Introductory Analysis Notes

Eugene Chau

Last updated May 2025

Contents

1 The Real and Complex Numbers

To motivate the study of the real numbers, we can consider how calculus would be different if we only had the rational numbers instead of all the real numbers. There is still a sensible notion of "limit" if we work purely in the rationals. We can "approach" any given $x \in \mathbb{Q}$ in many ways using only rational numbers—for example, the sequences $\{x+1, x+\frac{1}{2}, x+\frac{1}{3}, \ldots\}$ and $\{x-1, x-\frac{1}{2}, x-\frac{1}{3}, \ldots\}$ both converge to x. Thus, we can also think of "continuity" for functions $\mathbb{Q} \to \mathbb{Q}$. The function $f : \mathbb{Q} \to \mathbb{Q}$ defined by $f(x) = x^2 - 2$ would be a continuous function. Since $f(0) = -2 < 0$ and $f(2) = 2 > 0$, the Intermediate Value Theorem would suggest that there is some $0 < x < 2$ such that $f(x) = 0$. However, f has no rational roots, since no rational number x satisfies the equation $x^2 = 2$. (We will prove this in Proposition 1.15.) Therefore, if the rational numbers were the only numbers to exist, then f would have no roots at all, contradicting our expectations. Essentially, there would be "holes" in the range of f, despite f being continuous. Perhaps by inserting additional numbers into $\mathbb{Q}$ to form a superset $S \supset \mathbb{Q}$, we could ensure that every continuous function $f : S \to S$ has no holes in its range. What numbers would we need to add to $\mathbb{Q}$? It turns out that we would need to add enough numbers to form the real numbers $\mathbb{R}$. In this chapter, we will define $\mathbb{R}$ and introduce the key property of $\mathbb{R}$ that makes the Intermediate Value Theorem true.

1.1 Fields

Defining the real numbers requires specifying how the addition and multiplication operations behave. For example, addition should be commutative ($a + b = b + a$) and associative ($a + (b + c) = (a + b) + c$), and adding 0 to a number should not change that number. Our first goal is to define a minimal set of properties that addition and multiplication must satisfy in the real numbers.

There are many kinds of mathematical objects that can be combined with each other in some way. Integers can be added together, matrices can be multiplied, and functions can be composed. A set of objects with a combining operation that satisfies certain rules (which we shall now specify) is called a *group*.

Definition 1.1. A *group* is a set G with a binary operation $\star : G \times G \to G$ such that:

(i) For all $x, y, z \in G$, $(x \star y) \star z = x \star (y \star z)$.

(ii) There exists an *identity element* $e \in G$ such that $x \star e = x = e \star x$ for all $x \in G$.

(iii) For all $x \in G$, there exists an *inverse element* $y \in G$ such that $x \star y = e = y \star x$.

Additionally, if $x \star y = y \star x$ for all $x, y \in G$, we call G an *abelian group*.

The phrasing of this definition suggests that a group G may have multiple identity elements, and that an element $x \in G$ may have multiple inverses. However, we can prove from the definition that these scenarios are not possible.

Proposition 1.2. *Let G be a group. Then G has exactly one identity element, and every $x \in G$ has exactly one inverse.*

Proof. If $e_1, e_2 \in G$ are identity elements, then $e_1 = e_1 \star e_2 = e_2$. Hence, G only has one identity element. Fix $x \in G$. If $y_1, y_2 \in G$ are inverses of x, then

$$y_1 = y_1 \star e = y_1 \star (x \star y_2) = (y_1 \star x) \star y_2 = e \star y_2 = y_2.$$

Hence, x has only one inverse. $\square$

For brevity, we can write "$(G, \star, e)$" to denote a group G with binary operation $\star$ and identity element e. There are many examples of groups—let's list some of them:

- The set of integers $\mathbb{Z}$ is an abelian group under addition, with 0 as the identity element. The same can be said about the set of rational numbers $\mathbb{Q}$.

- The set of rational numbers minus 0 (denoted $\mathbb{Q} \setminus \{0\}$) is an abelian group under multiplication, with 1 as the identity element. Note that we have to exclude 0 because 0 has no inverse with respect to multiplication (i.e. there is no $x \in \mathbb{Q}$ such that $0x = 1$).

- For any $n \in \mathbb{N}$, let $\mathrm{GL}_n(\mathbb{Q})$ be the set of invertible $n \times n$ matrices where each entry is a rational number. Then $\mathrm{GL}_n(\mathbb{Q})$ is a group under matrix multiplication, where the identity element is the $n \times n$ matrix with 1's on the diagonal and 0's everywhere else.

- For any set X, the set of bijective functions $f : X \to X$ is a group under function composition. The identity element is the identity function $f(x) = x$.

Exercise 1.3.
(a) Let $\mathbb{Z}^2$ be the set of ordered pairs of integers. Define addition in $\mathbb{Z}^2$ by

$$(a, b) + (c, d) \coloneqq (a + c, b + d).$$

Prove that $\mathbb{Z}^2$ is an abelian group with identity element $(0, 0)$.
(b) Prove that $\mathrm{GL}_2(\mathbb{Q})$ is not abelian.

While $\mathbb{R}$ is an abelian group under addition, this description fails to capture the fact that $\mathbb{R}$ has two binary operations. We need the notion of a *field* for that.

Definition 1.4. A *field* is a set F equipped with binary operations $+ : F \times F \to F$ and $\cdot : F \times F \to F$ and distinct elements $0, 1 \in F$ such that
(i) $(F, +, 0)$ is an abelian group,
(ii) $(F \setminus \{0\}, \cdot, 1)$ is an abelian group, and
(iii) for all $x, y, z \in F$, $x \cdot (y + z) = x \cdot y + x \cdot z$.
We call 0 the *additive identity* of F and 1 the *multiplicative identity* of F. For any $x \in F$, the inverse of x in $(F, +, 0)$ is denoted $-x$. For any $x \in F \setminus \{0\}$, the inverse of x in $(F \setminus \{0\}, \cdot, 1)$ is denoted x^{-1}.

Remark. As we usually do with algebraic variables, we write $x \cdot y$ as xy.

Proposition 1.5. *Let F be a field. Then:*
(a) $0x = 0$ for all $x \in F$.
(b) 0 does not have a multiplicative inverse.
(c) $(-x)y = -xy = x(-y)$ for all $x, y \in F$.
(d) If $x \neq 0$ and $y \neq 0$, then $(xy)^{-1}$ exists and $(xy)^{-1} = x^{-1}y^{-1}$. As a corollary, the existence of $(xy)^{-1}$ implies that $xy \neq 0$ by part (b).

Proof. (a) Let $x \in F$. We have that $0x + 0x = (0+0)x = 0x$. Hence,

$$0x = 0x + 0 = 0x + (0x + (-0x)) = (0x + 0x) + (-0x) = 0x + (-0x) = 0.$$

(b) Suppose 0^{-1} exists. Then $0 = 0(0^{-1})$ by part (c). But $0(0^{-1}) = 1$ by definition of multiplicative inverse. Hence, $0 = 1$, which is a contradiction because $0 \neq 1$ in the definition of a field.

(c) Let $x, y \in F$. Note that

$$xy + (-x)y = (x + (-x))y = 0y = 0$$

and

$$xy + x(-y) = x(y + (-y)) = x(0) = 0.$$

Therefore, $-xy = (-x)y$ and $-xy = x(-y)$ since the additive inverse of xy is unique.

(d) If $x \neq 0$ and $y \neq 0$, then using commutativity and associativity gives

$$(xy)(x^{-1}y^{-1}) = (xx^{-1})(yy^{-1}) = (1)(1) = 1.$$

Thus, xy has a multiplicative inverse equal to $x^{-1}y^{-1}$. $\square$

Remark. Part (d) justifies the rule for multiplying fractions: $(\frac{a}{b})(\frac{c}{d}) = \frac{ac}{bd}$ if $b \neq 0$ and $d \neq 0$. Indeed, $(\frac{a}{b})(\frac{c}{d}) = (ab^{-1})(cd^{-1}) = (ac)(b^{-1}d^{-1}) = (ac)(bd)^{-1}$.

Exercise 1.6.

(a) Prove that $\mathbb{Q}$ (with its usual addition and multiplication operations) is a field.

(b) Prove that $\mathbb{Z}$ (with its standard addition and multiplication operations) is *not* a field.

1.2 Ordered Fields

In real analysis, we constantly need to determine whether a certain number is less than a given threshold. To do this, we need to specify how to compare two real numbers. Essentially, we need to define what "less than" means.

Definition 1.7. A *total order* on a set S is a relation "$<$" on S such that:

(i) For all $x, y \in S$, exactly one of the statements $x < y$, $x = y$, or $y < x$ is true.

(ii) For all $x, y, z \in S$, if $x < y$ and $y < z$, then $x < z$. In other words, $<$ is transitive.

If a total order exists on S, we say that S is *totally ordered*.

Definition 1.8. An *ordered field* is a field F equipped with a total order "$<$" such that:

(i) For all $x, y, z \in F$, if $y < z$, then $x + y < x + z$.

(ii) For all $x, y \in F$, if $x > 0$ and $y > 0$, then $xy > 0$.

We are already familiar with one ordered field, namely $\mathbb{Q}$. The real numbers will also be an ordered field.

Proposition 1.9. *Let F be an ordered field, and let $x, y, z \in F$.*

(a) If $x > 0$ and $y < z$, then $xy < xz$.

(b) If $x < 0$ and $y < z$, then $xy > xz$.

(c) $x^2 \geq 0$ with equality if and only if $x = 0$. In particular, $1 > 0$.

(d) If $0 < x < y$, then $0 < y^{-1} < x^{-1}$.

Proof. (a) Suppose $x > 0$ and $y < z$. Then $0 = y - y < z - y$. Hence, $xz - xy = x(z-y) > 0$, so $xy = xy + 0 < xy + (xz - xy) = xz$.

(b) Suppose $x < 0$ and $y < z$. Then $0 = -x + x < -x + 0 = -x$ and $0 = y - y < z - y$. Therefore, $xy - xz = -x(z-y) > 0$, so $xz = xz + 0 < xz + (xy - xz) = xy$.

(c) If $x > 0$, then $x^2 = xx > 0$. If $x < 0$, then $-x > 0$, so $x^2 = (-x)(-x) > 0$. Finally, if $x = 0$, then $x^2 = 0^2 = 0$. In particular, $1 = 1^2 > 0$ because $1 \neq 0$.

(d) Suppose $0 < x < y$. If $y^{-1} < 0$, then $1 = y^{-1}y < 0y = 0$ by part (a), contradicting part (c). If $y^{-1} = 0$, then $1 = yy^{-1} = 0$, which is a contradiction because the field axioms specify that $1 \neq 0$. Hence, $y^{-1} > 0$. By a similar argument, $x^{-1} > 0$ as well. It remains to prove that $y^{-1} < x^{-1}$. We know that $x^{-1}y^{-1} > 0$ since $x^{-1} > 0$ and $y^{-1} > 0$. Therefore,

$$y^{-1} = (x^{-1}y^{-1})x < (x^{-1}y^{-1})y = x^{-1}$$

by part (a). $\square$

Definition 1.10. Let F be an ordered field. For all $x \in F$, define

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

Proposition 1.11. *Let F be an ordered field, and let $x, y \in F$. Then:*

(a) $|x| \geq x$ and $|x| \geq 0$.

(b) $|x|^2 = x^2$.

(c) $x^2 < y^2$ if and only if $-|y| < x < |y|$ if and only if $|x| < |y|$.

(d) $x^2 = y^2$ if and only if $|x| = |y|$.

(e) $|xy| = |x| \cdot |y|$.

Proof. (a) If $x \geq 0$, then $|x| = x \geq 0$, and if $x < 0$, then $|x| = -x > 0 > x$.

(b) Since $|x| \in \{x, -x\}$ and $x^2 = (-x)^2$, we have that $|x|^2 = x^2$.

(c) Suppose $x^2 < y^2$. If $x \geq |y| \geq 0$, then $x^2 \geq |y|^2 = y^2$, which is a contradiction. If $x \leq -|y| \leq 0$, then $0 \leq |y| \leq -x$, so $x^2 = (-x)^2 \geq |y|^2 = y^2$, which is a contradiction. Hence, it is not the case that $x \geq |y|$ or that $x \leq -|y|$, so we must have that $-|y| < x < |y|$. Now suppose $-|y| < x < |y|$. Note that $-x < |y|$ because $-|y| < x$. If $x \geq 0$, then $|x| = x < |y|$. If $x < 0$, then $|x| = -x < |y|$. Hence, $|x| < |y|$ in both cases. Finally, suppose $|x| < |y|$. Then $x^2 = |x|^2 < |y|^2 = y^2$ because $0 \leq |x| < |y|$.

(d) If $|x| = |y|$, then $x^2 = |x|^2 = |y|^2 = y^2$. On the other hand, suppose $|x| \neq |y|$. Without loss of generality, we can assume that $|x| < |y|$. Since $0 \leq |x| < |y|$, we have that $x^2 = |x|^2 < |y|^2 = y^2$, so $x^2 \neq y^2$.

(e) Observe that $|xy|^2 = (xy)^2 = x^2y^2 = |x|^2|y|^2 = (|x| \cdot |y|)^2$. Hence, by part (d), $|xy| = |(|xy|)| = ||x| \cdot |y|| = |x| \cdot |y|$. $\square$

Theorem 1.12 (Triangle Inequality)**.** *Let F be an ordered field. Then $|x+y| \leq |x| + |y|$ for all $x, y \in F$.*

Proof. Fix $x, y \in F$. By Proposition 1.11, we have that

$$|x+y|^2 = (x+y)^2$$

$$\begin{aligned}
&= x^2 + 2xy + y^2 \\
&= |x|^2 + 2xy + |y|^2 \\
&\leq |x|^2 + 2|xy| + |y|^2 \\
&= |x|^2 + 2|x| \cdot |y| + |y|^2 \\
&= (|x| + |y|)^2,
\end{aligned}$$

so $|x + y| = |(|x + y|)| \leq ||x| + |y|| = |x| + |y|$. □

Definition 1.13. Let S be a totally ordered set, and let $E \subset S$. We say that E is *bounded above* if there exists $u \in S$ such that $x \leq u$ for all $x \in E$. We call u an *upper bound* of E. Similarly, E is *bounded below* if there exists $\ell \in S$ such that $\ell \leq x$ for all $x \in E$, and we call ℓ a *lower bound* of E. If E is bounded above and bounded below, we say that E is *bounded*.

If u is an upper bound of E such that $u \leq t$ for all upper bounds t of E, then u is the *least upper bound* or the *supremum* of E, and we denote $\sup(E) := u$. If ℓ is a lower bound of E such that $m \leq \ell$ for all lower bounds m of E, then ℓ is the *greatest lower bound* or the *infimum* of E, and we denote $\inf(E) := \ell$.

We say that S has the *least-upper-bound property* if every non-empty set in S that is bounded above has a least upper bound. Similarly, S has the *greatest-lower-bound property* if every non-empty set in S that is bounded below has a greatest lower bound. If S has the least-upper-bound property, we say that S is *complete*.

Proposition 1.14. *If an ordered field F has the least-upper-bound property, then F has the greatest-lower-bound property.*

Proof. For each subset $S \subset F$, denote $-S := \{-x \mid x \in S\}$. Note that $-(-S) = S$. Suppose S has an upper bound $u \in F$. If $t \in -S$, then $-t \in S$, so $-t \leq u$ and hence $-u \leq t$. Thus, $-u$ is a lower bound of $-S$. A similar argument shows that if S has a lower bound $\ell \in F$, then $-\ell$ is an upper bound of $-S$.

Suppose $u = \sup(U)$ exists for some $U \subset F$. Then $-u$ is a lower bound of $-U$. We claim that $-u$ is the greatest lower bound of $-U$. Suppose there exists a lower bound s of $-U$ such that $s > -u$. Then $-s$ is an upper bound of $-(-U) = U$, which means $-s \geq u$ since $u = \sup(U)$. However, since $s > -u$, we have that $-s < u$, contradicting the fact that $-s \geq u$. Therefore, $\inf(-U) = -u = -\sup(U)$. In particular, $\inf(-U)$ exists.

Suppose F has the least-upper-bound property. Let $S \subset F$ be non-empty and bounded below by $x \in F$. Then $-S$ is non-empty and bounded above by $-x$, so $\sup(-S)$ exists by the least-upper-bound property of F. Hence, $\inf(S) = \inf(-(-S))$ exists. □

Proposition 1.15. *There is no $q \in \mathbb{Q}$ such that $q^2 = 2$.*

Proof. Suppose there exists $q \in \mathbb{Q}$ such that $q^2 = 2$. We can write $q = \frac{m}{n}$ where $m \in \mathbb{Z}$, $n \in \mathbb{N}$, and at least one of m and n are odd. We have that $2 = q^2 = \frac{m^2}{n^2}$, so $m^2 = 2n^2$ is even. If m is odd, then $m = 2k + 1$ for some $k \in \mathbb{Z}$, so

$$m^2 = (2k + 1)^2 = 4k^2 + 4k + 1 = 2(2k^2 + 2k) + 1$$

is odd, contradicting that m^2 is even. Hence, m is even, so $m = 2\ell$ for some $\ell \in \mathbb{Z}$. Therefore, $2n^2 = m^2 = (2\ell)^2 = 4\ell^2$, so $n^2 = 2\ell^2$ is even. The same argument used to prove that m

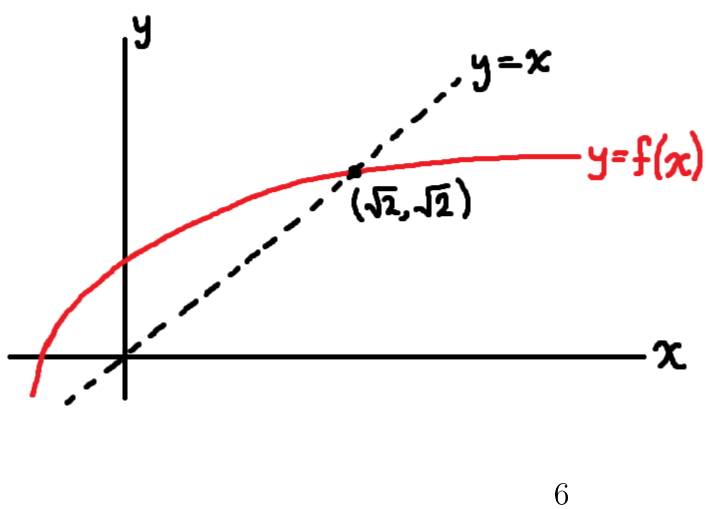
is even now proves that n is even. But at least one of m and n are odd, so we have a contradiction. $\square$

Theorem 1.16. *$\mathbb{Q}$ is not complete.*

Proof. Let $S = \{q \in \mathbb{Q} \mid q^2 < 2\}$. Since $1^2 = 1 < 2$, we have that $1 \in S$, so S is non-empty. If $q \in \mathbb{Q}$ and $q > 2$, then $q^2 > 2^2 > 2$, so $q \notin S$. Hence, S is bounded above by $2 \in \mathbb{Q}$. Suppose S has a least upper bound $u \in \mathbb{Q}$. Note that $u \geq 1$ because $1 \in S$. Let $t = \frac{2u+2}{u+2} \in \mathbb{Q}$. Suppose $u^2 < 2$. Then $u(u+2) = u^2 + 2u < 2 + 2u$, so $u < \frac{2u+2}{u+2} = t$. Also, $(2u+2)^2 = 4u^2 + 8u + 4 < 2u^2 + 8u + 8 = 2(u+2)^2$, so $t^2 = \frac{(2u+2)^2}{(u+2)^2} < 2$. Hence, $u < t$ and $t \in S$, which is impossible since u is an upper bound of S. Now suppose $u^2 > 2$. We can use a similar argument as in the "$u^2 < 2$" case to show that $u > t$ and $t^2 > 2$. Note that $t > 0$ since $u > 0$. It follows that t is an upper bound of S—if $q > t > 0$, then $q \notin S$ since $q^2 > t^2 > 2$. But we have that $u > t$, contradicting that u is the least upper bound of S. Therefore, $u^2 = 2$, contradicting Proposition 1.15. We conclude that S does not have a least upper bound in $\mathbb{Q}$ despite being non-empty and bounded above. $\square$

Returning to our thought experiment at the beginning of this chapter, $\mathbb{Q}$ not being complete is ultimately what makes the Intermediate Value Theorem fail for continuous functions $\mathbb{Q} \to \mathbb{Q}$. By drawing the graph of $f(x) = x^2 - 2$, it is apparent that if $S = \{q \in \mathbb{Q} \mid q^2 < 2\}$ had a supremum $s \in \mathbb{Q}$, then s would be a root of f. However, we proved that S has no supremum in $\mathbb{Q}$. The real numbers, which we will define shortly, will not have this issue.

How did we know to consider $t = \frac{2u+2}{u+2}$ in the proof? Let us briefly forget about "rigour" and assume standard facts about the real numbers. Intuitively, the supremum of $S = \{q \in \mathbb{Q} \mid q^2 < 2\}$ in the real numbers should be $\sqrt{2}$, which is irrational. In the proof, we want to assume that $u = \sup(S) \in \mathbb{Q}$ and obtain a contradiction. We know that $u \neq \sqrt{2}$ since u is rational, so either $u < \sqrt{2}$ or $u > \sqrt{2}$. If $u < \sqrt{2}$, we want to find $t \in \mathbb{Q}$ such that $u < t < \sqrt{2}$, and if $u > \sqrt{2}$, we want $t \in \mathbb{Q}$ such that $\sqrt{2} < t < u$. Hopefully, we can find a simple mapping f such that $t = f(u)$. We know that $\mathbb{Q}$ is a field, which means $\mathbb{Q}$ is closed under the basic arithmetic operations. Therefore, we should consider rational functions $f(x) = \frac{p(x)}{q(x)}$ where the polynomials p and q have integer coefficients—for such functions f, if $x \in \mathbb{Q}$ is in the domain of f, then $f(x) \in \mathbb{Q}$. For positive x, we want that $x < f(x) < \sqrt{2}$ if $0 < x < \sqrt{2}$ and $x > f(x) > \sqrt{2}$ if $x > \sqrt{2}$. At this point, we can sketch what the graph of f should look like, and we get the following figure:



The sketch tells us that $f(\sqrt{2}) = \sqrt{2}$, or equivalently, that $p(\sqrt{2}) = \sqrt{2}q(\sqrt{2})$. Since $\sqrt{2}(2+\sqrt{2}) = 2\sqrt{2}+2$, we can take $p(x) = 2x+2$ and $q(x) = 2+x$. We thus obtain the formula $f(x) = \frac{2x+2}{x+2}$. We can sketch the graph of this rational function and see that it qualitatively matches the graph shown in the figure. Therefore, we take $t = \frac{2u+2}{u+2}$.

1.3 Definition of the Real Numbers

Definition 1.17. $\mathbb{R}$ is the (unique) complete ordered field.

We should note that two facts are required for this definition to be valid:

1. A complete ordered field must actually exist (otherwise, $\mathbb{R}$ would not exist).

2. Any two complete ordered fields are actually the same up to the labelling of their elements (otherwise, this definition does not fully characterize $\mathbb{R}$).

We will prove the first claim by constructing the real numbers out of the rational numbers in Section 1.7. However, we will not prove the second claim.

We will also assume without proof that $\mathbb{Q}$ is a subset of $\mathbb{R}$. More precisely (for readers who know some abstract algebra), there is a unique field homomorphism $\mathbb{Q} \to \mathbb{R}$, and we identify $\mathbb{Q}$ with its image under this homomorphism.

We now begin our investigation of the completeness property of $\mathbb{R}$. This is the property that distinguishes $\mathbb{R}$ from $\mathbb{Q}$; it underpins all of real analysis. We will see later how completeness implies the Intermediate Value Theorem, answering our question from the beginning of this chapter. As a first application of completeness, we will prove shortly that every nonnegative real number has an n^{th} root for any $n \in \mathbb{N}$.

Theorem 1.18 (Archimedean Property)**.** *Let $x > 0$ and $y \in \mathbb{R}$. Then there exists $n \in \mathbb{N}$ such that $nx > y$.*

Proof. Let $S = \{nx \mid n \in \mathbb{N}\}$. Suppose $nx \leq y$ for all $n \in \mathbb{N}$. Then S is non-empty and bounded above by y, so $t = \sup(S)$ exists. Let $n_0 \in \mathbb{N}$ such that $t - n_0 x < \frac{x}{2}$. Then $(n_0+1)x \in S$, but $(n_0+1)x = n_0 x + x > n_0 x + \frac{x}{2} > t$. Hence, t is not an upper bound of S, which is a contradiction. Therefore, there exists $n \in \mathbb{N}$ such that $nx > y$. $\square$

Proposition 1.19. *Let $x \geq y \geq 0$. Then $x^n - y^n \leq nx^{n-1}(x-y)$ for all $n \in \mathbb{N}$.*

Proof. Recall that $x^n - y^n = (x-y)\sum_{k=0}^{n-1} x^k y^{n-1-k}$. For all integers $0 \leq k \leq n-1$, we have that $x^k y^{n-1-k} \leq x^k x^{n-1+k} = x^{n-1}$ since $0 \leq y \leq x$ and $n-1+k \geq 0$. Therefore, $x^n - y^n \leq (x-y)\sum_{k=0}^{n-1} x^{n-1} = nx^{n-1}(x-y)$ since $x - y \geq 0$. $\square$

Theorem 1.20. *Let $x \geq 0$ and $n \in \mathbb{N}$. Then there exists a unique $y \geq 0$ such that $y^n = x$.*

Proof. Let $S = \{t \geq 0 \mid t^n \leq x\}$. Note that $0 \in S$ since $0^n = 0 \leq x$, so S is non-empty. If $x \geq 1$, then $x^n \geq x$, so $t \leq x$ for all $t \in S$. If $0 \leq x < 1$, then $t \leq 1$ for all $t \in S$ since $t^n \leq x < 1 = 1^n$. Hence, $y = \sup(S)$ exists. Note that $y \geq 0$ since $0 \in S$. We want to show that $y^n = x$. Suppose $y^n < x$. By the Archimedean Property, there exists $k \in \mathbb{N}$ such that $n(y+1)^{n-1} < k(x - y^n)$. Then

$$\begin{aligned}
\left(y + \frac{1}{k}\right)^n &\leq y^n + n\left(y + \frac{1}{k}\right)^{n-1}\left(\frac{1}{k}\right) \quad \text{by Proposition 1.19} \\
&\leq y^n + n(y+1)^{n-1}\left(\frac{1}{k}\right) \quad \text{since } \frac{1}{k} \leq 1 \\
&< y^n + x - y^n \quad \text{by definition of } k \\
&= x,
\end{aligned}$$

so $y + \frac{1}{k} \in S$, contradicting that $y = \sup(S)$. Now suppose $y^n > x$. By the Archimedean Property, there exists $N \in \mathbb{N}$ such that $Ny > 1$ by the Archimedean Property, and there exists $M \in \mathbb{N}$ such that $ny^{n-1} < M(y^n - x)$. Let $k = \max\{M, N\}$. Then $ky \geq Ny > 1$, so $y > y - \frac{1}{k} > 0$. Also, $ny^{n-1} < M(y^n - x) \leq k(y^n - x)$, so $\frac{ny^{n-1}}{k} < y^n - x$. Therefore,

$$\begin{aligned}
\left(y - \frac{1}{k}\right)^n &\geq y^n - ny^{n-1}\left(\frac{1}{k}\right) \quad \text{by Proposition 1.19} \\
&> y^n - (y^n - x) \quad \text{since } \frac{ny^{n-1}}{k} < y^n - x \\
&= x,
\end{aligned}$$

so $y - \frac{1}{k}$ is an upper bound of S, contradicting that $y - \frac{1}{k} < y = \sup(S)$. Therefore, $y^n = x$.

Now we prove the uniqueness of y. Suppose $y_1^n = x$ where $y_1 \geq 0$. If $y_1 < y$, then $y_1^n < y^n = x$, and if $y < y_1$, then $x = y^n < y_1^n$. Both conclusions contradict the assumption that $y_1^n = x$, so it must be the case that $y_1 = y$. □

Notation. For $x \geq 0$ and $n \in \mathbb{N}$, we denote the unique nonnegative solution to the equation $y^n = x$ as $y = \sqrt[n]{x} = x^{1/n}$. Conventionally, if $n = 2$, we write $\sqrt{x}$ instead of $\sqrt[2]{x}$.

Note that $\sqrt{2}$ and $-\sqrt{2}$ are irrational by Proposition 1.15. Therefore, we have proved that irrational numbers exist.

Proposition 1.21. *For all $x \in \mathbb{R}$, there exists a unique $n \in \mathbb{Z}$ such that $x - 1 < n \leq x$.*

Proof. Let $S = \{n \in \mathbb{Z} \mid n > x - 1\}$. By the Archimedean Property, S is non-empty. The Archimedean Property also implies the existence of $m \in \mathbb{N}$ such that $m > -(x-1)$. Then $-m < x - 1$, and it follows that S is bounded below by $-m$. By the Well-Ordering Principle, S has a smallest element n. Of course, $n \in \mathbb{Z}$ and $n > x - 1$ since $n \in S$. To show that $n \leq x$, suppose $n > x$. Then $n - 1 > x - 1$, which means $n - 1 \in S$, contradicting that n is the smallest element of S. Therefore, $n \leq x$.

To prove uniqueness, suppose $m \in \mathbb{Z}$ satisfies $x - 1 < m \leq x$. If $m < n$, then $m \leq n - 1 \leq x - 1$, which is impossible. Similarly, if $n < m$, then $n \leq m - 1 \leq x - 1$, which is impossible. Hence, $m = n$. □

Definition 1.22. For all $x \in \mathbb{R}$, the *floor of x*, denoted $\lfloor x \rfloor$, is the unique integer such that $x - 1 < \lfloor x \rfloor \leq x$.

Exercise 1.23. Prove that for all $x \in \mathbb{R}$, there exists a unique integer $n \in \mathbb{Z}$ such that $x \leq n < x+1$. We call n the *ceiling of* x and denote it $\lceil x \rceil$.

Exercise 1.24 (Rational Exponents)**.** Fix $x > 0$ and $q \in \mathbb{Q}$. Write $q = \frac{m}{n}$ where $m \in \mathbb{Z}$ and $n \in \mathbb{N}$. We would like to define

$$x^q := (x^{1/n})^m.$$

(a) To make sure our definition is valid, we should check that x^q can have only one value. Prove that if $\frac{m}{n} = \frac{k}{\ell}$ where $k \in \mathbb{Z}$ and $\ell \in \mathbb{N}$, then

$$(x^{1/n})^m = (x^{1/\ell})^k.$$

Therefore, the value of x^q is independent of how we represent q as a fraction.

(b) Let $q_1, q_2 \in \mathbb{Q}$. Show that

$$x^{q_1} x^{q_2} = x^{q_1+q_2}$$

and

$$(x^{q_1})^{q_2} = x^{q_1 q_2}.$$

(c) Let $y > 0$. Show that

$$(xy)^q = x^q y^q.$$

1.4 Density of the Rationals and Irrationals in $\mathbb{R}$

Theorem 1.25. *Let $x, y \in \mathbb{R}$ such that $x < y$. Then there exists $q \in \mathbb{Q}$ such that $x < q < y$.*

Proof. Choose $n \in \mathbb{N}$ such that $n(y-x) > 1$. Then $nx + 1 < ny$. Let $m = \lfloor nx+1 \rfloor$. Then $nx < m \leq nx + 1 < ny$, so $x < \frac{m}{n} < y$. Therefore, we can choose $q = \frac{m}{n} \in \mathbb{Q}$ to satisfy the conclusion of the theorem. □

Theorem 1.26. *Let $x, y \in \mathbb{R}$ such that $x < y$. Then there exists $t \in \mathbb{R} \setminus \mathbb{Q}$ such that $x < t < y$.*

Proof. By Theorem 1.25, there exists $q_1 \in \mathbb{Q}$ such that $\frac{x}{\sqrt{2}} < q_1 < \frac{y}{\sqrt{2}}$. Again by Theorem 1.25, there exists $q_2 \in \mathbb{Q}$ such that $\frac{x}{\sqrt{2}} < q_2 < q_1$. Since $q_1 \neq q_2$, at least one of q_1, q_2 is non-zero. Therefore, there exists a non-zero $q \in \mathbb{Q}$ such that $\frac{x}{\sqrt{2}} < q < \frac{y}{\sqrt{2}}$. Then $x < q\sqrt{2} < y$. We claim that $q\sqrt{2}$ is irrational. Suppose $q\sqrt{2} \in \mathbb{Q}$. Then there exists $m \in \mathbb{Z}$ and $n \in \mathbb{N}$ such that $q\sqrt{2} = \frac{m}{n}$. Since $q \in \mathbb{Q}$, there exists $a \in \mathbb{Z}$ and $b \in \mathbb{N}$ such that $q = \frac{a}{b}$. Note that $a \neq 0$ since $q \neq 0$. Therefore, $\sqrt{2} = (\frac{b}{a})(\frac{m}{n}) = \frac{bm}{an} \in \mathbb{Q}$, which is impossible. We conclude that $q\sqrt{2}$ is irrational, so we can pick $t = q\sqrt{2}$. □

Exercise 1.27. Prove that every real number is the supremum of a set of rational numbers.

Exercise 1.28. Let A and B be non-empty sets of real numbers that are bounded above. Let

$$A + B = \{a + b \mid a \in A, b \in B\}.$$

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

Now assume further that A and B only contain nonnegative numbers. Let

$$AB = \{ab \mid a \in A, b \in B\}.$$

Prove that $\sup(AB) = \sup(A)\sup(B)$.

1.5 Complex Numbers

Definition 1.29. The set of complex numbers, denoted $\mathbb{C}$, is formed by equipping the set $\mathbb{R}^2$ with the addition operation

$$(a, b) + (c, d) \coloneqq (a + c, b + d)$$

and the multiplication operation

$$(a, b) \cdot (c, d) \coloneqq (ac - bd, ad + bc),$$

where $a, b, c, d \in \mathbb{R}$. The *imaginary unit* is $i \coloneqq (0, 1)$. Two complex numbers (a, b) and (c, d) are *equal* if $a = c$ and $b = d$.

By convention, we denote $(a, b) \in \mathbb{C}$ as $a + bi$. Then the addition and multiplication operations are

$$(a + bi) + (c + di) \coloneqq (a + c) + (b + d)i$$

and

$$(a + bi)(c + di) \coloneqq ac - bd + (ad + bc)i.$$

The real numbers naturally embed themselves into $\mathbb{C}$ via the injective map $x \mapsto x + 0i$ for $x \in \mathbb{R}$. Therefore, we can view $\mathbb{R}$ as a subset of $\mathbb{C}$, and we consider $x \in \mathbb{R}$ to be equal to $x + 0i \in \mathbb{C}$.

