convergent with

$$\lim_{n\to\infty} d_n = \lim_{n\to\infty} c_n\,.$$

2.7. **Limit Superior and Limit Inferior.** The power of the Monotonic Sequence Theorem (Proposition 2.10) lies in the fact that from every real sequence $\{a_k\}_{k\in\mathbb{N}}$ that is bounded above (bounded below), we can construct a nonincreasing (nondecreasing) sequence from its "tails". Specifically, we construct the sequence $\{\overline{a}_k\}$ ($\{\underline{a}_k\}$) with elements defined by

$$\overline{a}_k = \sup\big\{a_l \,:\, l \geq k\big\} \quad \big(\underline{a}_k = \inf\big\{a_l \,:\, l \geq k\big\}\big)\,.$$

(Here you should think of $\overline{a}$ and $\underline{a}$ as new letters that are related to a.) This sequence is clearly nonincreasing (nondecreasing). The convergence of such sequences is characterized by the Monotonic Sequence Theorem, which motivates the following definition.

Definition 2.8. *For every sequence $\{a_k\}_{k\in\mathbb{N}}$ in $\mathbb{R}$, its* limit superior *and* limit inferior *are*

$$\limsup_{k\to\infty} a_k \equiv \begin{cases} \lim\limits_{k\to\infty} \overline{a}_k & \text{if } \sup\{a_k\} < \infty\,, \\ \infty & \text{otherwise}\,; \end{cases}$$

$$\liminf_{k\to\infty} a_k \equiv \begin{cases} \lim\limits_{k\to\infty} \underline{a}_k & \text{if } \inf\{a_k\} > -\infty\,, \\ -\infty & \text{otherwise}\,. \end{cases}$$

These are called simply the "$\limsup$*" and "*$\liminf$*" for short.*

Remark. By the Monotonic Sequence Theorem (Proposition 2.10) we have that

$$\limsup_{k\to\infty} a_k \equiv \begin{cases} \inf\{\overline{a}_k\} & \text{if } \sup\{a_k\} < \infty\,, \\ \infty & \text{otherwise}\,; \end{cases}$$

$$\liminf_{k\to\infty} a_k \equiv \begin{cases} \sup\{\underline{a}_k\} & \text{if } \inf\{a_k\} > -\infty\,, \\ -\infty & \text{otherwise}\,. \end{cases}$$

Example. Consider the sequence $\{a_k\}$ given by

$$a_k = (-1)^k\,\frac{k+1}{k} \quad \text{for } k \in \mathbb{Z}_+\,.$$

The first eight terms of the sequences $\{\overline{a}_k\}$, $\{a_k\}$, and $\{\underline{a}_k\}$ are

$\overline{a}_k$	$\frac{3}{2}$	$\frac{3}{2}$	$\frac{5}{4}$	$\frac{5}{4}$	$\frac{7}{6}$	$\frac{7}{6}$	$\frac{9}{8}$	$\frac{9}{8}$
a_k	-2	$\frac{3}{2}$	$-\frac{4}{3}$	$\frac{5}{4}$	$-\frac{6}{5}$	$\frac{7}{6}$	$-\frac{8}{7}$	$\frac{9}{8}$
$\underline{a}_k$	-2	$-\frac{4}{3}$	$-\frac{4}{3}$	$-\frac{6}{5}$	$-\frac{6}{5}$	$-\frac{8}{7}$	$-\frac{8}{7}$	$-\frac{10}{9}$

Notice that $\{a_k\}$ diverges while $\{\overline{a}_k\}$ and $\{\underline{a}_k\}$ are both monotonic and converge to 1 and -1 respectively. Therefore

$$\limsup_{k\to\infty} a_k = 1\,, \qquad \liminf_{k\to\infty} a_k = -1\,.$$

Remark. Notice that, unlike the limit, the $\limsup$ and $\liminf$ are defined for every real sequence, taking values in $\mathbb{R}_{\text{ex}}$, and that in general

$$-\infty \leq \liminf_{k\to\infty} a_k \leq \limsup_{k\to\infty} a_k \leq \infty\,.$$

Example. The sequence $\{(-1)^k k\}$ is neither bounded above nor bounded below. Therefore

$$\limsup_{k\to\infty} (-1)^k k = \infty\,, \qquad \liminf_{k\to\infty} (-1)^k k = -\infty\,.$$

The key to mastering lim sup and lim inf is to understand the following characterizations. These should be compared with the characterization of the limit of a convergent sequence given by Proposition 2.7.

Proposition 2.17. *Let $\{b_k\}$ be a sequence in $\mathbb{R}$. Let $b \in \mathbb{R}$. Then*

$$\limsup_{k\to\infty} b_k = b \qquad \left(\liminf_{k\to\infty} b_k = b\right), \tag{2.15}$$

if and only if for every $c \in \mathbb{R}$

$$b < c \quad \Longrightarrow \quad b_k < c \quad \textit{eventually (frequently)}, \tag{2.16a}$$

and for every $a \in \mathbb{R}$

$$a < b \quad \Longrightarrow \quad a < b_k \quad \textit{frequently (eventually)}. \tag{2.16b}$$

Proof. We give the proof of the lim sup assertion. The lim inf assertion is proved similarly.
($\Rightarrow$) Suppose that (2.15) holds. Let $a < b$ and $c > b$. Then because

$$b = \limsup_{k\to\infty} b_k = \lim_{k\to\infty} \overline{b}_k\,,$$

where $\overline{b}_k = \sup\{b_l \,:\, l \geq k\}$, it follows that

$$a < \overline{b}_k < c \qquad \text{eventually}\,.$$

Because $b_k \leq \overline{b}_k$ for every k, we see directly that $b_k < c$ eventually, whereby (2.16a) holds. Moreover, $a < \overline{b}_k$ implies that for some $l \geq k$ we have $a < b_l$. (Otherwise a would be an upper bound for the set $\{b_l \,:\, l \geq k\}$, which contradicts the fact b is the least upper bound of this set.) Hence, $a < b_k$ frequently, whereby (2.16b) holds.
($\Leftarrow$) Suppose that (2.16) holds. Let $a < b$ and $c > b$ be arbitrary. Then (2.16a) implies that

$$\overline{b}_k = \sup\{b_l \,:\, l \geq k\} \leq c \quad \text{eventually}\,,$$

while (2.16b) implies that

$$a < \overline{b}_k = \sup\{b_l \,:\, l \geq k\} \quad \text{eventually}\,.$$

(If for some k we had $\overline{b}_k \leq a$ then $b_l \leq a$ for every $l \geq k$, which contradicts (2.16b).) Thus, we see

$$a \leq \inf\{\overline{b}_k\} \leq c\,.$$

But $a < b$ and $c > b$ were arbitrary, so that

$$\limsup_{k\to\infty} b_k = \inf\{\overline{b}_k\} = b\,,$$

whereby (2.15) holds. □

Convergent sequences may be characterized in terms of their lim sup and lim inf as follows.

Proposition 2.18. *Let $\{a_k\}$ be a sequence in $\mathbb{R}$. Then $\{a_k\}$ converges if and only if*

$$-\infty < \liminf_{k\to\infty} a_k = \limsup_{k\to\infty} a_k < \infty\,,$$

in which case

$$\lim_{k\to\infty} a_k = \liminf_{k\to\infty} a_k = \limsup_{k\to\infty} a_k\,.$$

Proof. Exercise. Hint: Use Propositions 2.7 and 2.17.

When adding and comparing sequences, lim sup and lim inf generally behave as follows.

Proposition 2.19. *Let $\{a_k\}$ and $\{b_k\}$ be sequences in $\mathbb{R}$. Then*

$$\limsup_{k\to\infty}(a_k + b_k) \leq \limsup_{k\to\infty} a_k + \limsup_{k\to\infty} b_k\,,$$
$$\liminf_{k\to\infty}(a_k + b_k) \geq \liminf_{k\to\infty} a_k + \liminf_{k\to\infty} b_k\,,$$

whenever the sum on the right-hand side is defined. Moreover, if $a_k \leq b_k$ eventually then

$$\limsup_{k\to\infty} a_k \leq \limsup_{k\to\infty} b_k\,, \qquad \liminf_{k\to\infty} a_k \leq \liminf_{k\to\infty} b_k\,.$$

Proof. Exercise.

When multiplying sequences, lim sup and lim inf generally behave as follows.

Proposition 2.20. *Let $\{a_k\}$ and $\{b_k\}$ be sequences in $\mathbb{R}$. If $\{a_k\}$ is convergent with*

$$\lim_{k\to\infty} a_k = a > 0\,,$$

then

$$\limsup_{k\to\infty} a_k b_k = a \limsup_{k\to\infty} b_k\,, \qquad \liminf_{k\to\infty} a_k b_k = a \liminf_{k\to\infty} b_k\,.$$

Proof. Exercise.

When they converge, the lim sup and lim inf of a sequence are actually limits of some of its subsequences.

Proposition 2.21. *Let $\{a_k\}$ be a sequence in $\mathbb{R}$. If $\{a_{n_k}\}$ is any subsequence of $\{a_k\}$ then*

$$\liminf_{k\to\infty} a_k \leq \liminf_{k\to\infty} a_{n_k} \leq \limsup_{k\to\infty} a_{n_k} \leq \limsup_{k\to\infty} a_k\,.$$

Moreover, there exist subsequences $\{a_{n_k}\}$ and $\{a_{m_k}\}$ such that

$$\lim_{k\to\infty} a_{n_k} = \limsup_{k\to\infty} a_k\,, \qquad \lim_{k\to\infty} a_{m_k} = \liminf_{k\to\infty} a_k\,.$$

Proof. Exercise. Hint: Use Propositions 2.7 and 2.17.

The following theorem associated with Bolzano and Weierstrass is an important consequence.

Proposition 2.22. (Bolzano-Weierstrass Theorem) *Every bounded sequence in $\mathbb{R}$ has a convergent subsequence.*

Proof. Let $\{b_k\}$ be a bounded sequence in $\mathbb{R}$. This implies that there exists $[a, c] \subset \mathbb{R}$ such that $\{b_k\} \subset [a, c]$. Then

$$-\infty < a \leq \liminf_{k\to\infty} b_k \leq \limsup_{k\to\infty} b_k \leq c < \infty\,.$$

The result then follows by Proposition 2.21. □

2.8. **Cauchy Criterion.** When a sequence is monotonic, just knowing that it is bounded tells you that it is convergent. When a sequence is not monotonic, determining whether it is convergent or divergent is generally much harder. For example, to establish convergence directly from Definition 2.5 you must first know the limit of the sequence. In 1821 Augustin Cauchy introduced a criterion for convergence that does not require knowledge of the limit.

Definition 2.9. *A sequence* $\{a_k\}_{k\in\mathbb{N}} \subset \mathbb{R}$ *is said to be* Cauchy *whenever for every* $\epsilon > 0$ *there exists* $N_\epsilon \in \mathbb{N}$ *such that*

$$k, l \geq N_\epsilon \quad \Longrightarrow \quad |a_k - a_l| < \epsilon \,. \tag{2.17}$$

In other words, a sequence is Cauchy if for every $\epsilon > 0$ we can find a tail of the sequence such that any two terms in the tail are within ϵ of each other. Roughly speaking, a Cauchy sequence is one whose terms generally cluster together.

Remark. In order to show that a sequence is Cauchy, it is not enough to show merely that successive terms get closer together. Consider the sequence $\{a_k\}$ given by $a_k = \sqrt{k}$ for every $k \in \mathbb{N}$. Its successive terms get closer together because

$$\lim_{k\to\infty} (a_{k+1} - a_k) = \lim_{k\to\infty} \left(\sqrt{k+1} - \sqrt{k}\right) = \lim_{k\to\infty} \frac{1}{\sqrt{k+1} + \sqrt{k}} = 0 \,.$$

However, $a_k \to \infty$ as $k \to \infty$, which implies $\{a_k\}$ is not Cauchy by Proposition 2.24 below.

The main result of this section is the so-called Cauchy criterion for convergence — namely, that a sequence in $\mathbb{R}$ is convergent if and only if it is Cauchy. The easier half of this criterion is established by the following.

Proposition 2.23. *A convergent sequence in* $\mathbb{R}$ *is Cauchy.*

Proof. Let $\{a_k\}$ be a convergent sequence in $\mathbb{R}$ with limit a. Let $\epsilon > 0$. Then by the definition of convergence there exists $N_\epsilon \in \mathbb{N}$ such that

$$k \geq N_\epsilon \quad \Longrightarrow \quad |a_k - a| < \frac{\epsilon}{2} \,.$$

It follows from the triangle inequality that if $k, l \geq N_\epsilon$ then

$$|a_k - a_l| = \big|(a_k - a) + (a - a_l)\big| \leq |a_k - a| + |a_l - a| < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon \,.$$

Hence, the sequence $\{a_k\}$ is Cauchy. □

We now take the first step toward establishing the harder half of the Cauchy criterion.

Proposition 2.24. *A Cauchy sequence in* $\mathbb{R}$ *is bounded.*

Proof. Let $\{a_k\}_{k\in\mathbb{N}}$ be a Cauchy sequence in $\mathbb{R}$. By Definition 2.9 there exists $N_1 \in \mathbb{N}$ such that

$$k, l \geq N_1 \quad \Longrightarrow \quad |a_k - a_l| < 1 \,.$$

This implies that for every $k \geq N_1$

$$|a_k| \leq |a_{N_1}| + |a_k - a_{N_1}| < |a_{N_1}| + 1 \,.$$

Hence, for every $k \in \mathbb{N}$ we have

$$|a_k| < \max\big\{|a_l| \,:\, l = 0, \cdots, N_1\big\} + 1 \,.$$

The sequence $\{a_k\}_{k\in\mathbb{N}}$ is therefore bounded. □

We are now ready to establish the Cauchy Criterion.

Proposition 2.25. (Cauchy Criterion) *A sequence in $\mathbb{R}$ is convergent if and only if it is Cauchy.*

Proof. Proposition 2.23 established that convergent sequences are Cauchy, so we only need to establish the other direction.

Let $\{a_k\}$ be a Cauchy sequence in $\mathbb{R}$. By Proposition 2.24 the sequence $\{a_k\}$ is bounded. By the Bolzano-Weierstrass Theorem (Proposition 2.22) it has a convergent subsequence $\{a_{n_k}\}$. Let a be the limit of this convergent subsequence. We will use the fact $\{a_k\}$ is a Cauchy sequence to show that it converges to a.

Let $\epsilon > 0$. Because the subsequence $\{a_{n_k}\}$ converges to a while the sequence $\{a_k\}$ is Cauchy, there exists an $N_\epsilon \in \mathbb{N}$ such that

$$k \geq N_\epsilon \quad \Longrightarrow \quad |a_{n_k} - a| < \frac{\epsilon}{2}\,,$$

and

$$k, l \geq N_\epsilon \quad \Longrightarrow \quad |a_k - a_l| < \frac{\epsilon}{2}\,.$$

Because $k \geq N_\epsilon$ implies that $n_k \geq N_\epsilon$, the line above implies that

$$k \geq N_\epsilon \quad \Longrightarrow \quad |a_k - a_{n_k}| < \frac{\epsilon}{2}\,.$$

It follows from the triangle inequality that if $k \geq N_\epsilon$ then

$$|a_k - a| \leq |a_k - a_{n_k}| + |a_{n_k} - a| < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon\,.$$

The sequence $\{a_k\}$ therefore converges to a. □

Remark. Cauchy devised his criterion to study the convergence of infinite series. We will see it applied extensively in that role in the next chapter.

2.9. **Contracting Sequences*.** We now use the Cauchy Criterion to establish the convergence for a class of sequences that often arise in applications — namely, contracting sequences.

Definition 2.10. *A sequence $\{a_k\}_{k\in\mathbb{N}} \subset \mathbb{R}$ is said to be* contracting *if there exists $r \in (0,1)$ such that*

$$|a_{k+1} - a_k| \leq r|a_k - a_{k-1}| \qquad \textit{eventually}\,. \tag{2.18}$$

In other words, a sequence is contracting if eventually the distance between its successive points is reduced by at least a factor $r < 1$ eventually.

Proposition 2.26. (Contracting Sequence) *Every contracting sequence in $\mathbb{R}$ is convergent.*

Remark. Because Proposition 2.25 showed that Cauchy sequences are convergent, it suffices to show that contracting sequences are Cauchy.

Proof. Let $\{a_k\}_{k\in\mathbb{N}} \subset \mathbb{R}$ be contracting. Then there exists $r \in (0,1)$ and $K \in \mathbb{N}$ such that (2.18) holds for every $k > K$. For every $k > K$ we can show by induction that

$$|a_k - a_{k+1}| < r|a_{k-1} - a_k| < r^2|a_{k-2} - a_{k-1}| < \cdots < r^{k-K}|a_K - a_{K+1}|\,. \tag{2.19}$$

The triangle inequality, the above inequality, and the difference of powers formula (1.1) then imply that for every $l > k > K$ we have the bound

$$\begin{aligned} |a_k - a_l| &\le |a_k - a_{k+1}| + |a_{k+1} - a_{k+2}| + \cdots + |a_{l-2} - a_{l-1}| + |a_{l-1} - a_l| \\ &< r^{k-K}\big(1 + r + \cdots + r^{l-k-2} + r^{l-k-1}\big)|a_K - a_{K+1}| \\ &= r^{k-K}\,\frac{1 - r^{l-k}}{1-r}\,|a_k - a_{k+1}| < \frac{r^{k-K}}{1-r}\,|a_K - a_{K+1}|\,. \end{aligned} \tag{2.20}$$

Now let $\epsilon > 0$. Because $r < 1$ we can pick $N_\epsilon > K$ such that $r^{N_\epsilon - K}|a_K - a_{K+1}| < (1-r)\epsilon$. Let $k, l > N_\epsilon$ with $k \neq l$. Without loss of generality we may assume that $k < l$. By bound (2.20) we have

$$|a_k - a_l| < \frac{r^{k-K}}{1-r}\,|a_K - a_{K+1}| < \frac{r^{N_\epsilon - K}}{1-r}\,|a_K - a_{K+1}| < \epsilon\,.$$

Therefore $\{a_k\}_{k\in\mathbb{N}}$ is Cauchy, whereby it is convergent by Proposition 2.25. □

While Proposition 2.26 says that every contracting real sequence is convergent, not every convergent sequence is contracting. To understand this better, we start with a characterization of contracting sequences.

Proposition 2.27. (Contracting Sequence Characterization) *Let $\{a_k\}$ be a real sequence such that $a_k \neq a_{k-1}$ eventually. Then $\{a_k\}$ is contracting if and only if*

$$\limsup_{k\to\infty} \frac{|a_{k+1} - a_k|}{|a_k - a_{k-1}|} < 1\,.$$

Proof. Exercise. □

Exercise. Prove Proposition 2.27.

Proposition 2.27 shows that the key to determining whether or not a given sequence $\{a_k\}$ is contracting is the quotient

$$\frac{|a_{k+1} - a_k|}{|a_k - a_{k-1}|}\,.$$

We illustrate this with two examples.

Example. The sequence $\{1/k\}_{k\in\mathbb{Z}_+}$ is not contracting. Indeed, for every $k > 1$ we have

$$\frac{|a_{k+1} - a_k|}{|a_k - a_{k-1}|} = \frac{|\frac{1}{k+1} - \frac{1}{k}|}{|\frac{1}{k} - \frac{1}{k-1}|} = \frac{\frac{1}{k} - \frac{1}{k+1}}{\frac{1}{k-1} - \frac{1}{k}} = \frac{\frac{1}{k(k+1)}}{\frac{1}{k(k-1)}} = \frac{k-1}{k+1}\,,$$

whereby

$$\limsup_{k\to\infty} \frac{|a_{k+1} - a_k|}{|a_k - a_{k-1}|} = \lim_{k\to\infty} \frac{k-1}{k+1} = 1\,.$$

Therefore Proposition 2.27 implies that $\{1/k\}_{k\in\mathbb{Z}_+}$ is not contracting.

Remark. The above example shows that not every convergent sequence is contracting.

Exercise. Show that $\{1/k^2\}_{k\in\mathbb{Z}_+}$ is not contracting.

Remark. The next example illustrates how to show a sequence is contracting by bounding the quotient above. We will see other ways to do this later in the course.

Example. Let $a_0 > 0$ and $a_{k+1} = 1/(2 + a_k^2)$ for every $k \in \mathbb{N}$. Show $\{a_k\}_{k\in\mathbb{N}}$ is contracting.

We first express $a_{k+1} - a_k$ in terms of a_k and a_{k-1} by using the defining relation twice as

$$a_{k+1} - a_k = \frac{1}{2 + a_k^2} - \frac{1}{2 + a_{k-1}^2} = \frac{a_{k-1}^2 - a_k^2}{(2 + a_k^2)(2 + a_{k-1}^2)} = \frac{a_{k-1} + a_k}{(2 + a_k^2)(2 + a_{k-1}^2)} (a_{k-1} - a_k) .$$

It follows that the quotient is

$$\frac{|a_{k+1} - a_k|}{|a_k - a_{k-1}|} = \frac{a_{k-1} + a_k}{(2 + a_k^2)(2 + a_{k-1}^2)} .$$

Next, we use the fact that $a_k \in (0, \frac{1}{2})$ for every $k \geq 1$ to obtain the bound

$$\frac{a_{k-1} + a_k}{(2 + a_k^2)(2 + a_{k-1}^2)} \leq \frac{\frac{1}{2} + \frac{1}{2}}{(2 + 0^2)(2 + 0^2)} = \frac{1}{4} \qquad \text{for every } k \geq 2 .$$

Here we have maximized the numerator and minimized the denominator over $a_{k-1}, a_k \in (0, \frac{1}{2})$. This crude bound yields

$$\frac{|a_{k+1} - a_k|}{|a_k - a_{k-1}|} \leq \frac{1}{4} \qquad \text{for every } k \geq 2 .$$

Therefore $\{a_k\}_{k\in\mathbb{N}}$ is contracting. (It is also convergent by Proposition 2.26.) □

Remark. More refined bounds are needed to do some of the exercises below. Such a bound might use the fact that a_{k-1} and a_k are related rather than treating them as independent.

Exercise. Let $a_0 > 0$ and $a_{k+1} = 1/(2 + a_k)$ for every $k \in \mathbb{N}$. Show $\{a_k\}_{k\in\mathbb{N}}$ is contracting.

Exercise. Let $a_0 > 0$ and $a_{k+1} = 1/(1 + a_k)$ for every $k \in \mathbb{N}$. Show $\{a_k\}_{k\in\mathbb{N}}$ is contracting.

Exercise. Let $a_0 > 0$ and $a_{k+1} = 1/(1 + a_k^2)$ for every $k \in \mathbb{N}$. Show $\{a_k\}_{k\in\mathbb{N}}$ is contracting.

Exercise. Let $a_0 > 0$ and $a_{k+1} = \frac{1}{2}(a_k + 2/a_k)$ for every $k \in \mathbb{N}$. Show $\{a_k\}_{k\in\mathbb{N}}$ is contracting.

Remark. A contracting sequence converges to its limit at an exponential rate or faster. Let $\{a_k\}$ be a contracting sequence that satisfies (2.18) for some $r \in (0, 1)$. Let a be its limit. Bound (2.20) says there exists $K \in \mathbb{N}$ such that for every $l > k > K$ we have

$$|a_k - a_l| < \frac{r^{k-K}}{1 - r} |a_K - a_{K+1}| ,$$

whereby the triangle inequality yields

$$|a_k - a| \leq |a_k - a_l| + |a_l - a| < \frac{r^{k-K}}{1 - r} |a_K - a_{K+1}| + |a_l - a| .$$

Because $|a_l - a| \to 0$ as $l \to \infty$, we see for every $k > K$ that

$$|a_k - a| \leq \frac{r^{k-K}}{1 - r} |a_K - a_{K+1}| .$$

This shows that $|a_k - a|$ vanishes at least as fast as r^k as $k \to \infty$, which is exponential in k. Converging sequences that converge at a slower rate cannot be contracting.

3. Series of Real Numbers

3.1. **Infinite Series.** Any finite set of real numbers can be summed. Here we study one way to make sense of the sum of an infinite sequence of real numbers.

3.1.1. *Convergence and Divergence.*

Definition 3.1. *Given any real sequence* $\{a_k\}_{k=0}^{\infty}$*, for every* $m, n \in \mathbb{N}$ *with* $m \leq n$ *define the* sigma notation*:*

$$\sum_{k=m}^{n} a_k \equiv a_m + a_{m+1} + \cdots + a_{n-1} + a_n \,.$$

Associated with the sequence of terms $\{a_k\}$ *is the sequence of* partial sums $\{s_n\}$ *defined by*

$$s_n \equiv \sum_{k=0}^{n} a_k \,.$$

It is convenient to encode $\{s_n\}$ *with the formal* infinite series

$$\sum_{k=0}^{\infty} a_k \,.$$

If the sequence $\{s_n\}$ *converges to a limit* s *then we say that the infinite series* converges*, and that* s *is the* sum *of the series. In that case we write*

$$\sum_{k=0}^{\infty} a_k = s \,.$$

If the sequence $\{s_n\}$ *diverges then we say that the infinite series* diverges*.*

Remark. It is clear that changing, adding, or removing a finite number of terms in a series does not affect whether the series converges or diverges, but if it converges, the sum would almost always be affected. For example,

$$\sum_{k=0}^{\infty} a_k \quad \text{converges} \qquad \Longleftrightarrow \qquad \sum_{k=5}^{\infty} a_k \quad \text{converges} \,,$$

but when they do converge the sums will generally differ — namely,

$$\text{in general} \quad \sum_{k=0}^{\infty} a_k \neq \sum_{k=5}^{\infty} a_k \,.$$

More specifically, these sums will be equal if and only if $a_0 + a_1 + a_2 + a_3 + a_4 = 0$.

Example. Consider the infinite series

$$\sum_{k=1}^{\infty} \frac{1}{k(k+1)} \,.$$

The n^{th} partial sum is given by

$$\begin{aligned} s_n &= \sum_{k=1}^{n} \frac{1}{k(k+1)} = \sum_{k=1}^{n} \left(\frac{1}{k} - \frac{1}{k+1} \right) \\ &= \left(1 - \frac{1}{2} \right) + \left(\frac{1}{2} - \frac{1}{3} \right) + \cdots + \left(\frac{1}{n} - \frac{1}{n+1} \right) \\ &= 1 + \left(-\frac{1}{2} + \frac{1}{2} \right) + \cdots + \left(-\frac{1}{n} + \frac{1}{n} \right) - \frac{1}{n+1} = 1 - \frac{1}{n+1} \,. \end{aligned}$$

We then see that the series is convergent with

$$\sum_{k=1}^{\infty} \frac{1}{k(k+1)} = \lim_{n\to\infty} s_n = \lim_{n\to\infty} \left(1 - \frac{1}{n+1} \right) = 1 \,.$$

3.1.2. *Telescoping Forms.* The previous example shows a series that can be put into so-called *telescoping form.*

Definition 3.2. *A formal inifinite series with terms $\{a_k\}_{k=m}^{\infty}$ is said to be in* telescoping form *if $a_k = c_{k-1} - c_k$ for some sequence $\{c_k\}_{k=m-1}^{\infty}$, so that the series is expressed as*

$$\sum_{k=m}^{\infty} (c_{k-1} - c_k) \,. \tag{3.1}$$

If a series can be put into telescoping form with a sequence $\{c_k\}$ that is known explicitly then the convergence or divergence of the series can be easily determined. Moreover, if it converges then its sum can be easily determined. This is because the sequence $\{c_k\}$ is simply related to the sequence of partial sums $\{s_n\}$. Indeed, for every $n \geq m$ we see that

$$\begin{aligned} s_n &= \sum_{k=m}^{n} a_k = \sum_{k=m}^{n} (c_{k-1} - c_k) \\ &= (c_{m-1} - c_m) + (c_m - c_{m+1}) + \cdots + (c_{n-2} - c_{n-1}) + (c_{n-1} - c_n) \\ &= c_{m-1} + (-c_m + c_m) + \cdots + (-c_{n-1} + c_{n-1}) - c_n \\ &= c_{m-1} - c_n \,. \end{aligned}$$

It follows immediately that the sequence $\{s_n\}$ converges if and only if the sequence $\{c_n\}$ converges, and that when these sequences converge we have that

$$\lim_{n\to\infty} s_n = c_{m-1} - \lim_{n\to\infty} c_n \,.$$

Hence, the following proposition holds.

Proposition 3.1. *Let $\{c_k\}_{k=m-1}^{\infty}$ be a real sequence. Then*

$$\sum_{k=m}^{\infty} (c_{k-1} - c_k) \quad \textit{converges} \qquad \Longleftrightarrow \qquad \{c_k\}_{k=m-1}^{\infty} \quad \textit{converges} \,.$$

Moreover, when these are convergent we have

$$\sum_{k=m}^{\infty} (c_{k-1} - c_k) = c_{m-1} - \lim_{k\to\infty} c_k \,.$$

Remark. The above considerations show that if a series is in the telescoping form (3.1) then there is a $c \in \mathbb{R}$ such that $c_k = c - s_k$ for every $k \geq m$, where $\{s_k\}_{k=m}^{\infty}$ is the sequence of partial sums. This means that finding an explicit telescoping form for a series is equivalent to finding an explicit expression for its partial sums. It should be clear that we can do this only in the rarest of cases.

3.1.3. *Divergence Test.* For most infinite series it is impossible to find an explicit expression for its partial sums. However, we can commonly determine whether a series is convergent or divergent without finding such an expression. The following proposition gives the simplest test for divergence.

Proposition 3.2. (Divergence Test) *Let $\{a_k\}$ be a real sequence.*

If the series $\displaystyle\sum_{k=0}^{\infty} a_k$ converges then $\lim_{k\to\infty} a_k = 0$.