Proposition 1.30. *$\mathbb{C}$ is a field with additive identity $0 = 0 + 0i$ and multiplicative identity $1 = 1 + 0i$.*

Proof. We will prove the existence of multiplicative inverses for non-zero complex numbers. The other field axioms are easy but tedious to verify, so we leave them to the reader. Suppose $z = x + yi \in \mathbb{C}$ is non-zero, where $x, y \in \mathbb{R}$. Then at least one of x and y is non-zero, so $x^2 + y^2 \neq 0$. Let $w = \frac{x}{x^2+y^2} - \frac{y}{x^2+y^2} i \in \mathbb{C}$. Then

$$\begin{aligned}
zw &= (x + yi)\left(\frac{x}{x^2 + y^2} - \frac{y}{x^2 + y^2} i\right) \\
&= \frac{x^2}{x^2 + y^2} - \left(-\frac{y^2}{x^2 + y^2}\right) + \left(-\frac{xy}{x^2 + y^2} + \frac{yx}{x^2 + y^2}\right) i \\
&= 1 + 0i.
\end{aligned}$$

The reader can check that multiplication is commutative in $\mathbb{C}$, so $zw = 1 + 0i = wz$. Therefore, z^{-1} exists and equals w. $\square$

Proposition 1.31. *$\mathbb{C}$ is not an ordered field under any total order "$<$".*

Proof. The crucial fact is that $i^2 = -1$. Indeed,

$$i^2 = (0 + 1i)(0 + 1i) = (0)(0) - (1)(1) + [(0)(1) + (1)(0)]i = -1 + 0i.$$

If $\mathbb{C}$ is an ordered field under some total order "$<$", then part (c) of Proposition 1.9 implies that $-1 = i^2 \geq 0$. But $-1 < 0$ because $1 > 0$ by part (c) of Proposition 1.9. We have a contradiction, so $\mathbb{C}$ cannot be an ordered field. $\square$

Definition 1.32. Let $z = x + yi \in \mathbb{C}$ where $x, y \in \mathbb{R}$.

(a) The *magnitude* of z is $|z| := \sqrt{x^2 + y^2}$.

(b) The *complex conjugate* of z is $\overline{z} := x - yi$.

(c) The *real part* of z is $\mathrm{Re}(z) := x$.

(d) The *imaginary part* of z is $\mathrm{Im}(z) := y$.

Notice that the notation for the magnitude of a complex number is the same as the notation for the absolute value of a real number. Fortunately, these two concepts do not clash—if $x \in \mathbb{R}$, then the magnitude of x is $\sqrt{x^2 + 0^2}$, which is equal to the absolute value of x.

Exercise 1.33. Let $z, w \in \mathbb{C}$. Prove that:

(a) $\overline{(\overline{z})} = z$.

(b) $\mathrm{Re}(z) = \frac{1}{2}(z + \overline{z})$.

(c) $\mathrm{Im}(z) = \frac{1}{2}(z - \overline{z})$.

(d) $|z|^2 = z\overline{z}$.

(e) $\overline{z + w} = \overline{z} + \overline{w}$.

(f) $\overline{zw} = (\overline{z})(\overline{w})$.

(g) $|zw| = |z| \cdot |w|$.

(h) $|\overline{z}| = |z|$.

1.6 The Cauchy–Schwarz Inequality

Definition 1.34. A *vector space* over a field F is a set V with a vector-additon operation $+ : V \times V \to V$ and a scalar-multiplication operation $\cdot : F \times V \to V$ such that

(i) $(V, +, \vec{0})$ is an abelian group (where $\vec{0} \in V$),

(ii) $1\vec{v} = \vec{v}$ for all $\vec{v} \in V$ (where 1 is the multiplicative identity of F),

(ii) $a(\vec{v} + \vec{w}) = a\vec{v} + a\vec{w}$ for all $a \in F$ and $\vec{v}, \vec{w} \in V$, and

(iv) $(a + b)\vec{v} = a\vec{v} + b\vec{v}$ for all $a, b \in F$ and $\vec{v} \in V$.

Given any field F and $n \in \mathbb{N}$, the set $F^n = \{(v_1, \ldots, v_n) : v_i \in F \text{ for all } 1 \le i \le n\}$ is a vector space over F under the operations

$$(v_1, \ldots, v_n) + (w_1, \ldots, w_n) = (v_1 + w_1, \ldots, v_n + w_n)$$

and

$$c(v_1, \ldots, v_n) = (cv_1, \ldots, cv_n)$$

where $c, v_i, w_i \in F$ for all $1 \le i \le n$.

Definition 1.35. Let V be a vector space over $F \in \{\mathbb{R}, \mathbb{C}\}$. A *semi-inner product* on V is a map $\langle \cdot, \cdot \rangle : V \times V \to F$ such that

(i) $\langle a\vec{v}_1 + \vec{v}_2, \vec{w} \rangle = a\langle \vec{v}_1, \vec{w} \rangle + \langle \vec{v}_2, \vec{w} \rangle$ for all $a \in F$ and $\vec{v}_1, \vec{v}_2, \vec{w} \in V$,

(ii) $\langle \vec{v}, \vec{w} \rangle = \overline{\langle \vec{w}, \vec{v} \rangle}$ for all $\vec{v}, \vec{w} \in V$, and

(iii) $\langle \vec{v}, \vec{v} \rangle \ge 0$ for all $\vec{v} \in V$. (Note that axiom (ii) already implies that $\langle \vec{v}, \vec{v} \rangle$ is a real number for all $\vec{v} \in V$ because $\langle \vec{v}, \vec{v} \rangle = \overline{\langle \vec{v}, \vec{v} \rangle}$.)

A *semi-inner product space* is a vector space equipped with a semi-inner product. If V is a semi-inner product space, then for any $\vec{v} \in V$, the *seminorm* of $\vec{v}$ is $||\vec{v}|| := \sqrt{\langle \vec{v}, \vec{v} \rangle}$.

Proposition 1.36. *Let V be a semi-inner product space. Then $\langle \vec{0}, \vec{v} \rangle = 0 = \langle \vec{v}, \vec{0} \rangle$ for all $\vec{v} \in V$.*

Proof. We have that $\langle \vec{0}, \vec{v} \rangle = \langle \vec{0} + \vec{0}, \vec{v} \rangle = \langle \vec{0}, \vec{v} \rangle + \langle \vec{0}, \vec{v} \rangle$, so $0 = \langle \vec{0}, \vec{v} \rangle$ and $\langle \vec{v}, \vec{0} \rangle = \overline{\langle \vec{0}, \vec{v} \rangle} = \overline{0} = 0$. □

Proposition 1.37. *Let V be a semi-inner product space and $\vec{v}, \vec{w} \in V$. If $||\vec{v}|| = ||\vec{w}|| = 0$, then $\langle \vec{v}, \vec{w} \rangle = 0$.*

Proof. Suppose $||\vec{v}|| = ||\vec{w}|| = 0$. Let $c = \langle \vec{v}, \vec{w} \rangle$. Then

$$\begin{aligned} 0 &\leq ||\vec{v} - c\vec{w}||^2 \\ &= \langle \vec{v} - c\vec{w}, \vec{v} - c\vec{w} \rangle \\ &= \langle \vec{v}, \vec{v} - c\vec{w} \rangle - \langle c\vec{w}, \vec{v} - c\vec{w} \rangle \\ &= \langle \vec{v}, \vec{v} \rangle - \overline{c}\langle \vec{v}, \vec{w} \rangle - c\overline{\langle \vec{v}, \vec{w} \rangle} + c\overline{c}\langle \vec{w}, \vec{w} \rangle \\ &= ||\vec{v}||^2 - \overline{c}c - c\overline{c} + |c|^2||\vec{w}||^2 \\ &= -2|c|^2 \quad \text{since } ||\vec{v}|| = ||\vec{w}|| = 0. \end{aligned}$$

Since $0 \leq -2|c|^2 \leq 0$, it follows that $-2|c|^2 = 0$, so $c = 0$. □

Theorem 1.38 (Pythagorean Theorem)**.** *Let V be a semi-inner product space and $\vec{v}, \vec{w} \in V$. If $\langle \vec{v}, \vec{w} \rangle = 0$, then $||\vec{v} + \vec{w}||^2 = ||\vec{v}||^2 + ||\vec{w}||^2$.*

Proof. We have

$$||\vec{v} + \vec{w}||^2 = ||\vec{v}||^2 + \langle \vec{v}, \vec{w} \rangle + \overline{\langle \vec{v}, \vec{w} \rangle} + ||\vec{w}||^2$$

for all $\vec{v}, \vec{w} \in V$, so if $\langle \vec{v}, \vec{w} \rangle = 0$, then $||\vec{v} + \vec{w}||^2 = ||\vec{v}||^2 + ||\vec{w}||^2$ indeed. □

Theorem 1.39 (Cauchy–Schwarz)**.** *Let V be a semi-inner product space. Then*

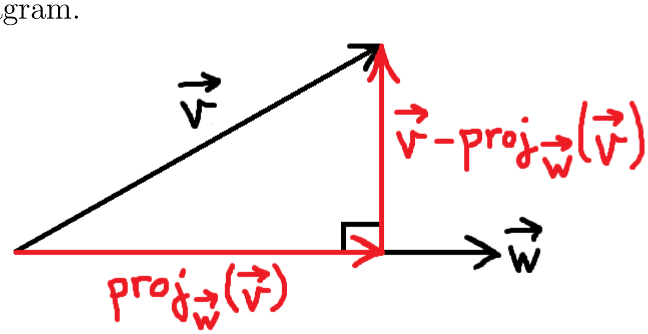
$$|\langle \vec{v}, \vec{w} \rangle| \leq ||\vec{v}|| \cdot ||\vec{w}||$$

for all $\vec{v}, \vec{w} \in V$.

Let us recall projections from linear algebra. In $\mathbb{R}^n$ with the standard inner product (i.e. the "dot" product), the *orthogonal projection* of a vector $\vec{v}$ onto a non-zero vector $\vec{w}$ is

$$\text{proj}_{\vec{w}}(\vec{v}) := \frac{\langle \vec{v}, \vec{w} \rangle}{||\vec{w}||^2}\vec{w}.$$

The vectors $\text{proj}_{\vec{w}}(\vec{v})$, $\vec{v} - \text{proj}_{\vec{w}}(\vec{v})$, and $\vec{v}$ form a right triangle, as shown in the following diagram.



Therefore, we can apply the Pythagorean Theorem and deduce that $||\mathrm{proj}_{\vec{w}}(\vec{v})|| \leq ||\vec{v}||$. Multiplying both sides of this inequality by $||\vec{w}||$ gives the desired inequality. We can generalize this intuition to any semi-inner product space.

Proof. If $||\vec{v}|| = ||\vec{w}|| = 0$, then $|\langle \vec{v}, \vec{w}\rangle| = 0$ by Proposition 1.37, so the inequality holds.

Now we consider the case where $||\vec{v}|| > 0$ or $||\vec{w}|| > 0$. Since the desired inequality is symmetric in $\vec{v}$ and $\vec{w}$, we can assume that $||\vec{w}|| > 0$ without loss of generality. Observe that

$$\begin{aligned}
\left\langle \frac{\langle \vec{v}, \vec{w}\rangle}{||\vec{w}||^2}\vec{w}, \vec{v} - \frac{\langle \vec{v}, \vec{w}\rangle}{||\vec{w}||^2}\vec{w} \right\rangle &= \left\langle \frac{\langle \vec{v}, \vec{w}\rangle}{||\vec{w}||^2}\vec{w}, \vec{v} \right\rangle - \left\langle \frac{\langle \vec{v}, \vec{w}\rangle}{||\vec{w}||^2}\vec{w}, \frac{\langle \vec{v}, \vec{w}\rangle}{||\vec{w}||^2}\vec{w} \right\rangle \\
&= \frac{\langle \vec{v}, \vec{w}\rangle}{||\vec{w}||^2}\overline{\langle \vec{v}, \vec{w}\rangle} - \frac{|\langle \vec{v}, \vec{w}\rangle|^2}{||\vec{w}||^4}\langle \vec{w}, \vec{w}\rangle \\
&= \frac{|\langle \vec{v}, \vec{w}\rangle|^2}{||\vec{w}||^2} - \frac{|\langle \vec{v}, \vec{w}\rangle|^2}{||\vec{w}||^2} \\
&= 0.
\end{aligned}$$

By the Pythagorean Theorem,

$$\begin{aligned}
||\vec{v}||^2 &= \left|\left| \frac{\langle \vec{v}, \vec{w}\rangle}{||\vec{w}||^2}\vec{w} + \vec{v} - \frac{\langle \vec{v}, \vec{w}\rangle}{||\vec{w}||^2}\vec{w} \right|\right|^2 \\
&= \left|\left| \frac{\langle \vec{v}, \vec{w}\rangle}{||\vec{w}||^2}\vec{w} \right|\right|^2 + \left|\left| \vec{v} - \frac{\langle \vec{v}, \vec{w}\rangle}{||\vec{w}||^2}\vec{w} \right|\right|^2 \\
&= \frac{|\langle \vec{v}, \vec{w}\rangle|^2}{||\vec{w}||^2} + \left|\left| \vec{v} - \frac{\langle \vec{v}, \vec{w}\rangle}{||\vec{w}||^2}\vec{w} \right|\right|^2,
\end{aligned}$$

so

$$||\vec{v}||^2 - \frac{|\langle \vec{v}, \vec{w}\rangle|^2}{||\vec{w}||^2} = \left|\left| \vec{v} - \frac{\langle \vec{v}, \vec{w}\rangle}{||\vec{w}||^2}\vec{w} \right|\right|^2 \geq 0.$$

Therefore,

$$|\langle \vec{v}, \vec{w}\rangle|^2 \leq ||\vec{v}||^2||\vec{w}||^2 = (||\vec{v}|| \cdot ||\vec{w}||)^2,$$

so

$$|\langle \vec{v}, \vec{w}\rangle| \leq ||\vec{v}|| \cdot ||\vec{w}||$$

because $|\langle \vec{v}, \vec{w}\rangle|$ and $||\vec{v}|| \cdot ||\vec{w}||$ are nonnegative. ∎

Corollary 1.39.1. *Let $z_1, \ldots, z_n, w_1, \ldots, w_n \in \mathbb{C}$. Then*

$$\left| \sum_{j=1}^n z_j \overline{w_j} \right| \leq \left(\sum_{j=1}^n |z_j|^2 \right)^{1/2} \left(\sum_{j=1}^n |w_j|^2 \right)^{1/2}.$$

Proof. For any $\vec{z} = (z_1, \ldots, z_n) \in \mathbb{C}^n$ and $\vec{w} = (w_1, \ldots, w_n) \in \mathbb{C}^n$, define

$$\langle \vec{z}, \vec{w}\rangle := \sum_{j=1}^n z_j \overline{w_j}. \tag{1}$$

Then $\langle \cdot, \cdot \rangle : \mathbb{C}^n \times \mathbb{C}^n \to \mathbb{C}$ is a semi-inner product, and the claimed inequality is simply the Cauchy–Schwarz Inequality applied to this semi-inner product. ∎

Corollary 1.39.2 (Triangle Inequality)**.** *Let V be a semi-inner product space and $\vec{v}, \vec{w} \in V$. Then $||\vec{v} + \vec{w}|| \leq ||\vec{v}|| + ||\vec{w}||$.*

Proof. We will show that $||\vec{v} + \vec{w}||^2 \leq (||\vec{v}|| + ||\vec{w}||)^2$. Indeed.

$$\begin{aligned}
||\vec{v} + \vec{w}||^2 &= ||\vec{v}||^2 + \langle \vec{v}, \vec{w} \rangle + \langle \vec{w}, \vec{v} \rangle + ||\vec{w}||^2 \\
&= ||\vec{v}||^2 + 2\mathrm{Re}(\langle \vec{v}, \vec{w} \rangle) + ||\vec{w}||^2 \\
&\leq ||\vec{v}||^2 + 2|\langle \vec{v}, \vec{w} \rangle| + ||\vec{w}||^2 \\
&\leq ||\vec{v}||^2 + 2||\vec{v}|| \cdot ||\vec{w}|| + ||\vec{w}||^2 \quad \text{by Cauchy–Schwarz} \\
&= (||\vec{v}|| + ||\vec{w}||)^2.
\end{aligned}$$

Since $||\vec{v} + \vec{w}|| \geq 0$ and $||\vec{v}|| + ||\vec{w}|| \geq 0$, we get that $||\vec{v} + \vec{w}|| \leq ||\vec{v}|| + ||\vec{w}||$. □

Remark. You may be curious about why the term "semi-inner product" has the prefix "semi". An *inner product* is a semi-inner product with the additional condition that $\langle \vec{v}, \vec{v} \rangle = 0$ if and only if $\vec{v} = \vec{0}$. The familiar "dot product" on $\mathbb{R}^n$, defined by

$$\vec{v} \cdot \vec{w} \coloneqq \sum_{j=1}^{n} v_j w_j$$

for all $\vec{v}, \vec{w} \in \mathbb{R}^n$, is an inner product in this sense. Equation (1) also defines an inner product, this time on $\mathbb{C}^n$. However, not all semi-inner products are inner products.

An *inner product space* is a vector space equipped with an inner product. It is possible to turn any semi-inner product space into an inner product space—the trick is to use equivalence classes and put all the zero-norm vectors into the same equivalence class. Let V be a semi-inner product space, and let $Z \subset V$ be the subset of vectors with zero norm. Define a relation $\sim$ on V such that $\vec{v} \sim \vec{w}$ if and only if $\vec{v} - \vec{w} \in Z$. Then $\sim$ is an equivalence relation (transitivity follows from the Triangle Inequality). Let X be the set of equivalence classes under $\sim$, and for each $\vec{v} \in V$, let $[\vec{v}]$ be the equivalence class of $\vec{v}$. Then in fact, X is a vector space under the operations

$$[\vec{v}] + [\vec{w}] = [\vec{v} + \vec{w}]$$

and

$$c[\vec{v}] = [c\vec{v}].$$

We can now define $\langle [\vec{v}], [\vec{w}] \rangle \coloneqq \langle \vec{v}, \vec{w} \rangle$, which is a true inner product on X since there is a unique equivalence class containing all the zero-norm elements of V. This construction is performed when we study square-integrable functions[1] (a future topic in analysis).

Exercise 1.40. Let V be a semi-inner product space. Prove that if $||\vec{v} - \vec{x}|| = 0$ and $||\vec{w} - \vec{y}|| = 0$, then $\langle \vec{v}, \vec{w} \rangle = \langle \vec{x}, \vec{y} \rangle$.

Exercise 1.41. Let V be an inner product space (this means if $||\vec{v}|| = 0$, then $\vec{v} = 0$). Determine which vectors $\vec{v}, \vec{w} \in V$ satisfy

$$|\langle \vec{v}, \vec{w} \rangle| = ||\vec{v}|| \cdot ||\vec{w}||.$$

[1] `https://en.wikipedia.org/wiki/Square-integrable_function`

1.7 Optional: Constructing the Real Numbers

In this section, we will construct the set of real numbers $\mathbb{R}$ using infinite sequences of rational numbers. We equip $\mathbb{R}$ with an addition operation, a multiplication operation, and a total order, and we prove that $\mathbb{R}$ is a complete ordered field. If the reader is not already familiar with sequences and "$\epsilon - N$" proofs, then the reader may wish to skip this section for now and come back after studying Chapter 3.

Definition 1.42. An infinite sequence $\{q_n\}$ of rational numbers is *Cauchy* if for all rational numbers $t > 0$, there exists $N \in \mathbb{N}$ such that $|q_n - q_m| < t$ for all $n, m \geq N$.

Let $\mathcal{C}$ be the set of Cauchy sequences of rational numbers.

Definition 1.43. Let $\{q_n\}, \{r_n\} \in \mathcal{C}$. We write $\{q_n\} \sim \{r_n\}$ if for all rational numbers $t > 0$, there exists $N \in \mathbb{N}$ such that $|q_n - r_n| < t$ for all $n \geq N$.

Lemma 1.44. *$\sim$ is an equivalence relation on $\mathcal{C}$.*

Proof. Let $\{q_n\}, \{r_n\}, \{s_n\} \in \mathcal{C}$. Fix a rational $t > 0$. We have that $|q_n - q_n| = 0 < t$ for all $n \geq 1$, so $\{q_n\} \sim \{q_n\}$. Suppose $\{q_n\} \sim \{r_n\}$. Then there exists $N_0 \in \mathbb{N}$ such that $|q_n - r_n| < t$ for all $n \geq N_0$. Note that $|r_n - q_n| = |q_n - r_n| < t$ for all $n \geq N_0$. Hence, $\{r_n\} \sim \{q_n\}$. Lastly, suppose $\{q_n\} \sim \{r_n\}$ and $\{r_n\} \sim \{s_n\}$. Then there exists $N_1, N_2 \in \mathbb{N}$ such that $|q_n - r_n| < \frac{t}{2}$ for all $n \geq N_1$ and $|r_n - s_n| < \frac{t}{2}$ for all $n \geq N_2$. By the Triangle Inequality (which holds in $\mathbb{Q}$ by Theorem 1.12), $|q_n - s_n| = |q_n - r_n + r_n - s_n| \leq |q_n - r_n| + |r_n - s_n| < \frac{t}{2} + \frac{t}{2} = t$ whenever $n \geq \max\{N_1, N_2\}$. Hence, $\{q_n\} \sim \{s_n\}$. $\square$

Definition 1.45. $\mathbb{R}$ is the set of equivalence classes of $\mathcal{C}$ under the equivalence relation $\sim$. We denote the equivalence class containing $\{q_n\}$ as $[\{q_n\}]$. For any $q \in \mathbb{Q}$, we denote $q_{\mathbb{R}} := [\{q, q, q, \dots\}]$.

Lemma 1.46. *Let $\{q_n\} \in \mathcal{C}$. Then there exists $C \in \mathbb{Q}$ such that $|q_n| \leq C$ for all $n \in \mathbb{N}$.*

Proof. Since $\{q_n\} \in \mathcal{C}$, there exists $N \in \mathbb{N}$ such that $|q_n - q_m| < 1$ for all $n, m \geq N$. In particular, for all $n \geq N$, we have that $|q_n - q_N| < 1$, so

$$|q_n| = |(q_n - q_N) + q_N| \leq |q_n - q_N| + |q_N| = 1 + |q_N|.$$

Now take $C = \max\{|q_1|, \dots, |q_{N-1}|, 1 + |q_N|\}$, which is in $\mathbb{Q}$ since each argument of the "max" function is in $\mathbb{Q}$. If $1 \leq n < N$, then $|q_n| \leq C$, and if $n \geq N$, then $|q_n| < 1 + |q_N| \leq C$. $\square$

Lemma 1.47. *If $\{q_n\}, \{r_n\} \in \mathcal{C}$, then $\{q_n + r_n\} \in \mathcal{C}$ and $\{q_n r_n\} \in \mathcal{C}$.*

Proof. Let $\{q_n\}, \{r_n\} \in \mathcal{C}$. By Lemma 1.46, there exist $C_q, C_r \in \mathbb{Q}$ such that $|q_n| \leq C_q$ and $|r_n| \leq C_r$ for all $n \in \mathbb{N}$. Note that C_q and C_r are nonnegative. Fix a rational $t > 0$. Let $\gamma = \min\{\frac{t}{2}, \frac{t}{1+C_q+C_r}\}$, which is a positive rational number since $1 + C_q + C_r \geq 1$. There exists $N_1, N_2 \in \mathbb{N}$ such that $|q_n - q_m| < \gamma$ for all $n, m \geq N_1$ and $|r_n - r_m| < \gamma$ for all $n, m \geq N_2$. Let $N = \max\{N_1, N_2\}$. If $n, m \geq N$, then

$$|(q_n + r_n) - (q_m + r_m)| \leq |q_n - q_m| + |r_n - r_m| < \gamma + \gamma \leq \frac{t}{2} + \frac{t}{2} = t$$

and

$$\begin{aligned}
|q_n r_n - q_m r_m| &= |q_n r_n - q_n r_m + q_n r_m - q_m r_m| \\
&= |q_n(r_n - r_m) + r_m(q_n - q_m)| \\
&\leq |q_n| \cdot |r_n - r_m| + |r_m| \cdot |q_n - q_m| \\
&\leq C_q\gamma + C_r\gamma \\
&< (1 + C_q + C_r)\gamma \\
&\leq t \quad [\text{since } \gamma \leq \frac{t}{1 + C_q + C_r}].
\end{aligned}$$

Therefore, $\{q_n + r_n\} \in \mathcal{C}$ and $\{q_n r_n\} \in \mathcal{C}$. $\square$

Lemma 1.48. *Let $\{q_n\}, \{r_n\}, \{s_n\}, \{t_n\} \in \mathcal{C}$. Suppose $\{q_n\} \sim \{s_n\}$ and $\{r_n\} \sim \{t_n\}$. Then $\{q_n + r_n\} \sim \{s_n + t_n\}$ and $\{q_n r_n\} \sim \{s_n t_n\}$.*

Proof. Suppose $\{q_n\} \sim \{s_n\}$ and $\{r_n\} \sim \{t_n\}$. Fix a rational $u > 0$. By Lemma 1.46, there exist $C_r, C_s \in \mathbb{Q}$ such that $|r_n| \leq C_r$ and $|s_n| \leq C_s$ for all $n \in \mathbb{N}$. Let $\gamma = \min\{\frac{u}{2}, \frac{u}{1+C_r+C_s}\}$. Then there exist $N_1, N_2 \in \mathbb{N}$ such that $|q_n - s_n| < \gamma$ for all $n \geq N_1$ and $|r_n - t_n| < \gamma$ for all $n \geq N_2$. If $n \geq \max\{N_1, N_2\}$, then

$$|(q_n + r_n) - (s_n + t_n)| \leq |q_n - s_n| + |r_n - t_n| < \gamma + \gamma \leq \frac{u}{2} + \frac{u}{2} = u$$

and

$$\begin{aligned}
|q_n r_n - s_n t_n| &= |q_n r_n - s_n r_n + s_n r_n - s_n t_n| \\
&\leq |r_n(q_n - s_n)| + |s_n(r_n - t_n)| \\
&= |r_n| \cdot |q_n - s_n| + |s_n| \cdot |r_n - t_n| \\
&\leq C_r\gamma + C_s\gamma \\
&< (1 + C_r + C_s)\gamma \\
&\leq u.
\end{aligned}$$

Therefore, $\{q_n + s_n\} \sim \{r_n + t_n\}$ and $\{q_n s_n\} \sim \{r_n t_n\}$. $\square$

Definition 1.49. For all $[\{q_n\}], [\{r_n\}] \in \mathbb{R}$, we define

$$[\{q_n\}] + [\{r_n\}] \coloneqq [\{q_n + r_n\}]$$

and

$$[\{q_n\}] \cdot [\{r_n\}] \coloneqq [\{q_n r_n\}].$$

Definition 1.50. Let $\{q_n\}, \{r_n\} \in \mathcal{C}$. We say that $\{q_n\} < \{r_n\}$ if there exists a rational $t > 0$ and $N \in \mathbb{N}$ such that $r_n - q_n > t$ for all $n \geq N$.

Lemma 1.51. *Let $\{q_n\}, \{r_n\}, \{s_n\}, \{t_n\} \in \mathcal{C}$. Suppose $\{q_n\} \sim \{s_n\}$, $\{r_n\} \sim \{t_n\}$, and $\{q_n\} < \{r_n\}$. Then $\{s_n\} < \{t_n\}$.*

Proof. Since $\{q_n\} < \{r_n\}$, there exists a rational $u > 0$ and $N_1 \in \mathbb{N}$ such that $r_n - q_n > u$ for all $n \geq N_1$. Since $\{q_n\} \sim \{s_n\}$ and $\{r_n\} \sim \{t_n\}$, there exist $N_2, N_3 \in \mathbb{N}$ such that $|q_n - s_n| < \frac{u}{3}$ if $n \geq N_2$ and $|r_n - t_n| < \frac{u}{3}$ if $n \geq N_3$. Therefore, for all $n \geq \max\{N_1, N_2, N_3\}$, we have that

$$t_n - s_n = (t_n - r_n) + (r_n - q_n) + (q_n - s_n) > -\frac{u}{3} + u - \frac{u}{3} = \frac{u}{3},$$

so $\{s_n\} < \{t_n\}$. $\square$

Definition 1.52. For all $[\{q_n\}], [\{r_n\}] \in \mathbb{R}$, we say that $[\{q_n\}] < [\{r_n\}]$ if $\{q_n\} < \{r_n\}$.

Theorem 1.53. $\mathbb{R}$ *is a field.*

Proof. First, note that if $[\{q_n\}] \in \mathbb{R}$, then

$$[\{q_n\}] + 0_{\mathbb{R}} = [\{q_n + 0\}] = [\{q_n\}] = [\{0 + q_n\}] = 0_{\mathbb{R}} + [\{q_n\}]$$

and

$$[\{q_n\}] \cdot 1_{\mathbb{R}} = [\{q_n \cdot 1\}] = [\{q_n\}] = [\{1 \cdot q_n\}] = 1_{\mathbb{R}} \cdot [\{q_n\}].$$

Hence, $\mathbb{R}$ has an additive identity $0_{\mathbb{R}}$ and a multiplicative identity $1_{\mathbb{R}}$. It is also clear that $0_{\mathbb{R}} \neq 1_{\mathbb{R}}$.

Since the operations on $\mathbb{R}$ reduce to term-wise operations on rational Cauchy sequences, most of the field axioms for $\mathbb{R}$ can be proved simply by passing to $\mathbb{Q}$ and using the fact that $\mathbb{Q}$ is a field. The proof that every non-zero element of $\mathbb{R}$ has a multiplicative inverse is more interesting. Let $[\{q_n\}] \in \mathbb{R}$ be non-zero. We claim that there exists $N \in \mathbb{N}$ such that $q_n \neq 0$ for all $n \geq N$. Indeed, suppose not. For all rational $t > 0$, there exists $N_1 \in \mathbb{N}$ such that $|q_n - q_m| < t$ for all $n, m \geq N_1$. Choose $N_2 \geq N_1$ such that $q_{N_2} = 0$. Then $|q_n - 0| = |q_n - q_{N_2}| < t$ for all $n \geq N_1$. Hence, $\{q_n\} \sim \{0, 0, 0, \dots\}$, so $[\{q_n\}] = 0_{\mathbb{R}}$, which is a contradiction. Thus, our claim is proved. Now let $r_n = 0$ if $1 \leq n < N$ and $r_n = q_n^{-1}$ if $n \geq N$. Then $q_n r_n = 1$ for all $n \geq N$, so $[\{q_n r_n\}] = 1_{\mathbb{R}}$. Therefore, $[\{q_n\}]$ has a multiplicative inverse $[\{r_n\}]$. $\square$

Theorem 1.54. $\mathbb{R}$ *is totally ordered.*

Proof. Let $x = [\{x_n\}]$ and $y = [\{y_n\}]$ be real numbers. Suppose it is not the case that $x < y$ or $y < x$. Fix a rational $t > 0$. Choose $N_1, N_2 \in \mathbb{N}$ such that $|x_n - x_m| < \frac{t}{3}$ for all $n, m \geq N_1$ and $|y_n - y_m| < \frac{t}{3}$ for all $n, m \geq N_2$. Let $N_3 = \max\{N_1, N_2\}$. Since $x \not< y$ and $y \not< x$, there exist $m_0, n_0 \geq N_3$ such that $x_{m_0} - y_{m_0} \leq \frac{t}{3}$ and $y_{n_0} - x_{n_0} \leq \frac{t}{3}$. Then for all $n \geq N_3$,

$$x_n - y_n = (x_n - x_{m_0}) + (x_{m_0} - y_{m_0}) + (y_{m_0} - y_n) < \frac{t}{3} + \frac{t}{3} + \frac{t}{3} = t$$

and

$$y_n - x_n = (y_n - y_{n_0}) + (y_{n_0} - x_{n_0}) + (x_{n_0} - x_n) < \frac{t}{3} + \frac{t}{3} + \frac{t}{3} = t,$$

so $|x_n - y_n| < t$. Therefore, $\{x_n\} \sim \{y_n\}$, so $x = y$. Hence, any two elements of $\mathbb{R}$ can be compared.

Now we need to show that no two of the statements "$x < y$", "$x = y$", and "$y < x$" can be simultaneously true. Suppose $x = y$. Fix a rational $t > 0$. Then there exists $N_4 \in \mathbb{N}$ such that $-t < x_n - y_n < t$ for all $n \geq N_4$. Therefore, there is no $N \in \mathbb{N}$ such that $x_n - y_n > t$ for all $n \geq N$ because for any given $N \in \mathbb{N}$, $n_0 = \max\{N, N_4\}$ satisfies $n_0 \geq N$ and $x_{n_0} - y_{n_0} < t$. Hence, $y \not< x$. By a similar argument, $x \not< y$. Lastly, suppose $x < y$. Then there exists a rational $t_0 > 0$ and $N_5 \in \mathbb{N}$ such that $y_n - x_n > t_0$ for all $n \geq N_5$. For any rational $t > 0$ and $N \in \mathbb{N}$, $n_1 = \max\{N, N_5\}$ satisfies $n_1 \geq N$ and $x_{n_1} - y_{n_1} < -t_0 < 0 < t$. Hence, there is no $N \in \mathbb{N}$ such that $x_n - y_n > t$ for all $n \geq N$, so $y \not< x$.

It remains to prove that "$<$" in $\mathbb{R}$ is transitive. Let $z = [\{z_n\}] \in \mathbb{R}$, and suppose that $x < y$ and $y < z$. Then there exist rational numbers $t_1, t_2 > 0$ and $N_6, N_7 \in \mathbb{N}$ such that $y_n - x_n > t_1$ for all $n \geq N_6$ and $z_n - y_n > t_2$ for all $n \geq N_7$. Let $N_8 = \max\{N_6, N_7\}$. Then for all $n \geq N_8$,

$$z_n - x_n = (z_n - y_n) + (y_n - x_n) > t_2 + t_1.$$

Therefore, $x < z$ since $t_2 + t_1$ is a positive rational number. $\square$

Theorem 1.55. $\mathbb{R}$ *is an ordered field.*

Proof. Theorems 1.53 and 1.54 say that $\mathbb{R}$ is a field with a total order "$<$". Now we need to prove that $<$ is compatible with the field operations. Let $x = [\{x_n\}]$, $y = [\{y_n\}]$, and $z = [\{z_n\}]$ be real numbers. Suppose $y < z$. Then there exists a rational $t_0 > 0$ and $N_0 \in \mathbb{N}$ such that $z_n - y_n > t_0$ for all $n \geq N_0$. It follows that $(x_n + z_n) - (x_n + y_n) = z_n - y_n > t_0$ for all $n \geq N_0$, so $x + y < x + z$.

Suppose $x > 0_{\mathbb{R}}$ and $y > 0_{\mathbb{R}}$. Then there exist rational numbers $t_1, t_2 > 0$ and $N_1, N_2 \in \mathbb{N}$ such that $x_n > t_1$ for all $n \geq N_1$ and $y_n > t_2$ for all $n \geq N_2$. If $n \geq \max\{N_1, N_2\}$, then $x_n y_n > t_1 t_2$, so $\{x_n y_n\} > \{0, 0, 0, \dots\}$ since $t_1 t_2$ is a positive rational number. Therefore, $xy > 0_{\mathbb{R}}$. $\square$

Lemma 1.56. *Let $S \subset \mathbb{R}$ be non-empty and bounded above. Let*

$$T = \{t \in \mathbb{Q} \mid t_{\mathbb{R}} \textit{ is not an upper bound of } S\}$$

and

$$U = \{u \in \mathbb{Q} \mid u_{\mathbb{R}} \textit{ is an upper bound of } S\}.$$

Then for all real numbers $x > 0_{\mathbb{R}}$, there exist $t \in T$ and $u \in U$ such that $(u - t)_{\mathbb{R}} < x$.

Proof. First, we show that T and U are non-empty. Since S is non-empty, there exists an element $a = \{[a_n]\} \in S$. Choose $N_1 \in \mathbb{N}$ such that $|a_n - a_{N_1}| < 1$ for all $n \geq N_1$. Then $a_{N_1} - 1 < a_n$ for all $n \geq N_1$, so $(a_{N_1} - 2)_{\mathbb{R}} < a$ because $a_n - (a_{N_1} - 2) > 1$ for all $n \geq N_1$. Hence, $a_{N_1} - 2 \in T$, so T is non-empty. We argue similarly to show that U is non-empty. Let $b = \{[b_n]\} \in \mathbb{R}$ be an upper bound of S. Then there exists $N_2 \in \mathbb{N}$ such that $|b_n - b_{N_2}| < 1$ for all $n \geq N_2$. Hence, $b_n < b_{N_2} + 1$ for all $n \geq N$, so $b < (b_{N_2} + 2)_{\mathbb{R}}$ because $(b_{N_2} + 2) - b_n > 1$ for all $n \geq N$. Therefore, $b_{N_2} + 2 \in U$, so U is non-empty.

Write $a_{N_1} - 2 = \frac{k_T}{m_T}$ and $b_{N_2} + 2 = \frac{k_U}{m_U}$ where $k_T, k_U \in \mathbb{Z}$ and $m_T, m_U \in \mathbb{N}$. Fix $x = \{[x_n]\} > 0_{\mathbb{R}}$. Then there exists a rational $q > 0$ and $N \in \mathbb{N}$ such that $x_n > q$ for all

$n \geq N$. Clearly $q_{\mathbb{R}} \leq x_n$ since $q - x_n$ is negative for large n. Since q is positive, we can write $q = \frac{k}{m}$ where $k, m \in \mathbb{N}$. Consider the set

$$V = \left\{ j \in \mathbb{Z} \;\middle|\; \frac{j}{2m} \in U \right\}.$$

Note that $\frac{2m|k_U|}{2m} = |k_U| \geq \frac{|k_U|}{m_U} \geq \frac{k_U}{m_U} \in U$, so $2m|k_U| \in V$. Hence, V is non-empty. Also, $\frac{-2m|k_T|}{2m} = -|k_T| \leq \frac{-|k_T|}{m_T} \leq \frac{k_T}{m_T} \in T$, so V is bounded below by $-2m|k_T|$. By the Well-Ordering Principle, V has a smallest element j_0. Then $\frac{j_0}{2m} \in U$ because $j_0 \in V$. The minimality of j_0 implies that $\frac{j_0-1}{2m} \in T$. Finally, $\frac{j_0}{2m} - \frac{j_0-1}{2m} = \frac{1}{2m} < \frac{k}{m} = q$, so letting $u = \frac{j_0}{2m}$ and $t = \frac{j_0-1}{2m}$, we have that $(u - t)_{\mathbb{R}} < q_{\mathbb{R}} \leq x$. $\square$

Theorem 1.57. $\mathbb{R}$ *is complete.*

Proof. Let $S \subset \mathbb{R}$ be non-empty and bounded above, and let T and U be the sets of the same names from Lemma 1.56. For all $n \in \mathbb{N}$, there exists $t_n \in T$ and $u_n \in U$ such that $u_n - t_n < \frac{1}{n}$ by Lemma 1.56. Let $m, n \in \mathbb{N}$ such that $m \leq n$. If $u_n \geq u_m + \frac{1}{m}$, then $u_n - t_n > u_n - u_m \geq \frac{1}{m} \geq \frac{1}{n}$, which is a contradiction. If $u_n \leq u_m - \frac{1}{m}$, then $u_m - t_m > u_m - u_n \geq \frac{1}{m}$, which is a contradiction. Therefore, we must have that $u_m - \frac{1}{m} < u_n < u_m + \frac{1}{m}$, or equivalently, that $|u_n - u_m| < \frac{1}{m}$.

Let $t > 0$ be a rational number, and write $t = \frac{k}{\ell}$ where $k, \ell \in \mathbb{N}$. Let $N_0 = \ell \in \mathbb{N}$. Then for all $n, m \geq N_0$, we have that

$$|u_n - u_m| < \frac{1}{\min\{n, m\}} \leq \frac{1}{N_0} = \frac{1}{\ell} \leq t$$

because $\min\{n, m\} \geq N_0$. Therefore, $\{u_n\} \in \mathcal{C}$.

Let $u = [\{u_n\}] \in \mathbb{R}$. Fix an arbitrary $x = [\{x_n\}] \in S$, and suppose $x \neq u$. Then there exists a rational $t > 0$ such that for all $N \in \mathbb{N}$, we have that $|x_n - u_n| \geq t$ for some $n \geq N$. Since $\{x_n\}, \{u_n\} \in \mathcal{C}$, there exists $N_1, N_2 \in \mathbb{N}$ such that $|x_n - x_m| < \frac{t}{3}$ for all $n, m \geq N_1$ and $|u_n - u_m| < \frac{t}{3}$ for all $n, m \geq N_2$. Let $N_3 = \max\{N_1, N_2\}$, and choose $n_0 \geq N_3$ such that $|x_{n_0} - u_{n_0}| \geq t$. Suppose $u_{n_0} < x_{n_0}$. Then for any $n \geq N_3$,

$$x_n - u_n = (x_n - x_{n_0}) + (x_{n_0} - u_{n_0}) + (u_{n_0} - u_n) > -\frac{t}{3} + t - \frac{t}{3} = \frac{t}{3}.$$

It follows that $(u_{n_0})_{\mathbb{R}} < x$, contradicting the fact that $u_{n_0} \in U$. Therefore, $x_{n_0} < u_{n_0}$. Now a similar argument shows that $u_n - x_n > \frac{t}{3}$ for all $n \geq N_3$. Therefore, $x < u$, so u is an upper bound of S.

Finally, we need to show that u is the least upper bound of S. Let $w = [\{w_n\}] < u$. There exists a rational $s > 0$ and $N_4 \in \mathbb{N}$ such that $u_n - w_n > s$ for all $n \geq N_4$. Choose $N_5, N_6 \in \mathbb{N}$ such that $|w_n - w_m| < \frac{s}{3}$ if $n, m \geq N_5$ and $|u_n - u_m| < \frac{s}{3}$ if $n, m \geq N_6$. Let $N_7 = \max\{N_5, N_6\}$. For any $n, m \geq N_7$, we have that

$$u_n - w_m = (u_n - w_n) + (w_n - w_m) > s - \frac{s}{3} = \frac{2s}{3}.$$

Write $s = \frac{p}{q}$ where $p, q \in \mathbb{N}$, and let $N_8 = \max\{N_7, 3q\}$. Note that $u_{N_8} - t_{N_8} < \frac{1}{N_8} \leq \frac{1}{3q} \leq \frac{s}{3}$. Hence,

$$t_{N_8} - w_n = (t_{N_8} - u_{N_8}) + (u_{N_8} - w_n) > -\frac{s}{3} + \frac{2s}{3} = \frac{s}{3}$$

for all $n \geq N_7$. It follows that $w < (t_{N_8})_{\mathbb{R}}$. But since $t_{N_8} \in T$, there exists $y \in S$ such that $(t_{N_8})_{\mathbb{R}} < y$. As a result, $w < y$, so w is not an upper bound of S. $\square$

Having performed this construction, we can now be satisfied that a complete ordered field exists. The fact that there is only one complete ordered field up to a unique isomorphism implies that *any* construction of a complete ordered field produces the same result as our construction. Therefore, in practice, we do not appeal to a specific construction of $\mathbb{R}$ when proving theorems about $\mathbb{R}$; all we need to know is that $\mathbb{R}$ is a complete ordered field.

2 Metric Spaces

The idea of "approaching" a real number in a limit is one of the fundamental concepts of calculus. However, to formally define what "approach" means, we need to have a concept of "distance" between two points. We know what "distance" means within Euclidean space $\mathbb{R}^n$, but we may want to take limits in other more abstract sets, like sets of functions. Moreover, the familiar Euclidean distance between two points in $\mathbb{R}^n$ is not the only sensible distance function we could define on $\mathbb{R}^n$. For instance, in $\mathbb{R}^2$, we could work with "taxicab distance" given by the distance function

$$d((x_1, x_2), (y_1, y_2)) = |x_1 - y_1| + |x_2 - y_2|.$$

A set equipped with a sensible distance function (which we will define in this chapter) is called a *metric space*. This chapter introduces metric spaces and some of the key concepts surrounding them, such as "neighbourhoods" and "limit points".

2.1 Countable and Uncountable Sets

Definition 2.1.

(a) If S is a finite set, then the *cardinality* of S, denoted $|S|$, is the number of elements in S.

(b) Two sets S and T (which may be infinite) have the same cardinality if there exists a bijection $f : S \to T$. In this case, we write $|S| = |T|$.

(c) If there exists an injection from S to T, we write $|S| \leq |T|$.

(d) If there exists an injection but no surjection from S to T, we write $|S| < |T|$.

Definition 2.2. A set S is *countable* if $|S| = |\mathbb{N}|$. If S is infinite and not countable, then S is *uncountable*. A set that is finite or countable is *denumerable*. A denumerable set is also called *at most countable*.

An *enumeration* of S is a (possibly infinite) list such that each element of S appears exactly once in the list at a finite position. For example,

$$\{0, 1, -1, 2, -2, 3, -3, \dots\}$$

gives an enumeration of $\mathbb{Z}$. However, the list

$$\{0, 1, 2, 3, \dots, -1, -2, -3, \dots\},$$

where all nonnegative integers are written before the negative integers, is not an enumeration of $\mathbb{Z}$ since -1 is not at a finite position—infinitely many integers appear before -1 in the list. Note that a bijection $f : \mathbb{N} \to S$ corresponds to an infinite enumeration $\{f(1), f(2), f(3), \dots\}$. On the other hand, an infinite enumeration $S = \{x_1, x_2, x_3, \dots\}$ induces a bijective map $n \mapsto x_n$ from $\mathbb{N}$ to S. Thus, S is countable if and only if S has an infinite enumeration. For example, we constructed an infinite enumeration of $\mathbb{Z}$ above, so $\mathbb{Z}$ is countable. More generally, S is denumerable if and only if S has an enumeration.

Theorem 2.3. *Let S be a set. If there exists an injection $f : S \to \mathbb{N}$, then S is denumerable.*

Proof. If S is finite, then S is denumerable, so suppose S is infinite. Let

$$T = \{f(x) \mid x \in S\} \subset \mathbb{N},$$

and note that T is infinite because $f : S \to T$ is a bijection and S is infinite. The Well-Ordering Principle implies that T has an m^{th} smallest element for all $m \in \mathbb{N}$. Define $g : \mathbb{N} \to S$ by setting $g(m)$ equal to the element $x \in S$ such that $f(x)$ is the m^{th} smallest element in T. If $g(i) = g(j)$, then $f(g(i)) = f(g(j))$ is the i^{th} smallest element in T as well as the j^{th} smallest element in T, so $i = j$. Hence, g is injective. For any $x \in S$, there must exist $m \in \mathbb{N}$ such that $f(x)$ is the m^{th} smallest element of T. Otherwise, there would be infinitely many elements of T smaller than $f(x)$, which is impossible since $f(x) \in \mathbb{N}$ cannot be larger than infinitely many natural numbers. Hence, $g(m) = x$ for some $m \in \mathbb{N}$, so g is surjective. Therefore, $|\mathbb{N}| = |S|$, so S is countable and hence denumerable. $\square$

Theorem 2.4. *Let E_n be a countable set for all $n \in \mathbb{N}$. Then $\bigcup_{n\in\mathbb{N}} E_n$ is countable.*

Proof. First, note that $\bigcup_{n\in\mathbb{N}} E_n$ is infinite since E_1 is infinite and $E_1 \subset \bigcup_{n\in\mathbb{N}} E_n$. It suffices to find an enumeration of $\bigcup_{n\in\mathbb{N}} E_n$. For each $n \in \mathbb{N}$, write $E_n = \{x_{n,1}, x_{n,2}, x_{n,3}, \dots\}$. We can write all the elements of $\bigcup_{n\in\mathbb{N}} E_n$ in an infinite grid like so:

$$\begin{matrix} x_{1,1} & x_{1,2} & x_{1,3} & \cdots \\ x_{2,1} & x_{2,2} & x_{2,3} & \cdots \\ x_{3,1} & x_{3,2} & x_{3,3} & \cdots \\ \vdots & \vdots & \vdots & \end{matrix}$$

For each integer $n \geq 2$, the elements $x_{i,j}$ such that $i + j = n$ form a diagonal D_n of this grid. Each diagonal has finitely many elements, and each element of $\bigcup_{n\in\mathbb{N}} E_n$ belongs in exactly one diagonal. Thus, we can list the elements of $\bigcup_{n\in\mathbb{N}} E_n$ by listing the elements in D_2, then the elements in D_3, and so on, making sure to skip all duplicate elements. Every element appears exactly once in this list because we skip duplicate elements. Also, every element appears at a finite position in the list because every diagonal is finite. Therefore, we have an enumeration of $\bigcup_{n\in\mathbb{N}} E_n$. $\square$

Corollary 2.4.1. *If S and T are countable, then $S \times T$ is countable.*

Proof. Write $S = \{s_1, s_2, s_3, \dots\}$ and $T = \{t_1, t_2, t_3, \dots\}$. Then

$$E_n = \{(s_n, t_j) \mid j \in \mathbb{N}\} = \{(s_n, t_1), (s_n, t_2), (s_n, t_3), \dots\}$$

is countable for each $n \in \mathbb{N}$. Hence, $S \times T = \bigcup_{n\in\mathbb{N}} E_n$ is countable by Theorem 2.4. $\square$

Corollary 2.4.2. $\mathbb{Q}$ *is countable.*

Proof. Every $x \in \mathbb{Q}$ has a unique "simplest form" $x = \frac{p}{q}$ where $p \in \mathbb{Z}$, $q \in \mathbb{N}$, and $\gcd(p, q) = 1$. Define $f : \mathbb{Q} \to \mathbb{Z} \times \mathbb{N}$ by $f(x) = (p, q)$ where $\frac{p}{q}$ is the simplest form of x. Note that f is injective because $f(x) = (p, q)$ if and only if $x = \frac{p}{q}$. Since $\mathbb{Z}$ and $\mathbb{N}$ are countable, $\mathbb{Z} \times \mathbb{N}$ is countable by Corollary 2.4.1, so there exists a bijection $g : \mathbb{Z} \times \mathbb{N} \to \mathbb{N}$. Then $g \circ f : \mathbb{Q} \to \mathbb{N}$ is an injection, so $\mathbb{Q}$ is denumerable by Theorem 2.3. Since $\mathbb{Q}$ is infinite, $\mathbb{Q}$ is countable. $\square$

Theorem 2.5 (Cantor)**.** *For any set S, we have $|S| < |\mathcal{P}(S)|$.*

Proof. Define $f : S \to \mathcal{P}(S)$ by $f(x) = \{x\}$. If $f(x) = f(y)$, then $\{x\} = \{y\}$, so $x = y$. Hence, f is injective.

Suppose there exists a surjection $g : S \to \mathcal{P}(S)$. Let $B = \{x \in S \mid x \notin g(x)\}$. Since g is surjective, $g(x_0) = B$ for some $x_0 \in S$. If $x_0 \in B$, then $x_0 \notin g(x_0) = B$, which is a contradiction. If $x_0 \notin B = g(x_0)$, then $x_0 \in B$ by definition of B, giving us a contradiction. Both cases give a contradiction, so g does not exist. $\square$

Theorem 2.6 (Cantor–Schröder–Bernstein)**.** *Let S and T be sets, and suppose there exist injective functions $f : S \to T$ and $g : T \to S$. Then there exists a bijection $h : S \to T$. In other words, if $|S| \le |T|$ and $|T| \le |S|$, then $|S| = |T|$.*

Proof. In this proof, we denote $(g \circ f)_n \coloneqq \underbrace{(g \circ f) \circ \cdots \circ (g \circ f)}_{n \text{ times}}$ and $(g \circ f)_0 \coloneqq \mathrm{Id}_S$. Let

$$V = \bigcup_{n=0}^{\infty} (g \circ f)_n(S \setminus g(T)).$$

We define $h : S \to T$ as follows. Fix $x \in S$. If $x \in V$, we set $h(x) = f(x)$. On the other hand, suppose $x \notin V$. Then $x \notin U_0 = S \setminus g(T)$, so $x \in g(T)$. Hence, $x = g(y)$ for some $y \in T$. Since g is injective, y is unique, so we can set $h(x) = y$. Written as a succinct formula, we have

$$h(x) = \begin{cases} f(x) & \text{if } x \in V \\ g^{-1}(x) & \text{if } x \notin V, \end{cases}$$

where $g^{-1}(x)$ denotes the unique element of T such that $g(g^{-1}(x)) = x$.

We now check that h is injective and surjective. Let $x_1, x_2 \in S$, and suppose $h(x_1) = h(x_2)$. If $x_1, x_2 \in V$, then $f(x_1) = h(x_1) = h(x_2) = f(x_2)$, so $x_1 = x_2$ because f is injective. If $x_1, x_2 \in S \setminus V$, then $x_1 = g(h(x_1)) = g(h(x_2)) = x_2$. Finally, suppose $x_1 \in V$ and $x_2 \notin V$. Then $x_2 = g(h(x_2)) = g(h(x_1)) = g(f(x_1))$. Since $x_1 \in V$, there exists $n \ge 0$ and $y \in S \setminus g(T)$ such that $x_1 = (g \circ f)_n(y)$. Then $x_2 = g(f(x_1)) = (g \circ f)_{n+1}(y) \in V$, which is a contradiction. Hence, x_1 and x_2 must both be in V or both be in $S \setminus V$, and we have seen that $x_1 = x_2$ in these cases. Therefore, h is injective.

Fix $y \in T$. If $y \in f(V)$, there exists $x \in V \subset S$ such that $y = f(x) = h(x)$. Suppose $y \notin f(V)$. We claim that $g(y) \notin V$. If $g(y) \in V$, then there exists $n \ge 0$ such that $g(y) \in (g \circ f)_n(S \setminus g(T))$. If $n = 0$, then $g(y) \in S \setminus g(T)$, which is a contradiction since $g(y) \in g(T)$. Hence, $n \ge 1$. Note that $(g \circ f)_n = g \circ [f \circ (g \circ f)_{n-1}]$, so there exists $z \in f \circ (g \circ f)_{n-1}(S \setminus g(T))$ such that $g(y) = g(z)$. Since g is injective, $y = z \in f \circ (g \circ f)_{n-1}(S \setminus g(T)) \subset f(V)$, which is a contradiction. Hence, $g(y) \notin V$, so $h(g(y)) = y$ by definition of h. Therefore, h is surjective. $\square$

We now apply the Cantor–Schröder–Bernstein Theorem to show that $|\mathcal{P}(\mathbb{N})| = |\mathbb{R}|$. We assume that the reader has at least an informal understanding of binary expansions and infinite series, and we omit rigorous discussion of these topics so as to focus on the application of Cantor–Schröder–Bernstein. The reader can supply the missing details after Chapter 3.

Lemma 2.7. $|(0,1)| = |\mathbb{R}|$.

Proof. Consider $f : \mathbb{R} \to (0,1)$ defined by

$$f(x) = \begin{cases} \frac{1}{2-2x} & \text{if } x < 0 \\ 1 - \frac{1}{2x+2} & \text{if } x \geq 0. \end{cases}$$

We first check that $0 < f(x) < 1$ for all $x \in \mathbb{R}$. If $x < 0$, then $2 - 2x > 2$, so $0 < \frac{1}{2-2x} = f(x) < \frac{1}{2}$. Suppose $x \geq 0$. Then $2x + 2 \geq 2$, so $0 < \frac{1}{2x+2} \leq \frac{1}{2}$. Hence, $-\frac{1}{2} \leq -\frac{1}{2x+2} < 0$, so $\frac{1}{2} \leq 1 - \frac{1}{2x+2} = f(x) < 1$.

We now show that f is bijective. Suppose $f(x_1) = f(x_2)$. Either $f(x_1) < \frac{1}{2}$ or $f(x_1) \geq \frac{1}{2}$. Assume first that $f(x_1) < \frac{1}{2}$. Then $x_1 < 0$ and $x_2 < 0$ because we showed above that $f(x) \geq \frac{1}{2}$ if $x \geq 0$. Hence, $\frac{1}{2-2x_1} = \frac{1}{2-2x_2}$, so $2 - 2x_2 = 2 - 2x_1$ and hence $x_1 = x_2$. Now assume that $f(x_1) \geq \frac{1}{2}$. Then $x_1 \geq 0$ and $x_2 \geq 0$ because $f(x) < 0$ if $x < 0$. Hence, $1 - \frac{1}{2x_1+2} = 1 - \frac{1}{2x_2+2}$, so $2x_2 + 2 = 2x_1 + 2$ and hence $x_1 = x_2$. Since $x_1 = x_2$ in both cases, f is injective.

Fix $y \in (0,1)$. Suppose $y < \frac{1}{2}$. Then $\frac{1}{y} > 2$. Let $x = 1 - \frac{1}{2y} < 0$. Then

$$f(x) = \frac{1}{1-2x} = \frac{1}{2 - 2(1 - \frac{1}{2y})} = \frac{1}{(\frac{1}{y})} = y.$$

Now suppose $y \geq \frac{1}{2}$. Then $0 < 1 - y \leq \frac{1}{2}$, so $\frac{1}{1-y} \geq 2$. Let $x = \frac{1}{2(1-y)} - 1 \geq 0$. Then

$$f(x) = 1 - \frac{1}{2x+2} = 1 - \frac{1}{2(\frac{1}{2(1-y)} - 1) + 2} = 1 - \frac{1}{(\frac{1}{1-y})} = 1 - (1-y) = y.$$

In both cases, we can find $x \in \mathbb{R}$ such that $f(x) = y$, so f is surjective. □

Theorem 2.8. $|\mathcal{P}(\mathbb{N})| = |\mathbb{R}|$.

Proof. It suffices to construct injections $f : (0,1) \to \mathcal{P}(\mathbb{N})$ and $g : \mathcal{P}(\mathbb{N}) \to (0,1)$. If we prove the existence of f and g, then by the Cantor–Schröder–Bernstein Theorem, $|\mathcal{P}(\mathbb{N})| = |(0,1)|$, and by Lemma 2.7, $|(0,1)| = |\mathbb{R}|$, so the desired result follows.

First, we construct f. Each $x \in (0,1)$ has at least one *binary expansion*, which is a sequence $\{a_n\}_{n \in \mathbb{N}}$ where $a_n \in \{0,1\}$ for each n and

$$x = \sum_{n=1}^{\infty} \frac{a_n}{2^n}.$$

The reader can prove this assertion later in Exercise 3.45.

Given $x \in (0,1)$, choose a binary expansion $\{a_n\}_{n \in \mathbb{N}}$ of x, and set

$$f(x) = \{n \in \mathbb{N} \mid a_n = 1\} \in \mathcal{P}(\mathbb{N}).$$

In this way, we assign each $x \in (0,1)$ to a unique value $f(x) \in \mathcal{P}(\mathbb{N})$.

Suppose $f(x) = f(y)$ where $x, y \in (0, 1)$. For each $n \in \mathbb{N}$, let

$$a_n = \begin{cases} 1 & \text{if } n \in f(x) \\ 0 & \text{otherwise.} \end{cases}$$

We call $\{a_n\}_{n \in \mathbb{N}}$ the *indicator sequence* of the set $f(x)$. We have that

$$x = \sum_{n=1}^{\infty} \frac{a_n}{2^n} = y,$$

so f is injective.

We turn to constructing g. For each $S \in \mathcal{P}(\mathbb{N})$, we form the indicator sequence $\{b_n\}_{n \in \mathbb{N}}$ of S defined by

$$b_n = \begin{cases} 1 & \text{if } n \in S \\ 0 & \text{otherwise,} \end{cases}$$

and we set

$$g(S) = \frac{1}{10} + \sum_{n=1}^{\infty} \frac{b_n}{10^n}.$$

Note that

$$0 < \frac{1}{10} \leq g(S) \leq \frac{1}{10} + \sum_{n=1}^{\infty} \frac{1}{10^n} = \frac{1}{10} + \frac{1}{9} < 1,$$

so g maps $\mathcal{P}(\mathbb{N})$ into $(0, 1)$. Suppose $g(S) = g(T)$ where $S, T \in \mathcal{P}(\mathbb{N})$. Let $\{b_n\}_{n \in \mathbb{N}}$ and $\{c_n\}_{n \in \mathbb{N}}$ be the indicator sequences of S and T, respectively. Then

$$\frac{1}{10} + \sum_{n=1}^{\infty} \frac{b_n}{10^n} = g(S) = g(T) = \frac{1}{10} + \sum_{n=1}^{\infty} \frac{c_n}{10^n},$$

so

$$0 = \sum_{n=1}^{\infty} \frac{b_n - c_n}{10^n}.$$

Suppose the sequences $\{b_n\}_{n \in \mathbb{N}}$ and $\{c_n\}_{n \in \mathbb{N}}$ are not equal. Then there exists a minimal $N \in \mathbb{N}$ such that $b_N \neq c_N$. Without loss of generality, suppose $b_N = 1$ and $c_N = 0$. Then

$$\frac{1}{10^N} = \frac{b_N - c_N}{10^N} = \sum_{n=N+1}^{\infty} \frac{c_n - b_n}{10^n} \leq \sum_{n=N+1}^{\infty} \frac{1}{10^n} = \frac{1}{9(10^N)},$$

so $1 \leq \frac{1}{9}$, which is a contradiction. Hence, $b_n = c_n$ for all $n \in \mathbb{N}$, so $S = T$. Therefore, g is injective. $\square$

Corollary 2.8.1. $\mathbb{R}$ *is uncountable.*

Proof. $|\mathbb{N}| < |\mathcal{P}(\mathbb{N})| = |\mathbb{R}|$. $\square$

Exercise 2.9.

(a) Prove that the set of finite integer sequences is countable.

(b) Prove that the set of infinite integer sequences is uncountable.

Exercise 2.10. Let A and B be sets. We define A^B as the set of all functions $B \to A$.

(a) Prove that if A and B are finite, then $|A^B| = |A|^{|B|}$.

(b) Suppose $|A| \geq 2$. Prove that $|A| < |A^A|$. Hint: use the same idea as the proof of Theorem 2.5 to show that a surjection $A \to A^A$ cannot exist.

(c) Let C be another set. Prove that $|(A^B)^C| = |A^{B \times C}|$.

(d) Suppose $|A_1| = |A_2|$ and $|B_1| = |B_2|$. Show that $|A_1^{B_1}| = |A_2^{B_2}|$.

Exercise 2.11 (Applications of Exercise 2.10)**.**

(a) Prove that $|\mathcal{P}(A)| = |\{0,1\}^A|$ for any set A.

(b) Prove that $|\mathbb{R} \times \mathbb{N}| = |\mathbb{R}|$. Hint: you may want to use the fact that $|\mathbb{R}| = |\mathcal{P}(\mathbb{N})|$.

(c) Prove that $|\mathbb{R}^\mathbb{R}| = |\mathcal{P}(\mathbb{R})|$.

2.2 Introduction to Metric Spaces

Definition 2.12. Let X be a set. A *metric* on X is a function $d : X \times X \to [0, \infty)$ such that for all $x, y, z \in X$,

(i) $d(x, y) = 0$ if and only if $x = y$,

(ii) $d(x, y) = d(y, x)$, and

(iii) $d(x, z) \leq d(x, y) + d(y, z)$.

The pair (X, d) is called a *metric space.*

Remark.

- Condition (iii) in the above definition is called the *Triangle Inequality.*
- For brevity, we often say "Let X be a metric space" instead of "Let (X, d) be a metric space." The metric on X is always denoted d or d_X, and the subscript is only necessary if we are working with multiple metric spaces at the same time.

Proposition 2.13 (Reverse Triangle Inequality)**.** *Let X be a metric space. Then $d(x, y) \geq |d(x, z) - d(y, z)|$ for all $x, y, z \in X$.*

Proof. By the Triangle Inequality,

$$d(x, z) \leq d(x, y) + d(y, z)$$

and

$$d(y, z) \leq d(y, x) + d(x, z) = d(x, y) + d(x, z).$$

Hence,

$$d(x, y) \geq d(x, z) - d(y, z)$$

and

$$d(x, y) \geq d(y, z) - d(x, z),$$

so $d(x, y) \geq |d(x, z) - d(y, z)|$ since $|d(x, z) - d(y, z)|$ is equal to one of $d(x, z) - d(y, z)$ or $d(y, z) - d(x, z)$. $\square$

Proposition 2.14 (The Euclidean Metric)**.** *Let $k \in \mathbb{N}$. Then $d : \mathbb{R}^k \times \mathbb{R}^k \to [0, \infty)$ defined by*

$$d(\vec{x}, \vec{y}) = \left(\sum_{i=1}^{k} (x_i - y_i)^2 \right)^{1/2}$$

is a metric, where x_i and y_i refer to the i^{th} component of $\vec{x}$ and $\vec{y}$, respectively.

Proof. Let $\vec{x}, \vec{y}, \vec{z} \in \mathbb{R}^k$. Then

$$d(\vec{x}, \vec{x}) = \left(\sum_{i=1}^{k} (x_i - x_i)^2 \right)^{1/2} = 0.$$

If $\vec{x} \neq \vec{y}$, then $x_j \neq y_j$ for some $1 \leq j \leq k$, so

$$d(\vec{x}, \vec{y}) = \left(\sum_{i=1}^{k} (x_i - y_i)^2 \right)^{1/2} \geq (x_j - y_j)^2 > 0.$$

Hence, condition (i) is satisfied. Condition (ii) also holds because $(x_i - y_i)^2 = (y_i - x_i)^2$ for all $1 \leq i \leq k$.

Now we prove the Triangle Inequality. We compute that

$$\begin{aligned}
d(\vec{x}, \vec{z})^2 &= \sum_{i=1}^{k} (x_i - z_i)^2 \\
&= \sum_{i=1}^{k} (x_i - y_i + y_i - z_i)^2 \\
&= \sum_{i=1}^{k} (x_i - y_i)^2 + 2 \sum_{i=1}^{k} (x_i - y_i)(y_i - z_i) + \sum_{i=1}^{k} (y_i - z_i)^2 \\
&\leq \sum_{i=1}^{k} (x_i - y_i)^2 + 2 \left| \sum_{i=1}^{k} (x_i - y_i)(y_i - z_i) \right| + \sum_{i=1}^{k} (y_i - z_i)^2 \\
&\leq \sum_{i=1}^{k} (x_i - y_i)^2 + 2 \left(\sum_{i=1}^{k} (x_i - y_i)^2 \right)^{1/2} \left(\sum_{i=1}^{k} (y_i - z_i)^2 \right)^{1/2} + \sum_{i=1}^{k} (y_i - z_i)^2 \\
&\quad \text{(by the Cauchy–Schwarz Inequality)} \\
&= \left(\left(\sum_{i=1}^{k} (x_i - y_i)^2 \right)^{1/2} + \left(\sum_{i=1}^{k} (y_i - z_i)^2 \right)^{1/2} \right)^2 \\
&= (d(\vec{x}, \vec{y}) + d(\vec{y}, \vec{z}))^2.
\end{aligned}$$

Therefore, $d(\vec{x}, \vec{z}) = |d(\vec{x}, \vec{z})| \leq |d(\vec{x}, \vec{y}) + d(\vec{y}, \vec{z})| = d(\vec{x}, \vec{y}) + d(\vec{y}, \vec{z})$ by parts (c) and (d) of Proposition 1.11. $\square$

From now on, unless otherwise specified, we will use the Euclidean metric when working in $\mathbb{R}^k$. Note that the Euclidean metric on $\mathbb{R}$ in particular satisfies

$$d_{\mathbb{R}}(x, y) = \sqrt{(x-y)^2} = |x - y|$$

for all $x, y \in \mathbb{R}$. A similar fact holds for $\mathbb{C}$. Let $z = x_1 + y_1 i \in \mathbb{C}$ and $w = x_2 + y_2 i \in \mathbb{C}$ where $x_1, y_1, x_2, y_2 \in \mathbb{R}$. By definition, $z = (x_1, y_1) \in \mathbb{R}^2$ and $w = (x_2, y_2) \in \mathbb{R}^2$. Hence,

$$d_{\mathbb{C}}(z, w) = d_{\mathbb{R}^2}((x_1, y_1), (x_2, y_2)) = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2} = |z - w|.$$

For the rest of this chapter, let X be a metric space with metric d unless otherwise specified.

Proposition 2.15. *Let $Y \subset X$. Then d restricted to $Y \times Y$ is a metric on Y. Thus, any subset of a metric space is itself a metric space.*

Proof. The three conditions for d to be a metric only involve universal quantifiers on X, so the conditions hold if X is replaced with a subset Y. If a statement holds for all $x \in X$, then the statement holds for all $x \in Y$. $\square$

Definition 2.16. Let $E \subset X$ and $x \in X$.

(a) Every element in E is called a *point* of E.

(b) A *neighbourhood* of x is a set $N_r(x) := \{y \in X \mid d(x, y) < r\}$ where $r > 0$ is a real number. We call r the *radius* of the neighbourhood.

(c) We call x an *interior point* of E if there exists $r > 0$ such that $N_r(x) \subset E$.

(d) We call x a *limit point* of E if for all $r > 0$, there exists $y \in E$ such that $d(x, y) < r$ and $y \neq x$.

(e) We call x an *isolated point* of E if there exists $r > 0$ such that $N_r(x) \cap E = \{x\}$.

Proposition 2.17. *Let $E \subset X$. Then every point of E is either an isolated point or a limit point of E, but not both.*

Proof. Fix $x \in E$. Suppose x is not an isolated point of E. Then given $r > 0$, we know that $N_r(x) \cap E \neq \{x\}$. Clearly, $\{x\} \subset N_r(x) \cap E$, so there must exist $y \in N_r(x) \cap E$ not equal to x. Equivalently, there exists $y \in E$ such that $d(x, y) < r$ and $y \neq x$. Hence, x is a limit point of E.

On the other hand, suppose x is an isolated point of E. Then there exists $r > 0$ such that $N_r(x) \cap E = \{x\}$. There cannot exist $y \in E$ such that $y \neq x$ and $d(x, y) < r$, for this would imply that $\{x, y\} \subset N_r(x) \cap E = \{x\}$, which is impossible. Hence, x is not a limit point of E. $\square$

Definition 2.18. Let $E \subset X$.

(a) E is *open* if every point of E is an interior point of E.

(b) The set of limit points of E is denoted E'.

(c) E is *closed* if every limit point of E is in E (i.e. if $E' \subset E$).

(d) The *complement* of E is $E^c := X \setminus E$.

(e) The *closure* of E is $\overline{E} := E \cup E'$.

(f) The *interior* of E is the set of interior points of E and is denoted E°.

(g) The *boundary* of E is $\partial E := \overline{E} \setminus E^\circ$.
(h) E is *dense* if $\overline{E} = X$.
(i) E is *bounded* if E is empty or if there exists $x \in E$ and $r > 0$ such that $E \subset N_r(x)$.
(j) E is *perfect* if E is closed and every point in E is a limit point of E.

The space $\mathbb{R}^2$ offers many examples of the kinds of sets described in Definition 2.18. When we prove general statements about metric spaces, it is often helpful to consider the special case $\mathbb{R}^2$ first, then generalize to arbitrary metric spaces.