Equivalently, if $\lim_{k\to\infty} a_k \neq 0$ then the series $\displaystyle\sum_{k=0}^{\infty} a_k$ diverges.

Proof. The proof is based on the fact that the k^{th} term in a formal infinite series can be expressed as $a_k = s_k - s_{k-1}$, where $s_{-1} = 0$ and $\{s_k\}_{k\in\mathbb{N}}$ is the sequence of partial sums. If the series converges then we know that

$$\lim_{k\to\infty} s_k = s\,, \qquad \lim_{k\to\infty} s_{k-1} = s\,,$$

where s is the sum of the series. It thereby follows that

$$\lim_{k\to\infty} a_k = \lim_{k\to\infty} s_k - \lim_{k\to\infty} s_{k-1} = s - s = 0\,. \quad \square$$

Remark. We can easily find examples of a series whose terms converge to zero, yet the series is divergent. One such example is the *harmonic series*:

$$\sum_{k=1}^{\infty} \frac{1}{k}\,.$$

Clearly $1/k \to 0$ as $k \to \infty$. However, we will soon show that this series diverges.

3.2. **Geometric Series.** An important example is that of geometric series.

Definition 3.3. *A formal infinite series of the form*

$$\sum_{k=0}^{\infty} ar^k$$

for some nonzero a and some $r \in \mathbb{R}$ is called a geometric *series.*

The convergence or divergence of a geometric series is easy to determine because it is one of those rare series where we can find an explicit experssion for its partial sums. For every $n \in \mathbb{N}$ let s_n denote the partial sum given by

$$s_n = \sum_{k=0}^{n} ar^k\,.$$

It is clear that if $r = 1$ then $s_n = (n+1)a$ and the series will diverge. So suppose that $r \neq 1$. We check that

$$s_n - rs_n = \sum_{k=0}^{n} ar^k - \sum_{k=0}^{n} ar^{k+1} = \sum_{k=0}^{n} ar^k - \sum_{k=1}^{n+1} ar^k = a - ar^{n+1}\,,$$

whereby the partial sum s_n is found to be

$$s_n = \frac{a - ar^{n+1}}{1-r}\,.$$

By letting n tend to ∞ in this expression we find that

$$\sum_{k=0}^{\infty} ar^k = \begin{cases} \dfrac{a}{1-r} & \text{if } |r| < 1\,, \\ \text{diverges} & \text{otherwise}\,. \end{cases}$$

Remark. The fact the geometric series diverges when $|r| \geq 1$ can also be seen easily from the Divergence Test. Indeed, in that case you see that

$$\lim_{k\to\infty} ar^k = \begin{cases} a & \text{(and hence is nonzero) if } r = 1\,, \\ \text{diverges} & \text{(and hence is nonzero) if } |r| \geq 1 \text{ and } r \neq 1\,, \end{cases}$$

whereby the Divergence Test (Proposition 3.2) shows that the geometric series diverges. Of course, the Divergence Test does not show the geometric series converges when $|r| < 1$.

Exercise. Consider a formal infinite series of the form

$$\sum_{k=1}^{\infty} kr^k$$

for some $r \in \mathbb{R}$. Find all the values of r for which this series converges and evaluate the sum. (Hint: Find an explicit expression for the partial sums and evaluate the limit. The explicit expression may be derived from the analogous expression for a geometric series.)

3.3. **Series with Nonnegative Terms.** If the terms of an infinite series are nonnegative then the associated sequence of partial sums will be nondecreasing. Hence, the least upper bound property can be employed in the guise of the Monotonic Sequence Theorem (Proposition 2.10) to show the convergence or divergence of the series. Specifically, we have the following proposition, which lies at the heart of most proofs about the convergence or divergence of series with nonnegative terms.

Proposition 3.3. (Series with Nonnegative Terms Convergence) *Let $\{a_k\}_{k=m}^{\infty}$ be a nonnegative sequence. Then*

$$\sum_{k=m}^{\infty} a_k \quad \textit{converges} \qquad \Longleftrightarrow \qquad \{s_k\}_{k=m}^{\infty} \quad \textit{is bounded above}\,,$$

where $\{s_k\}_{k=m}^{\infty}$ is the sequence of partial sums associated with the formal infinite series.

Proof. First show that the sequence $\{s_k\}_{k=m}^{\infty}$ is nondecreasing and then apply the Monotonic Sequence Theorem. The details are left as an exercise. □

One way to establish whether or not a sequence of partial sums is bounded above is to compare it with a sequence of partial sums for which the answer is known. This is often done with one of the following comparison tests.

Proposition 3.4. (Comparison Tests for Series with Nonnegative Terms) *Let* $\{a_k\}$ *and* $\{b_k\}$ *be nonnegative sequences that satisfy one of the following comparison conditions:*
(i) the direct comparison

$$\exists M \in \mathbb{R}_+ \quad \textit{such that} \quad a_k \le M\, b_k \quad \textit{eventually};$$

(ii) the limit comparison (if each b_k *is positive)*

$$\limsup_{k\to\infty} \frac{a_k}{b_k} < \infty\,;$$

(iii) the ratio comparison (if each a_k *and* b_k *is positive)*

$$\frac{a_{k+1}}{a_k} \le \frac{b_{k+1}}{b_k} < \infty \quad \textit{eventually}\,.$$

Then

$$\sum_{k=0}^{\infty} b_k \quad \textit{converges} \quad \Longrightarrow \quad \sum_{k=0}^{\infty} a_k \quad \textit{converges}\,, \tag{3.2}$$

$$\left(\sum_{k=0}^{\infty} a_k \quad \textit{diverges} \quad \Longrightarrow \quad \sum_{k=0}^{\infty} b_k \quad \textit{diverges}\,. \right)$$

Proof. First, condition (i) implies (3.2) because if $a_k \le M\, b_k$ for every $k \ge m$ then the fact that $\sum b_k$ converges yields the upper bound

$$\sum_{k=m}^{n} a_k \le M \sum_{k=m}^{n} b_k \le M \sum_{k=m}^{\infty} b_k < \infty\,.$$

Proposition 3.3 therefore implies that $\sum a_k$ converges. Next, condition (ii) implies condition (i) (and hence (3.2)) upon observing that

$$\limsup_{k\to\infty} \frac{a_k}{b_k} < M < \infty \quad \Longrightarrow \quad a_k \le M b_k \quad \text{eventually}\,.$$

Finally, condition (iii) implies condition (ii) (and hence (3.2)) upon observing that

$$\begin{aligned} \frac{a_{k+1}}{a_k} \le \frac{b_{k+1}}{b_k} < \infty \quad \text{eventually} \quad &\Longrightarrow \quad \frac{a_{k+1}}{b_{k+1}} \le \frac{a_k}{b_k} \quad \text{eventually} \\ &\Longrightarrow \quad \left\{ \frac{a_k}{b_k} \right\} \quad \text{is nonincreasing eventually} \\ &\Longrightarrow \quad \lim_{k\to\infty} \frac{a_k}{b_k} < \infty\,. \quad \square \end{aligned}$$

Exercise. The proof shows that the Direct Comparison Test works when the Limit Comparison Test works, and that the Limit Comparison Test works when the Ratio Comparison Test works. Are there examples where (a) the Direct Comparison Test works but the Limit Comparison Test fails, or (b) the Limit Comparison Test works but the Tatio Comparison Test fails?

Example. We can apply the Direct Comparison Test to show the harmonic series diverges. Consider the comparison

$$1 \geq 1\,, \qquad \tfrac{1}{2} \geq \tfrac{1}{2}\,, \qquad \tfrac{1}{3}+\tfrac{1}{4} \geq \tfrac{1}{4}+\tfrac{1}{4}\,, \qquad \tfrac{1}{5}+\tfrac{1}{6}+\tfrac{1}{7}+\tfrac{1}{8} \geq \tfrac{1}{8}+\tfrac{1}{8}+\tfrac{1}{8}+\tfrac{1}{8}\,, \quad \cdots\,.$$

Notice that for every $m \geq 1$ we have

$$\sum_{k=2^{m-1}+1}^{2^m} \frac{1}{k} = \frac{1}{2^{m-1}+1} + \frac{1}{2^{m-1}+2} + \cdots + \frac{1}{2^m} \geq 2^{m-1}\,\frac{1}{2^n} = \frac{1}{2}\,.$$

Summing both sides, we see that for every $n \in \mathbb{N}$ the partial sum s_{2^n} satisfies

$$s_{2^n} = \sum_{k=1}^{2^n} \frac{1}{k} = 1 + \sum_{m=1}^{n} \left(\sum_{k=2^{m-1}+1}^{2^m} \frac{1}{k} \right) \geq 1 + \frac{n}{2}\,.$$

Therefore the sequence of partial sums $\{s_n\}$ clearly diverges.

Easy consequences of Proposition 3.4 are the following tests for when two series converge or diverge together.

Proposition 3.5. (Two-Way Comparison Tests for Series with Nonnegative Terms) *Let $\{a_k\}$ and $\{b_k\}$ be nonnegative sequences that satisfy one of the following two-way comparison conditions:*
(i) the two-way direct comparison

$$\exists L, M \in \mathbb{R}_+ \quad \textit{such that} \quad L\,b_k \leq a_k \leq M\,b_k \quad \textit{eventually};$$

(ii) the two-way limit comparison (if each b_k is positive)

$$0 < \liminf_{k\to\infty} \frac{a_k}{b_k} \leq \limsup_{k\to\infty} \frac{a_k}{b_k} < \infty\,;$$

Then

$$\sum_{k=0}^{\infty} b_k \quad \textit{converges} \quad \Longleftrightarrow \quad \sum_{k=0}^{\infty} a_k \quad \textit{converges}\,, \tag{3.3}$$

$$\left(\sum_{k=0}^{\infty} a_k \quad \textit{diverges} \quad \Longleftrightarrow \quad \sum_{k=0}^{\infty} b_k \quad \textit{diverges}\,. \right)$$

Proof. Exercise.

Remark. The version of condition (ii) given in many elementary calculus courses is

$$0 < \lim_{k\to\infty} \frac{a_k}{b_k} < \infty\,.$$

This requires the above limit to exist, whereas our version does not make any such requirements. (Recall that the lim inf and lim sup exist in $\mathbb{R}_{\text{ex}}$ for every sequence, even ones that diverge.)

3.4. **A Series for *e*.** Recall that we defined e to be the real number such that

$$\lim_{n\to\infty}\left(1+\frac{1}{n}\right)^n = \lim_{n\to\infty}\left(1-\frac{1}{n}\right)^{-n} = e\,. \tag{3.4}$$

Here we will show that e is also the sum of an infinite series. What makes this example different from the previous ones is that here we will evaluate the sum without passing to the limit in an explicit expression for the partial sums. Rather, we will compare the partial sums with first sequence in (3.4).

Proposition 3.6. *The number e satisfies*

$$\sum_{k=0}^{\infty}\frac{1}{k!} = e\,. \tag{3.5}$$

Remark. This result is due to Euler. Newton had showed that the solution $y = \exp(x)$ of the initial-value problem $y' = y$, $y(0) = 1$ behaved like an exponential function and was given by

$$\exp(x) = \sum_{k=0}^{\infty}\frac{x^k}{k!}\,.$$

Its base, $\exp(1)$ is given by the series in (3.5). Euler showed this base is e as defined by (3.4).

Proof. Define the sequences $\{a_n\}_{n=1}^{\infty}$ and $\{s_n\}_{n=0}^{\infty}$ by

$$a_n = \left(1+\frac{1}{n}\right)^n\,, \qquad s_n = \sum_{k=0}^{n}\frac{1}{k!}\,.$$

Because $\{s_n\}_{n=0}^{\infty}$ is the sequence of parial sums associated with the infinite series in (3.5), the result will follow upon using the fact that $a_n \to e$ to show that $s_n \to e$.

For $n \geq 2$ the binomial formula yields

$$a_n = \left(1+\frac{1}{n}\right)^n = \sum_{k=0}^{n}\frac{n!}{k!\,(n-k)!}\,\frac{1}{n^k} = 1 + 1 + \sum_{k=2}^{n}\frac{1}{k!}\prod_{l=1}^{k-1}\left(1-\frac{l}{n}\right)\,.$$

Because the product in each term of the last sum is bounded above by 1, we see that $a_n \leq s_n$. Hence, the sequence of partial sums satisfies

$$e = \lim_{n\to\infty} a_n \leq \liminf_{n\to\infty} s_n\,. \tag{3.6}$$

Let $m \in \mathbb{Z}_+$. For every $n > m$ we have

$$a_n \geq \sum_{k=0}^{m}\frac{n!}{k!\,(n-k)!}\,\frac{1}{n^k} = 1 + 1 + \sum_{k=2}^{m}\frac{1}{k!}\prod_{l=1}^{k-1}\left(1-\frac{l}{n}\right)\,.$$

Letting $n \to \infty$ in this inequality while keeping m fixed shows that s_m is bounded above as

$$e = \lim_{n\to\infty} a_n \geq s_m\,.$$

Because the sequence $\{s_n\}$ is increasing and bounded above by e, by the Monotone Sequence Theorem (Proposition 2.10) it converges with

$$\lim_{m\to\infty} s_m \leq e\,. \tag{3.7}$$

This upper bound together with the lower bound (3.6) shows that $s_n \to e$ as $n \to \infty$, and thereby establishes (3.5). □

We can use this result to prove that e is not rational.

Proposition 3.7. *The number e is irrational.*

Proof. Suppose e is rational. Then $e = p/q$ for some p, $q \in \mathbb{Z}_+$. By Proposition 3.6

$$e = \sum_{k=0}^{\infty} \frac{1}{k!}$$

Let

$$s_q = \sum_{k=0}^{q} \frac{1}{k!}\,.$$

Then by subtracting s_q from e and comparing the resulting infinite series with a geometric series we see that

$$\begin{aligned} 0 < e - s_q = \sum_{k=q+1}^{\infty} \frac{1}{k!} &= \frac{1}{(q+1)!} + \frac{1}{(q+2)!} + \frac{1}{(q+3)!} + \cdots \\ &< \frac{1}{(q+1)!}\left(1 + \frac{1}{q+1} + \frac{1}{(q+1)^2} + \cdots\right) = \frac{1}{q!\,q}\,. \end{aligned}$$

Upon multiplying these inequalities by $q!$ we arrive at

$$0 < q!(e - s_q) < \frac{1}{q}\,.$$

But $q!e = (q-1)!p$ is an integer while $q!s_q$ is also an integer because

$$q!s_q = \sum_{k=0}^{q} \frac{q!}{k!}\,.$$

Therefore $q!(e - s_q)$ is an integer that lies in $(0, \frac{1}{q})$, which is a contradiction because $q \geq 1$. □

Remark. In fact, e is transcendental. A real number is said to be *algebraic* if it is a zero of some polynomial with integer coefficients. Real numbers that are not algebraic are said to be *transcendental*. Almost all real numbers are transcendental, but it is not easy to show that any given real number is transcendental.

3.5. **Series with Nonincreasing Positive Terms.** The harmonic series is a special case of the so-called p-series, which is formally given by

$$\sum_{k=1}^{\infty} \frac{1}{k^p}\,. \tag{3.8}$$

Because the terms of this series are nonincreasing and positive, the following convergence tests can be applied. Proposition 3.3 plays a central role in their proofs.

3.5.1. *Cauchy Condensation Tests.* The following convergence test appeared in 1821.

Proposition 3.8. (Cauchy 2^k Test) *Let $\{a_k\}$ be a nonincreasing, positive sequence. Then*

$$\sum_{k=1}^{\infty} a_k \quad \textit{converges} \quad \Longleftrightarrow \quad \sum_{k=0}^{\infty} 2^k a_{2^k} \quad \textit{converges}\,.$$

Proof. The result is a consequence of the direct comparisons

$$\begin{aligned}
a_2 &\le a_1 \le a_1\,,\\
a_4 &\le a_2 \le a_2\,,\\
a_4 &\le a_3 \le a_2\,,\\
a_8 &\le a_4 \le a_4\,,\\
a_8 &\le a_5 \le a_4\,,\\
a_8 &\le a_6 \le a_4\,,\\
a_8 &\le a_7 \le a_4\,,\\
&\vdots
\end{aligned}$$

These can be written in the general form

$$a_{2^{j+1}} \le a_k \le a_{2^j} \qquad \text{for } 2^j \le k < 2^{j+1}\,,$$

which yield the bounds

$$\sum_{j=0}^{n-1} 2^j a_{2^{j+1}} \le \sum_{k=1}^{2^n-1} a_k \le \sum_{j=0}^{n-1} 2^j a_{2^j} \qquad \text{for every } n \in \mathbb{Z}_+\,.$$

The details are left as an exercise. □

Example. Because $\{1/k^p\}$ is a nonincreasing, positive sequence, Proposition 3.8 implies that the p-series (3.8) converges or diverges as the series

$$\sum_{k=0}^{\infty} 2^k \frac{1}{2^{kp}} = \sum_{k=0}^{\infty} \left(2^{1-p}\right)^k.$$

But this is a geometric series that clearly converges for $p > 1$ and diverges for $p \le 1$.

Remark. The proof of the Cauchy 2^k Test outlined above extends the argument by which we showed the harmonic series diverges. Indeed, the harmonic series is just the p-series for $p = 1$.

The Cauchy 2^k Test can be generalized to subsequences of $\{a_k\}$ of the form $\{a_{n_k}\}$ where there exist constants $\underline{m}$ and $\overline{m}$ such that

$$0 < \underline{m} \le \frac{n_{k+1} - n_k}{n_k - n_{k-1}} \le \overline{m} < \infty\,.$$

This leads to other Cauchy Condensation Tests. For example, we can choose $n_k = m^k$ for some $m \in \mathbb{N}$ with $m > 1$. This satisfies

$$\frac{n_{k+1} - n_k}{n_k - n_{k-1}} = \frac{m^{k+1} - m^k}{m^k - m^{k-1}} = m\,,$$

whereby $\underline{m} = \overline{m} = m$. This leads to a "$m^k$ Test." Its statement and proof is left as an exercise.

Exercise. State and prove the "m^k Test."

3.5.2. *Integral Test.* The next convergence test goes back to Colin Maclaurin in 1742. It requires the use of integrals — in fact, the use of *improper* integrals. These will be developed rigorously later in the course. However, here we will assume you have some familiarity with them from your elementary calculus courses.

Proposition 3.9. (Integral Test) *Let $m \in \mathbb{N}$. Let $f : [m, \infty) \to (0, \infty)$ be a nonincreasing, positive, locally integrable (continuous, for example) function. Then*

$$\sum_{k=m}^{\infty} f(k) \quad \textit{converges} \quad \Longleftrightarrow \quad \int_m^{\infty} f(x)\,\mathrm{d}x \quad \textit{converges}\,.$$

Remark. The integral above is understood in the sense of an improper integral.

Proof. The key fact we need from integration theory is that the improper integral

$$\int_m^{\infty} f(x)\,\mathrm{d}x \quad \text{converges}$$

whenever the sequence $\{S_n\}$ converges, where each S_n is defined by

$$S_n = \int_m^n f(x)\,\mathrm{d}x\,.$$

Because $\{S_n\}$ is an increasing sequence, it will converge if it is bounded above.

The result will then be a consequence of the fact that

$$S_n = \sum_{k=m+1}^{n} \int_{k-1}^{k} f(x)\,\mathrm{d}x \quad \text{for every } n \in \mathbb{Z}_+\,,$$

and the direct comparisons

$$f(k) \le \int_{k-1}^{k} f(x)\,\mathrm{d}x \le f(k-1)\,, \quad \text{for every } k \in \mathbb{N} \text{ with } k > m\,.$$

These facts should be clear from your knowledge of definite integrals from elementary calculus. If not then a picture should help clarify things. We will establish them rigorously later. Here we will assume they are true and complete the proof. Summing these direct comparisons gives

$$\sum_{k=m+1}^{n} f(k) \le S_n \le \sum_{k=m+1}^{n} f(k-1) = \sum_{k=m}^{n-1} f(k)\,.$$

The remaining details are left as an exercise. □

Example. Because $\{1/k^p\}$ is a nonincreasing, positive sequence, Proposition 3.9 implies that the p-series (3.8) converges or diverges as the improper integral

$$\int_1^{\infty} \frac{1}{x^p}\,\mathrm{d}x\,.$$

we can easily check that

$$S_n = \int_1^n \frac{1}{x^p}\,\mathrm{d}x = \begin{cases} \dfrac{1}{p-1}\left(1 - \dfrac{1}{n^{p-1}}\right) & \text{for } p \neq 1\,, \\ \log(n) & \text{for } p = 1\,. \end{cases}$$

(Here $\log(\,\cdot\,)$ denotes the natural logarithm.) We then see that the sequence $\{S_n\}$ converges for $p > 1$ and diverges for $p \le 1$. The same is thereby true for the p-series.

3.6. **Log Test*.** This test bases its conclusions about the convergence of a series with positive terms upon a comparison with a *p-series.* Its earliest form was given by Cauchy in 1821. It involves the (natural) logarithm function, which will be developed rigorously later in the course. However, here we will assume that you have familiarity with logarithms from earlier courses.

Proposition 3.10. (Logarithm Test) *Let $\{a_k\}$ be a positive real sequence. Then*

$$\liminf_{k\to\infty} \frac{\log(a_k)}{\log(\frac{1}{k})} > 1 \quad \Longrightarrow \quad \sum_{k=0}^{\infty} a_k \quad \textit{converges}\,,$$

$$\frac{\log(a_k)}{\log(\frac{1}{k})} \leq 1 \quad \textit{eventually} \quad \Longrightarrow \quad \sum_{k=0}^{\infty} a_k \quad \textit{diverges}\,.$$

If neither condition is met then the series may either converge or diverge.

Remark. Some books give the divergence criterion of the Log Test as

$$\limsup_{k\to\infty} \frac{\log(a_k)}{\log(\frac{1}{k})} < 1 \quad \Longrightarrow \quad \sum_{k=0}^{\infty} a_k \quad \text{diverges}\,.$$

This is clearly a weaker result than the one we give. It is the one given by Cauchy.

Remark. Notice that Proposition 3.10 does not require the sequence $\{a_k\}$ to be nonincreasing, unlike the tests in the previous section.

Proof. We will begin the proof of the convergence assertion and leave its finish and the entire proof of the divergence assertion as exercises. Let $p \in \mathbb{R}$ such that

$$1 < p < \liminf_{k\to\infty} \frac{\log(a_k)}{\log(\frac{1}{k})}\,.$$

By Proposition 2.17 this implies that

$$1 < p < \frac{\log(a_k)}{\log(\frac{1}{k})} \quad \text{eventually}\,.$$

Because $\log(\frac{1}{k}) < 0$ for every $k > 1$, we see that

$$\log(a_k) < p\log\left(\frac{1}{k}\right) = \log\left(\frac{1}{k^p}\right) \quad \text{eventually}\,,$$

whereby

$$a_k < \frac{1}{k^p} \quad \text{eventually}\,.$$

The remainder of the proof of the convergence assertion and the entire proof of the divergence assertion are left as exercises. □

Exercise. Complete the proof of the convergence assertion of Proposition 3.10.

Exercise. Prove the divergence assertion of Proposition 3.10.

3.7. **Alternating Series.** Until now we have only studied convergence tests for nonnegative series. The underlying tool has been the Monotonic Sequence Theorem (Proposition 2.10), which was used to prove Proposition 3.3. Here we use the Monotonic Sequence Theorem to characterize convergence for a special class of series with alternating sign.

Proposition 3.11. (Alternating Series Test) *Let $\{a_k\}$ be a positive, nonincreasing sequence in $\mathbb{R}$. Then*

$$\sum_{k=0}^{\infty}(-1)^k a_k \quad \textit{converges} \qquad \Longleftrightarrow \qquad \lim_{k\to\infty} a_k = 0\,.$$

Proof. The direction $\Longrightarrow$ is just Proposition 3.2 (Divergence Test). To prove the other direction, let

$$s_n = \sum_{k=0}^{n}(-1)^k a_k\,.$$

First, the picture is that

$$\{s_{2k}\}_{k\in\mathbb{N}} \quad \text{is nonincreasing}\,, \quad \text{while} \quad \{s_{2k+1}\}_{k\in\mathbb{N}} \quad \text{is nondecreasing}\,,$$

and that

$$s_{2k} > s_{2j+1} \quad \text{for every } j,k\in\mathbb{N}\,.$$

Indeed, the first two assertions follow because

$$\begin{aligned} s_{2k+2} - s_{2k} &= a_{2k+2} - a_{2k+1} \le 0\,,\\ s_{2k+3} - s_{2k+1} &= -a_{2k+3} + a_{2k+2} \ge 0\,. \end{aligned}$$

Next, because $s_{2k} > s_{2k+1}$ for every $k\in\mathbb{N}$, for any $j\le k$ we have

$$s_{2k} > s_{2k+1} \ge s_{2j+1}\,, \qquad s_{2j} \ge s_{2k} > s_{2k+1}\,.$$

The result follows by exchanging j and k in the last inequality. The monotonic subsequences $\{s_{2k}\}$ and $\{s_{2k+1}\}$ are thereby bounded below and above respectively. By the Monotonic Sequence Theorem they therefore converge. Let

$$\overline{s} = \lim_{k\to\infty} s_{2k}\,, \qquad \underline{s} = \lim_{k\to\infty} s_{2k+1}\,.$$

Then

$$\overline{s} - \underline{s} = \lim_{k\to\infty}\big(s_{2k} - s_{2k+1}\big) = \lim_{k\to\infty} a_{2k+1} = 0\,,$$

whereby $\overline{s} = \underline{s}$. The last step is to show that this fact implies that $\{s_k\}$ converges. This is left as an exercise. □

Examples.

$$\sum_{k=1}^{\infty}\frac{(-1)^k}{k^p} \quad \text{converges if and only if } p>0\,, \qquad \sum_{k=2}^{\infty}\frac{(-1)^k}{\log(k)} \quad \text{converges}\,.$$

The first of these series is called the alternating p-series.

3.8. **Cauchy Criterion for Series.** The Monotonic Sequence Theorem has been the tool underlying all the convergence tests until now. The remaining convergence tests in this chapter are built upon the following characterization of convergent series due to Cauchy in 1821.

Proposition 3.12. (Cauchy Criterion for Series) *Let $\{a_k\}$ be a real sequence. Then*

$$\sum_{k=0}^{\infty} a_k \quad \textit{converges} \tag{3.9a}$$

if and only if for every $\epsilon > 0$ there exists $N_\epsilon \in \mathbb{N}$ such that for every $m, n \in \mathbb{N}$ we have

$$n \geq m > N_\epsilon \quad \Longrightarrow \quad \left| \sum_{k=m}^{n} a_k \right| < \epsilon \,. \tag{3.9b}$$

Proof. Let $\{s_n\}$ be the sequence of partial sums that are defined for any $n \in \mathbb{N}$ by

$$s_n = \sum_{k=0}^{n} a_k \,.$$

By definition the series convergence (3.9a) holds if and only if the sequence $\{s_n\}$ converges. By the Cauchy Criterion, Proposition 2.25, the sequence $\{s_n\}$ converges if and only if it is Cauchy. From Definition 2.9 of Cauchy we see that this holds if and only if for every $\epsilon > 0$ there exists $N_\epsilon \in \mathbb{N}$ such that for every $m, n \in \mathbb{N}$ we have

$$n > m \geq N_\epsilon \quad \Longrightarrow \quad |s_n - s_m| < \epsilon \,.$$

But this is equivalent to

$$n > m \geq N_\epsilon \quad \Longrightarrow \quad \left| \sum_{k=m+1}^{n} a_k \right| < \epsilon \,,$$

which is equivalent to (3.9b) after replacing m above by $m-1$ because for every $m, n \in \mathbb{N}$ the condition $n > m - 1 \geq N_\epsilon$ is equivalent to the condition $n \geq m > N_\epsilon$. □

3.9. **Absolute Convergence.** We now use Proposition 3.12, the Cauchy Criterion for Series, to establish a convergence test that does not require the terms of a series to be nonnegative.

Proposition 3.13. (Absolute Convergence Test) *Let $\{a_k\}$ be a real sequence. Then*

$$\sum_{k=0}^{\infty} |a_k| \quad \textit{converges} \quad \Longrightarrow \quad \sum_{k=0}^{\infty} a_k \quad \textit{converges} \,.$$

Proof. Let $\epsilon > 0$. Because the first series converges, Proposition 3.12 implies that there exists $N_\epsilon \in \mathbb{N}$ such that for every $m, n \in \mathbb{N}$ we have

$$n \geq m > N_\epsilon \quad \Longrightarrow \quad \sum_{k=m}^{n} |a_k| < \epsilon \,.$$

But the triangle inequality then shows that

$$n \geq m > N_\epsilon \quad \Longrightarrow \quad \left| \sum_{k=m}^{n} a_k \right| \leq \sum_{k=m}^{n} |a_k| < \epsilon \,,$$

whereby Proposition 3.12 implies that the second series converges. □

Proposition 3.13 motivates the following definition.

Definition 3.4. *If $\{a_k\}$ is a real sequence such that*

$$\sum_{k=0}^{\infty} |a_k| \quad \textit{converges}\,,$$

then we say that

$$\sum_{k=0}^{\infty} a_k \quad \text{converges absolutely } \textit{or is} \text{ absolutely convergent}\,.$$

Convergent series that are not absolutely convergent are called conditionally convergent*.*

Example. Consider the alternating p-series

$$\sum_{n=1}^{\infty} \frac{(-1)^n}{n^p}\,.$$

This converges for $p > 0$ by the Alternating Series Test, but it converges absolutely only for $p > 1$. This example shows that not every convergent series is absolutely convergent. In other words, there are conditionally convergent series.