Example 2.19. For any $r > 0$, $N_r(0) = \{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 < r^2\}$ is the open disk of radius r centred at the origin. As its name suggests, $N_r(0)$ is open. The closure of $N_r(0)$ is $\overline{N_r(0)} = \{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 \leq r^2\}$, which is the closed disk of radius r centred at the origin. Every point of $\overline{N_r(0)}$ is a limit point of $\overline{N_r(0)}$. Since $\overline{N_r(0)}$ is closed, $\overline{N_r(0)}$ is perfect. The boundary of $N_r(0)$ is the circle $\{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 = r^2\}$.

Example 2.20. Let $S = \{(\frac{1}{n}, \frac{1}{n}) \in \mathbb{R}^2 \mid n \in \mathbb{N}\}$. Then S is neither open nor closed. The interior of S is empty, and $(0, 0)$ is a limit point of S that is not in S. Every point of S is an isolated point. The closure of S is $\overline{S} = S \cup \{(0, 0)\}$.

Example 2.21. Let $H = \{(x, y) \in \mathbb{R}^2 \mid y \geq 0\}$. Then H is perfect and unbounded. The interior of H is $H^\circ = \{(x, y) \in \mathbb{R}^2 \mid y > 0\}$. Since $H \neq H^\circ$, we see that H is not open.

Example 2.22. Let $L = \{(n, 0) \in \mathbb{R}^2 \mid n \in \mathbb{Z}\}$. Then L is closed since L has no limit points. Indeed, the requirement for L to be closed is that $L' \subset L$, and the empty set is a subset of any set.

Proposition 2.23. *Let $E \subset X$. Then E is open if and only if E^c is closed.*

Proof. Suppose E is open. Let $x \in X$ be a limit point of E^c. Suppose $x \in E$. Then there exists $r > 0$ such that $N_r(x) \subset E$. But since x is a limit point of E^c, there exists $y \in E^c$ such that $y \in N_r(x)$. Hence, $N_r(x) \not\subset E$, which is a contradiction. Therefore, $x \notin E$, so $x \in E^c$. Since E^c contains all of its limit points, E^c is closed.

Conversely, suppose E^c is closed. Let $x \in E$. Then x is not a limit point of E^c, so there exists $r > 0$ such that no $y \in E^c$ satisfies $0 < d(x, y) < r$. Let $z \in N_r(x)$. If $z = x$, then clearly $z \in E$. If $z \neq x$, then $0 < d(x, z) < r$, so $z \in E$ because $z \notin E^c$. In both cases, we have $z \in E$, so $N_r(x) \subset E$. Therefore, E is open. □

Proposition 2.24. *Let $r > 0$ and $x \in X$. Then $N_r(x) \subset X$ is open.*

Proof. Let $y \in N_r(x)$ and $p = r - d(x, y) > 0$. Fix $z \in N_p(y)$. Then

$$d(x, z) \leq d(x, y) + d(y, z) < d(x, y) + p = r,$$

so $z \in N_r(x)$. Therefore, $N_p(y) \subset N_r(x)$, so $N_r(x)$ is open. □

Proposition 2.25. *Let $\mathcal{O}$ be a set of open subsets of X, and let $\mathcal{C}$ be a set of closed subsets of X.*
(a) $\bigcup_{E \in \mathcal{O}} E \subset X$ *is open.*
(b) $\bigcap_{E \in \mathcal{C}} E$ *is closed.*
(c) If $\mathcal{O}$ is finite, then $\bigcap_{E \in \mathcal{O}} E \subset X$ *is open.*
(d) If $\mathcal{C}$ is finite, then $\bigcup_{E \in \mathcal{C}} E \subset X$ *is closed.*

Proof. (a) Let $x \in \bigcup_{E \in \mathcal{O}} E$. Then there exists $E_x \in \mathcal{O}$ such that $x \in E_x$. Since E_x is open, there exists $r > 0$ such that $N_r(x) \subset E_x \subset \bigcup_{E \in \mathcal{O}} E$. Therefore, $\bigcup_{E \in \mathcal{O}} E$ is open.

(b) Note that $(\bigcap_{E \in \mathcal{C}} E)^c = \bigcup_{E \in \mathcal{C}} E^c$ is an arbitrary union of open sets because the complement of a closed set is open. Hence, $(\bigcap_{E \in \mathcal{C}} E)^c$ is open by part (a), so $\bigcap_{E \in \mathcal{C}} E = ((\bigcap_{E \in \mathcal{C}} E)^c)^c$ is closed.

(c) Suppose $\mathcal{O}$ is finite. Write $\mathcal{O} = \{E_1, \ldots, E_n\}$ where $n = |\mathcal{O}| \geq 0$. Fix $x \in \bigcap_{E \in \mathcal{O}} E$. Then for all $1 \leq i \leq n$, there exists $r_i > 0$ such that $N_{r_i}(x) \subset E_i$. Let $r = \min_{1 \leq i \leq n} r_i > 0$. Then $N_r(x) \subset N_{r_i}(x) \subset E_i$ for all $1 \leq i \leq n$, so $N_r(x) \subset \bigcap_{i=1}^n E_i = \bigcap_{E \in \mathcal{O}} E$. Therefore, $\bigcap_{E \in \mathcal{O}} E$ is open.

(d) Suppose $\mathcal{C}$ is finite. Then $(\bigcup_{E \in \mathcal{C}} E)^c = \bigcap_{E \in \mathcal{C}} E^c$ is a finite intersection of open sets and is therefore open by part (c). Hence, the complement $\bigcup_{E \in \mathcal{C}} E$ is closed. □

Proposition 2.26. *Let $E \subset X$.*

(a) $E^\circ \subset X$ is open.

(b) $\overline{E} \subset X$ is closed.

(c) If E is open, then $E^\circ = E$.

(d) If E is closed, then $\overline{E} = E$.

(e) If $E \subset F \subset X$, then $E' \subset F'$. Moreover, if F is closed, then $\overline{E} \subset F$.

Proof. (a) Let $x \in E^\circ$. Then there exists $r > 0$ such that $N_r(x) \subset E$. For any $y \in N_r(x)$, there exists $s > 0$ such that $N_s(y) \subset N_r(x) \subset E$, so $y \in E^\circ$. Therefore, $N_r(x) \subset E^\circ$, so E° is open.

(b) Let $x \in X$ be a limit point of $\overline{E}$. We want to show that $x \in \overline{E}$. If $x \in E$, then we are done since $E \subset \overline{E}$. Suppose $x \notin E$, and fix $r > 0$. We want to show that $x \in E'$, which means we must prove that there exists $y \in E$ satisfying $0 < d(x, y) < r$. Since x is a limit point of $\overline{E}$, there exists $z \in \overline{E}$ such that $0 < d(x, z) < r$. If $z \in E$, then we can take $y = z$. Otherwise, $z \in E'$, so there exists $y \in E$ such that $0 < d(z, y) < \min\{d(x, z), r - d(x, z)\}$. Then

$$d(x, y) \geq d(x, z) - d(y, z) > d(x, z) - d(x, z) = 0$$

and

$$d(x, y) \leq d(x, z) + d(z, y) < d(x, z) + r - d(x, z) = r,$$

so $0 < d(x, y) < r$. In both cases, we have found $y \in E$ such that $0 < d(x, y) < r$, so $x \in E' \subset \overline{E}$.

(c) Suppose E is open. Let $x \in E^\circ$. Then there exists $r > 0$ such that $N_r(x) \subset E$. Since $x \in N_r(x)$, we have that $x \in E$, so $E^\circ \subset E$. On the other hand, every point of E is an interior point by definition of E being open, so $E \subset E^\circ$.

(d) Suppose E is closed. Then E contains all of its limit points, so $E' \subset E$. Therefore, $\overline{E} = E \cup E' = E$.

(e) Let $E \subset F \subset X$. Suppose $x \in E'$. Then for any $r > 0$, there exists $y \in E \subset F$ such that $0 < d(x, y) < r$. Hence, $x \in F'$, so $E' \subset F'$. Suppose F is also closed. Then since $E \subset F$ and $E' \subset F'$, we have that $\overline{E} = E \cup E' \subset F \cup F' = \overline{F}$. By part (d), $\overline{F} = F$, so $\overline{E} \subset F$. □

Proposition 2.27. *Let $x, y \in \mathbb{R}$ such that $x < y$. Then $(x, y) \subset \mathbb{R}$ is open.*

Proof. Let $z = \frac{x+y}{2}$ and $r = \frac{y-x}{2}$. Then

$$N_r(z) = (z - r, z + r) = (x, y),$$

so (x, y) is open since $N_r(z)$ is open. $\square$

Proposition 2.28. *Let $x, y \in \mathbb{R}$ such that $x < y$. Then $\overline{(x, y)} = [x, y]$. Hence, $[x, y]$ is closed in $\mathbb{R}$.*

Proof. Fix $\epsilon > 0$, and let $z = \min\{\frac{y-x}{2}, \frac{\epsilon}{2}\}$. Then $x + z \in (x, y)$ because

$$x < x + z \leq x + \frac{y - x}{2} = \frac{x + y}{2} < y.$$

Also, $0 < |(x + z) - x| = z \leq \frac{\epsilon}{2} < \epsilon$. Hence, x is a limit point of (x, y). By a similar argument, y is a limit point of (x, y). Therefore, $[x, y] \subset \overline{(x, y)}$.

It now suffices to prove that $[x, y]$ is closed, or equivalently, that $[x, y]^c = (-\infty, x) \cup (y, \infty)$ is open. Let $v \in (-\infty, x) \cup (y, \infty)$. Suppose $v < x$. Let $r = \frac{x-v}{2} > 0$, and fix $w \in N_r(v)$. Then $w - v \leq |w - v| < r$, so $w < r + v = \frac{x+v}{2} < x$, which means that $w \in (-\infty, x)$. Therefore, $N_r(v) \subset (-\infty, x) \cup (y, \infty)$. If $v > y$, then we let $r = \frac{v-y}{2} > 0$ and argue similarly to show that $N_r(v) \subset (-\infty, x) \cup (y, \infty)$. Hence, $(-\infty, x) \cup (y, \infty)$ is open. $\square$

Theorem 2.29. *Suppose $E \subset X$ has a limit point $x \in X$. Then there is an infinite subset $S \subset E$ such that for all $r > 0$, all but finitely many elements of S are in $N_r(x) \setminus \{x\}$.*

Proof. Let $i \geq 1$ be any integer. Since x is a limit point of E, there exists $y_i \in E$ such that $y_i \in N_{1/i}(x)$ and $y_i \neq x$. Let $S = \{y_n \mid n \in \mathbb{N}\}$, and fix $r > 0$. By the Archimedean Property, there exists $m \in \mathbb{N}$ such that $m < \frac{1}{r}$, so that $\frac{1}{m} < r$. Then $r_m \leq \frac{1}{n} < r$. Now for any $k \geq m$, we have that $0 < d(x, y_k) < \frac{1}{k} \leq \frac{1}{m} < r$. Therefore, the finite set $\{y_1, \ldots, y_{m-1}\}$ contains all points of S that are outside $N_r(x) \setminus \{x\}$.

To prove that S is infinite, suppose S is finite. Then we can define $p = \min_{y \in S} d(x, y) > 0$. But then there is no $y \in S$ such that $y \in N_{p/2}(x)$, which contradicts what we proved in the previous paragraph. $\square$

Corollary 2.29.1. *Suppose $E \subset X$ has a limit point $x \in X$. Then $E \cap N_r(x)$ is infinite for any $r > 0$.*

Proof. Let $S \subset E$ be an infinite subset such that only finitely many points of S are outside $N_r(x) \setminus \{x\}$ for any $r > 0$. Then $S \cap (N_r(x) \setminus \{x\})$ must be infinite for any $r > 0$. Since $S \cap (N_r(x) \setminus \{x\}) \subset S \cap N_r(x) \subset E \cap N_r(x)$, the result follows. $\square$

Proposition 2.30. *Let $E \subset X$. Then E' is closed.*

Proof. We will prove that $(E')^c$ is open. Let $x \in (E')^c$. Then there exists $r > 0$ such that $N_r(x) \setminus \{x\} \subset E^c$. Fix $y \in N_r(x)$. If $y = x$, then $y \notin E'$. Now suppose $y \neq x$. Let $s = \min\{d(x, y), r - d(x, y)\} > 0$. If $z \in N_s(y)$, then

$$d(x, z) \geq d(x, y) - d(z, y) > d(x, y) - s \geq d(x, y) - d(x, y) = 0$$

and

$$d(x, z) \leq d(x, y) + d(y, z) < d(x, y) + s \leq d(x, y) + (r - d(x, y)) = r,$$

so $z \in N_r(x) \setminus \{x\} \subset E^c$. Therefore, $N_s(y) \subset E^c$, so $y \notin E'$. It follows that $N_r(x) \subset (E')^c$, so $(E')^c$ is open. $\square$

Theorem 2.31. *Let $E \subset Y \subset X$. Then E is open relative to Y if and only if there exists $F \subset X$ that is open relative to X and $F \cap Y = E$.*

Proof. For all $y \in E$ and $r > 0$, let $N_r^Y(y) := \{z \in Y \mid d_X(y,z) < r\}$ and $N_r^X(y) := \{z \in X \mid d_X(y,z) < r\}$. Note that $N_r^Y(y) = N_r^X(x) \cap Y$.

Suppose E is open relative to Y. Then for all $y \in E$, there exists $r_y > 0$ such that $N_{r_y}^Y(y) \subset E$. Let $F = \bigcup_{y \in E} N_{r_y}^X(y)$, which is open relative to X. Then

$$F \cap Y = \left(\bigcup_{y \in E} N_{r_y}^X(y) \right) \cap Y = \bigcup_{y \in E} (N_{r_y}^X(y) \cap Y) = \bigcup_{y \in E} N_{r_y}^Y(y) = E,$$

where the last equality follows because $N_{r_y}^Y(y) \subset E$ for all $y \in E$.

Conversely, suppose there is an open subset F of X such that $F \cap Y = E$. Let $y \in E$. Then $y \in F$, so there exists $r_y > 0$ such that $N_{r_y}^X(y) \subset F$. Therefore,

$$N_{r_y}^Y(y) = N_{r_y}^X(y) \cap Y \subset F \cap Y = E,$$

so E is open relative to Y. ☐

Exercise 2.32. Let $(X_1, d_1), \dots, (X_n, d_n)$ be metric spaces. For each $x = (x_1, \dots, x_n) \in X_1 \times \cdots \times X_n$ and $y = (y_1, \dots, y_n) \in X_1 \times \cdots \times X_n$, set

$$d(x,y) = \sqrt{d_1(x_1,y_1)^2 + \cdots + d_n(x_n,y_n)^2}.$$

Prove that d is a metric on $X_1 \times \cdots \times X_n$.

Exercise 2.33. Let $E \subset X$.

(a) Prove that $(\overline{E})^c = (E^c)^\circ$ and that $(E^\circ)^c = \overline{(E^c)}$.

(b) We say that E is *nowhere dense* if the interior of $\overline{E}$ is empty. Prove that E is nowhere dense if and only if the interior of E^c is dense (i.e. $\overline{(E^c)^\circ} = X$).

Exercise 2.34. Prove that every open subset of $\mathbb{R}$ can be written as a countable union of disjoint open intervals.

Exercise 2.35. Let (X, d) be a metric space and $\tau(X, d)$ be the set of open subsets of X under the metric d. Let T be the set consisting of $\mathbb{R}$, the empty set, and all open intervals $(-x, x)$ where $x > 0$. Prove that there is no metric d on $\mathbb{R}$ such that $T = \tau(\mathbb{R}, d)$.

Exercise 2.36. Let $M_n(\mathbb{R})$ be the set of $n \times n$ matrices with real entries.

(a) For any $A \in M_n(\mathbb{R})$, prove that

$$||A|| := \sup_{||x||=1} ||Ax||$$

is finite. We call $||A||$ the *operator norm* of A. Prove that $||Ax|| \leq ||A|| \cdot ||x||$ for all $x \in \mathbb{R}^n$.

(b) Let $d(T_1, T_2) = ||T_1 - T_2||$ for each $T_1, T_2 \in M_n(\mathbb{R})$. Prove that d is a metric on $M_n(\mathbb{R})$.

(c) Let $\mathrm{GL}_n(\mathbb{R})$ be the set of invertible matrices in $M_n(\mathbb{R})$. Prove that $\mathrm{GL}_n(\mathbb{R})$ is dense in $M_n(\mathbb{R})$. Hint: fix $A \in M_n(\mathbb{R})$, and consider $A + \lambda I_n$ for small λ. Recall that $B \in M_n(\mathbb{R})$ is non-invertible if and only if the characteristic polynomial $c_B(x) = \det(B - xI_n)$ has a root at $x = 0$.

(d) Prove that if $T \in \mathrm{GL}_n(\mathbb{R})$, then

$$||Tx|| \geq \frac{1}{||T^{-1}||}$$

for all $||x|| = 1$. Use this fact to prove that $\mathrm{GL}_n(\mathbb{R})$ is open in $M_n(\mathbb{R})$.

2.3 Compactness

Definition 2.37. Let $E \subset X$. A set $\{\mathcal{O}_\alpha\}_\alpha$ of open subsets of X is called an *open cover* of E (relative to X) if $E \subset \bigcup_\alpha \mathcal{O}_\alpha$. If there is a finite subset $\{\mathcal{O}_{\alpha_1}, \ldots, \mathcal{O}_{\alpha_n}\}$ of $\{\mathcal{O}_\alpha\}_\alpha$ that covers E, we say that $\{\mathcal{O}_{\alpha_1}, \ldots, \mathcal{O}_{\alpha_n}\}$ is a *finite subcover* of $\{\mathcal{O}_\alpha\}$.

Definition 2.38. $K \subset X$ is *compact* (relative to X) if every open cover of K has a finite subcover.

It is easier to work with finite sets than infinite sets. Compactness allows us to extract a finite subset E from an infinite set K, then use E to learn about K. One of the main challenges of working with compact sets is identifying a useful open cover to consider. If we choose the "correct" open cover, we can draw powerful conclusions from the associated finite subcover.

A common choice of an open cover for a compact set K is the set of neighbourhoods $\{N_r(x) : x \in K\}$ for some fixed $r > 0$. Many proofs proceed by forming this open cover, extracting a finite subcover $\{N_r(x_1), \ldots, N_r(x_n)\}$, and doing something with each x_i using the fact that there are only finitely many of them. Typically, we compute a value $f(x_i)$ for each i and then work with $\min_{1\leq i\leq n} f(x_i)$ or $\max_{1\leq i\leq n} f(x_i)$. More generally, if each $x \in K$ is associated with some $r_x > 0$, then $\{N_{r_x}(x) : x \in K\}$ is an open cover, and we can obtain a finite subcover $\{N_{r_{x_i}}(x_i)\}_{i=1}^n$. It is often useful to let $r = \min_{1\leq i\leq n} r_{x_i}$ and note that $r > 0$ since there are only finitely many radii r_{x_i}.

Theorem 2.39. *Let $K \subset Y \subset X$. Then K is compact relative to X if and only if K is compact relative to Y.*

Proof. Suppose K is compact relative to X. Let $\{\mathcal{O}_\alpha^Y\}_\alpha$ be an open cover of K relative to Y. By Theorem 2.31, for each $\mathcal{O}_\alpha^Y$, there exists an open set $\mathcal{O}_\alpha^X$ relative to X such that $\mathcal{O}_\alpha^X \cap Y = \mathcal{O}_\alpha^Y$. Then $K \subset \bigcup_\alpha \mathcal{O}_\alpha^Y \subset \bigcup_\alpha \mathcal{O}_\alpha^X$, so $\{\mathcal{O}_\alpha^X\}$ is an open cover of K relative to X. Since K is compact relative to X, there is a finite subcover $\{\mathcal{O}_{\alpha_1}^X, \ldots, \mathcal{O}_{\alpha_n}^X\}$. Now

$$K \subset \left(\bigcup_{j=1}^n \mathcal{O}_{\alpha_j}^X\right) \cap Y = \bigcup_{j=1}^n (\mathcal{O}_{\alpha_j}^X \cap Y) = \bigcup_{j=1}^n \mathcal{O}_\alpha^Y,$$

so $\{\mathcal{O}_{\alpha_1}^Y, \ldots, \mathcal{O}_{\alpha_n}^Y\}$ is a finite subcover of $\{\mathcal{O}_\alpha^Y\}_\alpha$. Therefore, K is compact relative to Y.

Conversely, suppose K is compact relative to Y. Let $\{\mathcal{O}_\alpha\}_\alpha$ be an open cover of K relative to X. For each index α, $\mathcal{O}_\alpha \cap Y$ is open relative to Y by Theorem 2.31. Hence, $\{\mathcal{O}_\alpha \cap Y\}_\alpha$ is an open cover of K relative to Y because

$$K \subset \left(\bigcup_\alpha \mathcal{O}_\alpha\right) \cap Y = \bigcup_\alpha (\mathcal{O}_\alpha \cap Y).$$

By compactness of K in Y, there exists a finite subcover $\{\mathcal{O}_{\alpha_1} \cap Y, \ldots, \mathcal{O}_{\alpha_n} \cap Y\}$. Therefore, $\{\mathcal{O}_{\alpha_1}, \ldots, \mathcal{O}_{\alpha_n}\}$ is a finite subcover of $\{\mathcal{O}_\alpha\}_\alpha$ because $K \subset \bigcup_{j=1}^n (\mathcal{O}_{\alpha_j} \cap Y) \subset \bigcup_{j=1}^n \mathcal{O}_{\alpha_j}$. Hence, K is compact relative to X. □

Theorem 2.39 says that the compactness of a set K depends only on the metric defined on K; we do not need to say that K is compact "relative to" some secondary metric space X. A metric space is either compact or not compact. However, we should note that a set may be compact in one metric but not in another. For example, $[0, 1]$ is compact when equipped with the Euclidean metric (as a consequence of the Heine–Borel Theorem, which we will see soon), but $[0, 1]$ is not compact when equipped with the "discrete metric"

$$d(x, y) = \begin{cases} 1 & \text{if } x \neq y \\ 0 & \text{if } x = y. \end{cases}$$

Proposition 2.40. *Every finite subset of a metric space is compact.*

Proof. Let $K \subset X$ be finite, and let $\{\mathcal{O}_\alpha\}_\alpha$ be an open cover of K. Write $K = \{x_1, \ldots, x_n\}$ where $n = |K| \geq 0$. For every $1 \leq j \leq k$, there is an open set $\mathcal{O}_{\alpha_j}$ such that $x_j \in \mathcal{O}_{\alpha_j}$. Then $\{\mathcal{O}_{\alpha_1}, \ldots, \mathcal{O}_{\alpha_n}\}$ is a finite subcover. □

Theorem 2.41. *If K is compact, then K is bounded.*

Proof. Suppose $K \subset X$ is compact. If K is empty, then we are done, so suppose K is non-empty. Since $\{N_1(x)\}_{x \in K}$ is an open cover of K, there exists a finite subcover $\{N_1(x_1), \ldots, N_1(x_n)\}$ where $n \geq 1$. Let $r = 1 + \max_{1 \leq i,j \leq n} d(x_i, x_j) > 0$, and fix $x \in X$. Then there exists $1 \leq m \leq n$ such that $x \in N_1(x_m)$, which means that $d(x_m, x) < 1$. Therefore,

$$d(x_1, x) \leq d(x_1, x_m) + d(x_m, x) < \max_{1 \leq i,j \leq n} d(x_i, x_j) + 1 = r,$$

so $K \subset N_r(x_1)$. Therefore, K is bounded. □

Theorem 2.42. *If K is compact, then K is closed.*

Proof. Let $K \subset X$ be compact. We can prove that K is closed by proving that K^c is open. Let $x \in K^c$. For each $y \in K$, let $r_y = \frac{1}{2} d(x, y) > 0$. Then $\{N_{r_y}(y)\}_{y \in K}$ is an open cover of K. Since K is compact, we can extract a finite subcover $\{N_{r_{y_1}}(y_1), \ldots, N_{r_{y_n}}(y_n)\}$. Let $r = \min\{r_{y_1}, \ldots, r_{y_n}\} > 0$. Suppose $z \in N_r(x)$. Then for each $1 \leq j \leq n$, we have that

$$d(y_j, z) \geq d(y_j, x) - d(z, x) > 2r_j - r \geq 2r_j - r_j = r_j,$$

so $z \notin N_{r_j}(y_j)$. Hence, $z \notin \bigcup_{j=1}^n N_{r_j}(y_j)$, so $z \notin K$ since $K \subset \bigcup_{j=1}^n N_{r_j}(y_j)$. We conclude that $N_r(x) \subset K^c$, so K^c is open. □

Theorem 2.43. *If K is compact and $E \subset K$ is closed (relative to K), then E is compact.*

Proof. Let $\{\mathcal{O}_\alpha\}_\alpha$ be an open cover of E. Since E is closed, E^c is open. Hence, $\{\mathcal{O}_\alpha\}_\alpha \cup \{E^c\}$ is an open cover of K. Since K is compact, we can extract a finite subcover $S = \{S_1, \ldots, S_n\}$. Then $S \setminus \{E^c\}$ is a finite subcover of $\{\mathcal{O}_\alpha\}_\alpha$. Therefore, E is compact. $\square$

Corollary 2.43.1. *Let $\mathcal{F}$ be a family of compact sets. Then $\bigcap_{K \in \mathcal{F}} K$ is compact.*

Proof. If $\mathcal{F}$ is empty, then the intersection $\bigcap_{K \in \mathcal{F}} K$ is empty and hence compact. Suppose $\mathcal{F}$ is not empty. Each $K \in \mathcal{F}$ is closed, so $\bigcap_{K \in \mathcal{F}} K$ is closed. Also, $\bigcap_{K \in \mathcal{F}} K$ is a subset of any compact set in $\mathcal{F}$. Hence, $\bigcap_{K \in \mathcal{F}} K$ is compact by Theorem 2.43. $\square$

Theorem 2.44. *If K is compact and $E \subset K$ is infinite, then E has a limit point in K.*

Proof. We prove the contrapositive. Suppose $E \subset K$ is infinite and has no limit point in K. For each $x \in K$, there exists $r_x > 0$ such that $N_{r_x}(x) \cap E \subset \{x\}$ because x is not a limit point of E. Consider the open cover $\{N_{r_x}(x)\}_x$ of K, and suppose there exists a finite subcover $\{N_{r_{x_1}}(x_1), \ldots, N_{r_{x_n}}(x_n)\}$. Then

$$E = K \cap E \subset \left(\bigcup_{j=1}^{n} N_{r_{x_j}}(x_j) \right) \cap E = \bigcup_{j=1}^{n} (N_{r_{x_j}}(x_j) \cap E) \subset \{x_1, \ldots, x_n\},$$

which is impossible because E is infinite. Hence, $\{N_{r_x}(x)\}_x$ has no finite subcover, so K is not compact. $\square$

Theorem 2.45. *Let $\mathcal{F}$ be a non-empty family of compact sets with the "finite-intersection property" (which means that any finite intersection of sets from $\mathcal{F}$ is non-empty). Then $\bigcap_{K \in \mathcal{F}} K$ is non-empty.*

Proof. Suppose, by way of contradiction, that $\bigcap_{K \in \mathcal{F}} K$ is empty. Choose a compact set $S \in \mathcal{F}$. Fix $x \in S$. There exists $K_x \in \mathcal{F}$ such that $x \notin K_x$ because $\bigcap_{K \in \mathcal{F}} K$ is empty. Since K_x is compact, K_x^c is open, so there exists $r_x > 0$ such that $N_{r_x}(x) \subset K_x^c$.

We now have an open cover $\{N_{r_x}(x)\}_{x \in S}$ of S. By compactness of S, there exists a finite subcover $\{N_{r_{x_1}}(x_1), \ldots, N_{r_{x_n}}(x_n)\}$. But

$$\begin{aligned}
S \cap K_{x_1} \cap \cdots \cap K_{x_n} &\subset \left(\bigcup_{j=1}^{n} N_{r_{x_j}}(x_j) \right) \cap (K_{x_1} \cap \cdots \cap K_{x_n}) \\
&= \bigcup_{j=1}^{n} (N_{r_{x_j}} \cap (K_{x_1} \cap \cdots \cap K_{x_n})) \\
&\subset \bigcup_{j=1}^{n} (N_{r_{x_j}} \cap K_{x_j}) \\
&= \emptyset,
\end{aligned}$$

contradicting the assumption that $\mathcal{F}$ has the finite-intersection property. $\square$

Corollary 2.45.1. *Suppose $\{K_n\}_{n\in\mathbb{N}}$ is a "decreasing" family of non-empty compact sets (which means that $K_{n+1} \subset K_n$ for all $n \in \mathbb{N}$). Then $\bigcap_{n\in\mathbb{N}} K_n$ is non-empty.*

Proof. Choose a finite number of sets $K_{n_1}, \ldots, K_{n_m}$ in $\{K_n\}_{n\in\mathbb{N}}$. Let $n_0 = \max\{n_1, \ldots, n_m\}$. Then $\bigcap_{j=1}^{m} K_{n_j} = K_{n_0}$ is non-empty. Hence, $\{K_n\}_{n\in\mathbb{N}}$ has the finite-intersection property. By Theorem 2.45, $\bigcap_{n\in\mathbb{N}} K_n$ is non-empty. □

Exercise 2.46 (Cardinality of Compact Sets)**.** Let K be a compact set. In this exercise, we will show that $|K| \leq |\mathbb{R}|$.

(a) Prove that for all $n \in \mathbb{N}$, there exists a finite set $\{x_{n,1}, \ldots, x_{n,m_n}\} \subset K$ such that

$$K = \bigcup_{i=1}^{m_n} N_{1/n}(x_{n,i}).$$

(b) Fix $x \in K$. Construct an infinite sequence of natural numbers $\{i_1, i_2, i_3, \ldots\}$ such that

$$\{x\} = \bigcap_{n=1}^{\infty} N_{1/n}(x_{n,i_n}).$$

Note: the sequence $\{i_1, i_2, i_3, \ldots\}$ is identical to a function $s : \mathbb{N} \to \mathbb{N}$ where $s(n) = i_n$ for all n. Therefore, the set of sequences of natural numbers is $\mathbb{N}^{\mathbb{N}}$ (see Exercise 2.10).

(c) Hence, we can define $f : K \to \mathbb{N}^{\mathbb{N}}$ by $f(x) = \{i_1, i_2, i_3, \ldots\}$. Note that f is injective because $\bigcap_{n=1}^{\infty} N_{1/n}(x_{n,i_n})$ is a singleton set, so $|K| \leq |\mathbb{N}^{\mathbb{N}}|$. To conclude that $|K| \leq |\mathbb{R}|$, prove that $|\mathbb{N}^{\mathbb{N}}| \leq |\mathbb{R}|$.

2.4 The Heine–Borel Theorem

In general, determining whether a given set is compact is difficult. Fortunately, there is a simple characterization of all compact subsets of $\mathbb{R}^k$. This section is dedicated to proving the Heine–Borel Theorem, which says that a subset $E \subset \mathbb{R}^k$ is compact if and only if E is closed and bounded.

Definition 2.47. Let $k \in \mathbb{N}$. A *k-cell* is a subset of $\mathbb{R}^k$ of the form

$$[a_1, b_1] \times \cdots \times [a_k, b_k]$$

where $a_i \leq b_i$ for all $1 \leq i \leq k$.

Proposition 2.48. *Every bounded subset of $\mathbb{R}^k$ is a subset of a k-cell.*

Proof. Let $E \subset \mathbb{R}^k$. If E is empty, then E is a subset of $[0,0] \times \cdots \times [0,0] \in \mathbb{R}^k$. Now suppose E is non-empty. Then there exists $\vec{v} \in E$ and $r > 0$ such that

$$d(\vec{x}, \vec{v}) = \left(\sum_{i=1}^{n} (x_i - v_i)^2 \right)^{1/2} < r$$

for all $\vec{x} \in E$. For any given integer $1 \leq j \leq k$, let $S_j = \{x_j \mid \vec{x} \in E\} \subset \mathbb{R}$. Then

$$|x_j - v_j| = ((x_j - v_j)^2)^{1/2} \leq \left(\sum_{i=1}^{n} (x_i - v_i)^2 \right)^{1/2} < r$$

for all $x_j \in S_j$. Let $a_j = v_j - r$ and $b_j = v_j + r$, so that $S_j \subset [a_j, b_j]$. Now fix any $\vec{x} \in E$. Then $x_j \in [a_j, b_j]$ for all $1 \leq j \leq k$, so $\vec{x} \in [a_1, b_1] \times \cdots \times [a_k, b_k]$. Therefore, $E \subset [a_1, b_1] \times \cdots \times [a_k, b_k]$. □

Theorem 2.49 (Nested-Interval Property). *Suppose $\{\ell_1, \ell_2, \ell_3, \dots\}$ and $\{u_1, u_2, u_3, \dots\}$ are sequences of real numbers such that $\ell_i \leq \ell_j \leq u_j \leq u_i$ for all $j \geq i \geq 1$. Then $\bigcap_{n=1}^{\infty} [\ell_n, u_n]$ is non-empty.*

Proof. Consider the set $L = \{\ell_1, \ell_2, \ell_3, \dots\}$. Then L is non-empty and bounded above since $\ell_n \leq u_1$ for all $n \geq 1$. Let $x = \sup(L)$. Fix $i \geq 1$. Then $x \geq \ell_i$ since x is an upper bound of S. If $j \geq i$, then $\ell_j \leq u_j \leq u_i$, and if $1 \leq j < i$, then $\ell_j \leq \ell_i \leq u_i$. Therefore, u_i is an upper bound of S, so $x \leq u_i$. It follows that $x \in [\ell_i, u_i]$. But $i \geq 1$ is arbitrary, so $x \in \bigcap_{n=1}^{\infty} [\ell_n, u_n]$. □

Theorem 2.50. *Every k-cell is compact.*

Proof. Let $k \in \mathbb{N}$. For each k-cell $C = [a_1, b_1] \times \cdots \times [a_k, b_k] \subset \mathbb{R}^k$, define

$$f(C) = \left(\sum_{i=1}^{k} (b_i - a_i)^2 \right)^{1/2}.$$

For any $\vec{x} = (x_1, x_2, \dots, x_k) \in C$ and $\vec{y} = (y_1, y_2, \dots, y_k) \in C$, we have that

$$d(\vec{x}, \vec{y}) = \left(\sum_{i=1}^{k} (y_i - x_i)^2 \right)^{1/2} \leq \left(\sum_{i=1}^{k} (b_i - a_i)^2 \right)^{1/2} = f(C)$$

because $x_i, y_i \in [a_i, b_i]$ for all $1 \leq i \leq k$.