When the definition of absolute convergence is combined with the Comparison Tests for series with nonnegative terms (Proposition 3.4), we get an array of new comparison tests for absolute convergence that can be applied to general series.

Proposition 3.14. (Absolute Comparison Tests) *Let $\{a_k\}$ and $\{b_k\}$ be real sequences that satisfy one of the following comparison conditions:*
(i) the direct comparison

$$\exists M \in \mathbb{R}_+ \quad \textit{such that} \quad |a_k| \le M\, b_k \quad \textit{eventually};$$

(ii) the limit comparison (if each b_k is positive)

$$\limsup_{k\to\infty} \frac{|a_k|}{b_k} < \infty\,;$$

(iii) the ratio comparison (if each $|a_k|$ and b_k is positive)

$$\frac{|a_{k+1}|}{|a_k|} \le \frac{b_{k+1}}{b_k} < \infty \quad \textit{eventually}\,.$$

Then

$$\sum_{k=0}^{\infty} b_k \quad \textit{converges} \quad \Longrightarrow \quad \sum_{k=0}^{\infty} a_k \quad \textit{converges absolutely}\,. \tag{3.10}$$

Proof. Exercise. □

Example. Because $|\cos(kx)| \le 1$ for every $x \in \mathbb{R}$ and $k \in \mathbb{Z}_+$, direct comparison with the p-series shows that the series

$$\sum_{k=1}^{\infty} \frac{\cos(kx)}{k^p} \quad \text{converges absolutely for } p > 1\,.$$

3.10. **Root and Ratio Tests.** The Root and Ratio Tests both draw their conclusions about the convergence of a series based upon absolute comparisons with a *geometric series.*

3.10.1. *Root Test.* This test goes back to Cauchy in 1821. Here is a modern version.

Proposition 3.15. (Root Test) *Let $\{a_k\}$ be a real sequence. Then*

$$\limsup_{k\to\infty} \sqrt[k]{|a_k|} < 1 \quad \Longrightarrow \quad \sum_{k=0}^{\infty} a_k \quad \textit{converges absolutely}\,,$$

$$\sqrt[k]{|a_k|} \geq 1 \quad \textit{frequently} \quad \Longrightarrow \quad \sum_{k=0}^{\infty} a_k \quad \textit{diverges}\,.$$

If neither condition is met then the series may either converge or diverge.

Remark. Some books give the divergence criterion of the Root Test as

$$\limsup_{k\to\infty} \sqrt[k]{|a_k|} > 1 \quad \Longrightarrow \quad \sum_{k=0}^{\infty} a_k \quad \text{diverges}\,.$$

This is clearly a weaker result than the one we give. It is the one given by Cauchy.

Proof. The convergence criterion follows by a direct comparison of the series with a convergent geometric series. Specifically, by Proposition 2.17 we have that

$$\limsup_{k\to\infty} \sqrt[k]{|a_k|} < r < 1 \quad \Longrightarrow \quad |a_k| < r^k \quad \text{eventually}\,.$$

The absolute convergence follows from the Direct Comparison Test of Proposition 3.14.

The divergence criterion when follows by showing that $\limsup |a_k| > 0$. Because

$$\sqrt[k]{|a_k|} \geq 1 \quad \text{frequently} \quad \Longleftrightarrow \quad |a_k| \geq 1 \quad \text{frequently}\,,$$

there exists a subsequence $\{a_{n_k}\}_k$ of $\{a_k\}_k$ such that

$$|a_{n_k}| \geq 1 \quad \text{eventually}\,.$$

Because $\limsup |a_k| \geq \limsup |a_{n_k}| \geq 1$, the sequence $\{a_k\}$ does not converge to zero. The Divergence Test (Proposition 3.2) then implies the associated series diverges.

The task of finding both a convergent and a divergent series that satisfy neither condition is left as an exercise. □

Remark. The version of the Root Test found in most elementary calculus texts makes a stronger hypothesis than the above version. It is the following.

Proposition 3.16. (Elementary Root Test) *Let $\{a_k\}$ be a nonzero real sequence such that*

$$\rho = \lim_{k\to\infty} \sqrt[k]{|a_k|} \quad \textit{exists}\,. \tag{3.11}$$

Then

$$\rho < 1 \quad \Longrightarrow \quad \sum_{k=0}^{\infty} a_k \quad \textit{converges absolutely}\,,$$

$$\rho > 1 \quad \Longrightarrow \quad \sum_{k=0}^{\infty} a_k \quad \textit{diverges}\,.$$

If $\rho = 1$ the series may either converge or diverge.

Hypothesis (3.11) above requires the existence of a limit, whereas Proposition 3.15 has no such requirement. (Recall that the lim sup exists in $\mathbb{R}_{\text{ex}}$ for every sequence, even ones that diverge.)

3.10.2. *Ratio Test.* This test goes back to Jean d'Alembert in 1768. Here is a modern version.

Proposition 3.17. (Ratio Test) *Let $\{a_k\}$ be a nonzero real sequence. Then*

$$\limsup_{k\to\infty} \frac{|a_{k+1}|}{|a_k|} < 1 \quad \Longrightarrow \quad \sum_{k=0}^{\infty} a_k \quad \textit{converges absolutely}\,,$$

$$\frac{|a_{k+1}|}{|a_k|} \geq 1 \quad \textit{eventually} \quad \Longrightarrow \quad \sum_{k=0}^{\infty} a_k \quad \textit{diverges}\,.$$

If neither condition is met then the series may either converge or diverge.

Remark. Some books give the divergence criterion of the Ratio Test as

$$\liminf_{k\to\infty} \frac{|a_{k+1}|}{|a_k|} > 1 \quad \Longrightarrow \quad \sum_{k=0}^{\infty} a_k \quad \text{diverges}\,.$$

This is clearly a weaker result than the one we give.

Proof. As with the proof of the Root Test, the convergence conclusion follows by a direct comparison of the series with a convergent geometric series, while the divergence conclusion follows by showing that $\limsup |a_k| > 0$. Specifically, by Proposition 2.17 we have that

$$\rho < r < 1 \quad \Longrightarrow \quad \limsup_{k\to\infty} \frac{|a_{k+1}|}{|a_k|} < r \quad \Longrightarrow \quad \frac{|a_{k+1}|}{|a_k|} < r \quad \text{eventually}\,.$$

An induction argument can then be used to show that for some $m \in \mathbb{N}$ we have

$$|a_k| \leq |a_m|\, r^{k-m} \quad \text{for every } k \geq m\,.$$

Because the geometric series

$$\sum_{k=m}^{\infty} \frac{|a_m|}{r^m}\, r^k \quad \text{converges}\,,$$

the comparison theorem implies

$$\sum_{k=0}^{\infty} a_k \quad \text{converges}\,.$$

We leave as exercises the proof of the divergence assertion and the task of finding examples of both a convergent and a divergent series for which neither condition is satisfied. □

Remark. The version of the Ratio Test found in most elementary calculus texts makes a stronger hypothesis than the above version. It is the following.

Proposition 3.18. (D'Alembert Ratio Test) *Let $\{a_k\}$ be a nonzero real sequence such that*

$$\rho = \lim_{k\to\infty} \frac{|a_{k+1}|}{|a_k|} \quad \textit{exists}\,. \tag{3.12}$$

Then

$$\rho < 1 \quad \Longrightarrow \quad \sum_{k=0}^{\infty} a_k \quad \textit{converges absolutely}\,,$$

$$\rho > 1 \quad \Longrightarrow \quad \sum_{k=0}^{\infty} a_k \quad \textit{diverges}\,.$$

If $\rho = 1$ then the series may either converge or diverge.

Hypothesis (3.12) above requires the existence of a limit, whereas Proposition 3.17 has no such requirement. (Recall that the lim sup exists in $\mathbb{R}_{\text{ex}}$ for every sequence, even ones that diverge.)

Example. Find the least upper bound and greatest lower bound of the set

$$S = \left\{ x \in \mathbb{R} \,:\, \sum_{n=0}^{\infty} \frac{(3n)!}{n!}\frac{(2n)!}{(4n)!}\, x^n \quad \text{converges} \right\}.$$

This can be easily done employing the Ratio Test. Indeed, because

$$\frac{|a_{n+1}|}{|a_n|} = \frac{\dfrac{(3n+3)!}{(n+1)!}\dfrac{(2n+2)!}{(4n+4)!}\,|x|^{n+1}}{\dfrac{(3n)!}{n!}\dfrac{(2n)!}{(4n)!}\,|x|^{n}} = \frac{(3n+3)(3n+2)(3n+1)(2n+2)(2n+1)}{(n+1)(4n+4)(4n+3)(4n+2)(4n+1)}\,|x|\,.$$

we find that

$$\lim_{n\to\infty} \frac{|a_{n+1}|}{|a_n|} = \frac{3^3\times 2^2}{4^4}\,|x| = \frac{3^3}{4^3}\,|x|\,.$$

The elementary Ratio Test, Proposition 3.18, implies that the series converges when $|x| < (4/3)^3$ and diverges when $|x| > (4/3)^3$. Therefore $\sup\{S\} = (4/3)^3$ while $\inf\{S\} = -(4/3)^3$. The Root Test can be applied with the aid of the factorial-root limit (2.14) to obtain the same result.

3.10.3. *Contrasting the Root and Ratio Tests.* The Root Test is sometimes harder to apply, but as the following indicates, its convergence assertion is generally sharper.

Proposition 3.19. *Let $\{a_k\}$ be a positive sequence. Then*

$$\liminf_{k\to\infty} \frac{a_{k+1}}{a_k} \le \liminf_{k\to\infty} \sqrt[k]{a_k} \le \limsup_{k\to\infty} \sqrt[k]{a_k} \le \limsup_{k\to\infty} \frac{a_{k+1}}{a_k}\,. \tag{3.13}$$

Proof. Exercise. (The middle inequality is obvious, so just prove the other two.) □

Exercise. Find one series for which all the inequalities in (3.13) are strict.

Remark. Proposiition 3.19 shows that if hypothesis (3.12) of the D'Alembert Ratio Test holds then hypothesis (3.11) of the Elementary Root Test does too. In other words, if the D'Alembert Ratio Test can be applied to a series then so can the Elementary Root Test. However, as the next exercise shows, sometimes the Elementary Root Test can be applied to a series but the D'Alembert Ratio Test cannot.

Exercise. Find a series for which hypothesis (3.11) holds but hypothesis (3.12) fails to hold.

Remark. Because both the Root and Ratio Tests draw their convergence conclusions based on comparison with a geometric series, they should only be used when such a comparison makes sense. For example, these tests can be used to assert the absolute convergence of series like

$$\sum_{k=1}^{\infty} k^4 2^{-k}\,, \qquad \sum_{l=0}^{\infty} e^{-l^2} 4^l\,, \qquad \sum_{m=0}^{\infty} \frac{(m!)^2}{(2m)!}(-3)^m\,,$$

but will yield no information about the convergence of series like

$$\sum_{k=2}^{\infty} \frac{\log(k)}{k^2}\,, \qquad \sum_{l=0}^{\infty} \left(\frac{3l+2}{l^4+2}\right)^{\frac{1}{2}}, \qquad \sum_{m=2}^{\infty} \frac{(-1)^m}{m(\log(m))^2}\,.$$

3.11. **Abel and Dirichlet Tests*.** We now apply the Cauchy criterion to establish two tests that, like the Alternating Series Test, can be applied to series that do not converge absolutely. We will establish the Dirichlet Test first and then use it to establish the Abel Test.

3.11.1. *Dirichlet Test.* This test arose in the 1829 work of Peter Gustav Lejuene-Dirichlet on the convergence of Fourier series. Here we state it and give an example of its use. We postpone its proof to a later subsection.

Proposition 3.20. (Dirichlet Test) *Let $\{a_k\}_{k\in\mathbb{N}}$ be a nonnegative, nonincreasing sequence in $\mathbb{R}$ such that*

$$\lim_{k\to\infty} a_k = 0\,.$$

Let $\{b_k\}_{k\in\mathbb{N}}$ be a sequence in $\mathbb{R}$ for which there exists M such that

$$\left|\sum_{k=0}^{n} b_k\right| \le M \quad \textit{for every } n\in\mathbb{N}\,. \tag{3.14}$$

Then

$$\sum_{k=0}^{\infty} a_k b_k \quad \textit{converges}\,.$$

Remark. The Dirichlet Test implies the convergence conclusion of the Alternating Series Test. Indeed, if we set $b_k = (-1)^k$ then

$$\sum_{k=0}^{n} b_k = \sum_{k=0}^{n} (-1)^k = \begin{cases} 1 & \text{for } n \text{ even}\,, \\ 0 & \text{for } n \text{ odd}\,. \end{cases}$$

Hence, bound (3.14) holds with $M = 1$. The Dirichlet Test then tells us that for every positive, nonincreasing real sequence $\{a_k\}_{k\in\mathbb{N}}$ such that $a_k \to 0$ as $k \to \infty$ the series

$$\sum_{k=0}^{\infty} (-1)^k a_k \quad \text{converges}\,.$$

As the next example illustrates, the Dirichlet Test is far more powerful than the Alternating Series Test.

Example. Consider the problem of determining all $x, p \in \mathbb{R}$ for which the Fourier p-series

$$\sum_{k=1}^{\infty} \frac{\cos(kx)}{k^p} \quad \text{converges}\,.$$

Because $|\cos(kx)| \le 1$ for every $k \in \mathbb{Z}_+$, direct comparison with the regular p-series shows that this series converges absolutely for $p > 1$. However this argument says nothing about what happens when $p \le 1$.

First observe that when $x \in \{2m\pi : m \in \mathbb{Z}\}$ we have $\cos(kx) = 1$ for every $k \in \mathbb{Z}_+$. In this case the Fourier p-series reduces to a regular p-series, which diverges for every $p \le 1$.

Next, observe that when $x \in \{(2m+1)\pi : m \in \mathbb{Z}\}$ we have $\cos(kx) = (-1)^k$ for every $k \in \mathbb{Z}_+$. In this case the Fourier p-series reduces to an alternating p-series, which (by the Alternating Series Test) converges for every $p > 0$.

We now use the Dirichlet Test to analyze the more general case when $x \notin \{2m\pi : m \in \mathbb{Z}\}$. Let $a_k = 1/k^p$ and $b_k = \cos(kx)$. Clearly the sequence $\{a_k\}$ is positive, decreasing, and vanishes

as $k \to \infty$. The hard step is to show that the partial sums associated with the sequence $\{b_k\}$ satisfy (3.14). To do this we use the trigonometric identity

$$2\sin(\tfrac{1}{2}x)\cos(kx) = \sin((k+\tfrac{1}{2})x) - \sin((k-\tfrac{1}{2})x)\,,$$

and the fact $\sin(\frac{1}{2}x) \neq 0$ when $x \notin \{2m\pi : m \in \mathbb{Z}\}$ to obtain (by a telescoping sum) the formula

$$\sum_{k=1}^{n} b_k = \sum_{k=1}^{n} \cos(kx) = \sum_{k=1}^{n} \frac{\sin((k+\frac{1}{2})x) - \sin((k-\frac{1}{2})x)}{2\sin(\frac{1}{2}x)} = \frac{\sin((n+\frac{1}{2})x) - \sin(\frac{1}{2}x)}{2\sin(\frac{1}{2}x)}\,.$$

It is clear from this formula that

$$\left|\sum_{k=1}^{n} b_k\right| = \left|\frac{\sin((n+\frac{1}{2})x) - \sin(\frac{1}{2}x)}{2\sin(\frac{1}{2}x)}\right| \leq \frac{|\sin((n+\frac{1}{2})x)| + |\sin(\frac{1}{2}x)|}{2\,|\sin(\frac{1}{2}x)|} \leq \frac{1}{|\sin(\frac{1}{2}x)|}\,.$$

Hence, bound (3.14) holds with $M = 1/|\sin(\frac{1}{2}x)|$. The Dirichlet Test then implies that when $x \notin \{2m\pi : m \in \mathbb{Z}\}$ the Fourier p-series converges for every $p > 0$.

Finally, we can use the Divergence Test to show that the Fourier p-series diverges for every $p \leq 0$. This follows easily once we know that

$$\limsup_{k\to\infty} \cos(kx) > 0 \quad \text{for every } x \in \mathbb{R}\,.$$

The details are left as an exercise. □

Remark. When applying the Dirichlet Test to a given series, we must identify the sequences $\{a_k\}$ and $\{b_k\}$, and check that all the hypotheses on them are satisfied. The hypotheses on $\{a_k\}$ are easy to check, so do that first: the sequence $\{a_k\}$ must be positive, nonincreasing, and vanish as $k \to \infty$. The hypothesis on $\{b_k\}$ is typically much harder to check: the associated partial sums must satisfy (3.14). The key to checking this in the above example was to write $b_k = c_{k+1} - c_k$ (by using a trigonometric identity) for some bounded sequence $\{c_k\}$, whereby the partial sums telescoped as

$$\sum_{k=0}^{n} b_k = \sum_{k=0}^{n} (c_{k+1} - c_k) = c_{n+1} - c_0\,.$$

This telescoping approach can be taken for a variety of other $\{b_k\}$ too.

3.11.2. *Summation-By-Parts Identity.* Our proof of the Dirichlet Test uses an identity that is a discrete analog of the integration-by-parts formula from calculus. Because it has many other applications, this identity gets its own proposition.

Proposition 3.21. (Summation-by-Parts Identity) *Let $\{a_k\}_{k\in\mathbb{N}}$ and $\{b_k\}_{k\in\mathbb{N}}$ be sequences in $\mathbb{R}$. Let $B_{-1} = 0$ and*

$$B_n = \sum_{k=0}^{n} b_k \quad \textit{for every } n \in \mathbb{N}\,.$$

Then for every $m, n \in \mathbb{N}$ with $m \leq n$ we have the identity

$$\sum_{k=m}^{n} a_k b_k = a_n B_n - a_m B_{m-1} + \sum_{k=m}^{n-1} (a_k - a_{k+1}) B_k\,, \tag{3.15}$$

with the understanding that the last sum is zero when $m = n$.

Remark. This is called the summation-by-parts identity because it is a discrete analog of the integration-by-parts formula

$$\int_m^n a(x)b(x)\,\mathrm{d}x = a(x)B(x)\Big|_m^n - \int_m^n a'(x)B(x)\,\mathrm{d}x\,,$$

where $B'(x) = b(x)$.

Proof. Because $b_k = B_k - B_{k-1}$ we have

$$\begin{aligned}\sum_{k=m}^{n} a_k b_k = \sum_{k=m}^{n} a_k(B_k - B_{k-1}) &= \sum_{k=m}^{n} a_k B_k - \sum_{k=m}^{n} a_k B_{k-1} \\ &= \sum_{k=m}^{n} a_k B_k - \sum_{k=m-1}^{n-1} a_{k+1} B_k \\ &= a_n B_n - a_m B_{m-1} + \sum_{k=m}^{n-1}(a_k - a_{k+1})B_k\,.\end{aligned}$$

To get from the first to the second line above we re-indexed the last sum. All the other steps are straightforward algebra. □

3.11.3. *Proof of Dirichlet Test.* We now turn to the proof of the Dirichlet Test.

Proof of Dirichlet Test. The proof is hinges on the following inequality. If $m, n \in \mathbb{N}$ with $n \geq m$ then

$$\begin{aligned}\left|\sum_{k=m}^{n} a_k b_k\right| &= \left|a_n B_n - a_m B_{m-1} + \sum_{k=m}^{n-1}(a_k - a_{k+1})B_k\right| \\ &\leq a_n|B_n| + a_m|B_{m-1}| + \sum_{k=m+1}^{n-1}(a_k - a_{k+1})|B_k| \\ &\leq a_n M + a_m M + \sum_{k=m}^{n-1}(a_k - a_{k+1})M = 2a_m M\,.\end{aligned}$$

Here we used the summation-by-parts identity (3.15) in the first step, the triangle inequality and the fact that $\{a_k\}_{k\in\mathbb{N}}$ is nonnegative and nonincreasing in the second step, the bound $|B_k| \leq M$ from (3.14) in the third step, and evaluated the telescoping sum in the last step.

Let $\epsilon > 0$. Because $a_k \to 0$ as $k \to \infty$, there exists $N_\epsilon \in \mathbb{N}$ such that for every $m \in \mathbb{N}$ we have

$$m > N_\epsilon \quad \Longrightarrow \quad 2a_m M < \epsilon\,.$$

Then for every $m, n \in \mathbb{N}$ we have

$$n \geq m > N_\epsilon \quad \Longrightarrow \quad \left|\sum_{k=m}^{n} a_k b_k\right| \leq 2a_m M < \epsilon\,.$$

The Cauchy Criterion for Series, Propisition 3.12, then implies that

$$\sum_{k=0}^{\infty} a_k b_k \quad \text{converges}\,,$$

whereby the Dirichlet Test, Propsition 3.20, is established. □

3.11.4. *Abel Test.* We now use the Dirichlet Test to establish a test that, like the Alternating Series Test and the Dirichlet Test, can be applied to series that do not converge absolutely. The test appeared in an article by Niels Hendrick Abel in 1841, long after his death in 1829. Its derivation from the Dirichlet Test appeared in an article by Dirichlet in 1862, shortly after his death in 1859.

Proposition 3.22. (Abel Test) *Let $\{a_k\}_{k\in\mathbb{N}}$ be a sequence in $\mathbb{R}$ such that the series*

$$\sum_{k=0}^{\infty} a_k \qquad \textit{converges}\,.$$

Let $\{b_k\}_{k\in\mathbb{N}}$ be a bounded, monotonic sequence in $\mathbb{R}$. Then

$$\sum_{k=0}^{\infty} a_k b_k \qquad \textit{converges}\,.$$

Proof. Without loss of generality we may assume the sequence $\{b_k\}_{k\in\mathbb{N}}$ is nonincreasing. Indeed, if it is nondecreasing then it can be replaced with its negative, which is nonincreasing.

Because $\{b_k\}_{k\in\mathbb{N}}$ is a bounded, nonincreasing sequence, it is convergent. Let b be its limit. Let $c_k = b_k - b$. Then $\{c_k\}_{k\in\mathbb{N}}$ is a nonincreasing sequence that converges to zero.

Because $\sum_{k=0}^{\infty} a_k$ converges, we have

$$\sum_{k=0}^{\infty} a_k b \qquad \text{converges}\,.$$

Because the partial sums of $\sum_{k=0}^{\infty} a_k$ are bounded and because $\{c_k\}_{k\in\mathbb{N}}$ is a nonincreasing sequence that converges to zero, the Dirichlet Test implies that

$$\sum_{k=0}^{\infty} a_k c_k \qquad \text{converges}\,.$$

The result follows because the sum of two converging series is a converging series. □

3.12. **Kummer and Related Tests*.** We now study tests for the absolute convergence of series that are often not covered in elementary calculus courses. These tests apply to series that are not covered either by the Cauchy 2^k and Integral Tests of Section 3.5, by the Log Test of Section 3.6, or by the Root and Ratio Tests of Section 3.10.

3.12.1. *Absolute Convergence Criterion.* All of the tests in this section rest upon the following absolute converence criterion.

Proposition 3.23. (Absolute Convergence Criterion) *Let $\{a_k\}_{k\in\mathbb{N}}$ be a sequence in $\mathbb{R}$. If there exists a nonnegative sequence $\{c_k\}_{k\in\mathbb{N}}$ in $\mathbb{R}$ that satisfies*

$$|a_k| \le c_k - c_{k+1} \qquad \textit{eventually}\,, \tag{3.16}$$

then

$$\sum_{k=0}^{\infty} a_k \qquad \textit{converges absolutely}\,. \tag{3.17}$$

Proof. By criterion (3.16) there exists $m \in \mathbb{N}$ such that

$$|a_k| \le c_k - c_{k+1} \qquad \text{for every } k \ge m\,.$$

Then for every $n \ge m$ we have

$$\sum_{k=m}^{n} |a_k| \le \sum_{k=m}^{n} (c_k - c_{k+1}) = c_m - c_{n+1} \le c_m\,.$$

Here we have used the fact that the second series is telescoping to evaluate it. Because its partial sums are bounded above by c_m, the series

$$\sum_{k=m}^{\infty} |a_k| \qquad \text{converges}\,.$$

Therefore (3.17) holds. □

Remark. The converse of this lemma is also true. Indeed, if (3.17) holds then criterion (3.16) is satisfied for every $k \in \mathbb{N}$ by the nonnegative sequence $\{c_k\}_{k\in\mathbb{N}}$ defined by

$$c_k = \sum_{j=k}^{\infty} |a_j| \qquad \text{for every } k \in \mathbb{N}\,.$$

Remark. This lemma could have been proved by applying the Direct Comparison Test of Proposition 3.14. Indeed, criterion (3.16) is condition (i) of that proposition with $M = 1$ and $b_k = c_k - c_{k+1}$ while the hypothesis of that proposition that $\sum_{k=0}^{\infty} b_k$ converges can be verified by using the fact that this series is telescoping as was done in the proof above.

Remark. It is evident from criterion (3.16) that the sequence $\{c_k\}_{k\in\mathbb{N}}$ must be nonincreasing, but this fact was not used explicitly in our proof.

3.12.2. *Kummer Tests.* The difficulty of applying Proposition 3.23 to a given sequence $\{a_k\}_{k\in\mathbb{N}}$ is that it gives us little guidance about how to chose the sequence $\{c_k\}_{k\in\mathbb{N}}$, which clearly has to depend upon $\{a_k\}_{k\in\mathbb{N}}$. The Kummer Tests are a family of tests, each member of which is specified by a positive sequence $\{d_k\}_{k\in\mathbb{N}}$ that does not depend upon $\{a_k\}_{k\in\mathbb{N}}$.

Proposition 3.24. (Kummer Tests) *Let $\{d_k\}_{k\in\mathbb{N}}$ be a positive sequence in $\mathbb{R}$. Let $\{a_k\}_{k\in\mathbb{N}}$ be a sequence in $\mathbb{R}$ with every a_k nonzero. Then*

$$0 < \liminf_{k\to\infty} \left(\frac{1}{d_k} \frac{|a_k|}{|a_{k+1}|} - \frac{1}{d_{k+1}} \right) \quad \Longrightarrow \quad \sum_{k=0}^{\infty} a_k \quad \textit{converges absolutely}\,. \tag{3.18}$$

If $\sum_{k=0}^{\infty} d_k$ diverges then

$$\frac{1}{d_k} \frac{|a_k|}{|a_{k+1}|} - \frac{1}{d_{k+1}} \le 0 \quad \textit{eventually} \quad \Longrightarrow \quad \sum_{k=0}^{\infty} |a_k| \quad \textit{diverges}\,. \tag{3.19a}$$

If $\limsup_{k\to\infty} d_k > 0$ then

$$\frac{1}{d_k} \frac{|a_k|}{|a_{k+1}|} - \frac{1}{d_{k+1}} \le 0 \quad \textit{eventually} \quad \Longrightarrow \quad \sum_{k=0}^{\infty} a_k \quad \textit{diverges}\,. \tag{3.19b}$$

Remark. If $\lim_{k\to\infty} d_k = 0$ and $\sum_{k=0}^{\infty} d_k$ diverges then a series $\sum_{k=0}^{\infty} a_k$ that satisfies the hypothesis of implication (3.19a) might either diverge or converge, but cannot convege absolutely.

Remark. Some books give the hypothesis of divergence implications (3.19) as

$$\limsup_{k\to\infty} \left(\frac{1}{d_k} \frac{|a_k|}{|a_{k+1}|} - \frac{1}{d_{k+1}} \right) < 0 \,.$$

The resulting implications are clearly weaker than the ones we give.

Proof. The hypothesis of implication (3.18) implies there exists $\delta > 0$ such that

$$\delta \le \frac{1}{d_k} \frac{|a_k|}{|a_{k+1}|} - \frac{1}{d_{k+1}} \quad \text{eventually} \,.$$

But then

$$|a_k| \le \frac{1}{\delta} \frac{|a_{k-1}|}{d_{k-1}} - \frac{1}{\delta} \frac{|a_k|}{d_k} \quad \text{eventually} \,.$$

Thus, the absolute convergence criterion (3.16) holds with $c_k = |a_{k-1}|/(\delta\, d_{k-1})$. Therefore implication (3.18) holds by Proposition 3.23.

The hypothesis of implications (3.19) implies there exists an $m \in \mathbb{N}$ such that

$$\frac{|a_k|}{d_k} \le \frac{|a_{k+1}|}{d_{k+1}} \quad \text{for every } k \ge m \,.$$

By induction we can show that

$$\frac{|a_m|}{d_m} \le \frac{|a_k|}{d_k} \quad \text{for every } k \ge m \,,$$

which implies that

$$\frac{|a_m|}{d_m}\, d_k \le |a_k| \quad \text{for every } k \ge m \,. \tag{3.20}$$

The Direct Comparision Test of Proposition 3.14 and (3.20) show that

$$\sum_{k=0}^{\infty} d_k \quad \text{diverges} \quad \Longrightarrow \quad \sum_{k=0}^{\infty} |a_k| \quad \text{diverges} \,,$$

whereby implication (3.19a) holds. Similarly, Proposition 2.19 and (3.20) show that

$$\limsup_{k\to\infty} d_k > 0 \quad \Longrightarrow \quad \limsup_{k\to\infty} |a_k| > 0 \,,$$

whereby the Divergence Test (Proposition 3.2) implies that implication (3.19b) holds. □

Remark. The Kummer Test associated with the sequence $d_k = 1$ is just the Ratio Test (Proposition 3.17). Therefore Kummer Tests are generalizations of the Ratio Test. Subsequent subsections will present other classical convergence tests that arise as the Kummer Test with other choices of d_k.