Suppose there is a k-cell $C_1 \subset \mathbb{R}^k$ that is not compact. Then there is an open cover S of C_1 that does not have a finite subcover. Inductively, for $j \geq 1$, let C_j be a k-cell such that no finite subset of S covers C_i. Write $C_j = [a_1, b_1] \times \cdots \times [a_k, b_k]$. Note that $[a_i, b_i] = [a_i, \frac{1}{2}(a_i + b_i)] \cup [\frac{1}{2}(a_i + b_i), b_i]$ for each $1 \leq i \leq k$. Hence, by splitting each $[a_i, b_i]$ into two subintervals, we can write C_j as a finite union of "k-subcells" of the form $[\ell_1, u_1] \times \cdots \times [\ell_k, u_k]$ where $u_i - \ell_i = \frac{1}{2}(b_i - a_i)$. Let $E = [\ell_1, u_1] \times \cdots \times [\ell_k, u_k]$ be any one of these k-subcells, and notice that

$$f(E) = \left(\sum_{i=1}^{k} (u_i - \ell_i)^2 \right)^{1/2} = \left(\sum_{i=1}^{k} \left(\frac{1}{2}(b_i - a_i) \right)^2 \right)^{1/2} = \frac{1}{2} \left(\sum_{i=1}^{k} (b_i - a_i)^2 \right)^{1/2} = \frac{1}{2} f(C_j).$$

If S had a finite subcover for each k-subcell, then the union of these subcovers would be a finite subcover of C_j. But by assumption, no finite subset of S covers C_j, so there must be a k-subcell C_{j+1} such that no finite subset of S covers C_{j+1}.

This recursive process produces k-cells $C_j = [a_1^{(j)}, b_1^{(j)}] \times \cdots \times [a_k^{(j)}, b_k^{(j)}]$ such that no finite subset of S covers C_j for any $j \geq 1$. Also, $f(C_{j+1}) = \frac{1}{2} f(C_j)$ and $[a_i^{(j+1)}, b_i^{(j+1)}] \subset [a_i^{(j)}, b_i^{(j)}]$ for all integers $j \geq 1$ and $1 \leq i \leq k$. Then by induction, $f(C_j) = (\frac{1}{2})^{j-1} f(C_1)$ for all $j \geq 1$. Note also that $f(C_1) > 0$. Indeed, if $f(C_1) = 0$, then $a_i^{(1)} = b_i^{(1)}$ for all $1 \leq i \leq k$, which makes C_1 a singleton set $\{\vec{x}\}$. But since S covers C_1, there is an open set $\mathcal{O} \in S$ such that $\vec{x} \in \mathcal{O}$, so $\{\mathcal{O}\}$ is a finite subset of S that covers C_1. This is a contradiction, so $f(C_1) > 0$.

Now consider

$$T = \bigcap_{j=1}^{\infty} C_j = \left(\bigcap_{j=1}^{\infty} [a_1^{(j)}, b_1^{(j)}] \right) \times \cdots \times \left(\bigcap_{j=1}^{\infty} [a_k^{(j)}, b_k^{(j)}] \right).$$

By the Nested-Interval Property, there exists an element $\vec{x} = (x_1, \ldots, x_k) \in T \subset C_1$. Since S covers C_1, there is an open set $\mathcal{O} \in S$ such that $\vec{x} \in \mathcal{O}$. Hence, there exists $r > 0$ such that $N_r(\vec{x}) \subset \mathcal{O}$. Recall that $f(C_1) > 0$, so we can divide by $f(C_1)$. Using the Archimedean Property, we can show that there exists $m \in \mathbb{N}$ such that $(\frac{1}{2})^m < \frac{r}{f(C_1)}$. Let $\vec{y} \in C_{m+1}$. Then $d(\vec{x}, \vec{y}) \leq f(C_{m+1}) = (\frac{1}{2})^m f(C_1) < r$, so $\vec{y} \in N_r(\vec{x}) \subset \mathcal{O}$. It follows that $C_{m+1} \subset \mathcal{O}$, which is a contradiction because no finite subset of S covers C_{m+1}. $\square$

Corollary 2.50.1 (The Heine–Borel Theorem)**.** *$K \subset \mathbb{R}^k$ is compact if and only if K is closed and bounded.*

Proof. Theorems 2.41 and 2.42 say that every compact set is closed and bounded. Conversely, suppose K is closed and bounded (relative to $\mathbb{R}^k$). Since K is bounded, there is a k-cell C that contains K by Proposition 2.48. Theorem 2.50 says that C is compact. Note that K is closed relative to C because K contains all of its limit points in $\mathbb{R}^k$, which C is a subset of. Therefore, K is a closed subset of a compact set, so K is compact by Theorem 2.43. $\square$

Theorem 2.51. *Let $E \subset \mathbb{R}^k$. Then the following are equivalent:*
(i) E is compact.
(ii) E is closed and bounded.
(iii) Every infinite subset of E has a limit point in E.

Proof. The Heine–Borel Theorem says that (ii) implies (i), and Theorem 2.44 says that (i) implies (iii). To finish the proof, we just need to show that (iii) implies (ii).

Suppose (iii) holds. Let $x \in \mathbb{R}^k$ be a limit point of E. By Theorem 2.29, there is an infinite subset $S \subset E$ such that for any $r > 0$, all but finitely many points of S are in $N_r(x) \setminus \{x\}$. Since S is infinite, S has a limit point $y \in E$. Let $D = d(x, y)$, and suppose $D > 0$. If $z \in N_{D/2}(y)$, then $d(x, z) \geq d(x, y) - d(z, y) > D - \frac{D}{2} = \frac{D}{2}$, so $z \notin N_{D/2}(x)$. Hence, $N_{D/2}(y) \subset (N_{D/2}(x))^c$. Since $S \cap (N_{D/2}(x))^c$ is finite, we know that $S \cap N_{D/2}(y)$ is also finite. But since y is a limit point of S, Corollary 2.29.1 implies that $S \cap N_{D/2}(y)$ is infinite, which is a contradiction. Therefore, $D = 0$, so $x = y \in E$, which means that E contains all of its limit points and is therefore closed.

Now suppose E is not bounded. Then E is non-empty. To obtain a contradiction, we construct a particular infinite subset $T \subset E$ which has no limit point. First, pick any point $t_1 \in E$. Recursively, for all integers $n \geq 2$, we pick $t_n \in E$ such that $d(t_1, t_n) > d(t_1, t_{n-1}) + 1$; this is possible because E is not bounded. Now let $T = \{t_n \mid n \in \mathbb{N}\} \subset E$. Let $i, j \in \mathbb{N}$

be distinct, and without loss of generality, assume that $i > j$. By induction, we have that $d(t_1, t_i) > d(t_1, t_j) + 1$. Hence,

$$d(t_i, t_j) \geq d(t_1, t_i) - d(t_1, t_j) > 1.$$

Therefore, all the $t_n \in T$ are distinct, so T is infinite. By assumption (iii), T has a limit point $w \in E$. Then by Corollary 2.29.1, $T \cap N_{1/2}(w)$ is infinite, so there exist two distinct elements $t_i, t_j \in T \cap N_{1/2}(w)$. But notice that $d(t_i, t_j) \leq d(t_i, w) + d(w, t_j) < \frac{1}{2} + \frac{1}{2} = 1$, which is impossible because we proved that $d(t_i, t_j) > 1$. Therefore, E must be bounded. $\square$

Theorem 2.52. *If $E \subset \mathbb{R}^k$ is a non-empty perfect set, then E is uncountable.*

Proof. Any non-empty perfect set must be infinite since finite sets do not have limit points by Corollary 2.29.1. Suppose E is countable, and write $E = \{x_1, x_2, x_3, \ldots\}$. To obtain a contradiction, we will use the Heine–Borel Theorem and Corollary 2.45.1 to construct a limit point $x_n \in E$ where n must be larger than any natural number. Let $n_0 = 1$, $r_0 = 1$, and $V_0 = N_{r_0}(x_{n_0})$. Recursively, suppose we have chosen n_i and r_i for some $i \geq 0$. Since x_{n_i} is a limit point of E, any neighbourhood of x_{n_i} contains infinitely many points of E. Hence, there exists $n_{i+1} > n_i$ such that $x_{n_{i+1}} \in V_i := N_{r_i}(x_{n_i})$. Choose $r_{i+1} > 0$ such that

$$r_{i+1} < \frac{1}{2}[r_i - d(x_{n_i}, x_{n_{i+1}})] \tag{1}$$

and

$$r_{i+1} < \min_{1 \leq j \leq n_i} d(x_j, x_{n_{i+1}}). \tag{2}$$

Condition (1) ensures that $V_{i+1} := N_{r_{i+1}}(x_{n_{i+1}}) \subset V_i$. Indeed, if $x \in V_{i+1}$, then

$$d(x, x_{n_i}) \leq d(x, x_{n_{i+1}}) + d(x_{n_{i+1}}, x_{n_i}) < r_{i+1} + (r_i - r_{i+1}) = r_i,$$

so $x \in V_i$. Condition (2) ensures that if $x_j \in \overline{V_{i+1}}$, then $j \geq n_{i+1}$. Indeed, if $x_j \in \overline{V_{i+1}}$, then $d(x_j, x_{n_{i+1}}) \leq r_{i+1} < d(x_k, x_{n_{i+1}})$ for all $1 \leq k \leq n_i$, so $j \geq n_{i+1}$ because $j \notin \{1, \ldots, n_i\}$.

Note that $r_{i+1} < \frac{1}{2} r_i$ by Condition (1). By induction, we have that $r_i < (\frac{1}{2})^i r_0 = (\frac{1}{2})^i$ for all $i \geq 1$. Since $\overline{V_i}$ is closed and bounded in $\mathbb{R}^k$, $\overline{V_i}$ is compact by the Heine–Borel Theorem. Observe that $\overline{V_{i+1}} \subset \overline{V_i}$ since $V_{i+1} \subset V_i$. By Corollary 2.45.1, $\bigcap_{i=1}^{\infty} \overline{V_i}$ is non-empty. Choose $x \in \bigcap_{i=1}^{\infty} \overline{V_i}$. For any $r > 0$, there exists $i_0 \in \mathbb{N}$ such that $(\frac{1}{2})^{i_0} < r$. Since $x \in \overline{V_{i_0}}$, we have that $d(x_{n_{i_0}}, x) \leq r_{i_0} < (\frac{1}{2})^{i_0} < r$, so $x_{n_{i_0}} \in N_r(x)$. Therefore, x is a limit point of E, so $x \in E$ since E is closed. Hence, $x = x_n$ for some $n \in \mathbb{N}$. But $x_n = x \in \overline{V_{n+1}}$, so $n \geq n + 1$, which is a contradiction. $\square$

Hence, we deduce (again) that $\mathbb{R}$ is uncountable because $\mathbb{R}$ is a non-empty perfect set.

Example 2.53 (The Cantor Set)**.** Let $C_0 = [0, 1]$ and

$$C_n = C_{n-1} \setminus \bigcup_{j=1}^{3^{n-1}} \left(\frac{3j-2}{3^n}, \frac{3j-1}{3^n} \right).$$

for all $n \geq 1$. For example, $C_1 = [0, \frac{1}{3}] \cup [\frac{2}{3}, 1]$ and $C_2 = [0, \frac{1}{9}] \cup [\frac{2}{9}, \frac{1}{3}] \cup [\frac{2}{3}, \frac{7}{9}] \cup [\frac{8}{9}, 1]$. Essentially, each C_n is a union of (one or more) closed intervals, and we form the next set C_{n+1} by removing the *middle third* of each closed interval. The *Cantor set* is the set

$$C \coloneqq \bigcap_{n=0}^{\infty} C_n,$$

which we can think of as the "limit" of the decreasing sequence $\{C_0, C_1, C_2, \dots\}$. This set C has many strange properties, making it a great source for constructing counterexamples in analysis. We will show that C is compact, perfect, and uncountable, all while having empty interior.

Clearly, C is bounded in $\mathbb{R}$ because C is a subset of $C_0 = [0, 1]$, which is bounded. We now prove that C_n is closed for all $n \geq 0$ by induction. First, $C_0 = [0, 1]$ is closed. Suppose that C_{n-1} is closed for some particular $n \geq 1$. To show that C_n is closed, we use the following lemma.

Lemma. Let X be a metric space. Let $E \subset X$ be closed and $F \subset X$ be open. Then $E \setminus F$ is closed in X.

Proof. Let $x \in X$ be a limit point of $E \setminus F$. By Proposition 2.26, $(E \setminus F)' \subset E'$ because $E \setminus F \subset E$. Hence, $x \in E' \subset E$ since E is closed. Suppose, by way of contradiction, that $x \in F$. Since F is open, there exists $r > 0$ such that $N_r(x) \subset F$. Since x is a limit point of $E \setminus F$, there exists $y \in E \setminus F$ such that $0 < d(x, y) < r$. But then $y \in N_r(x) \subset F$, so $y \in F$, contradicting that $y \in E \setminus F$. Therefore, $x \notin F$, so $x \in E \setminus F$. $\square$

By Proposition 2.25, $\bigcup_{j=1}^{3^{n-1}} (\frac{3j-2}{3^n}, \frac{3j-1}{3^n})$ is open since $(\frac{3j-2}{3^n}, \frac{3j-1}{3^n})$ is open for each j. Our inductive hypothesis says that C_{n-1} is closed. Hence, $C_n = C_{n-1} \setminus \bigcup_{j=1}^{3^{n-1}} (\frac{3j-2}{3^n}, \frac{3j-1}{3^n})$ is closed by the lemma. This completes the inductive step. By Proposition 2.25, C is closed since each C_n is closed. Therefore, C is compact by the Heine–Borel Theorem.

Fix $x \in C$. We want to show that $x \in C'$, so fix $r > 0$. By the Archimedean Property, there exists $n \in \mathbb{N}$ such that $\frac{1}{3^n} < r$. Let $k = \lfloor 3^n x \rfloor$. Then $3^n x - 1 < k \leq 3^n$, which implies that $\frac{k}{3^n} \leq x < \frac{k+1}{3^n}$. Since $|x - \frac{k}{3^n}| = x - \frac{k}{3^n} < \frac{k+1}{3^n} - \frac{k}{3^n} = \frac{1}{3^n} < r$, we just need to show that $\frac{k}{3^n} \in C$. Suppose $\frac{k}{3^n} \notin C$. Then there is a minimal $m \geq 0$ such that $\frac{k}{3^n} \notin C_m$. Clearly, $\frac{k}{3^n} \in [0, 1]$ since $3^n x \geq 0$ and $x \leq 1$, so $m \geq 1$. Hence, there exists $0 \leq j \leq 3^{m-1}$ such that $\frac{k}{3^n} \in (\frac{3j-2}{3^m}, \frac{3j-1}{3^m})$ for some $0 \leq j \leq 3^{m-1}$. By way of contradiction, suppose $m \leq n$. Then $\frac{k}{3^n} < \frac{3j-1}{3^m}$ implies that $k < 3^{n-m}(3j-1)$. Hence, $k + 1 \leq 3^{n-m}(3j-1)$ since k and $3^{n-m}(3j-1)$ are integers. It follows that

$$\frac{3j-2}{3^m} < \frac{k}{3^n} \leq x < \frac{k+1}{3^n} \leq \frac{3j-1}{3^m},$$

so $x \notin C_m$, which is a contradiction. Therefore, $m > n$. The inequality $\frac{3j-2}{3^m} < \frac{k}{3^n} < \frac{3j-1}{3^m}$ implies that $3j - 2 < 3^{m-n}k < 3j - 1$. Note that $3^{m-n}k$ is an integer because $m > n$. But no integer can be strictly between two consecutive integers, so we have a contradiction. Therefore, $\frac{k}{3^n} \in C$, so $x \in C'$. We proved previously that C is closed, and we just proved that every point of C is a limit point of C, so C is perfect.

Note that $0 \in C$ since $0 < \frac{1}{3^n}$ for all $n \geq 1$. Hence, C is non-empty, so C is uncountable by Theorem 2.52.

Finally, we prove that C has empty interior. Suppose C has an interior point x. Then there exists $\delta > 0$ such that $(x-\delta, x+\delta) \subset C$. Let $n \in \mathbb{N}$ such that $\frac{1}{3^n} < \frac{\delta}{4}$, and let $k = \lfloor 3^n x \rfloor$. Then $\frac{k-1}{3^n} < x \leq \frac{k}{3^n}$. It follows that

$$x - \delta \leq \frac{k}{3^n} - \delta < \frac{k-4}{3^n} < \frac{k-1}{3^n} < x.$$

Hence, the intervals $(\frac{k-4}{3^n}, \frac{k-3}{3^n})$, $(\frac{k-3}{3^n}, \frac{k-2}{3^n})$, and $(\frac{k-2}{3^n}, \frac{k-1}{3^n})$ are all in $(x-\delta, x)$, which is contained in $(0,1)$. It follows that

$$1 \leq k-4 < k-3 < k-2 \leq 3^n - 2. \tag{3}$$

Since $k-4$, $k-3$, and $k-2$ are three consecutive integers, one of them must be of the form $3j-2$ where $j \in \mathbb{Z}$. Moreover, $1 \leq j \leq 3^{n-1}$ because of the inequality (3). Hence, $(\frac{3j-2}{3^n}, \frac{3j-1}{3^n}) \subset C \subset C_n$, which is a contradiction, so x is not an interior point of C.

Exercise 2.54 (Self-Similarity of the Cantor Set)**.** For any set $S \subset \mathbb{R}$ and constants $a, b \in \mathbb{R}$, denote

$$aS + b \coloneqq \{ax + b \mid x \in S\}.$$

Prove that the sets C_n in the construction of the Cantor set satisfy the "recurrence relation"

$$C_n = \begin{cases} [0,1] & \text{if } n = 0 \\ \frac{1}{3}C_{n-1} \cup \left(\frac{1}{3}C_{n-1} + \frac{2}{3}\right) & \text{if } n \geq 1. \end{cases}$$

Hence, prove that $C = \frac{1}{3}C \cup (\frac{1}{3}C + \frac{2}{3})$.

2.5 Connectedness

Definition 2.55. Two sets $A \subset X$ and $B \subset X$ are *separated* (relative to X) if $\overline{A} \cap B = \emptyset = A \cap \overline{B}$.

Definition 2.56. $E \subset X$ is *connected* (relative to X) if E is not the union of two non-empty separated subsets of X.

Theorem 2.57. *Let $Y \subset X$ and $A, B \subset Y$. Then A and B are separated relative to Y if and only if A and B are separated relative to X.*

Proof. For any set $S \subset Y$, let $\overline{S}^Y$ denote the closure of S relative to Y, and let $\overline{S}^X$ denote the closure of S relative to X. Note that $\overline{S}^Y = \overline{S}^X \cap Y$.

Suppose A and B are separated relative to Y. Then $\overline{A}^X \cap B \subset Y$ since $B \subset Y$. Hence,

$$\overline{A}^X \cap B = \overline{A}^X \cap B \cap Y = \overline{A}^Y \cap B = \emptyset.$$

By a similar argument, $A \cap \overline{B}^X = \emptyset$, so A and B are separated relative to X.

Conversely, suppose A and B are separated relative to X. Then

$$\overline{A}^Y \cap B = \overline{A}^X \cap B \cap Y = \emptyset = A \cap \overline{B}^X \cap Y = A \cap \overline{B}^Y,$$

so A and B are separated relative to Y. $\square$

Corollary 2.57.1. *Let $E \subset Y \subset X$. Then E is connected relative to X if and only if E is connected relative to Y.*

Proof. If E is not connected relative to X, then $E = A \cup B$ where $A, B \subset X$ are non-empty and separated relative to X. But notice that $A, B \subset Y$ since $A \cup B = E \subset Y$. Hence, A and B are separated relative to Y by Theorem 2.57, so E is not connected relative to Y. The proof of the reverse implication is similar. $\square$

Therefore, separatedness and connectedness are in fact intrinsic properties of a metric space, just like compactness (cf. Theorem 2.39).

Lemma 2.58. *If $S \subset \mathbb{R}$ is non-empty and bounded above, then $\sup(S) \in \overline{S}$. Similarly, if $S \subset \mathbb{R}$ is non-empty and bounded below, then $\inf(S) \in \overline{S}$.*

Proof. If $\sup(S) \in S$, then $\sup(S) \subset \overline{S}$ since $S \subset \overline{S}$. Now suppose $\sup(S) \notin S$. Fix $\epsilon > 0$. By definition of supremum, there exists $x \in S$ such that $\sup(S) - \epsilon < x \leq \sup(S)$. But $\sup(S) \notin S$, so $x < \sup(S)$. It follows that $0 < \sup(S) - x = |\sup(S) - x| < \epsilon$, so $\sup(S) \in S' \subset \overline{S}$. The proof that $\inf(S) \in \overline{S}$ when $\inf(S)$ exists is essentially the same with some signs and inequalities flipped. $\square$

Theorem 2.59. *Let $E \subset \mathbb{R}$. Then E is connected if and only if $(x, y) \subset E$ for all $x, y \in E$ such that $x < y$.*

Proof. Suppose there exists $x, y \in E$ such that $x < y$ and $(x, y) \not\subset E$. Then there exists $z \in (x, y)$ such that $z \notin E$. Let $A = \{w \in E \mid w < z\}$ and $B = \{w \in E \mid w > z\}$. Then A and B are non-empty because $x \in A$ and $y \in B$, and $E = A \cup B$. Also, since $A \subset (-\infty, z)$ and $B \subset (z, \infty)$, we have that $A \cap B \subset (-\infty, z) \cap (z, \infty) = \emptyset$. For any $w \in A$, let $r = \frac{z-w}{2} > 0$. If $v \in N_r(w)$, then $v - w \leq |v - w| < r$, so $v < r + w = \frac{z+w}{2} < z$. Hence, $N_r(w) \cap B = \emptyset$, so w is not a limit point of B. Therefore, $A \cap \overline{B} = \emptyset$, and by a similar argument, $\overline{A} \cap B = \emptyset$. Thus, we have shown that A and B are separated, so E is not connected.

Conversely, suppose E is not connected, so that $E = C \cup D$ where C, D are non-empty and $\overline{C} \cap D = \emptyset = C \cap \overline{D}$. Pick an element $x \in C$ and an element $y \in D$. Clearly, $x \in E$ and $y \in E$ since $C \subset E$ and $D \subset E$. Without loss of generality, suppose that $x < y$. Let $S = [x, y] \cap C$. Then S is non-empty (since $x \in S$) and bounded above by y, so $\sup(S)$ exists. Note that $\overline{S} \subset \overline{C}$ and $\overline{S} \subset \overline{[x, y]} = [x, y]$. Since $\sup(S) \in \overline{S}$ by Lemma 2.58, we have that $x \leq \sup(S) \leq y$. If $\sup(S) = y \in D$, then $\sup(S) \in \overline{C} \cap D$, contradicting that $\overline{C} \cap D$ is empty. Hence, $\sup(S) < y$. We now consider two cases: either $\sup(S) \in C$ or $\sup(S) \notin C$. Suppose $\sup(S) \in C$. Then $\sup(S) \notin \overline{D}$, so since $\overline{D}^c$ is open, there exists $0 < r < y - \sup(S)$ such that $N_r(\sup(S)) \subset \overline{D}^c \subset D^c$. Pick $z = \sup(S) + \frac{r}{2}$, and note that $z \in (x, y)$. Then $z \notin C$ because $z > \sup(S)$, and $z \in D^c$ because $z \in N_r(\sup(S))$. Therefore, $z \notin E$. Now suppose $\sup(S) \notin C$. Then $\sup(S) \neq x$, so $\sup(S) \in (x, y)$. We know that $\sup(S) \notin D$ because $\overline{C} \cap D$ is empty and $\sup(S) \in \overline{C}$. Therefore, $\sup(S) \notin C \cup D = E$. In both cases, we have shown that $(x, y) \cap E^c$ is non-empty, so $(x, y) \not\subset E$. $\square$

Exercise 2.60. Prove that X is connected if and only if the only subsets of X that are both open and closed relative to X are the empty set and X itself.

Exercise 2.61. Prove that the only connected subsets of the Cantor set C from Example 2.53 are the empty set and the singleton sets $\{x\}$ where $x \in C$. Thus, we say that C is *totally disconnected*.

Exercise 2.62. Prove that every connected metric space X with at least two points is uncountable. (Hint: choose distinct points $x, y \in X$, and let $D = d(x, y) > 0$. Then show that for all $c \in (0, D)$, there exists $z \in X$ such that $d(x, z) = c$.)

Exercise 2.63. Prove that $E \subset X$ is connected if and only if E is not the union of two non-empty disjoint open subsets of X.

3 Sequences and Series

Having familiarized ourselves with distances in metric spaces, we are now ready to define limits in metric spaces. We will briefly study sequences and the notion of *convergence* in general metric spaces, but the main focus of this chapter is on sequences in $\mathbb{R}$ and $\mathbb{C}$. We will also study infinite series in $\mathbb{R}$ and $\mathbb{C}$.

Throughout this chapter, X is a metric space with metric d.

3.1 Sequences and Subsequences

Definition 3.1. A *sequence in* X is a function $a : \mathbb{N} \to X$. We let $a_n := a(n)$ for each $n \in \mathbb{N}$, and we denote the sequence by $\{a_n\}_n$. A *subsequence* of $\{a_n\}_n$ is a sequence $\{b_k\}_k$ with a strictly-increasing function $\phi : \mathbb{N} \to \mathbb{N}$ such that $b_k = a_{\phi(k)}$ for all $k \in \mathbb{N}$. We define $n_k := \phi(k)$ and write $\{b_k\}_k$ as $\{a_{n_k}\}_k$.

The subscript outside the brackets $\{\}$ indicates the symbol that the sequence is indexed on. Usually, the sequence's indexing is clear, so we often omit this subscript.

Definition 3.2. Let $\{a_n\}$ be a sequence in X.

(a) $\{a_n\}$ is *convergent* if there exists $L \in X$ such that for all $\epsilon > 0$, there exists $N \in \mathbb{N}$ such that $d(a_n, L) < \epsilon$ for all $n \geq N$. In this case, we say that $\{a_n\}$ *converges to* L and that L is a *limit* of $\{a_n\}$. On the other hand, if $\{a_n\}$ is not convergent, then $\{a_n\}$ is *divergent*.

(b) $\{a_n\}$ is *Cauchy* if for all $\epsilon > 0$, there exists $N \in \mathbb{N}$ such that $d(a_n, a_m) < \epsilon$ for all $n, m \geq N$.

(c) $\{a_n\}$ is *bounded* if the set $\{a_n \mid n \in \mathbb{N}\}$ is a bounded subset of X.

The following proposition says that if a sequence has a limit, then this limit is unique. As a result, we can use the familiar notation "$\lim_{n\to\infty} a_n = L$" to say that L is *the* limit of a sequence $\{a_n\}$.

Proposition 3.3. *Let $\{a_n\}$ be a sequence in X. If $\{a_n\}$ converges to L_1 and L_2, then $L_1 = L_2$.*

Proof. Fix $\epsilon > 0$. Then there exist $N_1, N_2 \in \mathbb{N}$ such that $d(a_n, L_1) < \epsilon$ for all $n \geq N_1$ and $d(a_n, L_2) < \epsilon$ for all $n \geq N_2$. Let $N = \max\{N_1, N_2\}$. Then

$$d(L_1, L_2) \leq d(L_1, a_N) + d(a_N, L_2) < \epsilon + \epsilon = 2\epsilon.$$

Hence, $d(L_1, L_2) < 2\epsilon$ for all $\epsilon > 0$, so it must be the case that $d(L_1, L_2) = 0$. Therefore, $L_1 = L_2$. □

Proposition 3.4. *Let $\{a_n\}$ be a sequence in X. Then $\lim_{n\to\infty} a_n = L$ if and only if $\lim_{n\to\infty} d_X(a_n, L) = 0$.*

Proof. Both of the statements "$\lim_{n\to\infty} a_n = L$" and "$\lim_{n\to\infty} d_X(a_n, L) = 0$" are equivalent to the statement that for all $\epsilon > 0$, there exists $N \in \mathbb{N}$ such that $d_X(a_n, L) = |d_X(a_n, L) - 0| < \epsilon$ for all $n \geq N$. □

Proposition 3.5. *Every convergent sequence is Cauchy.*

Proof. Let $\{a_n\}$ be a convergent sequence in X, and let $L = \lim_{n\to\infty} a_n$. Fix $\epsilon > 0$. Then there exists $N \in \mathbb{N}$ such that $d(a_n, L) < \epsilon$ for all $n \geq N$. Therefore, if $n, m \geq N$, then

$$d(a_n, a_m) \leq d(a_n, L) + d(L, a_n) < \epsilon + \epsilon = 2\epsilon,$$

so $\{a_n\}$ is Cauchy. □

Remark. The factor of 2 appearing in the final inequality of the previous proof does not affect the conclusion that $\{a_n\}$ is Cauchy. Since 2 is independent of ϵ, we can just replace each instance of "ϵ" with "$\frac{\epsilon}{2}$" if we want the right-hand side of the final inequality to be "ϵ". Thus, there is no need to force the final inequality to be of the form "$d(a_n, a_m) < \epsilon$"; it suffices to obtain an inequality of the form "$d(a_n, a_m) < C\epsilon$" where $C > 0$ is a constant independent of ϵ. A similar remark applies if we want to prove that a sequence converges to a limit L.

Proposition 3.6. *Every Cauchy sequence is bounded.*

Proof. Let $\{a_n\}$ be Cauchy. Then there exists $N \in \mathbb{N}$ such that $d(a_n, a_m) < 1$ for all $n, m \geq N$. Let $M = 1 + \max_{1\leq j\leq N} d(a_j, a_N)$. If $1 \leq j \leq N$, then $d(a_j, a_N) < M$, and if $j > N$, then $d(a_j, a_N) < 1 \leq M$. Therefore, $d(a_j, a_N) < M$ for all $j \in \mathbb{N}$, so the set $\{a_n \mid n \in \mathbb{N}\}$ is bounded. □

Proposition 3.7. *Let $\{a_n\}$, $\{b_n\}$, and $\{c_n\}$ be sequences in $\mathbb{R}$ such that $a_n \leq b_n \leq c_n$ for all n large enough. Suppose $\{a_n\}$ and $\{c_n\}$ both converge to $L \in \mathbb{R}$. Then $\{b_n\}$ also converges to L.*

Proof. Let $N_1 \in \mathbb{N}$ be such that $a_n \leq b_n \leq c_n$ for all $n \geq N_1$. Fix $\epsilon > 0$. Then there exist $N_2, N_3 \in \mathbb{N}$ such that $|a_n - L| < \epsilon$ for all $n \geq N_2$ and $|c_n - L| < \epsilon$ for all $n \geq N_3$. Let $N = \max\{N_1, N_2, N_3\}$ and suppose $n \geq N$. Then $0 \leq b_n - a_n \leq c_n - a_n$ since $n \geq N_1$. Hence,

$$c_n - a_n = |c_n - a_n| \leq |c_n - L| + |L - a_n| = 2\epsilon,$$

so

$$|b_n - L| \leq |b_n - a_n| + |a_n - L| = (b_n - a_n) + |a_n - L| \leq (c_n - a_n) + |a_n - L| < 3\epsilon.$$

Therefore, $\{b_n\}$ converges to L. □

Proposition 3.8. *For any $x \in X$, we have $\lim_{n\to\infty} x = x$.*

Proof. Let $x_n = x$ for all $n \geq 1$. For any $\epsilon > 0$, we have $d(x_n, x) = 0 < \epsilon$ for all $n \geq 1$. Hence, $\{x_n\}$ converges to x. □

Proposition 3.9. *Let $\{a_n\}$ and $\{b_n\}$ be convergent sequences of complex numbers. Let $L_1 = \lim_{n\to\infty} a_n$ and $L_2 = \lim_{n\to\infty} b_n$. Then:*

(a) $\lim_{n\to\infty}(a_n + b_n) = L_1 + L_2$.

(b) $\lim_{n\to\infty} a_n b_n = L_1 L_2$.

(c) $\lim_{n\to\infty} \frac{a_n}{b_n} = \frac{L_1}{L_2}$ *if* $L_2 \neq 0$ *and* $b_n \neq 0$ *for all* $n \in \mathbb{N}$.

(d) $\lim_{n\to\infty}(a_n - b_n) = L_1 - L_2$.

(e) $\lim_{n\to\infty} |a_n| = |L_1|$.

Proof. Fix $\epsilon > 0$. Choose $N_1, N_2 \in \mathbb{N}$ such that $|a_n - L_1| < \epsilon$ for all $n \geq N_1$ and $|b_n - L_2| < \epsilon$ for all $n \geq N_2$.

(a) Let $N = \max\{N_1, N_2\}$. Then for all $n \geq N$,

$$|(a_n + b_n) - (L_1 + L_2)| \leq |a_n - L_1| + |b_n - L_2| < \epsilon + \epsilon = 2\epsilon.$$

(b) Every convergent sequence is Cauchy and hence bounded, so let $M \geq 0$ be such that $|b_n| \leq M$ for all $n \in \mathbb{N}$. Let $N = \max\{N_1, N_2\}$. Then

$$\begin{aligned}
|a_n b_n - L_1 L_2| &= |a_n b_n - b_n L_1 + b_n L_1 - L_1 L_2| \\
&\leq |a_n b_n - b_n L_1| + |b_n L_1 - L_1 L_2| \\
&= |b_n|(|a_n - L_1|) + |L_1|(|b_n - L_2|) \\
&\leq M\epsilon + |L_1|\epsilon \\
&< (1 + M + |L_1|)\epsilon.
\end{aligned}$$

Hence, $\{a_n b_n\}$ converges to $L_1 L_2$ because $1 + M + |L_1|$ is independent of ϵ.

(c) Choose $N_3 \in \mathbb{N}$ such that $|b_n - L_2| < \frac{1}{2}|L_2|$ for all $n \geq N_3$. Then

$$|b_n| = |L_2 - (L_2 - b_n)| \geq |L_2| - |L_2 - b_n| > \frac{1}{2}|L_2|$$

for all $n \geq N_3$. Let $N = \max\{N_2, N_3\}$. If $n \geq N$, then

$$\left|\frac{1}{b_n} - \frac{1}{L_2}\right| = \left|\frac{L_2 - b_n}{b_n L_2}\right| = \frac{|L_2 - b_n|}{(|b_n|)|L_2|} < \frac{\epsilon}{(\frac{1}{2}|L_2|)|L_2|},$$

so $\lim_{n\to\infty} \frac{1}{b_n} = \frac{1}{L_2}$. Now by using part (b), we see that

$$\lim_{n\to\infty} \frac{a_n}{b_n} = \lim_{n\to\infty} a_n \left(\frac{1}{b_n}\right) = L_1 \left(\frac{1}{L_2}\right) = \frac{L_1}{L_2}.$$

(d) Let $x_n = x$ for all $n \geq 1$. Then $d(x_n, x) = 0 < \epsilon$ for all $n \geq 1$. Hence, $\{x_n\}$ converges to x.

(e) By part (b), $\lim_{n\to\infty} -b_n = \lim_{n\to\infty}(-1)b_n = (-1)L_2 = -L_2$. Hence,

$$\lim_{n\to\infty} (a_n - b_n) = \lim_{n\to\infty} (a_n + (-b_n)) = L_1 - L_2$$

by part (a).

(f) By the Reverse Triangle Inequality (Proposition 2.13),

$$||a_n| - |L_1|| = ||a_n - 0| - ||L_1 - 0|| \leq |a_n - L| < \epsilon$$

for all $n \geq N_1$. $\square$

Proof. Suppose $\lim_{n\to\infty} |a_n| = 0$. Fix $\epsilon > 0$. Then there exists $N \in \mathbb{N}$ such that $||a_n| - 0| < \epsilon$ for all $n \geq N$. But $||a_n| - 0| = |a_n| = |a_n - 0|$. Therefore, $|a_n - 0| < \epsilon$ for all $n \geq N$, so $\lim_{n\to\infty} a_n = 0$. $\square$

Proposition 3.10.

(a) If $-1 < c < 1$, then $\lim_{n\to\infty} c^n = 0$.

(b) For all $k \in \mathbb{N}$ and $c > 1$, $\lim_{n\to\infty} \frac{n^k}{c^n} = 0$.

(c) For all $a > 0$, $\lim_{n\to\infty} a^{1/n} = 1$.

(d) $\lim_{n\to\infty} n^{1/n} = 1$.

Proof. (a) Fix $\epsilon > 0$. If $c = 0$, then the result is immediate. Now suppose $c \neq 0$. By the Archimedean Property, there exists $M \in \mathbb{N}$ such that $\frac{1}{M} < \epsilon$. Let $r = \frac{1}{|c|} > 1$ and $k = r - 1 > 0$. Then

$$r^n = (1+k)^n = \sum_{j=0}^{n} \binom{n}{j} k^j \geq 1 + \binom{n}{1} k = 1 + nk$$

for all $n \in \mathbb{N}$. By the Archimedean Property, there exists $N \in \mathbb{N}$ such that $Nk > M - 1$. For all $n \geq N$, we have that

$$r^n \geq 1 + nk > M,$$

so

$$|c^n - 0| = |c|^n = \frac{1}{r^n} < \frac{1}{M} < \epsilon.$$

Therefore, $\lim_{n\to\infty} c^n = 0$.

(b) Let $k \in \mathbb{N}$ and $c > 1$. Then $c^{1/k} > 1$. By the Archimedean Property, there exists $N \in \mathbb{N}$ such that $\frac{1}{N} < c^{1/k} - 1$, or equivalently, $\frac{N+1}{N} < c^{1/k}$. Let $r = \frac{1}{c}(\frac{N+1}{N})^k$, and note that $0 < r < 1$. Also, if $j \geq N$, then $\frac{j+1}{j} = 1 + \frac{1}{j} \leq 1 + \frac{1}{N} = \frac{N+1}{N}$, so $\frac{1}{c}(\frac{j+1}{j})^k \leq r$. Hence, for all $n \geq N$, we have that

$$0 \leq \frac{n^k}{c^n} = \frac{N^k}{c^N} \prod_{j=N}^{n-1} \frac{(\frac{j+1}{j})^k}{c} \leq \frac{N^k}{c^N} \prod_{j=N}^{n-1} r = \frac{N^k}{c^N} r^{n-N} = \frac{N^k}{c^N r^N} r^n.$$

Since $0 < r < 1$, part (a) says that $\lim_{n\to\infty} r^n = 0$. Hence, $\lim_{n\to\infty} \frac{N^k}{c^N r^N} r^n = 0$ because $\frac{N^k}{c^N r^N}$ is a constant. By Proposition 3.7, $\lim_{n\to\infty} \frac{n^k}{c^n} = 0$.

(c) If $a = 1$, the result is immediate. Suppose that $a > 1$. Let $b_n = a^{1/n} - 1 > 0$. Then for all $n \in \mathbb{N}$,

$$a = (b_n + 1)^n = \sum_{k=1}^{n} \binom{n}{k} b_n^k \geq 1 + \binom{n}{1} b_n = 1 + nb_n,$$

so $0 < b_n \leq \frac{a-1}{n}$. The Archimedean Property implies that $\lim_{n\to\infty} \frac{a-1}{n} = 0$. Hence, $\lim_{n\to\infty} b_n = 0$ by Proposition 3.7, so $\lim_{n\to\infty} a^{1/n} = \lim_{n\to\infty}(b_n + 1) = 0 + 1 = 1$. Finally, suppose that $0 < a < 1$. Then $\frac{1}{a} > 1$, so

$$\lim_{n\to\infty} a^{1/n} = \lim_{n\to\infty} \frac{1}{(\frac{1}{a})^{1/n}} = \frac{1}{\lim_{n\to\infty}(\frac{1}{a})^{1/n}} = \frac{1}{1} = 1.$$

(d) For each $n \in \mathbb{N}$, let $c_n = n^{1/n} - 1 > 0$. Then

$$n = (c_n + 1)^n = \sum_{k=0}^{n} \binom{n}{k} c_n^k \geq 1 + \binom{n}{2} c_n^2 = 1 + \frac{n(n-1)}{2} c_n^2.$$

If $n \geq 2$, then

$$c_n^2 \leq (n-1)\left(\frac{2}{n(n-1)}\right) = \frac{2}{n},$$

so $0 < c_n \leq \sqrt{\frac{2}{n}}$. Fix $\epsilon > 0$ and choose $N \in \mathbb{N}$ such that $N > \frac{2}{\epsilon^2}$. Then for all $n \geq \max\{N, 2\}$, we have that

$$|c_n - 0| = c_n \leq \sqrt{\frac{2}{n}} \leq \sqrt{\frac{2}{N}} < \sqrt{\epsilon^2} = \epsilon.$$

Hence, $\lim_{n\to\infty} c_n = 0$, so $\lim_{n\to\infty} n^{1/n} = \lim_{n\to\infty}(c_n + 1) = 0 + 1 = 1$. □

Definition 3.11. Let $\{a_n\}$ be a sequence in X. We say that $L \in X$ is a *subsequential limit* of $\{a_n\}$ if there exists a subsequence $\{a_{n_k}\}$ that converges to L.

Theorem 3.12. *Let $\{a_n\}$ be a sequence in X, and let $E = \{a_n \mid n \in \mathbb{N}\} \subset X$. Then L is a subsequential limit of $\{a_n\}$ if and only if L occurs infinitely often in $\{a_n\}$ or $L \in E'$.*

Proof. Let $L \in X$ be a subsequential limit of $\{a_n\}$. Suppose L does not occur infinitely often in $\{a_n\}$. Let $\{a_{n_k}\}$ be a subsequence that converges to L. Fix $\epsilon > 0$. Then there exists $N \in \mathbb{N}$ such that $d(a_{n_k}, L) < \epsilon$ for all $k \geq N$. Since L does not occur infinitely often in $\{a_n\}$, there must exist $\ell \geq N$ such that $a_{n_\ell} \neq L$. It follows that $0 < d(a_{n_\ell}, L) < \epsilon$, so $L \in E'$.

Conversely, suppose that L occurs infinitely often in $\{a_n\}$ or that $L \in E'$. If L occurs infinitely often in $\{a_n\}$, then $\{L, L, L, \dots\}$ is a subsequence that converges to L. Now suppose $L \in E'$. Pick a_{n_1} such that $0 < d(a_{n_1}, L) < 1$. Suppose inductively that we have picked a_{n_k} for some $k \in \mathbb{N}$. Then $N_{1/(k+1)}(L) \cap E$ is infinite by Corollary 2.29.1, so we can pick $a_{n_{k+1}}$ such that $n_{k+1} > n_k$ and $0 < d(a_{n_{k+1}}, L) < \frac{1}{k+1}$. By induction, $\{a_{n_k}\}$ is a subsequence such that $0 < d(a_{n_k}, L) < \frac{1}{k}$ for all $k \in \mathbb{N}$. Fix $\epsilon > 0$. By the Archimedean Property, there exists $N \in \mathbb{N}$ such that $\frac{1}{N} < \epsilon$. Then for all $k \geq N$, we have that $d(a_{n_k}, L) < \frac{1}{k} \leq \frac{1}{n_N} \leq \frac{1}{N} < \epsilon$. Therefore, $\{a_{n_k}\}$ converges to L. □

Corollary 3.12.1. *There exists a real sequence whose set of subsequential limits is $\mathbb{R}$.*

Proof. Recall that $\mathbb{Q}$ is countable, so $\mathbb{Q}$ has an enumeration $\{a_1, a_2, a_3, \dots\}$, which is a real sequence. Since $\mathbb{Q}' = \mathbb{R}$ by Theorem 1.25, the result follows from Theorem 3.12. □

Definition 3.13. We say that X is *complete* if every Cauchy sequence in X is convergent.

Lemma 3.14. *Any Cauchy sequence with a convergent subsequence is itself convergent.*

Proof. Let $\{a_n\}$ be a Cauchy sequence in a metric space X, and suppose $\{a_n\}$ has a convergent subsequence $\{a_{n_k}\}$. Let $L = \lim_{k\to\infty} a_{n_k}$, and fix $\epsilon > 0$. Then there exists $N_1 \in \mathbb{N}$ such that $d(a_{n_k}, L) < \epsilon$ for all $k \geq N_1$. Since $\{a_n\}$ is Cauchy, there exists $N_2 \in \mathbb{N}$ such that $d(a_n, a_m) < \epsilon$ for all $n, m \geq N_2$. Let $N = \max\{N_1, N_2\}$. Then for all $n \geq N$,

$$d(a_n, L) \leq d(a_n, a_{n_N}) + d(a_{n_N}, L) < \epsilon + \epsilon = 2\epsilon,$$

so $\{a_n\}$ converges to L. □

Theorem 3.15 (Bolzano–Weierstrass)**.**

(a) Every sequence in a compact set K has a convergent subsequence.

(b) For each $k \in \mathbb{N}$, every bounded sequence in $\mathbb{R}^k$ has a convergent subsequence.

Proof. (a) If $\{x_n\}$ has infinitely many occurrences of the same value x_*, then $\{x_*, x_*, x_*, \dots\}$ is a subsequence which converges to x_*. Now suppose no value of $\{x_n\}$ occurs infinitely often. Then the set $E = \{x_n \mid n \in \mathbb{N}\}$ is an infinite subset of the compact set K, so E has a limit point $x \in K$ by Theorem 2.44. By Theorem 3.12, x is a subsequential limit of $\{x_n\}$, proving that $\{x_n\}$ has a convergent subsequence.

(b) Every bounded sequence in $\mathbb{R}^k$ is contained in a k-cell by Proposition 2.48, which is compact by Theorem 2.50. The result now follows from part (a). $\square$

Corollary 3.15.1.

(a) Every compact metric space is complete.

(b) For all $k \in \mathbb{N}$, $\mathbb{R}^k$ is complete.

Proof. We can prove both parts of this corollary simultaneously. Any Cauchy sequence in a compact metric space has a convergent subsequence by the Bolzano–Weierstrass Theorem. In $\mathbb{R}^k$, every Cauchy sequence is bounded and hence has a convergent subsequence by the Bolzano–Weierstrass Theorem. Lemma 3.14 now implies the desired results. $\square$

Exercise 3.16. For any sequence $\{a_n\}$ in X, prove that the set of subsequential limits of $\{a_n\}$ is a closed subset of X.

Exercise 3.17. Let $\{a_n\}$ be a sequence in X. Prove that $\{a_n\}$ is convergent if and only if every subsequence of $\{a_n\}$ is convergent.

Exercise 3.18.

(a) Let $\{a_n\}$ be a sequence in a compact metric space K. Prove that $\{a_n\}$ is convergent if every convergent subsequence of $\{a_n\}$ has the same limit.

(b) What if we do not assume that K is compact?

3.2 Limit Superior and Limit Inferior

Even if a sequence is divergent, we may want to know how the sequence is bounded as $n \to \infty$. This section introduces the limit superior ("lim sup") and the limit inferior ("lim inf"), which are like the limits of the supremum and the infimum of the tail end of the sequence. For example, the sequence $\{(-1)^n(1+\frac{1}{n})\}_n$ certainly diverges, but $\limsup_{n\to\infty}(-1)^n(1+\frac{1}{n}) = 1$ and $\liminf_{n\to\infty}(-1)^n(1+\frac{1}{n}) = -1$.

Theorem 3.19 (Monotone Convergence Theorem)**.** *Let $\{x_n\}$ be a sequence in $\mathbb{R}$ that is monotonically increasing and bounded above. Then $\{x_n\}$ converges to $\sup_{n\in\mathbb{N}} x_n$. Similarly, if $\{y_n\}$ is a sequence in $\mathbb{R}$ that is monotonically decreasing and bounded below, then $\{y_n\}$ converges to $\inf_{n\in\mathbb{N}} y_n$.*

Proof. Let $s = \sup_{n\in\mathbb{N}} x_n$ and fix $\epsilon > 0$. Then there exists $N \in \mathbb{N}$ such that $s - \epsilon < x_N \leq s$. Since $\{x_n\}$ is monotonically increasing, we have that $s - \epsilon < x_N \leq x_n \leq s$ for all $n \geq N$. Hence, $0 \leq s - x_n = |s - x_n| < \epsilon$ for all $n \geq N$, so $\lim_{n\to\infty} x_n = s$. A similar argument applies for a monotonically-decreasing sequence $\{y_n\}$ which is bounded below. $\square$

Definition 3.20. Let $S \subset \mathbb{R}$ be non-empty. We say that $\sup(S) = \infty$ if S is not bounded above, and we say that $\inf(S) = -\infty$ if S is not bounded below.

Definition 3.21. Let $\{a_n\}$ be a sequence in $\mathbb{R}$. Then

$$\limsup_{n\to\infty} a_n := \inf_{n\in\mathbb{N}}(\sup_{m\geq n} a_m)$$

and

$$\liminf_{n\to\infty} a_n := \sup_{n\in\mathbb{N}}(\inf_{m\geq n} a_m).$$

Definition 3.22. Let $\{a_n\}$ be a sequence in $\mathbb{R}$.

(a) $\lim_{n\to\infty} a_n = \infty$ if for all $x \in \mathbb{R}$, there exists $N \in \mathbb{N}$ such that $a_n \geq x$ for all $n \geq N$.

(b) $\lim_{n\to\infty} a_n = -\infty$ if for all $x \in \mathbb{R}$, there exists $N \in \mathbb{N}$ such that $a_n \leq x$ for all $n \geq N$.

Proposition 3.23. *Let $\{a_n\}$ be a real sequence.*

(a) $\limsup_{n\to\infty} a_n = \infty$ if and only if $\{a_n\}$ is not bounded above.

(b) $\limsup_{n\to\infty} a_n = -\infty$ if and only if $\lim_{n\to\infty} a_n = -\infty$.

Similar statements apply for $\liminf$.

Proof. (a) Suppose $\limsup_{n\to\infty} a_n = \infty$. Then $\sup_{m\geq n} a_m = \infty$ for all $n \in \mathbb{N}$. In particular, $\sup_{m\geq 1} a_m = \infty$, so $\{a_n\}$ is not bounded above. Conversely, suppose $\{a_n\}$ is not bounded above. By way of contradiction, suppose that $\limsup_{n\to\infty} a_n < \infty$. Then there exists $x \in \mathbb{R}$ such that $\limsup_{n\to\infty} a_n = \inf_{n\in\mathbb{N}}(\sup_{m\geq n} a_m) < x$. By definition of infimum, there exists $N \in \mathbb{N}$ such that $\sup_{m\geq N} a_m < x$. Let $M = \max\{a_1, \ldots, a_N, x\}$, and note that $a_n \leq M$ for all $n \in \mathbb{N}$. Therefore, $\{a_n\}$ is bounded above, contradicting our initial assumption. Hence, $\limsup_{n\to\infty} a_n = \infty$.

(b) Suppose $\limsup_{n\to\infty} a_n = -\infty$. Fix $x \in \mathbb{R}$. Then there exists $N \in \mathbb{N}$ such that $\sup_{m\geq N} a_m < x$. It follows that $a_n \leq x$ for all $n \geq N$. Hence, $\lim_{n\to\infty} a_n = -\infty$. Conversely, suppose $\lim_{n\to\infty} a_n = -\infty$. For all $x \in \mathbb{R}$, there exists $M \in \mathbb{N}$ such that $a_m \leq x$ for all $m \geq M$. Hence, $\limsup_{n\to\infty} a_n = \inf_{n\in\mathbb{N}}(\sup_{m\geq n} a_m) \leq x$ for all $x \in \mathbb{R}$, so it must be the case that $\limsup_{n\to\infty} a_n = -\infty$. $\square$

Proposition 3.24. *Let $\{a_n\}$ be a real sequence. Then* $\limsup_{n\to\infty} a_n = \lim_{n\to\infty}(\sup_{m\geq n} a_m)$ *and* $\liminf_{n\to\infty} a_n = \lim_{n\to\infty}(\inf_{m\geq n} a_m)$.

Proof. Suppose $\limsup_{n\to\infty} a_n$ is finite. For each $n \in \mathbb{N}$, let $b_n = \sup_{m\geq n} a_m$, so that $\limsup_{n\to\infty} a_n = \inf_{n\in\mathbb{N}} b_n$. Note that $\{a_m \mid m \geq n+1\} \subset \{a_m \mid m \geq n\}$ for each $n \in \mathbb{N}$, so any upper bound of the latter set must also be an upper bound of the former set. Hence, $b_{n+1} \leq b_n$ for all $n \in \mathbb{N}$. Since $\inf_{n\in\mathbb{N}} b_n$ is finite by assumption, $\{b_n\}$ is a monotonically-decreasing sequence that is bounded below, so the Monotone Convergence Theorem says that

$$\lim_{n\to\infty}(\sup_{m\geq n} a_m) = \lim_{n\to\infty} b_n = \inf_{n\in\mathbb{N}} b_n = \limsup_{n\to\infty} a_n.$$

Suppose $\limsup_{n\to\infty} a_n = \infty$. Then $\{a_n\}$ is unbounded, so for any $x \in \mathbb{R}$, we have that $\sup_{m\geq n} a_m > x$ for all $n \geq 1$. Hence, $\lim_{n\to\infty}(\sup_{m\geq n} a_m) = \infty$. Finally, suppose $\limsup_{n\to\infty} a_n = -\infty$. Then $\lim_{n\to\infty} a_n = -\infty$, so for any $x \in \mathbb{R}$, there exists $N \in \mathbb{N}$

such that $a_n \le x$ for all $n \ge N$. Hence, $\sup_{m\ge n} a_m \le x$ for all $n \ge N$. Therefore, $\lim_{n\to\infty}(\sup_{m\ge n} a_m) = -\infty$. We conclude that $\limsup_{n\to\infty} a_n = \lim_{n\to\infty}(\sup_{m\ge n} a_m)$ in all cases.

The proof that $\liminf_{n\to\infty} a_n = \lim_{n\to\infty}(\inf_{m\ge n} a_m)$ is similar. □

Proposition 3.25. *For any real sequence $\{a_n\}$, we have that $\liminf_{n\to\infty} a_n \le \limsup_{n\to\infty} a_n$.*

Proof. Suppose there is a real sequence $\{a_n\}$ such that $\liminf_{n\to\infty} a_n > \limsup_{n\to\infty} a_n$. Pick $x \in \mathbb{R}$ such that $\liminf_{n\to\infty} a_n > x > \limsup_{n\to\infty} a_n$. Then there exists $N_1, N_2 \in \mathbb{N}$ such that $\inf_{m\ge N_1} a_m > x > \sup_{m\ge N_2} a_m$. Let $N = \max\{N_1, N_2\}$. Then $a_N \ge \inf_{m\ge N_1} a_m > x > \sup_{m\ge N_2} a_m \ge a_N$, which is a contradiction. □

Theorem 3.26. *A real sequence $\{a_n\}$ is convergent if and only if*

$$-\infty < \liminf_{n\to\infty} a_n = \limsup_{n\to\infty} a_n < \infty,$$

in which case $\lim_{n\to\infty} a_n = \liminf_{n\to\infty} a_n$.

Proof. Suppose $-\infty < \liminf_{n\to\infty} a_n = \limsup_{n\to\infty} a_n < \infty$. Let $L = \liminf_{n\to\infty} a_n$, and fix $\epsilon > 0$. Then there exist $N_1, N_2 \in \mathbb{N}$ such that $L-\epsilon < \inf_{m\ge N_1} a_m \le L$ and $L \le \sup_{m\ge N_2} a_m < L+\epsilon$. Let $N = \max\{N_1, N_2\}$. If $n \ge N$, then

$$L - \epsilon < \inf_{m\ge N_1} a_m \le a_n \le \sup_{m\ge N_2} a_m < L + \epsilon,$$

so $|a_n - L| < \epsilon$. Hence, $\lim_{n\to\infty} a_n = L = \liminf_{n\to\infty} a_n$.

Conversely, suppose $\{a_n\}$ converges to some $L \in \mathbb{R}$. Fix $\epsilon > 0$, and choose $N \in \mathbb{N}$ such that $L - \epsilon < a_n < L + \epsilon$ for all $n \ge N$. Then

$$L - \epsilon \le \inf_{m\ge N} a_m \le \sup_{n\in\mathbb{N}}(\inf_{m\ge n} a_m) = \liminf_{n\to\infty} a_n$$

and

$$\limsup_{n\to\infty} a_n = \inf_{n\in\mathbb{N}}(\sup_{m\ge n} a_m) \le \sup_{m\ge N} a_m \le L + \epsilon.$$

Since $\epsilon > 0$ can be arbitrarily small, it follows that $L \le \liminf_{n\to\infty} a_n$ and $\limsup_{n\to\infty} a_n \le L$. By Proposition 3.25, we have that

$$L \le \liminf_{n\to\infty} a_n \le \limsup_{n\to\infty} a_n \le L,$$

so $-\infty < L = \liminf_{n\to\infty} a_n = \limsup_{n\to\infty} a_n < \infty$. □

Theorem 3.27. *Let $\{a_n\}$ be a positive sequence in $\mathbb{R}$. Then*

$$\liminf_{n\to\infty} \frac{a_{n+1}}{a_n} \le \liminf_{n\to\infty} \sqrt[n]{a_n} \le \limsup_{n\to\infty} \sqrt[n]{a_n} \le \limsup_{n\to\infty} \frac{a_{n+1}}{a_n}.$$

Proof. First, note that each of these four limits must be nonnegative because $\{a_n\}$ is a positive sequence. Proposition 3.25 implies that $\liminf_{n\to\infty} \sqrt[n]{a_n} \le \limsup_{n\to\infty} \sqrt[n]{a_n}$. We now prove that $\liminf_{n\to\infty} \frac{a_{n+1}}{a_n} \le \liminf_{n\to\infty} \sqrt[n]{a_n}$. If $\liminf_{n\to\infty} \frac{a_{n+1}}{a_n} = 0$, then we are done since $\sqrt[n]{a_n} \ge 0$ for all $n \in \mathbb{N}$. Suppose $\liminf_{n\to\infty} \frac{a_{n+1}}{a_n} > 0$ (this includes the case where $\liminf_{n\to\infty} \frac{a_{n+1}}{a_n} = \infty$). Let $0 < x < \liminf_{n\to\infty} \frac{a_{n+1}}{a_n}$. Then there exists $N \in \mathbb{N}$ such that $\inf_{m\ge N} \frac{a_{m+1}}{a_m} > x$. That is, $a_{m+1} > xa_m$ for all $m \ge N$. By induction, it follows that $a_m > x^{m-N}a_N$ and hence $\sqrt[m]{a_m} > x\sqrt[m]{x^{-N}a_N}$ for all $m \ge N+1$. Fix $\epsilon > 0$. Part (c) of Proposition 3.10 says that $\lim_{m\to\infty} \sqrt[m]{x^{-N}a_N} = 1$, so there exists $M \in \mathbb{N}$ such that $|1 - \sqrt[m]{x^{-N}a_N}| < \frac{\epsilon}{x}$ for all $m \ge M$. Let $K = \max\{M, N+1\}$. Then $\sqrt[m]{a_m} > x(1-\frac{\epsilon}{x}) = x - \epsilon$ for all $m \ge K$, so

$$\liminf_{n\to\infty} \sqrt[n]{a_n} = \sup_{n\in\mathbb{N}}(\inf_{m\ge n} \sqrt[m]{a_m}) \ge \inf_{m\ge K} \sqrt[m]{a_m} \ge x - \epsilon.$$

Since ϵ can be arbitrarily small, we see that $\liminf_{n\to\infty} \sqrt[n]{a_n} \ge x$. But x is any arbitrary number in the interval $(0, \liminf_{n\to\infty} \frac{a_{n+1}}{a_n})$, so $\liminf_{n\to\infty} \sqrt[n]{a_n} \ge \liminf_{n\to\infty} \frac{a_{n+1}}{a_n}$.

If $\limsup_{n\to\infty} \frac{a_{n+1}}{a_n}$ is finite, a similar argument as the one in the previous paragraph proves that $\limsup_{n\to\infty} \sqrt[n]{a_n} \le \limsup_{n\to\infty} \frac{a_{n+1}}{a_n}$. If $\limsup_{n\to\infty} \frac{a_{n+1}}{a_n} = \infty$, then we trivially have that $\limsup_{n\to\infty} \sqrt[n]{a_n} \le \infty = \limsup_{n\to\infty} \frac{a_{n+1}}{a_n}$. □

We must note that the limit laws of Proposition 3.9 do not necessarily hold for lim sup and lim inf. We leave the details to the reader in Exercise 3.30. However, under some additional assumptions, we can obtain a "sum law" and a "product law" for lim sup and lim inf.

Proposition 3.28. *Suppose $\{a_n\}$ and $\{b_n\}$ are real sequences such that $\limsup_{n\to\infty} a_n$ is finite and $\lim_{n\to\infty} b_n$ exists. Then*

$$\limsup_{n\to\infty}(a_n + b_n) = \limsup_{n\to\infty} a_n + \limsup_{n\to\infty} b_n.$$

If moreover $\lim_{n\to\infty} b_n > 0$, then

$$\limsup_{n\to\infty} a_n b_n = \left(\limsup_{n\to\infty} a_n\right)\left(\limsup_{n\to\infty} b_n\right).$$

Proof. Let $L_1 = \limsup_{n\to\infty} a_n = \inf_{m\in\mathbb{N}} \sup_{n\ge m} a_n$ and $L_2 = \lim_{n\to\infty} b_n$. Fix $\epsilon > 0$. Then there exists $N_1 \in \mathbb{N}$ such that $|b_n - L_2| < \epsilon$ for all $n \ge N_1$. Fix $m \in \mathbb{N}$. Since $\sup_{n\ge m} a_n \ge L_1$, there exists $n_0 \ge \max\{m, N_1\}$ such that $a_{n_0} > L_1 - \epsilon$. Then since $b_{n_0} > L_2 - \epsilon$, we have that $a_{n_0} + b_{n_0} > L_1 + L_2 - \epsilon$. Therefore, $\sup_{n\ge m}(a_n + b_n) > L_1 + L_2 - 2\epsilon$. Since m is arbitrary, we have that $\inf_{m\in\mathbb{N}} \sup_{n\ge m}(a_n + b_n) \ge L_1 + L_2 - 2\epsilon$. On the other hand, there exists $N_2 \in \mathbb{N}$ such that $\sup_{n\ge N_2} a_n < L_1 + \epsilon$. Let $N = \max\{N_1, N_2\}$. Then for all $n \ge N$, we have $a_n + b_n < (L_1 + \epsilon) + (L_2 + \epsilon) = L_1 + L_2 + 2\epsilon$, so $\sup_{n\ge N}(a_n + b_n) \le L_1 + L_2 + 2\epsilon$. Therefore,

$$L_1 + L_2 - 2\epsilon \le \inf_{m\in\mathbb{N}} \sup_{n\ge m}(a_n + b_n) \le L_1 + L_2 + 2\epsilon.$$

Since $\epsilon > 0$ is arbitrary, we have

$$\limsup_{n\to\infty}(a_n + b_n) = \inf_{m\in\mathbb{N}} \sup_{n\ge m}(a_n + b_n) = L_1 + L_2.$$

Now suppose $L_2 > 0$, and fix $0 < \epsilon < L_2$. As before, there exists $N_1 \in \mathbb{N}$ such that $|b_n - L_2| < \epsilon$ for all $n \geq N_1$, and for any fixed $m \in \mathbb{N}$, there exists $n_0 \geq \max\{m, N_1\}$ such that $a_{n_0} > L_1 - \epsilon$. Then $a_{n_0} b_{n_0} > (L_1 - \epsilon) b_{n_0}$ since $b_{n_0} > L_2 - \epsilon > 0$. If $L_1 - \epsilon \geq 0$, then $(L_1 - \epsilon) b_{n_0} \geq (L_1 - \epsilon)(L_2 - \epsilon)$. If $L_1 - \epsilon < 0$, then $(L_1 - \epsilon) b_{n_0} > (L_1 - \epsilon)(L_2 + \epsilon)$ since $b_{n_0} < L_2 + \epsilon$. Therefore,

$$\sup_{n \geq m} a_n b_n \geq a_{n_0} b_{n_0} > \min\{(L_1 - \epsilon)(L_2 - \epsilon), (L_1 - \epsilon)(L_2 + \epsilon)\}.$$

Since m is arbitrary,

$$\inf_{m \in \mathbb{N}} \sup_{n \geq m} a_n b_n \geq \min\{(L_1 - \epsilon)(L_2 - \epsilon), (L_1 - \epsilon)(L_2 + \epsilon)\}$$

for any $0 < \epsilon < L_2$. Both $(L_1 - \epsilon)(L_2 - \epsilon)$ and $(L_1 - \epsilon)(L_2 + \epsilon)$ get arbitrarily close to $L_1 L_2$ when ϵ is arbitrarily small, so $\inf_{m \in \mathbb{N}} \sup_{n \geq m} a_n b_n \geq L_1 L_2$. Since $\inf_{m \in \mathbb{N}} \sup_{n \geq m} a_n = L_1$, there exists $N_2 \in \mathbb{N}$ with $\sup_{n \geq N_2} a_n < L_1 + \epsilon$. For $N = \max\{N_1, N_2\}$ and $n \geq N$, we have $a_n b_n < (L_1 + \epsilon) b_n$ since $b_n > L_2 - \epsilon > 0$. If $L_1 + \epsilon \geq 0$, then $(L_1 + \epsilon) b_n \leq (L_1 + \epsilon)(L_2 + \epsilon)$, and if $L_1 + \epsilon < 0$, then $(L_1 + \epsilon) b_n < (L_1 + \epsilon)(L_2 - \epsilon)$. Therefore,

$$a_n b_n < \max\{(L_1 + \epsilon)(L_2 + \epsilon), (L_1 + \epsilon)(L_2 - \epsilon)\}$$

for all $n \geq N$. Hence,

$$\inf_{m \in \mathbb{N}} \sup_{n \geq m} a_n b_n \leq \sup_{n \geq N} a_n b_n \leq \max\{(L_1 + \epsilon)(L_2 + \epsilon), (L_1 + \epsilon)(L_2 - \epsilon)\}.$$

If ϵ is arbitrarily small, both $(L_1 + \epsilon)(L_2 + \epsilon)$ and $(L_1 + \epsilon)(L_2 - \epsilon)$ get arbitrarily close to $L_1 L_2$, so $\inf_{m \in \mathbb{N}} \sup_{n \geq m} a_n b_n \leq L_1 L_2$. Therefore,

$$\limsup_{n \to \infty} a_n b_n = \inf_{m \in \mathbb{N}} \sup_{n \geq m} a_n b_n = L_1 L_2.$$

□

Proposition 3.28 also holds if we replace lim sup with lim inf, and the proof is substantially the same.

Exercise 3.29. Let $\{a_n\}$ be a bounded sequence in $\mathbb{R}$, and let S be the set of subsequential limits of $\{a_n\}$. Prove that $\limsup_{n \to \infty} a_n = \sup(S)$ and $\liminf_{n \to \infty} a_n = \inf(S)$.

Exercise 3.30.

(a) Find real sequences $\{a_n\}$, $\{b_n\}$ such that

$$\limsup_{n \to \infty} (a_n + b_n) \neq \limsup_{n \to \infty} a_n + \limsup_{n \to \infty} b_n.$$

(b) Find real sequences $\{a_n\}$, $\{b_n\}$ such that

$$\limsup_{n \to \infty} a_n b_n \neq \left(\limsup_{n \to \infty} a_n \right) \left(\limsup_{n \to \infty} b_n \right).$$

3.3 Introduction to Series Convergence

Definition 3.31. Let $\{a_n\}$ be a sequence of complex numbers. For each $N \in \mathbb{N}$, the N^{th} *partial sum* of $\{a_n\}$ is $s_n := \sum_{n=1}^{N} a_n$. We say that $\sum_{n=1}^{\infty} a_n$ is *convergent* if $\lim_{n\to\infty} s_n$ exists. If $\sum_{n=1}^{\infty} a_n$ is not convergent, then it is *divergent*.

If $\sum_{n=1}^{\infty} |a_n|$ is convergent, then $\sum_{n=1}^{\infty} a_n$ is said to be *absolutely convergent*. If $\sum_{n=1}^{\infty} a_n$ is convergent but $\sum_{n=1}^{\infty} |a_n|$ is not, then $\sum_{n=1}^{\infty} a_n$ is *conditionally convergent*.

It follows that if $\sum_{n=1}^{\infty} a_n$ converges, then for all $\epsilon > 0$, there exists $N \in \mathbb{N}$ such that

$$\left|\sum_{n=m+1}^{\infty} a_n\right| = \left|\sum_{n=1}^{\infty} a_n - \sum_{n=1}^{m} a_n\right| < \epsilon$$

whenever $m \geq N$.