Exercise. Show that the Kummer Test associated with the sequence $d_k = 1$ is the Ratio Test.

3.12.3. *Raabe Test.* The Kummer Test associated with the sequence $d_k = 1/k$ for $k \geq 1$ is the Raabe Test.

Proposition 3.25. (Raabe Test) *Let $\{a_k\}_{k\in\mathbb{N}}$ be a sequence in $\mathbb{R}$ with every a_k nonzero. Then*

$$1 < \liminf_{k\to\infty} \left[k \left(\frac{|a_k|}{|a_{k+1}|} - 1 \right) \right] \quad \Longrightarrow \quad \sum_{k=0}^{\infty} a_k \quad \textit{converges absolutely}\,, \tag{3.21}$$

and

$$k \left(\frac{|a_k|}{|a_{k+1}|} - 1 \right) \leq 1 \quad \textit{eventually} \quad \Longrightarrow \quad \sum_{k=0}^{\infty} |a_k| \quad \textit{diverges}\,. \tag{3.22}$$

Remark. Some books give the divergence implication of the Raabe Test as

$$\limsup_{k\to\infty} \left[k \left(\frac{|a_k|}{|a_{k+1}|} - 1 \right) \right] < 1 \quad \Longrightarrow \quad \sum_{k=0}^{\infty} |a_k| \quad \text{diverges}\,.$$

This clearly weaker than the one given in (3.22).

Proof. Exercise. □

Example. Apply the Raabe Test to the p-series

$$\sum_{k=1}^{\infty} \frac{1}{k^p} \quad \text{for } p > 0\,.$$

When $1 < p$ we have

$$\frac{|a_k|}{|a_{k+1}|} = \left(\frac{k+1}{k} \right)^p \geq 1 + p \left(\frac{k+1}{k} - 1 \right) = 1 + \frac{p}{k}\,,$$

whereby

$$\liminf_{k\to\infty} \left[k \left(\frac{|a_k|}{|a_{k+1}|} - 1 \right) \right] \geq p > 1\,.$$

Therefore the p-series converges when $1 < p$.

When $0 < p \leq 1$ we have

$$\frac{|a_k|}{|a_{k+1}|} = \left(\frac{k+1}{k} \right)^p \leq 1 + p \left(\frac{k+1}{k} - 1 \right) = 1 + \frac{p}{k}\,,$$

whereby

$$k \left(\frac{|a_k|}{|a_{k+1}|} - 1 \right) \leq p \leq 1\,.$$

Therefore the p-series diverges when $0 < p \leq 1$ and converges when $1 < p$. □

Remark. Our conclusion in the above example for the case $p = 1$ required the strong version of the Raabe divergence implication given in (3.22).

3.12.4. *Bertrand Test.* The Kummer Test associated with the sequence $d_k = (k\log(k))^{-1}$ for $k \geq 2$ is the Bertrand Test.

Proposition 3.26. (Bertrand Test) *Let $\{a_k\}_{k\in\mathbb{N}}$ be a sequence in $\mathbb{R}$ with every a_k nonzero. Then*

$$1 < \liminf_{k\to\infty} \left[\log(k)\left(k\left(\frac{|a_k|}{|a_{k+1}|} - 1\right) - 1\right)\right] \quad \Longrightarrow \quad \sum_{k=0}^{\infty} a_k \quad \textit{converges absolutely}\,, \tag{3.23}$$

and

$$k\log(k)\frac{|a_k|}{|a_{k+1}|} - (k+1)\log(k+1) \leq 0 \quad \textit{eventually} \quad \Longrightarrow \quad \sum_{k=0}^{\infty} |a_k| \quad \textit{diverges}\,. \tag{3.24}$$

Remark. Some books give the divergence implication of the Bertand Test as

$$\limsup_{k\to\infty} \left[\log(k)\left(k\left(\frac{|a_k|}{|a_{k+1}|} - 1\right) - 1\right)\right] < 1 \quad \Longrightarrow \quad \sum_{k=0}^{\infty} |a_k| \quad \text{diverges}\,. \tag{3.25}$$

This is weaker than the one given in (3.24). However, it is often easier to use (3.25) than (3.24).

Proof. Exercise. □

3.12.5. *Gauss Test.* The Raabe and Bertrand Tests combine to yield a classical test that Gauss used to check the convergence of certain hypergeometric series.

Proposition 3.27. (Gauss Test) *Let $\{a_k\}_{k=1}^{\infty}$ be a real sequence with every $a_k \neq 0$ such that*

$$\frac{|a_k|}{|a_{k+1}|} = 1 + \frac{r}{k} + \frac{b_k}{k^s}\,,$$

where $r \in \mathbb{R}$, $s > 1$, and $\{b_k\}_{k=1}^{\infty}$ is a bounded sequence. Then

$$\sum_{k=1}^{\infty} a_k \quad \textit{converges absolutely} \quad \Longleftrightarrow \quad r > 1\,.$$

Proof. Because we have

$$\lim_{k\to\infty}\left[k\left(\frac{|a_k|}{|a_{k+1}|} - 1\right)\right] = \lim_{k\to\infty}\left[r + \frac{b_k}{k^{s-1}}\right] = r\,,$$

the Raabe Test implies that the series converges absolutely when $r > 1$ and does not converge absolutely when $r < 1$. The Raabe Test can say nothing about the case when $r = 1$ because we do not know the sign of the b_k.

Because when $r = 1$ we have

$$\lim_{k\to\infty}\left[\log(k)\left(k\left(\frac{|a_k|}{|a_{k+1}|} - 1\right) - 1\right)\right] = \lim_{k\to\infty}\left[\log(k)\,\frac{b_k}{k^{s-1}}\right] = 0 < 1\,,$$

the Bertrand Test implies that the series does not converge absolutely when $r < 1$. □

4. Sets of Real Numbers

4.1. **Closure, Closed, and Dense.** The notions of *closure, closed,* and *dense* pertain to sets as they relate to the limit process. The closure of a set is all points that are the limit of some convergent sequence that lies within the set. If the closure of a set is the set itself then the set is said to be closed. Simply put, limits do not get out of closed sets. If the closure of a set is everything then the set is said to be dense. Simply put, limits can go anywhere from dense sets. Here we make these notions precise for subsets of $\mathbb{R}$.

4.1.1. *Closure.* We begin with the definition.

Definition 4.1. *Given any $A \subset \mathbb{R}$ its* closure *is given by*

$$A^{\mathrm{c}} = \Big\{ a \in \mathbb{R} \,:\, a \text{ is the limit of a sequence in } A \Big\} .$$

It is clear that $A \subset A^{\mathrm{c}}$ for every $A \subset \mathbb{R}$. Indeed, every $a \in A$ is the limit of the constant sequence $\{a_k\}$ with $a_k = a$ for every $k \in \mathbb{N}$. As we will now see, sometimes $A^{\mathrm{c}} = A$, but in general A^{c} will be larger than A.

Examples. It is easy to show that $\emptyset^{\mathrm{c}} = \emptyset$ and $\mathbb{R}^{\mathrm{c}} = \mathbb{R}$.

Examples. If $a < b$ then the closures of the intervals (a, b), $(a, b]$, $[a, b)$, $[a, b]$, (a, ∞), $[a, \infty)$, $(-\infty, b)$, and $(-\infty, b]$ are given by

$$(a, b)^{\mathrm{c}} = (a, b]^{\mathrm{c}} = [a, b)^{\mathrm{c}} = [a, b]^{\mathrm{c}} = [a, b] ,$$
$$(a, \infty)^{\mathrm{c}} = [a, \infty)^{\mathrm{c}} = [a, \infty) , \qquad (-\infty, b)^{\mathrm{c}} = (-\infty, b]^{\mathrm{c}} = (-\infty, b] .$$

You should be able to prove these facts.

We have the following propositions.

Proposition 4.1. *If $A \subset \mathbb{R}$ is nonempty and bounded above (below) then* $\sup\{A\} \in A^{\mathrm{c}}$ *(*$\inf\{A\} \in A^{\mathrm{c}}$*).*

Proof. We have to show that if $\sup\{A\} \notin A$ then there exists a sequence $\{a_k\} \subset A$ such that $a_k \to \sup\{A\}$ as $k \to \infty$. The details are left as an exercise. □

Proposition 4.2. *For every A, $B \subset \mathbb{R}$ we have that*

$$(i) \quad A \subset B \implies A^{\mathrm{c}} \subset B^{\mathrm{c}} ,$$
$$(ii) \quad (A \cup B)^{\mathrm{c}} = A^{\mathrm{c}} \cup B^{\mathrm{c}} ,$$
$$(iii) \quad (A \cap B)^{\mathrm{c}} \subset A^{\mathrm{c}} \cap B^{\mathrm{c}} .$$

Proof. Exercise.

An important fact is that the closure of $\mathbb{Q}$ is $\mathbb{R}$. In other words, every real number is the limit of a sequence of rational numbers.

Proposition 4.3. $\mathbb{Q}^{\mathrm{c}} = \mathbb{R}$.

Proof. Let $a \in \mathbb{R}$. Consider the sequence of intervals $\{I_k\}_{k \in \mathbb{N}}$ where each I_k is given by

$$I_k = \big(a - \tfrac{1}{2^k} \,,\, a + \tfrac{1}{2^k} \big) .$$

For each $k \in \mathbb{N}$ the third assertion of Proposition 1.19 implies there exists a $a_k \in I_k \cap \mathbb{Q}$. The step of showing that $a_k \to a$ as $k \to \infty$ is left as an exercise. It then follows that $a \in \mathbb{Q}^{\mathrm{c}}$, whereby the assertion follows. □

4.1.2. *Closed.* We are ready for the next definition.

Definition 4.2. *A subset A of $\mathbb{R}$ is said to be* closed *when $A = A^{\mathrm{c}}$.*

Examples. The empty set $\emptyset$ is closed.

Examples. If $a < b$ then intervals of the form $[a, a]$, $[a, b]$, $[a, \infty)$, $(-\infty, b]$, and $\mathbb{R} = (-\infty, \infty)$ are closed, while intervals of the form (a, b), $(a, b]$, $[a, b)$, (a, ∞), and $(-\infty, b)$ are not.

Remark. Earlier we had used the word "closed" to describe those intervals in $\mathbb{R}$ that contain their endpoints. The foregoing examples show that our new usage of "closed" coincides with that old usage. The new usage can also apply to sets that are not intervals, so is more general.

Our terminology seems to demand that closures should be closed. This is indeed the case.

Proposition 4.4. *Let $A \subset \mathbb{R}$. Then A^{c} is closed (i.e. $(A^{\mathrm{c}})^{\mathrm{c}} = A^{\mathrm{c}}$).*

Proof. Let $a \in (A^{\mathrm{c}})^{\mathrm{c}}$. We must show that $a \in A^{\mathrm{c}}$. Because $a \in (A^{\mathrm{c}})^{\mathrm{c}}$ there exists a sequence $\{b_i\}_{i\in\mathbb{N}}$ in A^{c} such that $b_i \to a$ as $i \to \infty$. If $b_i = a$ for some $i \in \mathbb{N}$ then $a = b_i \in A^{\mathrm{c}}$.

On the other hand, if $b_i \neq a$ for every $i \in \mathbb{N}$ then because $b_i \in A^{\mathrm{c}}$ for each $i \in \mathbb{N}$ there exists a sequence $\{b_{(i,j)}\}_{j\in\mathbb{N}}$ in A such that $b_{(i,j)} \to b_i$ as $j \to \infty$. The picture is

$$\begin{array}{cccccccc}
b_{(0,0)}\,, & b_{(0,1)}\,, & b_{(0,2)}\,, & \cdots & b_{(0,j)}\,, & \cdots & \to & b_0 \\
b_{(1,0)}\,, & b_{(1,1)}\,, & b_{(1,2)}\,, & \cdots & b_{(1,j)}\,, & \cdots & \to & b_1 \\
b_{(2,0)}\,, & b_{(2,1)}\,, & b_{(2,2)}\,, & \cdots & b_{(2,j)}\,, & \cdots & \to & b_2 \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & & \vdots \\
b_{(i,0)}\,, & b_{(i,1)}\,, & b_{(i,2)}\,, & \cdots & b_{(i,j)}\,, & \cdots & \to & b_i \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & & \vdots \\
 & & & & & & & \downarrow \\
 & & & & & & & a\,,
\end{array}$$

Because for each $i \in \mathbb{N}$ we have $b_{(i,j)} \to b_i$ as $j \to \infty$ and $|b_i - a| > 0$, there exists a $j_i \in \mathbb{N}$ such that

$$\left|b_{(i,j_i)} - b_i\right| < |b_i - a|\,.$$

Set $a_i = b_{(i,j_i)}$ for each $i \in \mathbb{N}$. It is clear that the sequence $\{a_i\}_{i\in\mathbb{N}}$ lies within A. The step of showing that $a_i \to a$ as $i \to \infty$ is left as an exercise. It then follows that $a \in A^{\mathrm{c}}$, whereby the assertion follows. □

Proposition 4.5. *Let $A \subset \mathbb{R}$. Then A^{c} is the smallest closed set that contains A.*

Proof. The previous proposition shows that A^{c} is closed. Earlier we showed that $A \subset A^{\mathrm{c}}$. Now let B be any closed set that contains A. We see from (i) of Proposition 4.2 that $A \subset B$ implies $A^{\mathrm{c}} \subset B^{\mathrm{c}}$. Because B is closed we know that $B^{\mathrm{c}} = B$. It follows that $A^{\mathrm{c}} \subset B^{\mathrm{c}} = B$. Therefore A^{c} is the smallest closed set that contains A. □

The property of being closed is preserved by certain set operations.

Proposition 4.6. *If A and B are closed subsets of $\mathbb{R}$ then $A \cap B$ and $A \cup B$ are closed. If $\{A_k\}_{k\in\mathbb{N}}$ is a sequence of closed subsets of $\mathbb{R}$ then*

$$\bigcap_{k\in\mathbb{N}} A_k \quad \textit{is closed}\,.$$

Proof. Exercise.

Remark. By repeated application of the first assertion above, we see that the union and intersection of any finite collection of closed sets is again closed. The second assertion states that the intersection of any countable collection of closed sets is again closed. The analogous statment for unions is generally false. Indeed, consider the countable collection of closed intervals $\{I_k\}_{k\in\mathbb{N}}$ where each I_k is given by

$$I_k = \left[-1 + \tfrac{1}{2^k} \, , \, 1 - \tfrac{1}{2^k} \right] .$$

We can easily show that

$$\bigcup_{k\in\mathbb{N}} I_k = (-1, 1) \, ,$$

which is not closed.

4.1.3. *Dense.* Finally, we have the concept of a set being dense in a larger one.

Definition 4.3. *Let $A \subset B \subset \mathbb{R}$. Then A is said to be* dense *in B if $B \subset A^c$.*

Examples. Proposition 4.3 states the $\mathbb{Q}$ is dense in $\mathbb{R}$. In a similar manner we can show that $(a, b) \cap \mathbb{Q}$ is dense in $[a, b]$, that $(a, \infty) \cap \mathbb{Q}$ is dense in $[a, \infty)$, and $(-\infty, b) \cap \mathbb{Q}$ is dense in $(-\infty, b]$.

Proposition 4.7. *If $A \subset B \subset C \subset D \subset \mathbb{R}$ and A is dense in D then B is dense in C.*

Proof. Exercise.

Proposition 4.8. *Let $A \subset \mathbb{R}$. Then A is dense in $\mathbb{R}$ if and only if for every interval (a, b) we have $A \cap (a, b) \neq \emptyset$.*

Proof. Exercise.

4.2. **Completeness.** Completeness is a central notion regarding sets in analysis. As such, it arises in many settings. Here we introduce it in the setting of subsets of $\mathbb{R}$, where it is easily characterized. The basic notion of a set being complete is as follows.

Definition 4.4. *A set $S \subset \mathbb{R}$ is said to be* complete *if every Cauchy sequence contained in S has a limit that is in S.*

The Cauchy Criterion, Proposition 2.25, immediately implies that $\mathbb{R}$ is complete. Moreover, it easily yields the following characterization of all complete subsets of $\mathbb{R}$.

Proposition 4.9. *A subset of $\mathbb{R}$ is complete if and only if it is closed.*

Proof. Exercise.

Remark. In more general settings the notions of complete and closed do not coincide. For example, consider the set $\mathbb{Q}$ equipped with the usual notion of distance. A sequence in $\mathbb{Q}$ is said to be convergent in $\mathbb{Q}$ if it is convergent as a sequence in $\mathbb{R}$ and its limit is in $\mathbb{Q}$. A sequence in $\mathbb{Q}$ is said to be Cauchy in $\mathbb{Q}$ if it is Cauchy as a sequence in $\mathbb{R}$. Because there are sequences in $\mathbb{Q}$ that are Cauchy in $\mathbb{Q}$ but not convergent in $\mathbb{Q}$ the set $\mathbb{Q}$ is not complete. On the other hand, the set $\mathbb{Q}$ is closed because it contains all possible limit points of sequences that are convergent in $\mathbb{Q}$.

4.3. **Connectedness*.** Connectedness is another central notion regarding sets in analysis. As such, it comes in many varieties. Fortunately, these varieties coincide in the setting of subsets of $\mathbb{R}$. The basic notion of a set being connected is as follows.

Definition 4.5. *A set $S \subset \mathbb{R}$ is said to be* disconnected *if there exists nonempty A, $B \subset S$ such that*

$$A \cup B = S\,, \qquad A^{\mathrm{c}} \cap B = A \cap B^{\mathrm{c}} = \emptyset\,. \tag{4.1}$$

Otherwise S is said to be connected.

If a set S is disconnected then the nonempty sets A and B that arise in Definition 4.5 have the property that any convergent sequence that lies within one of them will have a limit that is not in the other. In other words, if $\{x_n\} \subset A$ is convergent and $x_n \to x$ then $x \notin B$.

Example. The set $(-\infty, 0) \cup (0, \infty)$ is disconnected because (4.1) is satisfied by the sets $A = (-\infty, 0)$ and $B = (0, \infty)$.

It is clear from Definition 4.5 that every disconnected set has at least two points in it. In particular, the empty set or any singleton set (a set containing only a single point) is connected. The following proposition shows that the connected subsets of $\mathbb{R}$ are precisely the intervals. Our proof will use the Interval Characterization Theorem, Proposition 1.20.

Proposition 4.10. *A subset of $\mathbb{R}$ is connected if and only if it is an interval.*

Remark. This shows that Definition 4.5 of a connected set leads to a result in accord with what we might naively expect connectedness to mean for subsets of $\mathbb{R}$.

Proof. ($\Longrightarrow$) Let $S \subset \mathbb{R}$ be connected. We will show that S is then an interval by using the Interval Characterization Theorem, Proposition 1.20.

Let x, $y \in S$ such that $x < y$. We must show that $(x, y) \subset S$. Let $z \in (x, y)$. We must show that $z \in S$. Suppose not. Let $A = (-\infty, z] \cap S$ and $B = [z, \infty) \cap S$. Because $x \in A$ and $y \in B$, these sets are nonempty. It is easy to check that A and B satisfy (4.1), whereby S is disconnected. But this contradicts the fact S is connected. Hence, $z \in S$. Therefore $(x, y) \subset S$. Because this is true for every x, $y \in S$ such that $x < y$, the Interval Characterization Theorem, Proposition 1.20, implies that S is an interval.

($\Longleftarrow$) Let $S \subset \mathbb{R}$ be an interval. Suppose that S is disconnected. Then there exists nonempty sets A, $B \subset S$ that satisfy (4.1). Let $x \in A$ and $y \in B$. Because we can always relabel the sets A and B, we can assume without loss of generality that $x < y$. Because x, $y \in S$ while S is an interval, we know from the Interval Characterization Theorem, Proposition 1.20, that $[x, y] \subset S$. Because $A \cup B = S$, we have $(A \cap [x, y]) \cup (B \cap [x, y]) = [x, y]$.

Now consider the point $z = \sup\{A \cap [x, y]\}$. By Propositions 4.1 and 4.2 we have

$$z \in (A \cap [x, y])^{\mathrm{c}} \subset A^{\mathrm{c}} \cap [x, y]\,.$$

Because $A^{\mathrm{c}} \cap B = \emptyset$ and $y \in B$, it follows that $z \neq y$, which implies that $z < y$. Because $z = \sup\{A \cap [x, y]\} < y$ while $(A \cap [x, y]) \cup (B \cap [x, y]) = [x, y]$, we see that $(z, y] \subset B$. But then $z \in [z, y] \subset B^{\mathrm{c}}$ by Proposition 4.2, so that $z \notin A$ because $A \cap B^{\mathrm{c}} = \emptyset$. On the other hand, because $A^{\mathrm{c}} \cap B = \emptyset$ while $(A \cap [x, y]) \cup (B \cap [x, y]) = [x, y]$, it follows that $z \in A \cap [x, y] \subset A$. This contradicts the conclusion of the sentence before it. Therefore S must be connected. $\square$

4.4. **Sequential Compactness.** Compactness is a central notion regarding sets in analysis that plays a crucial role in many existence proofs. As such, it also comes in many varieties. Fortunately, these varieties coincide in the setting of subsets of $\mathbb{R}$. We will take advantage of this coincidence by presenting only the concept of sequential compactness, for which we have all the tools at hand.

Definition 4.6. *A set $A \subset \mathbb{R}$ is said to be* sequentially compact *if every sequence in A has a subsequence that converges to a limit in A.*

Example. The interval $[0, \infty)$ is not sequentially compact because the increasing sequence $\{k\}_{k\in\mathbb{N}}$ diverges to ∞, and therefore has no convergent subseqence.

Example. The interval $(0, 1)$ is not sequentially compact because the limit of the convergent sequence $\{2^{-k}\}_{k\in\mathbb{N}}$ is 0, which is not in $(0, 1)$.

4.4.1. *Characterization.* It is clear that every sequentially compact set must be closed, for otherwise there would be a convergent sequence within it whose limit lies outside it. Intuitively, a sequentially compact set must also be "small enough" that every sequence within it has a convergent subsequence. The following characterization of sequentially compact subsets of $\mathbb{R}$ uses the Bolzano-Weierstrass Theorem to show that "small enough" is simply that the set is bounded.

Proposition 4.11. *A set $A \subset \mathbb{R}$ is sequentially compact if and only if A is closed and bounded.*

Proof. ($\Longrightarrow$) Suppose that A is either not bounded or not closed.

If A is not bounded then for every $k \in \mathbb{N}$ there exists $a_k \in A$ such that $|a_k| > k$. But then the sequence $\{|a_k|\}$ diverges to ∞, whereby every subsequence of $\{|a_k|\}$ diverges to ∞. But then Propositions 2.8 and 2.9 imply that every subsequence of $\{a_k\}$ diverges. Hence, $\{a_k\}$ is a sequence in A that has no convergent subsequence. Therefore A is not sequentially compact.

If A is not closed then there exists a sequence $\{a_k\}$ in A and a point $a \notin A$ such that $a_k \to a$ as $k \to \infty$. By Proposition 2.8 every subsequence of $\{a_k\}$ also converges to the point a, which is not in A. Therefore A is not sequentially compact.

($\Longleftarrow$) Let $A \subset \mathbb{R}$ be closed and bounded. Let $\{a_k\}$ be an arbitrary sequence in A. Because A is bounded, the sequence $\{a_k\}$ is bounded. By the Bolzano-Weierstrass Theorem, $\{a_k\}$ has a converging subsequence $\{a_{n_k}\}$. Let a be the limit of this subsequence. Because A is closed and $\{a_{n_k}\}$ is in A, the limit a must also be in A. By the arbitrariness of $\{a_k\}$, we conclude that every sequence in A has a subsequence that converges to a limit in A. Therefore A is sequentially compact. □

An immediate consequence of this characterizarion is the following.

Proposition 4.12. *Let $A \subset \mathbb{R}$ be sequentially compact and $B \subset \mathbb{R}$ be closed. Then $A \cap B$ is sequentially compact. In particular, every closed subset of A is sequentially compact.*

Proof. Exercise.

It also follows that the property of being sequentially compact is preserved by unions.

Proposition 4.13. *If A and B are sequentially compact subsets of $\mathbb{R}$ then so is $A \cup B$.*

Proof. Exercise.

Remark. By repeated application of this proposition, we see that the union of any finite collection of sequentially compact sets is again sequentially compact.

4.4.2. *Open Interval Covering Property.* Sequentially compact subsets of $\mathbb{R}$ have many covering properties, the simplest of which is the following.

Proposition 4.14. *Let $A \subset \mathbb{R}$ be sequentially compact. Let $\{(a_k, b_k)\}_{k\in\mathbb{N}}$ be a countable collection of open intervals that covers A — i.e. such that*

$$A \subset \bigcup_{k\in\mathbb{N}} (a_k, b_k)\,.$$

Then there exist $n \in \mathbb{N}$ such that

$$A \subset \bigcup_{k=0}^{n} (a_k, b_k)\,.$$

Remark. This property is often stated as: every countable covering of A by open intervals has a finite subcovering. Many other covering properties have a similar flavor — namely, that every covering of A by a certain class of sets has a smaller subcovering.

Proof. Suppose not. Then for every $n \in \mathbb{N}$ there exists an $x_n \in A$ such that

$$x_n \notin \bigcup_{k=0}^{n} (a_k, b_k)\,. \tag{4.2}$$

Because $\{x_n\}_{n\in\mathbb{N}} \subset A$ while A is sequentially compact, there exists a subsequence $\{x_{n_k}\}_{k\in\mathbb{N}}$ that converges to a limit $x \in A$. Because $\{(a_k, b_k)\}_{k\in\mathbb{N}}$ covers A, there exists $m \in \mathbb{N}$ such that $x \in (a_m, b_m)$. Because $x_{n_k} \to x$ as $k \to \infty$, this implies that $x_{n_k} \in (a_m, b_m)$ eventually as $k \to \infty$. But this contradicts the fact seen from (4.2) that $x_{n_k} \notin (a_m, b_m)$ for every $n_k \geq m$. □

There is a converse of the previous proposition that we state without proof.

Proposition 4.15. *Let $A \subset \mathbb{R}$ such that for every countable collection of open intervals $\{(a_k, b_k)\}_{k\in\mathbb{N}}$ that covers A there exist $n \in \mathbb{N}$ such that*

$$A \subset \bigcup_{k=0}^{n} (a_k, b_k)\,.$$

Then A is sequentially compact.

Remark. By combining Propositions 4.14 and 4.15, we see that the covering property stated in the hypothesis of Proposition 4.15 characterizes sequential compactness for subsets of $\mathbb{R}$. This property is closely related to the properties of *countable compactness* and *compactness*, which are also covering properties that characterize sequential compactness for subsets of $\mathbb{R}$. In more general settings these notions of campactness can differ from each other.

4.4.3. *Cantor Intersection Theorem.* Sequential compactness lies behind the following useful generalization by Cantor of the Nested Interval Theorem.

Proposition 4.16. (Cantor Intersection Theorem) *Let $\{A_n\}_{n\in\mathbb{N}}$ be a nested countable collection of nonempty, closed, bounded subsets of $\mathbb{R}$ — i.e. nonempty, closed, bounded sets such that*

$$A_0 \supset A_1 \supset \cdots \supset A_n \supset A_{n+1} \supset \cdots\,.$$

Then

$$\bigcap_{n\in\mathbb{N}} A_n \neq \emptyset\,.$$

Proof. Because each A_n is nonempty, for every $n \in \mathbb{N}$ there exists $x_n \in A_n$. Consider the sequence $\{x_n\}$. Because the collection $\{A_n\}$ is nested, for every $m, n \in \mathbb{N}$ such that $n \geq m$ we have $A_n \subset A_m$, whereby $x_n \in A_m$. The sequence $\{x_n\}$ is thereby in each A_m eventually.

Because A_0 is closed and bounded, it is sequentially compact by Proposition 4.11. Because A_0 is sequentially compact while $\{x_n\} \subset A_0$, there exists a subsequence $\{x_{n_k}\}_{k\in\mathbb{N}}$ that converges to a limit $x \in A_0$. Because the subsequence $\{x_{n_k}\}$ is in each A_m eventually while each A_m is closed, it follows that x is in each A_m. Hence, $x \in \cap_{m\in\mathbb{N}} A_m$, whereby the intersection is nonempty. □

Remark. The hypothesis that the A_n are bounded is essential. For example, consider the nested collection of nonempty, closed sets $\{[n, \infty)\}_{n\in\mathbb{N}}$. Clearly,

$$\bigcap_{n\in\mathbb{N}} [n, \infty) = \emptyset\,.$$

The next proposition will explore this further.

Propsition 4.16 provides half of the following characterization of sequentially compactness.

Proposition 4.17. *Let $A \subset \mathbb{R}$ be closed. Then A is sequentially compact if and only if every countable collection of nested, nonempty, closed subsets of A has a nonempty intersection.*

Proof. ($\Longrightarrow$) This follows from Propsitions 4.11 and 4.16.

($\Longleftarrow$) Suppose that A is not sequentially compact. Because A is closed and is not sequentially compact, Proposition 4.11 implies that A is not bounded. Therefore there exists a sequence $\{x_n\}_{n\in\mathbb{N}} \subset A$ such that $|x_n| > n$ for every $n \in \mathbb{N}$. For each $m \in \mathbb{N}$ define $B_m = \{x_n : n \geq m\}$. We leave as an exercise the proof that $\{B_m\}_{m\in\mathbb{N}}$ is a countable collection of nested, nonempty, closed subsets of A that has an empty intersection. □

Exercise. Show that the $\{B_m\}_{m\in\mathbb{N}}$ constructed above is a countable collection of nested, nonempty, closed subsets of A that has an empty intersection.