Proposition 3.32 (Cauchy Criterion)**.** *$\sum_{n=1}^{\infty} a_n$ converges if and only if for all $\epsilon > 0$, there exists $N \in \mathbb{N}$ such that $|\sum_{k=m}^{n} a_k| < \epsilon$ whenever $n \geq m \geq N$.*

Proof. $\mathbb{C}$ and $\mathbb{R}^2$ are identical as metric spaces, so $\mathbb{C}$ is complete by Corollary 3.15.1. Let $s_n = \sum_{k=1}^{n} a_k$ for all $n \in \mathbb{N}$, and let $s_0 = 0$. Then $\sum_{n=1}^{\infty} a_n$ converges if and only if the sequence $\{s_n\}$ converges. Since $\mathbb{C}$ is complete, $\{s_n\}$ converges if and only if $\{s_n\}$ is Cauchy. Finally, $\{s_n\}$ is Cauchy if and only if for all $\epsilon > 0$, there exists $N \in \mathbb{N}$ such that

$$|s_n - s_{m-1}| = \left|\sum_{k=1}^{n} a_k - \sum_{k=1}^{m-1} a_k\right| = \left|\sum_{k=m}^{n} a_k\right| < \epsilon$$

for all $n \geq m \geq N$. □

Proposition 3.33. *Any absolutely convergent series is convergent.*

Proof. Suppose $\sum_{n=1}^{\infty} |a_n|$ converges. Fix $\epsilon > 0$. By the Cauchy Criterion, there exists $N \in \mathbb{N}$ such that $|\sum_{k=m}^{n} |a_k|| < \epsilon$ for all $n \geq m \geq N$. Hence, $|\sum_{k=m}^{n} a_k| \leq \sum_{k=m}^{n} |a_k| < \epsilon$ for all $n \geq m \geq N$, so $\sum_{n=1}^{\infty} a_n$ converges by the Cauchy Criterion. □

Proposition 3.34 (Triangle Inequality for Series)**.** *Let $\sum_{n=1}^{\infty} a_n$ be absolutely convergent. Then $|\sum_{n=1}^{\infty} a_n| \leq \sum_{n=1}^{\infty} |a_n|$.*

Proof. Let $\sum_{n=1}^{\infty} a_n = A$ and $\sum_{n=1}^{\infty} |a_n| = B$. By way of contradiction, suppose $|A| > B$. Since $|A| - B > 0$, there exists $N \in \mathbb{N}$ such that $|A - \sum_{n=1}^{N} a_n| < |A| - B$. Then

$$\begin{aligned}
|A| - B &= |A| - \sum_{n=1}^{\infty} |a_n| \\
&\leq |A| - \sum_{n=1}^{N} |a_n| \quad \text{since } |a_n| \geq 0 \text{ for all } n \\
&\leq |A| - \left|\sum_{n=1}^{N} a_n\right| \quad \text{by the Triangle Inequality}
\end{aligned}$$

$$\begin{aligned}
&\leq \left| A - \sum_{n=1}^{N} a_n \right| \quad \text{by the Triangle Inequality} \\
&< |A| - B,
\end{aligned}$$

which is a contradiction. Therefore, $|A| \leq B$, which is the desired result. $\square$

3.4 Convergence Tests

Theorem 3.35 (Divergence Test)**.** *If $\sum_{n=1}^{\infty} a_n$ converges, then $\lim_{n\to\infty} a_n = 0$.*

Proof. Suppose $\sum_{n=1}^{\infty} a_n$ converges. Fix $\epsilon > 0$. Then by the Cauchy Criterion, there exists $N \in \mathbb{N}$ such that

$$|a_n - 0| = \left| \sum_{k=n}^{n} a_k \right| < \epsilon$$

for all $n \geq N$. Hence, $\lim_{n\to\infty} a_n = 0$. $\square$

Proposition 3.36. *Let $x \in \mathbb{C}$. Then $\sum_{n=1}^{\infty} x^n$ converges if and only if $|x| < 1$, in which case we have that $\sum_{n=1}^{\infty} x^n = \frac{x}{1-x}$.*

Proof. Suppose $|x| \geq 1$. Then $|x^n| = |x|^n \geq 1$ for all $n \in \mathbb{N}$, so $\{x^n\}$ does not converge to 0 as $n \to \infty$. Hence, $\sum_{n=1}^{\infty} x^n$ diverges by the Divergence Test. Conversely, suppose $|x| < 1$. We claim that

$$\sum_{n=1}^{N} x^n = \frac{x - x^{N+1}}{1-x}$$

for all $N \in \mathbb{N}$. If $N = 1$, then $\sum_{n=1}^{N} x^n = x = \frac{x(1-x)}{1-x} = \frac{x - x^{N+1}}{1-x}$, so the claim holds for $N = 1$. Now let $M \in \mathbb{N}$ be arbitrary, and suppose the claim holds for $N = M$. Then

$$\sum_{n=1}^{M+1} x^n = x^{M+1} + \sum_{n=1}^{M} x^n = x^{M+1} + \frac{x - x^{M+1}}{1-x} = \frac{x^{M+1} - x^{M+2}}{1-x} + \frac{x - x^{M+1}}{1-x} = \frac{x - x^{M+2}}{1-x},$$

so the claim holds for $N = M + 1$. Therefore, the claim holds for all $N \in \mathbb{N}$ by induction. Using the limit laws from Proposition 3.9, we see that

$$\sum_{n=1}^{\infty} x^n = \lim_{N\to\infty} \sum_{n=1}^{N} x^n = \lim_{N\to\infty} \frac{x - x^{N+1}}{1-x} = \frac{x}{1-x} \lim_{N\to\infty} (1 - x^N) = \frac{x}{1-x}(1 - \lim_{N\to\infty} x^N) = \frac{x}{1-x},$$

where the last equality follows from part (a) of Proposition 3.10 because $|x| < 1$. $\square$

Theorem 3.37 (Comparison Test)**.** *Let $\{a_n\}, \{b_n\}$ be real sequences such that $0 \leq a_n \leq b_n$ for all n large enough. If $\sum_{n=1}^{\infty} b_n$ converges, then $\sum_{n=1}^{\infty} a_n$ converges.*

Proof. Let $N \in \mathbb{N}$ be such that $0 \leq a_n \leq b_n$ for all $n \geq N$. Suppose $\sum_{n=1}^{\infty} b_n$ converges. Let $B = \sum_{n=1}^{\infty} b_n$, $\beta = \sum_{n=1}^{N-1} b_n$, and $\alpha = \sum_{n=1}^{N-1} a_n$. For all $m \geq N$, let $s_m = \sum_{n=1}^{m} a_n$, and note that $\{s_m\}_{m\geq N}$ is a monotonically increasing sequence since $s_{m+1} - s_m = a_{m+1} \geq 0$ for all $m \geq N$. Also,

$$s_m = \sum_{n=1}^{m} a_n = \alpha + \sum_{n=N}^{m} a_n \leq \alpha + \sum_{n=N}^{m} b_n \leq \alpha + \sum_{n=N}^{\infty} b_n = \alpha + B - \beta$$

for all $m \geq N$, so $\{s_m\}_{m\geq N}$ is bounded above. By the Monotone Convergence Theorem (Theorem 3.19), $\lim_{m\to\infty} s_m$ exists, so $\sum_{n=1}^{\infty} a_n$ converges. □

Theorem 3.38 (Cauchy Condensation Test)**.** *Let $\{a_n\}$ be a nonnegative decreasing sequence. Then $\sum_{n=1}^{\infty} a_n$ converges if and only if $\sum_{n=1}^{\infty} 2^n a_{2^n}$ converges.*

Proof. Suppose $\sum_{n=1}^{\infty} a_n$ converges to A. Then for all $N \in \mathbb{N}$,

$$\begin{aligned}
\sum_{n=1}^{N} 2^n a_{2^n} &= 2\sum_{n=1}^{N} \sum_{m=0}^{2^{n-1}-1} a_{2^n} \\
&\leq 2\sum_{n=1}^{N} \sum_{m=0}^{2^{n-1}-1} a_{2^{n-1}+m} \quad \text{since } 2^{n-1} + m < 2^n \text{ whenever } 0 \leq m \leq 2^{n-1} - 1 \\
&= 2\sum_{n=1}^{2^N-1} a_n \\
&\leq 2A,
\end{aligned}$$

so $\sum_{n=1}^{\infty} 2^n a_{2^n}$ converges by the Monotone Convergence Theorem. On the other hand, suppose $\sum_{n=1}^{\infty} 2^n a_{2^n}$ converges to B. Then for all $N \in \mathbb{N}$,

$$\begin{aligned}
\sum_{n=1}^{N} a_n &\leq \sum_{n=1}^{2^N-1} a_n \quad \text{since } N \leq 2^N - 1 \\
&= \sum_{n=1}^{N} \sum_{m=0}^{2^{n-1}-1} a_{2^{n-1}+m} \\
&\leq \sum_{n=1}^{N} \sum_{m=0}^{2^{n-1}-1} a_{2^{n-1}} \quad \text{since } 2^{n-1} \leq 2^{n-1} + m \text{ for all } 0 \leq m \leq 2^{n-1} - 1 \\
&= \sum_{n=1}^{N} 2^{n-1} a_{2^{n-1}} \\
&= a_1 + \sum_{n=1}^{N-1} 2^n a_{2^n} \\
&\leq a_1 + B,
\end{aligned}$$

so $\sum_{n=1}^{\infty} a_n$ converges by the Monotone Convergence Theorem. □

Corollary 3.38.1. *For all $q \in \mathbb{Q}$, $\sum_{n=1}^{\infty} \frac{1}{n^q}$ converges if and only if $q > 1$.*

Proof. Let $q \in \mathbb{Q}$. By the Cauchy Condensation Test, $\sum_{n=1}^{\infty} \frac{1}{n^q}$ converges if and only if

$$\sum_{n=1}^{\infty} 2^n \left(\frac{1}{(2^n)^q} \right) = \sum_{n=1}^{\infty} 2^{n-nq} = \sum_{n=1}^{\infty} (2^{1-q})^n$$

converges, which happens if and only if $|2^{1-q}| = 2^{1-q} < 1$. Finally, $2^{1-q} < 1$ if and only if $q > 1$. □

Theorem 3.39 (Alternating Series Test). *Let $\{a_n\}$ be a nonnegative and monotonically decreasing sequence. Then*

$$\left| \sum_{k=m}^{n} (-1)^k a_k \right| \leq a_m$$

whenever $n \geq m \geq 1$. Consequently, if $\lim_{n\to\infty} a_n = 0$, then $\sum_{n=1}^{\infty} (-1) a^n$ converges and

$$\left| \sum_{k=m}^{\infty} (-1)^k a_k \right| \leq a_m$$

for all $m \in \mathbb{N}$.

Proof. Suppose $n \geq m \geq 1$. If m is even and n is odd, then

$$\sum_{k=m}^{n} (-1)^k a_k = (a_m - a_{m+1}) + \cdots + (a_{n-1} - a_n) \geq 0$$

and

$$\sum_{k=m}^{n} (-1)^k a_k = a_m - (a_{m+1} - a_{m+2}) - \cdots - (a_{n-2} - a_{n-1}) - a_n \leq a_m.$$

If m is even and n is even, then

$$\sum_{k=m}^{n} (-1)^k a_k = (a_m - a_{m+1}) + \cdots + (a_{n-2} - a_{n-1}) + a_n \geq 0$$

and

$$\sum_{k=m}^{n} (-1)^k a_k = a_m - (a_{m+1} - a_{m+2}) - \cdots - (a_{n-1} - a_n) \leq a_m.$$

Therefore, $|\sum_{k=m}^{n} (-1)^k a_k| \leq a_m$ if m is even. Now suppose m is odd. Let $b_1 = a_1$ and $b_{k+1} = a_k$ for all $k \geq 1$. Then $\{b_k\}$ is a nonnegative and monotonically decreasing sequence. Now

$$\left| \sum_{k=m}^{n} (-1)^k a_k \right| = \left| \sum_{k=m}^{n} (-1)^{k+1} a_k \right| = \left| \sum_{k=m}^{n} (-1)^{k+1} b_{k+1} \right| = \left| \sum_{k=m+1}^{n+1} (-1)^k b_k \right| \leq b_{m+1} = a_m$$

since $m + 1$ is even. Thus, we have shown that $|\sum_{k=m}^{n} (-1)^k a_k| \leq a_m$ for all integers $n \geq m \geq 1$.

Suppose $\lim_{n\to\infty} a_n = 0$. Fix $\epsilon > 0$, and choose $N \in \mathbb{N}$ such that $a_N < \epsilon$. Then

$$\left|\sum_{k=m}^{n} (-1)^k a_k\right| \leq a_m \leq a_N < \epsilon$$

for all $n \geq m \geq N$, so $\sum_{n=1}^{\infty} a_n$ converges by the Cauchy Criterion. Now fix $m \in \mathbb{N}$, and define the sequence $\{c_n\}$ where $c_n = \left|\sum_{k=m}^{n} (-1)^k a_k\right|$ for all $n \geq m$. Then

$$\left|\sum_{k=m}^{\infty} (-1)^k a_k\right| = \lim_{n\to\infty} c_n = \limsup_{n\to\infty} c_n \leq a_m$$

because $c_n \leq a_m$ for all $n \geq m$. $\square$

Corollary 3.39.1. $\sum_{n=1}^{\infty} (-1)^n \frac{1}{n}$ *is conditionally convergent.*

Proof. Since $\{\frac{1}{n}\}$ is a nonnegative decreasing sequence and $\lim_{n\to\infty} \frac{1}{n} = 0$, $\sum_{n=1}^{\infty} (-1)^n \frac{1}{n}$ converges. On the other hand, $\sum_{n=1}^{\infty} |(-1)^n \frac{1}{n}| = \sum_{n=1}^{\infty} \frac{1}{n}$ diverges by Corollary 3.38.1. $\square$

Lemma 3.40. *Let $\{a_n\}$ be a complex sequence, and suppose there exists $M \geq 0$ such that $\left|\sum_{n=1}^{N} a_n\right| \leq M$ for all $N \in \mathbb{N}$. Let $\{b_n\}$ be a monotonically decreasing nonnegative sequence. Then $\left|\sum_{n=1}^{N} a_n b_n\right| \leq M b_1$ for all $N \in \mathbb{N}$.*

To motivate the proof, observe that

$$a_1 b_1 + a_2 b_2 = a_1(b_1 - b_2) + (a_1 + a_2) b_2$$

and

$$a_1 b_1 + a_2 b_2 + a_3 b_3 = a_1(b_1 - b_2) + (a_1 + a_2)(b_2 - b_3) + (a_1 + a_2 + a_3) b_3.$$

The proof generalizes the above pattern to the sum $a_1 b_1 + \cdots + a_N b_N$, then uses the Triangle Inequality to obtain the desired result.

Proof. Fix $N \in \mathbb{N}$. Define $s_N = b_N$ and $s_n = b_n - b_{n+1}$ for all $1 \leq n \leq N - 1$. Then $b_n = \sum_{k=n}^{N} s_k$ for all $1 \leq n \leq N$. Let $S = \{(n,k) \in \mathbb{N}^2 \mid 1 \leq n \leq k \leq N\}$, and observe that

$$\sum_{n=1}^{N} \sum_{k=n}^{N} a_n s_k = \sum_{(n,k)\in S} a_n s_k = \sum_{k=1}^{N} \sum_{n=1}^{k} a_n s_k.$$

Hence,

$$\sum_{n=1}^{N} a_n b_n = \sum_{n=1}^{N} a_n \sum_{k=n}^{N} s_k = \sum_{n=1}^{N} \sum_{k=n}^{N} a_n s_k = \sum_{k=1}^{N} \sum_{n=1}^{k} a_n s_k = \sum_{k=1}^{N} s_k \sum_{n=1}^{k} a_n.$$

Note that $s_n \geq 0$ for all $1 \leq n \leq N$ because $\{b_n\}$ is a monotonically decreasing sequence and $b_N \geq 0$. By the Triangle Inequality,

$$\left|\sum_{k=1}^{N} s_k \sum_{n=1}^{k} a_n\right| \leq \sum_{k=1}^{N} \left|s_k \sum_{n=1}^{k} a_n\right| = \sum_{k=1}^{n} s_k \left|\sum_{n=1}^{k} a_n\right| \leq \sum_{k=1}^{n} s_k M = M \sum_{k=1}^{n} s_k = M b_1.$$

$\square$

Theorem 3.41 (Dirichlet's Test)**.** *Suppose $\{a_n\}$ is a complex sequence and $\{b_n\}$ is a monotonically decreasing sequence such that*

(i) the sequence $\{\sum_{n=1}^N a_n\}_N$ is bounded, and

(ii) $\lim_{n\to\infty} b_n = 0$.

Then $\sum_{n=1}^\infty a_n b_n$ converges.

Proof. Since $\{\sum_{n=1}^N a_n\}_N$ is bounded, there exists $M \geq 0$ such that $|\sum_{n=1}^N a_n| \leq M$ for all $N \in \mathbb{N}$. Hence, for any $1 \leq i \leq j$,

$$\left|\sum_{n=i}^{j} a_n\right| = \left|\sum_{n=1}^{j} a_n - \sum_{n=1}^{i-1} a_n\right| \leq \left|\sum_{n=1}^{j} a_n\right| + \left|\sum_{n=1}^{i-1} a_n\right| \leq 2M.$$

By the Monotone Convergence Theorem, $0 = \lim_{n\to\infty} b_n = \inf_{n\in\mathbb{N}} b_n$, which means $b_n \geq 0$ for all $n \in \mathbb{N}$. Fix $\epsilon > 0$, and pick $N \in \mathbb{N}$ such that $b_N = |b_N - 0| < \epsilon$. For any $m \geq N$, we have that

$$\left|\sum_{n=N}^{m} a_n\right| = \left|\sum_{n=1}^{m} a_n - \sum_{n=1}^{N-1} a_n\right| \leq \left|\sum_{n=1}^{m} a_n\right| + \left|\sum_{n=1}^{N-1} a_n\right| \leq 2M.$$

Therefore, by Lemma 3.40,

$$\left|\sum_{n=N}^{m} a_n b_n\right| \leq 2M b_N \leq 2M\epsilon$$

for all $m \geq N$. Hence, if $n \geq m \geq N$, then

$$\left|\sum_{k=m}^{n} a_k b_k\right| = \left|\sum_{k=N}^{n} a_k b_k - \sum_{k=N}^{m-1} a_k b_k\right| \leq \left|\sum_{k=N}^{n} a_k b_k\right| + \left|\sum_{k=N}^{m-1} a_k b_k\right| \leq 2M\epsilon + 2M\epsilon = 4M\epsilon,$$

so $\sum_{n=1}^\infty a_n b_n$ converges by the Cauchy Criterion. □

Theorem 3.42 (Ratio Test)**.** *Let $\{a_n\}$ be a complex sequence whose terms are all non-zero.*

(a) If $\limsup_{n\to\infty} \left|\frac{a_{n+1}}{a_n}\right| < 1$, then $\sum_{n=1}^\infty a_n$ converges absolutely.

(b) If $\liminf_{n\to\infty} \left|\frac{a_{n+1}}{a_n}\right| > 1$, then $\sum_{n=1}^\infty a_n$ diverges.

Proof. (a) Let $x \in \mathbb{R}$ such that $0 \leq \limsup_{n\to\infty} \left|\frac{a_{n+1}}{a_n}\right| < x < 1$. Then there exists $N \in \mathbb{N}$ such that $\sup_{n\geq N} \left|\frac{a_{n+1}}{a_n}\right| < x$, which means $|a_{n+1}| < x|a_n|$ for all $n \geq N$. By induction, we can show that $|a_n| \leq x^{n-N}|a_N|$ for all $n \geq N$. Since $0 < x < 1$, the geometric series

$$\sum_{n=1}^{\infty} x^{n-N}|a_N| = x^{-N}|a_N| \sum_{n=1}^{\infty} x^n$$

converges, so $\sum_{n=1}^\infty |a_n|$ converges by the Comparison Test. Hence, $\sum_{n=1}^\infty a_n$ is absolutely convergent.

(b) Suppose $\liminf_{n\to\infty} \left|\frac{a_{n+1}}{a_n}\right| > 1$. Then there exists $N \in \mathbb{N}$ such that $1 < \inf_{n\geq N} \left|\frac{a_{n+1}}{a_n}\right|$, which means $|a_n| < |a_{n+1}|$ for all $n \geq N$. It follows by induction that $|a_n| \geq |a_N|$ for all $n \geq N$, so $\liminf_{n\to\infty} |a_n| \geq |a_N|$. But $a_N \neq 0$ by hypothesis, so $\liminf_{n\to\infty} |a_n| > 0$. Hence, it is impossible that $\lim_{n\to\infty} a_n = 0$, for this would imply that $\liminf_{n\to\infty} |a_n| = \lim_{n\to\infty} |a_n| = 0$. Therefore, $\sum_{n=1}^\infty a_n$ diverges by the Divergence Test. □

Theorem 3.43 (Root Test)**.** *Let $\{a_n\}$ be a complex sequence.*

(a) If $\limsup_{n\to\infty} \sqrt[n]{|a_n|} < 1$, then $\sum_{n=1}^{\infty} a_n$ converges absolutely.

(b) If $\limsup_{n\to\infty} \sqrt[n]{|a_n|} > 1$, then $\sum_{n=1}^{\infty} a_n$ diverges.

Proof. (a) Let $L = \limsup_{n\to\infty} \sqrt[n]{|a_n|} \geq 0$, and suppose $L < 1$. Let $x = \frac{L+1}{2}$, and notice that $0 \leq L < x < 1$. Then there exists $N \in \mathbb{N}$ such that $\sup_{n\geq N} \sqrt[n]{|a_n|} < x$. Hence, $|a_n| < x^n$ for all $n \geq N$. Since $|x| < 1$, the geometric series $\sum_{n=1}^{\infty} x^n$ converges. By the Comparison Test, $\sum_{n=1}^{\infty} |a_n|$ also converges, so $\sum_{n=1}^{\infty} a_n$ is absolutely convergent.

(b) Suppose $\limsup_{n\to\infty} \sqrt[n]{|a_n|} > 1$. Then for all $N \in \mathbb{N}$, there exists $n \geq N$ such that $\sqrt[n]{|a_n|} > 1$, which means $|a_n| > 1$. Therefore, there is a subsequence $\{a_{n_k}\}$ such that $|a_{n_k}| > 1$ for all k. Since $\{a_{n_k}\}$ cannot converge to 0, neither can $\{a_n\}$, so $\sum_{n=1}^{\infty} a_n$ diverges by the Divergence Test. □

Remark. If $\limsup_{n\to\infty} \sqrt[n]{|a_n|} = 1$, then the Root Test is inconclusive. In this case, the Ratio Test is also inconclusive by Theorem 3.27.

Corollary 3.43.1. *The series $\sum_{n=1}^{\infty} \frac{z^n}{n}$ converges if and only if $|z| \leq 1$ and $z \neq 1$.*

Proof. Observe that

$$\lim_{n\to\infty} \sqrt[n]{\left|\frac{z^n}{n}\right|} = \lim_{n\to\infty} \frac{|z|}{n^{1/n}} = \frac{|z|}{\lim_{n\to\infty} n^{1/n}} = |z|,$$

so by the Root Test, the series converges absolutely if $|z| < 1$ and diverges if $|z| > 1$. By Corollary 3.38.1, the series diverges if $z = 1$. Suppose $|z| = 1$ and $z \neq 1$. Then

$$\left|\sum_{n=1}^{N} z^n\right| = \left|\frac{z(1-z^N)}{1-z}\right| = |z| \cdot \left|\frac{1-z^N}{1-z}\right| \leq \frac{1+|z|^N}{|1-z|} = \frac{2}{|1-z|}$$

for all $N \in \mathbb{N}$. That is, the sequence $\{\sum_{n=1}^{N} z^n\}_N$ is bounded. Since $\{\frac{1}{n}\}_n$ is a monotonically decreasing sequence and $\lim_{n\to\infty} \frac{1}{n} = 0$, Dirichlet's Test says that $\sum_{n=1}^{\infty} z^n \frac{1}{n}$ converges. □

Theorem 3.44. *Let $\sum_{n=0}^{\infty} a_n$ be absolutely convergent and $\sum_{n=0}^{\infty} b_n$ be convergent. Then $\sum_{n=0}^{\infty} \sum_{k=0}^{n} a_k b_{n-k}$ converges and*

$$\sum_{n=0}^{\infty} \sum_{k=0}^{n} a_k b_{n-k} = \left(\sum_{n=0}^{\infty} a_n\right)\left(\sum_{n=0}^{\infty} b_n\right).$$

If $\sum_{n=0}^{\infty} b_n$ is absolutely convergent, then so is $\sum_{n=0}^{\infty} \sum_{k=0}^{n} a_k b_{n-k}$.

Proof. We first claim that

$$\sum_{n=0}^{N} \sum_{k=0}^{n} a_k b_{n-k} = \sum_{k=0}^{N} a_k \sum_{n=0}^{N-k} b_n$$

for all $N \geq 0$. Fix $N \geq 0$, and let

$$S_1 = \{(k, n-k) \mid 0 \leq k \leq n \leq N \text{ where } k, n \in \mathbb{Z}\}$$

and

$$S_2 = \{(k, n) \mid 0 \le k \le N \text{ and } 0 \le n \le N - k \text{ where } k, n \in \mathbb{Z}\}.$$

Then

$$\sum_{n=0}^{N} \sum_{k=0}^{n} a_k b_{n-k} = \sum_{(i,j) \in S_1} a_i b_j$$

and

$$\sum_{k=0}^{N} a_k \sum_{n=0}^{N-k} b_n = \sum_{k=0}^{N} \sum_{n=0}^{N-k} a_k b_n = \sum_{(i,j) \in S_2} a_i b_j,$$

so it suffices to prove that $S_1 = S_2$. Suppose $(i, j) = (i, (i + j) - i) \in S_1$. Then $0 \le i \le i + j \le N$ implies that $0 \le i \le N$ and $0 \le j \le i + j \le N - i$, so $(i, j) \in S_2$. Conversely, suppose $(i, j) \in S_2$. Then $0 \le i \le N$ and $0 \le j \le N - i$. Hence, $0 \le i \le i + j \le N$, so $(i, j) = (i, (i + j) - i) \in S_1$. This proves the claim.

Let $A = \sum_{n=0}^{\infty} a_n$, $B = \sum_{n=0}^{\infty} b_n$, and $C = \sum_{n=0}^{\infty} |a_n|$. Fix $\epsilon > 0$. Then there exists $N_1 \ge 0$ such that $\sum_{n=m+1}^{\infty} |a_n| < \epsilon$ for all $m \ge N_1$. Notice that if $m \ge N_1$, then $|A - \sum_{n=0}^{m} a_n| = |\sum_{n=m+1}^{\infty} a_n| \le \sum_{n=m+1}^{\infty} |a_n| < \epsilon$. We can also choose $N_2 \ge 0$ such that $|B - \sum_{n=0}^{m} b_n| < \epsilon$ for all $m \ge N_2$. Now suppose $N \ge N_1 + N_2$. Then

$$\begin{aligned}
\left| \sum_{n=0}^{N} \sum_{k=0}^{n} a_k b_{n-k} - AB \right| &= \left| \sum_{k=0}^{N} a_k \sum_{n=0}^{N-k} b_n - AB \right| \\
&= \left| \sum_{k=0}^{N-N_2} a_k \sum_{n=0}^{N-k} b_n + \sum_{k=N-N_2+1}^{N} a_k \sum_{n=0}^{N-k} b_n - AB \right| \\
&\le \left| \sum_{k=0}^{N-N_2} a_k \sum_{n=0}^{N-k} b_n - AB \right| + \left| \sum_{k=N-N_2+1}^{N} a_k \sum_{n=0}^{N-k} b_n \right|.
\end{aligned}$$

Now we bound both of the magnitudes on the right-hand side. First, we compute that

$$\begin{aligned}
\left| \sum_{k=0}^{N-N_2} a_k \sum_{n=0}^{N-k} b_n - AB \right| &\le \left| \sum_{k=0}^{N-N_2} a_k \sum_{n=0}^{N-k} b_n - \sum_{k=0}^{N-N_2} a_k B \right| + \left| \sum_{k=0}^{N-N_2} a_k B - AB \right| \\
&= \left| \sum_{k=0}^{N-N_2} a_k (\sum_{n=0}^{N-k} b_n - B) \right| + |B| \cdot \left| \sum_{k=0}^{N-N_2} a_k - A \right| \\
&\le \sum_{k=0}^{N-N_2} |a_k| \cdot \left| \sum_{n=0}^{N-k} b_n - B \right| + |B| \cdot \left| \sum_{k=0}^{N-N_2} a_k - A \right| \\
&\quad \text{(by the Triangle Inequality)} \\
&\le \sum_{k=0}^{N-N_2} |a_k| \epsilon + |B| \cdot \left| \sum_{k=0}^{N-N_2} a_k - A \right| \\
&\quad \text{(since } N - k \ge N_2 \text{ whenever } 0 \le k \le N - N_2\text{)} \\
&\le \sum_{k=0}^{N-N_2} |a_k| \epsilon + |B| \epsilon \quad \text{(since } N - N_2 \ge N_1\text{)}
\end{aligned}$$

$$\leq C\epsilon + |B|\epsilon.$$

For the other magnitude, first define $M = \sup_{m\geq 0} |\sum_{n=0}^{m} b_n|$, which exists because every convergent sequence is bounded. Then

$$\begin{aligned}
\left|\sum_{k=N-N_2+1}^{N} a_k \sum_{n=0}^{N-k} b_n\right| &\leq \sum_{k=N-N_2+1}^{N} |a_k| \cdot \left|\sum_{n=0}^{N-k} b_n\right| \quad \text{(by the Triangle Inequality)}\\
&\leq \sum_{k=N-N_2+1}^{N} |a_k| M\\
&\leq M \sum_{k=N-N_2+1}^{\infty} |a_k|\\
&\leq M\epsilon \quad (\text{since } N - N_2 \geq N_1).
\end{aligned}$$

Therefore,

$$\left|\sum_{n=0}^{N}\sum_{k=0}^{n} a_k b_{n-k} - AB\right| \leq C\epsilon + |B|\epsilon + M\epsilon = (C + |B| + M)\epsilon$$

for all $N \geq N_1 + N_2$, so $\sum_{n=0}^{\infty}\sum_{k=0}^{n} a_k b_{n-k} = AB$ since $C + |B| + M$ is independent of ϵ.

Suppose $\sum_{n=0}^{\infty} b_n$ is absolutely convergent. Then $\sum_{n=0}^{\infty} |a_n|$ is absolutely convergent and $\sum_{n=0}^{\infty} |b_n|$ is convergent, so the result we have just proved says that $\sum_{n=0}^{\infty}\sum_{k=0}^{n} |a_k b_{n-k}| = \sum_{n=0}^{\infty}\sum_{k=0}^{n} |a_k| \cdot |b_{n-k}|$ is convergent. Hence, $\sum_{n=0}^{\infty}\sum_{k=0}^{n} a_k b_{n-k}$ is absolutely convergent. $\square$

Exercise 3.45. Fix $x \in [0, 1]$, and let $b \geq 2$ be an integer. Construct a sequence $\{c_n\}$ where $c_n \in \{0, 1, \ldots, b-1\}$ for each $n \in \mathbb{N}$ and

$$x = \sum_{n=1}^{\infty} \frac{c_n}{b^n}.$$

Hint: be greedy!

We call $\{c_n\}$ a *base-b expansion* of x. If $b = 2$, then $\{c_n\}$ is a *binary expansion*, and if $b = 10$, then $\{c_n\}$ is a *decimal expansion*.

Exercise 3.46 (Limit Comparison Test)**.** Let $\{a_n\}$ and $\{b_n\}$ be positive sequences such that $\lim_{n\to\infty} \frac{a_n}{b_n}$ is a nonnegative real number. Show that if $\sum_{n=1}^{\infty} b_n$ converges, then $\sum_{n=1}^{\infty} a_n$ also converges.

Exercise 3.47. Let ℓ_1 be the set of sequences $\{x_n\}$ in $\mathbb{C}$ such that $\sum_{n=1}^{\infty} |x_n|$ converges. For all $x, y \in \ell_1$, let $d(x, y) = \sum_{n=1}^{\infty} |x_n - y_n|$.

(a) Show that d is a metric on ℓ_1.

(b) Show that ℓ_1 is a complete metric space.

(c) Find a subset of ℓ_1 that is closed and bounded but not compact.

Exercise 3.48 (Baire Category Theorem)**.**

(a) Let X be a complete metric space and $\mathcal{O}_1, \mathcal{O}_2, \mathcal{O}_3, \ldots$ be open dense subsets of X. Prove that $\bigcap_{n=1}^{\infty} \mathcal{O}_n$ is dense in X. This result is called the *Baire Category Theorem* (BCT).

(b) Write $\mathbb{R} \setminus \mathbb{Q}$ as a countable intersection of open dense subsets of $\mathbb{R}$.

(c) Hence, prove that $\mathbb{Q}$ is *not* a countable intersection of open dense subsets of $\mathbb{R}$. (Note: the word "dense" is actually redundant here—since $\mathbb{Q}$ is dense, any open set containing $\mathbb{Q}$ must be dense.)

(d) Let $\mathcal{C}_1, \mathcal{C}_2, \mathcal{C}_3, \dots$ be closed subsets of X with empty interior. Prove that $\bigcup_{n=1}^{\infty} \mathcal{C}_n$ has empty interior. Hint: this result is equivalent to the BCT as stated in part (a). Use Exercise 2.33 to see the connection.

(e) Prove that if X is a complete metric space and $E \subset X$ is a non-empty perfect set, then E is uncountable. This is a stronger version of Theorem 2.52.

Exercise 3.49. Euclid showed that there are infinitely many primes. For each $n \in \mathbb{N}$, let p_n be the n^{th} prime number (e.g. $p_1 = 2$, $p_2 = 3$, and $p_3 = 5$). In this exercise, we will show that

$$\sum_{n=1}^{\infty} \frac{1}{p_n} = \infty.$$

(a) Recall that each integer $n \geq 2$ has a unique prime factorization. Show that

$$\sum_{n=1}^{N} \frac{1}{n} \leq \prod_{n=1}^{N} \left(1 + \frac{1}{p_n} + \frac{1}{p_n^2} + \cdots + \frac{1}{p_n^N}\right)$$

for all $N \geq 1$. Hint: if we expand the product on the right-hand side, we get a sum where each term is $\frac{1}{m}$ for some $m \in \mathbb{N}$.

It follows that

$$\sum_{n=1}^{N} \frac{1}{n} \leq \prod_{n=1}^{N} \left(1 + \frac{1}{p_n} + \frac{1}{p_n^2} + \cdots\right) = \prod_{n=1}^{N} \frac{1}{1 - \frac{1}{p_n}}$$

for all $N \geq 1$.

(b) Show that $(1 - \frac{1}{n})^n \geq \frac{1}{4}$ for all integers $n \geq 2$. Hint: if $n \geq 3$, write

$$\left(1 - \frac{1}{n}\right)^n = \sum_{k=0}^{n} \binom{n}{k} \frac{(-1)^k}{n^k} = \frac{n^2 - 1}{3n^2} + \sum_{k=4}^{n} \binom{n}{k} \frac{(-1)^k}{n^k}.$$

using the Binomial Theorem, then argue that $\sum_{k=4}^{n} \binom{n}{k} \frac{(-1)^k}{n^k} \geq 0$.

(c) Hence, show that

$$\prod_{n=1}^{N} \frac{1}{1 - \frac{1}{p_n}} \leq 4^{\sum_{n=1}^{N} \frac{1}{p_n}}$$

for all $N \geq 1$, and conclude that $\sum_{n=1}^{\infty} \frac{1}{p_n} = \infty$.

3.5 Series Rearrangements

Definition 3.50. A *rearrangement* of a sequence $\{a_n\}$ is a sequence $\{a_{\phi(n)}\}$ where $\phi : \mathbb{N} \to \mathbb{N}$ is a bijection.

Theorem 3.51. *Let $\sum_{n=1}^{\infty} a_n$ be absolutely convergent. Then $\sum_{n=1}^{\infty} a_{\phi(n)} = \sum_{n=1}^{\infty} a_n$ for any rearrangement $\{a_{\phi(n)}\}$ of $\{a_n\}$.*

Proof. Fix $\epsilon > 0$, and choose $N \in \mathbb{N}$ such that $\sum_{n=N+1}^{\infty} |a_n| < \epsilon$. Let

$$M = \max\{\phi^{-1}(a_1), \ldots, \phi^{-1}(a_N)\} \in \mathbb{N},$$

and observe that $\{a_1, \ldots, a_N\} \subset \{\phi(a_1), \ldots, \phi(a_M)\}$. Therefore, for all $m \geq M$,

$$\begin{aligned}
\left|\sum_{n=1}^{\infty} a_n - \sum_{n=1}^{m} a_{\phi(n)}\right| &= \left|\sum_{n=1}^{\infty} a_n - \sum_{n=1}^{N} a_n - \sum_{\substack{1\leq n\leq m \\ \phi(n)>N}} a_{\phi(n)}\right| \\
&\leq \left|\sum_{n=1}^{\infty} a_n - \sum_{n=1}^{N} a_n\right| + \left|\sum_{\substack{1\leq n\leq m \\ \phi(n)>N}} a_{\phi(n)}\right| \\
&\leq \left|\sum_{n=N+1}^{\infty} a_n\right| + \sum_{\substack{1\leq n\leq m \\ \phi(n)>N}} |a_{\phi(n)}| \\
&\leq \sum_{n=N+1}^{\infty} |a_n| + \sum_{n=N+1}^{\infty} |a_n| \\
&= 2\epsilon.
\end{aligned}$$

Hence, $\sum_{n=1}^{\infty} a_{\phi(n)} = \lim_{m\to\infty} \sum_{n=1}^{m} a_{\phi(n)} = \sum_{n=1}^{\infty} a_n$. $\square$

Theorem 3.52 (Riemann Rearrangement Theorem)**.** *Let $\{a_n\}$ be a real sequence such that $\sum_{n=1}^{\infty} a_n$ is conditionally convergent. Let $-\infty \leq x \leq y \leq \infty$. Then there is a rearrangement $\{a_{\phi(n)}\}$ of $\{a_n\}$ such that $\liminf_{N\to\infty} \sum_{n=1}^{N} a_{\phi(n)} = x$ and $\limsup_{N\to\infty} \sum_{n=1}^{N} a_{\phi(n)} = y$.*

The idea of the proof is simple. Since $\sum_{n=1}^{\infty} a_n$ is conditionally convergent, the nonnegative terms must sum to ∞ whereas the negative terms must sum to $-\infty$. If x and y are finite, then we can construct the desired rearrangement of $\{a_n\}$ by picking enough nonnegative terms until the running sum is greater than x, then picking enough negative terms until the running sum is less than y, and repeating this process. If x or y are infinite, then we need to slightly tweak the algorithm to make the running sum eventually diverge to ∞ or $-\infty$.

Proof. For all $m \geq 1$, let P_m be the m^{th} nonnegative term in the sequence $\{a_n\}$, and let N_m be the m^{th} negative term in the sequence $\{a_n\}$. If both $\sum_{m=1}^{\infty} P_m$ and $\sum_{n=1}^{\infty} N_m$ were finite, then $\sum_{n=1}^{\infty} a_n$ would converge absolutely, and if exactly one of $\sum_{m=1}^{\infty} P_m$ and $\sum_{m=1}^{\infty} N_m$ were finite, then $\sum_{n=1}^{\infty} a_n$ would diverge. These scenarios are impossible, so we must have $\sum_{m=1}^{\infty} P_m = \infty$ and $\sum_{m=1}^{\infty} N_m = -\infty$. Notice also that $\lim_{n\to\infty} a_n = 0$ by the Divergence Test, so $\lim_{m\to\infty} P_m = 0 = \lim_{m\to\infty} N_m$.

Suppose first that x and y are finite. Let $p_0 = 0$ and $n_0 = 0$. Suppose we have constructed p_{i-1} and n_{i-1} for some integer $i \geq 1$. We let p_i be the smallest integer such that $p_i > p_{i-1}$ and

$$\sum_{m=1}^{p_i} P_m + \sum_{m=1}^{n_{i-1}} N_m > x. \tag{1}$$

Then, we let n_i be the smallest integer such that $n_i > n_{i-1}$ and

$$\sum_{m=1}^{p_i} P_m + \sum_{m=1}^{n_i} N_m < y. \tag{2}$$

We know that p_i and n_i exist because $\sum_{m=1}^{\infty} P_m = \infty$ and $\sum_{m=1}^{\infty} N_m = -\infty$.

Now that we have the infinite increasing sequences $\{p_i\}$ and $\{n_i\}$, we form a rearrangement of $\{a_n\}$ by:

- picking the first p_1 terms from $\{P_m\}$,
- then picking the first n_1 terms from $\{N_m\}$,
- then picking the next $p_2 - p_1$ terms from $\{P_m\}$ (so that we have picked p_2 terms in total from $\{P_m\}$),
- then picking the next $n_2 - n_1$ terms from $\{N_m\}$ (so that we have picked n_2 terms in total from $\{N_m\}$),
- and so on.

Call the rearrangement $\{b_n\}$. Since the inequalities (1) and (2) hold infinitely often, we know that $\liminf_{N\to\infty} \sum_{n=1}^{N} b_n \leq y$ and $\limsup_{N\to\infty} \sum_{n=1}^{N} b_n \geq x$.

Now fix $\epsilon > 0$. Since $\lim_{m\to\infty} P_m = 0$, there exists $M \geq p_2$ such that $P_m < \epsilon$ for all $m \geq M$. Fix i such that $p_i \geq M$. By construction, $\sum_{n=1}^{p_i+n_{i-1}-1} b_n \leq x$, so $\sum_{n=1}^{p_i+n_{i-1}} b_n < x+\epsilon$ because $P_{p_i} < \epsilon$. Then, for all k such that $p_i+n_{i-1} < k \leq p_i+n_i$, b_k is negative, so $\sum_{n=1}^{k} b_n < x + \epsilon$. Next, for all k such that $p_i + n_i < k < p_{i+1} + n_i$, we have $\sum_{n=1}^{k} b_n \leq x < x + \epsilon$ by construction. Therefore, $\sum_{n=1}^{k} b_n < x+\epsilon$ for all $p_i+n_{i-1} \leq k \leq p_{i+1}+n_i$. Since i is arbitrary, it follows that $\sum_{n=1}^{k} b_n < x + \epsilon$ for all k large enough, so $\limsup_{N\to\infty} \sum_{n=1}^{N} b_n \leq x + \epsilon$. Therefore, $\limsup_{N\to\infty} \sum_{n=1}^{N} b_n = x$ since $\epsilon > 0$ is arbitrary. A similar argument shows that $\liminf_{N\to\infty} \sum_{n=1}^{N} b_n = y$.

If $x = \infty$, we need to modify (1) to instead say

$$\sum_{m=1}^{p_i} P_m + \sum_{m=1}^{n_{i-1}} N_m > i.$$

Then $\sum_{n=1}^{p_i+n_{i-1}} b_n > i$ for all i, so $\limsup_{N\to\infty} \sum_{n=1}^{N} b_n = \infty$. If $x = -\infty$, then we modify (1) and (2) to

$$\sum_{m=1}^{p_i} P_m + \sum_{m=1}^{n_{i-1}} N_m > -i + 2$$

and

$$\sum_{m=1}^{p_i} P_m + \sum_{m=1}^{n_i} N_m < -i - 2,$$

respectively. In this case, the rearranged sum diverges to $-\infty$. We can handle the cases $y = \pm\infty$ in a similar way. $\square$

Theorem 3.53. *Let $\{a_{n,m}\}$ be a double complex sequence (i.e. a function $\mathbb{N} \times \mathbb{N} \to \mathbb{C}$) such that $\sum_{n=1}^{\infty} \sum_{m=1}^{\infty} |a_{n,m}| < \infty$. Then $\sum_{n=1}^{\infty} \sum_{m=1}^{\infty} a_{n,m} = \sum_{m=1}^{\infty} \sum_{n=1}^{\infty} a_{n,m}$.*

Proof. By assumption, $\sum_{n=1}^{\infty} \sum_{m=1}^{\infty} a_{n,m}$ is absolutely convergent and hence convergent. Fix $\epsilon > 0$. Choose $N_0 \geq 1$ such that $\sum_{n=N+1}^{\infty} \sum_{m=1}^{\infty} |a_{n,m}| < \epsilon$ for any $N \geq N_0$. For each integer n such that $1 \leq n \leq N_0$, $\sum_{m=1}^{\infty} |a_{n,m}|$ converges by assumption, so we can choose $T_n \geq 1$ such that $\sum_{m=M+1}^{\infty} |a_{n,m}| < 2^{-n}\epsilon$ for any $M \geq T_n$. Let $M_1 = \max\{T_1, \ldots, T_{N_0}\} \geq 1$. Suppose $M \geq M_1$ and $N \geq N_0$. Notice that

$$\begin{aligned}
\sum_{n=1}^{\infty} \sum_{m=M+1}^{\infty} |a_{n,m}| &\leq \sum_{n=1}^{\infty} \sum_{M_1+1}^{\infty} |a_{n,m}| \\
&= \sum_{n=1}^{N_0} \sum_{m=M_1+1}^{\infty} |a_{n,m}| + \sum_{n=N_0+1}^{\infty} \sum_{m=M_1+1}^{\infty} |a_{n,m}| \\
&< \sum_{n=1}^{N_0} 2^{-n}\epsilon + \sum_{n=N_0+1}^{\infty} \sum_{m=1}^{\infty} |a_{n,m}| \\
&< \epsilon \sum_{n=1}^{\infty} 2^{-n} + \epsilon \\
&= \epsilon \left(\frac{1/2}{1 - 1/2} \right) + \epsilon \\
&= 2\epsilon.
\end{aligned}$$

Hence,

$$\begin{aligned}
\left| \sum_{n=1}^{\infty} \sum_{m=1}^{\infty} a_{n,m} - \sum_{m=1}^{M} \sum_{n=1}^{N} a_{n,m} \right| &= \left| \sum_{n=1}^{\infty} \sum_{m=1}^{\infty} a_{n,m} - \sum_{n=1}^{N} \sum_{m=1}^{M} a_{n,m} \right| \\
&\leq \left| \sum_{n=1}^{\infty} \sum_{m=1}^{\infty} a_{n,m} - \sum_{n=1}^{\infty} \sum_{m=1}^{M} a_{n,m} \right| + \left| \sum_{n=1}^{\infty} \sum_{m=1}^{M} a_{n,m} - \sum_{n=1}^{N} \sum_{m=1}^{M} a_{n,m} \right| \\
&= \left| \sum_{n=1}^{\infty} \sum_{m=M+1}^{\infty} a_{n,m} \right| + \left| \sum_{n=N+1}^{\infty} \sum_{m=1}^{M} a_{n,m} \right| \\
&\leq \sum_{n=1}^{\infty} \sum_{m=M+1}^{\infty} |a_{n,m}| + \sum_{n=N+1}^{\infty} \sum_{m=1}^{M} |a_{n,m}| \\
&< 2\epsilon + \sum_{n=N+1}^{\infty} \sum_{m=1}^{\infty} |a_{n,m}|
\end{aligned}$$

$$
\begin{aligned}
&< 2\epsilon + \epsilon \quad \text{since } N \geq N_0 \\
&= 3\epsilon.
\end{aligned}
$$

Given any $m \geq 1$, $\sum_{n=1}^{\infty} |a_{n,M}|$ converges by the Comparison Test since $\sum_{n=1}^{\infty} \left(\sum_{m=1}^{\infty} |a_{n,m}|\right)$ converges and $0 \leq |a_{n,M}| \leq \sum_{m=1}^{\infty} |a_{n,m}|$ for all $n \geq 1$. Hence, for all $1 \leq m \leq M$, there exists $U_m \geq 1$ such that $\sum_{n=N_1+1}^{\infty} |a_{n,m}| < 2^{-m}\epsilon$ for all $N_1 \geq U_m$. Let $N_1 = \max\{U_1, \ldots, U_M, N_0\}$, so that

$$\sum_{m=1}^{M} \sum_{n=N_1+1}^{\infty} |a_{n,m}| < \sum_{m=1}^{M} 2^{-m}\epsilon < \epsilon \sum_{m=1}^{\infty} 2^{-m} = \epsilon.$$

Then

$$
\begin{aligned}
\left| \sum_{n=1}^{\infty} \sum_{m=1}^{\infty} a_{n,m} - \sum_{m=1}^{M} \sum_{n=1}^{\infty} a_{n,m} \right| &\leq \left| \sum_{n=1}^{\infty} \sum_{m=1}^{\infty} a_{n,m} - \sum_{m=1}^{M} \sum_{n=1}^{N_1} a_{n,m} \right| + \left| \sum_{m=1}^{M} \sum_{n=1}^{N_1} a_{n,m} - \sum_{m=1}^{M} \sum_{n=1}^{\infty} a_{n,m} \right| \\
&< 3\epsilon + \left| \sum_{m=1}^{M} \sum_{n=N_1+1}^{\infty} a_{n,m} \right| \quad \text{since } N_1 \geq N_0 \\
&\leq 3\epsilon + \sum_{m=1}^{M} \sum_{n=N_1+1}^{\infty} |a_{n,m}| \\
&< 3\epsilon + \epsilon \\
&= 4\epsilon.
\end{aligned}
$$

Since $M \geq M_1$ is arbitrary, $\sum_{m=1}^{\infty} \sum_{n=1}^{\infty} a_{n,m}$ converges to $\sum_{n=1}^{\infty} \sum_{m=1}^{\infty} a_{n,m}$. □

Exercise 3.54. Let

$$\zeta(n) = \sum_{k=1}^{\infty} \frac{1}{k^n}$$

for all integers $n \geq 2$. Find the value of

$$\sum_{n=1}^{\infty} \frac{\zeta(2n)}{4^n}.$$

3.6 Euler's Number

Since

$$\lim_{n\to\infty} \left| \frac{1}{(n+1)!} \bigg/ \frac{1}{n!} \right| = \lim_{n\to\infty} \frac{1}{n+1} = 0 < 1,$$

the series $\sum_{n=0}^{\infty} \frac{1}{n!}$ converges by the Ratio Test.

Definition 3.55. *Euler's number* is $e := \sum_{n=0}^{\infty} \frac{1}{n!}$.

Theorem 3.56. *e is irrational.*

Proof. For any $N \geq 0$, we have that

$$
\begin{aligned}
0 < e - \sum_{n=0}^{N} \frac{1}{n!} &= \sum_{n=N+1}^{\infty} \frac{1}{n!} \\
&= \frac{1}{(N+1)!} \sum_{n=0}^{\infty} \prod_{k=1}^{n} \frac{1}{N+1+k} \\
&\leq \frac{1}{(N+1)!} \sum_{n=0}^{\infty} \frac{1}{(N+2)^n} \\
&= \frac{1}{(N+1)!} \left(\frac{1}{1 - \frac{1}{N+2}} \right) \\
&= \frac{1}{(N+1)!} \left(\frac{N+2}{N+1} \right) \\
&\leq \frac{1}{(N+1)!} \left(\frac{2N+2}{N+1} \right) \\
&= \frac{2}{(N+1)!}.
\end{aligned}
$$

Hence,

$$
0 < N!e - \sum_{n=0}^{N} \frac{N!}{n!} \leq \frac{2}{N+1} \tag{3}
$$

for all $N \geq 0$.

Suppose e is rational. Then there exists $p \in \mathbb{Z}$ and $q \in \mathbb{N}$ such that $e = \frac{p}{q}$. For all $N \geq 0$, equation (3) says that

$$
0 < N! \left(\frac{p}{q} \right) - \sum_{n=0}^{N} \frac{N!}{n!} \leq \frac{2}{N+1}.
$$

Hence,

$$
0 < N!p - q \sum_{n=0}^{N} \frac{N!}{n!} \leq \frac{2q}{N+1}.
$$

In particular, setting $N = 2q$ gives

$$
0 < (2q)!p - q \sum_{n=0}^{2q} \frac{(2q)!}{n!} \leq \frac{2q}{2q+1} < 1.
$$

Now $(2q)!p - q \sum_{n=0}^{2q} \frac{(2q)!}{n!}$ is an integer since $\frac{(2q)!}{n!}$ is an integer for all $0 \leq n \leq 2q$. But there is no integer strictly between 0 and 1, so we have a contradiction. Therefore, e is irrational. $\square$

Theorem 3.57. $e = \lim_{n \to \infty} (1 + \frac{1}{n})^n$.

Proof. By the Binomial Theorem,

$$
\left(1 + \frac{1}{n}\right)^n = \sum_{k=0}^{n} \binom{n}{k} \left(\frac{1}{n}\right)^k
$$

$$
\begin{aligned}
&= 1 + \sum_{k=1}^{n} \frac{n(n-1)\cdots(n-k+1)}{k!n^k} \\
&= 1 + \sum_{k=1}^{n} \frac{1}{k!}\left[\left(\frac{n}{n}\right)\left(\frac{n-1}{n}\right)\cdots\left(\frac{n-k+1}{n}\right)\right] \\
&= 1 + \sum_{k=1}^{n} \frac{1}{k!}\left[\left(1-\frac{1}{n}\right)\cdots\left(1-\frac{k-1}{n}\right)\right]
\end{aligned}
$$

for all $n \in \mathbb{N}$. Set $A_{k,n} = (1-\frac{1}{n})\cdots(1-\frac{k-1}{n})$ for all $1 \leq k \leq n$, and notice that

$$0 < A_{n,n} < A_{n-1,n} < \cdots < A_{1,n} = 1.$$

Fix $\epsilon > 0$. Then there exists $N_1 \in \mathbb{N}$ such that $\left|\sum_{k=n+1}^{\infty} \frac{1}{n!}\right| < \epsilon$ whenever $n \geq N_1$. Since $\lim_{n\to\infty} A_{N_1,n} = 1$, there exists $N_2 \geq N_1$ such that $1 - A_{N_1,n} < \epsilon$ for all $n \geq N_2$. Suppose $n \geq N_2$. Then

$$
\begin{aligned}
\left|\left(1+\frac{1}{n}\right)^n - \sum_{k=0}^{\infty}\frac{1}{k!}\right| &= \left|\sum_{k=1}^{n}\frac{1}{k!}A_{k,n} - \sum_{k=1}^{\infty}\frac{1}{k!}\right| \\
&\leq \left|\sum_{k=1}^{n}\frac{1}{k!}A_{k,n} - \sum_{k=1}^{n}\frac{1}{k!}\right| + \left|\sum_{k=n+1}^{\infty}\frac{1}{k!}\right| \\
&< \left|\sum_{k=1}^{n}\frac{1}{k!}(A_{k,n}-1)\right| + \epsilon \quad \text{since } n \geq N_2 \geq N_1 \\
&\leq \sum_{k=1}^{n}\frac{1}{k!}(1-A_{k,n}) + \epsilon \quad \text{since } A_{k,n} < 1 \\
&\leq \sum_{k=1}^{N_1}\frac{1}{k!}(1-A_{N_1,n}) + \sum_{k=N_1+1}^{n}\frac{1}{k!}(1-A_{k,n}) + \epsilon \\
&\qquad \text{since } A_{N_1,n} \leq A_{k,n} \text{ for all } 1 \leq k \leq N_1 \\
&< \sum_{k=1}^{N_1}\frac{1}{k!}\epsilon + \sum_{k=N_1+1}^{n}\frac{1}{k!} + \epsilon \quad \text{since } n \geq N_2 \text{ and } 1 - A_{k,n} \leq 1 \\
&< \epsilon\sum_{k=1}^{\infty}\frac{1}{k!} + \sum_{k=N_1+1}^{\infty}\frac{1}{k!} + \epsilon \\
&< \epsilon e + \epsilon + \epsilon \quad \text{by definition of } N_1 \\
&= (e+2)\epsilon.
\end{aligned}
$$

Since $e+2$ is independent of ϵ, the desired result follows. $\square$

Exercise 3.58. Let $b_n \in \{0,1\}$ for all $n \in \mathbb{N}$. Prove that $\sum_{n=1}^{\infty} \frac{b_n}{n!}$ is irrational if and only if there are infinitely many integers $n \geq 1$ such that $b_n = 1$.

Exercise 3.59. Fix $z \in \mathbb{C}$. Adapt the proof of Theorem 3.57 to show that

$$\lim_{n\to\infty}\left(1+\frac{z}{n}\right)^n = \sum_{k=0}^{\infty}\frac{z^k}{k!}.$$

Exercise 3.60. Let S be the set of real numbers $c > 0$ such that

$$\sum_{n=1}^{\infty} \frac{n^n}{c^n n!}$$

converges. Find $\inf(S)$, and compute the famous limit

$$\lim_{n\to\infty} \frac{n}{\sqrt[n]{n!}}.$$

3.7 Characterizations of Compactness

In this section, we look at a generalization of the Heine–Borel Theorem. In general, a closed and bounded subset of a metric space is not necessarily compact—Exercise 3.65 gives an example. In this section, we will give some necessary and sufficient conditions for a metric space to be compact.

Fix a metric space (K, d).

Definition 3.61. K is *limit-point compact* if every infinite subset of K has a limit point.

Definition 3.62. K is *sequentially compact* if every sequence in K has a convergent subsequence.

Definition 3.63. K is *totally bounded* if for all $r > 0$, there exists $x_1, \ldots, x_n \in K$ such that $\{N_r(x_1), \ldots, N_r(x_n)\}$ covers K.

If $K \subset X$, $r > 0$, and $x_1, \ldots, x_n \in K$, then $K \subset \bigcup_{j=1}^{n} N_r^K(x_j)$ if and only if $K \subset \bigcup_{j=1}^{n} N_r^X(x_j)$ since $N_r^K(x_j) = N_r^X(x_j) \cap K$. Therefore, total boundedness is an intrinsic property of a metric space; we do not need to specify that K is totally bounded "relative to" an ambient metric space X.

Theorem 3.64. *The following are equivalent:*

(i) K is compact.
(ii) K is limit-point compact.
(iii) K is sequentially compact.
(iv) K is complete and totally bounded.

Proof. (i) $\implies$ (ii): this is Theorem 2.44.

(ii) $\implies$ (iii): let K be limit-point compact, and let $\{x_n\}$ be a sequence in K. Let $E = \{x_n \mid n \in \mathbb{N}\}$. If E is finite, then $\{x_n\}$ has infinitely occurrences of some term $L \in K$, which means $\{x_n\}$ has a convergent subsequence $\{L, L, L, \ldots\}$. If E is infinite, then E' is non-empty since K is limit-point compact, so $\{x_n\}$ has a convergent subsequence by Theorem 3.12.

(iii) $\implies$ (iv): if $\{s_n\}$ is a Cauchy sequence, then $\{s_n\}$ has a convergent subsequence since K is sequentially compact, so $\{s_n\}$ converges by Lemma 3.14. Hence, K is complete. Suppose, by way of contradiction, that K is not totally bounded. Pick $r > 0$ such that no finite collection of neighbourhoods $\{N_r(x_1), \ldots, N_r(x_n)\}$ (where $x_1, \ldots, x_n \in K$) covers K. Then K is non-empty (since an empty collection of neighbourhoods would cover the empty set), so pick $t_1 \in K$. After picking $t_i \in K$ for some $i \geq 1$, pick $t_{i+1} \in K$ such that $t_{i+1} \notin \bigcup_{j=1}^{i} N_r(t_j)$; this is always possible since $\{N_r(t_1), \ldots, N_r(t_i)\}$ does not cover K by assumption. This recursive process produces a sequence $\{t_n\}$ in K such that $d(t_i, t_j) \geq r$ whenever $i > j$. Observe that no subsequence of $\{t_n\}$ is Cauchy, so $\{t_n\}$ has no convergent subsequence, contradicting the assumption that K is sequentially compact. Therefore, K must be totally bounded.

(iv) $\implies$ (i): this proof will be conceptually similar to the proof of the Heine–Borel Theorem. Suppose, by way of contradiction, that K is complete and totally bounded but not compact. In this argument, all neighbourhoods are relative to K. Let $\{\mathcal{O}_\alpha\}_\alpha$ be a cover of K that has no finite subcover. We claim that there exists a sequence $\{y_n\}$ in K such that for all $n \in \mathbb{N}$,

- $d(y_{n+1}, y_n) < \frac{3}{2^{n+1}}$ and
- no finite subcover of $\{\mathcal{O}_\alpha\}_\alpha$ covers $N_{1/(2^n)}(y_n)$.

Since K is totally bounded, there exist $x_1^{(1)}, \ldots, x_m^{(1)} \in K$ such that $\{N_{1/2}(x_j^{(1)})\}_{j=1}^m$ covers K. Here, the subscript "(1)" is just a second index. Suppose, by way of contradiction, that for every $1 \leq j \leq m$, some finite $S_j \subset \{\mathcal{O}_\alpha\}_\alpha$ covers $N_{1/2}(x_j^{(1)}) \cap K$. Then $\bigcup_{j=1}^m S_j$ covers $\bigcup_{j=1}^m [N_{1/2}(x_j^{(1)}) \cap K] = K$. But $\bigcup_{j=1}^m S_j$ is a finite subcover of $\{\mathcal{O}_\alpha\}_\alpha$, so we have a contradiction. Therefore, there must be a neighbourhood $N_{1/2}(x_{j_1}^{(1)})$ such that no finite subcover of $\{\mathcal{O}_\alpha\}_\alpha$ covers $N_{1/2}(x_{j_1}^{(1)}) \cap K$. Let $y_1 = x_{j_1}^{(1)} \in K$. Then no finite subcover of $\{\mathcal{O}_\alpha\}_\alpha$ covers $N_{1/2}(y_1)$.

Suppose that for some $n \in \mathbb{N}$, we have chosen $y_n \in K$ such that no finite subcover of $\{\mathcal{O}_\alpha\}_\alpha$ covers $N_{1/(2^n)}(y_n)$. Since K is totally bounded, there exist $x_1^{(n+1)}, \ldots, x_\ell^{(n+1)} \in K$ such that $\{N_{1/(2^{n+1})}(x_j^{(n+1)})\}_{j=1}^\ell$ covers K. By the same argument as the above paragraph, there must be a neighbourhood $N_{1/(2^{n+1})}(x_{j_{n+1}}^{(n+1)})$ such that no finite subset of $\{\mathcal{O}_\alpha\}_\alpha$ covers $N_{1/(2^{n+1})}(x_{j_{n+1}}^{(n+1)}) \cap N_{1/(2^n)}(y_n)$, since no finite subset of $\{\mathcal{O}_\alpha\}_\alpha$ covers $N_{1/(2^n)}(y_n)$. Let $y_{n+1} = x_{j_{n+1}}^{(n+1)} \in K$. Then no finite subset of $\{\mathcal{O}_\alpha\}_\alpha$ can cover $N_{1/(2^{n+1})}(y_{n+1})$. Note that $N_{1/(2^{n+1})}(y_{n+1}) \cap N_{1/(2^n)}(y_n)$ must be non-empty since the empty set is trivially covered by an empty subcover. Pick $t \in N_{1/(2^{n+1})}(y_{n+1}) \cap N_{1/(2^n)}(y_n)$. Then

$$d(y_{n+1}, y_n) \leq d(y_{n+1}, t) + d(t, y_n) < \frac{1}{2^{n+1}} + \frac{1}{2^n} = \frac{3}{2^{n+1}}.$$

By induction, the sequence $\{y_n\}$ exists, as claimed. Now we want to show that $\{y_n\}$ is Cauchy. For any integers $m \geq n \geq 1$, we have that

$$d(y_n, y_m) \leq \sum_{j=n}^{m-1} d(y_j, y_{j+1}) < \sum_{j=n}^{m-1} \frac{3}{2^{j-1}} \leq \sum_{j=n}^{\infty} \frac{3}{2^{j-1}} = \frac{3}{2^{n-2}}.$$

Fix $\epsilon > 0$. Then there exists $N \in \mathbb{N}$ such that $\frac{3}{2^{N-2}} < \epsilon$ since $\lim_{n\to\infty} \frac{3}{2^{n-2}} = 0$. Suppose $m \geq n \geq N$. Then $n - N \geq 0$, so $\frac{1}{2^{n-N}} \leq 1$. Hence,

$$d(y_n, y_m) \leq \frac{3}{2^{n-2}} = \frac{1}{2^{n-N}} \left(\frac{3}{2^{N-2}} \right) \leq \frac{3}{2^{N-2}} < \epsilon,$$

so $\{y_n\}$ is Cauchy. Since K is complete, $\lim_{n\to\infty} y_n$ exists. Let $y = \lim_{n\to\infty} y_n \in K$. Then $y \in \mathcal{O}_{\alpha_0}$ for some index α_0, so there exists $r > 0$ such that $N_r(y) \subset \mathcal{O}_{\alpha_0}$. Pick $m \in \mathbb{N}$ such that $\frac{1}{2^m} < \frac{r}{2}$ and $d(y_m, y) < \frac{r}{2}$; this is possible since $\lim_{n\to\infty} \frac{1}{2^n} = 0$ and $\lim_{n\to\infty} y_n = y$. Let $t \in N_{1/(2^m)}(y_m)$. Then

$$d(t, y) \le d(t, y_m) + d(y_m, y) < \frac{1}{2^m} + \frac{r}{2} < \frac{r}{2} + \frac{r}{2} = r.$$

Hence, $N_{1/(2^m)}(y_m) \subset N_r(y) \subset \mathcal{O}_{\alpha_0}$. This is a contradiction since $N_{1/(2^m)}(y_m)$ is not covered by any finite subcover of $\{\mathcal{O}_\alpha\}_\alpha$. $\square$

Exercise 3.65. Let $\mathbb{R}_{\text{discrete}}$ be the set $\mathbb{R}$ equipped with the discrete metric

$$d(x, y) = \begin{cases} 1 & \text{if } x \neq y \\ 0 & \text{otherwise.} \end{cases}$$

(a) Prove that every subset of $\mathbb{R}_{\text{discrete}}$ is closed, bounded, and complete.
(b) Prove that no infinite subset of $\mathbb{R}_{\text{discrete}}$ is compact.

Exercise 3.66. Without using Theorem 3.64 directly, prove that the set

$$S = \{q \in \mathbb{Q} \mid q \ge 0 \text{ and } q^2 < 2\}$$

is closed and totally bounded in $\mathbb{Q}$ but is not compact.

4 Continuity

Fix metric spaces (X, d_X), (Y, d_Y), and (Z, d_Z).

4.1 Limits of Functions Between Metric Spaces

Definition 4.1. Let $E \subset X$, $f : E \to Y$ and $p \in E'$. We write "$\lim_{x \to p} f(x) = q$" (where $q \in Y$) if for all $\epsilon > 0$, there exists $\delta > 0$ such that $d_Y(f(x), q) < \epsilon$ whenever $0 < d_X(x, p) < \delta$.

We say that $\lim_{x \to p} f(x)$ *exists* if there exists $q \in Y$ such that $\lim_{x \to p} f(x) = q$. If there is no such q, then $\lim_{x \to p} f(x)$ does not exist.

Theorem 4.2. *Let $E \subset X$, $f : E \to Y$, $p \in E'$, and $q \in Y$. Then $\lim_{x \to p} f(x) = q$ if and only if $\lim_{n \to \infty} f(x_n) = q$ for every sequence $\{x_n\}$ in E such that $\lim_{n \to \infty} x_n = p$.*

Proof. Suppose $\lim_{x \to p} f(x) = q$. Fix $\epsilon > 0$, and choose $\delta > 0$ such that $d_Y(f(x), q) < \epsilon$ whenever $0 < d_X(x, p) < \delta$. Let $\{x_n\}$ be a sequence in E such that $\lim_{n \to \infty} x_n = p$. Since $\delta > 0$, there exists $N \in \mathbb{N}$ such that $d_X(x_n, p) < \delta$ for all $n \geq N$. Hence, $d_Y(f(x_n), q) < \epsilon$ for all $n \geq N$, so $\lim_{n \to \infty} f(x_n) = q$.

Conversely, suppose $\lim_{n \to \infty} f(x_n) = q$ for every sequence $\{x_n\}$ in E that converges to p. Suppose by way of contradiction that $\lim_{x \to p} f(x) \neq q$. Then there exists $\epsilon > 0$ such that for all $\delta > 0$, there exists $x \in E$ such that $0 < d_X(x, p) < \delta$ and $d_Y(f(x), q) \geq \epsilon$. Therefore, for all $n \in \mathbb{N}$, there exists $x_n \in E$ such that $0 < d_X(x_n, p) < \frac{1}{n}$ and $d_Y(f(x_n), q) \geq \epsilon$. Note that $\lim_{n \to \infty} x_n = p$ since $\lim_{n \to \infty} \frac{1}{n} = 0$, but $\lim_{n \to \infty} f(x_n) \neq q$ since $d_Y(f(x_n), q) \geq \epsilon$ for all $n \in \mathbb{N}$. This is a contradiction, so $\lim_{x \to p} f(x) = q$. $\square$

Because of this connection between limits of functions and limits of sequences, many results from Chapter 3 about limits of sequences have analogous results in terms of functions. For example, the analogue of Proposition 3.3 is the following:

Proposition 4.3. *Let $E \subset X$; $f : E \to Y$; $p \in E'$; and $q_1, q_2 \in Y$. If $\lim_{x \to p} f(x) = q_1$ and $\lim_{x \to p} f(x) = q_2$, then $q_1 = q_2$.*

The "Squeeze Theorem" is the analogue of Proposition 3.7:

Proposition 4.4 (Squeeze Theorem)**.** *Let $E \subset \mathbb{R}$; $f, g, h : E \to \mathbb{R}$; $p \in E'$; and $L \in \mathbb{R}$. Suppose there exists $r > 0$ such that $f(x) \leq g(x) \leq h(x)$ for all $x \in E \cap N_r(p)$, and suppose $\lim_{x \to p} f(x) = L = \lim_{x \to p} h(x)$. Then $\lim_{x \to p} g(x) = L$.*

We also have the familiar "limit laws" from calculus (the analogue of Proposition 3.9):

Proposition 4.5. *Let $E \subset \mathbb{C}$; $f, g : E \to \mathbb{C}$; $p \in E'$; and $L_1, L_2 \in \mathbb{C}$. Suppose $\lim_{x \to p} f(x) = L_1$ and $\lim_{x \to p} g(x) = L_2$. Then:*

(a) $\lim_{x \to p}[f(x) + g(x)] = L_1 + L_2$.

(b) $\lim_{x \to p}[f(x) - g(x)] = L_1 - L_2$.

(c) $\lim_{x \to p} f(x)g(x) = L_1 L_2$.

(d) $\lim_{x \to p} \frac{f(x)}{g(x)} = \frac{L_1}{L_2}$ *if* $L_2 \neq 0$.

Each of these results can be proved by using Theorem 4.2 and appealing to the analogous result for sequences.

4.2 Continuity, Limits, and Open Sets

Definition 4.6. Let $f : X \to Y$ and $x_0 \in X$. We say that f is *continuous at* x_0 if for all $\epsilon > 0$, there exists $\delta > 0$ such that $d_Y(f(x), f(x_0)) < \epsilon$ whenever $d_X(x, x_0) < \delta$. If f is continuous at every point in X, then f is *continuous* (on X).

Theorem 4.7. *Let $f : X \to Y$ and $x_0 \in X$. Then f is continuous at x_0 if and only if either x_0 is an isolated point or $\lim_{x\to x_0} f(x) = f(x_0)$.*

Proof. Suppose f is continuous at x_0, and suppose x_0 is not an isolated point. Then $x_0 \in X'$. Fix $\epsilon > 0$, and choose $\delta > 0$ such that $d_Y(f(x), f(x_0)) < \epsilon$ whenever $d_X(x, x_0) < \delta$. It is clear that if $0 < d_X(x, x_0) < \delta$, then $d_X(x, x_0) < \delta$, so $d_Y(f(x), f(x_0)) < \epsilon$. Therefore, $\lim_{x\to x_0} f(x) = f(x_0)$.

Conversely, suppose x_0 is either an isolated point or $\lim_{x\to x_0} f(x) = f(x_0)$. We consider the two cases separately. Suppose x_0 is an isolated point. Then there exists $\delta > 0$ such that the only $x \in X$ satisfying $d_X(x, x_0) < \delta$ is $x = x_0$. Hence, for any fixed $\epsilon > 0$, if $d_X(x, x_0) < \delta$, then $d_Y(f(x), f(x_0)) = d_Y(f(x_0), f(x_0)) = 0 < \epsilon$. Therefore, f is continuous at x_0 . Now suppose $\lim_{x\to x_0} f(x) = f(x_0)$. For any fixed $\epsilon > 0$, there exists $\delta > 0$ such that $d_Y(f(x), f(x_0)) < \epsilon$ whenever $0 < d_X(x, x_0) < \delta$. Now clearly $d_Y(f(x_0), f(x_0)) = 0 < \epsilon$, so we have that $d_Y(f(x), f(x_0)) < \epsilon$ whenever $d_X(x, x_0) < \delta$. Therefore, f is continuous at x_0. □

Theorem 4.8 (The Topological Definition of Continuity)**.** *A function $f : X \to Y$ is continuous on X if and only if the pre-image of every open subset of Y under f is an open subset of X.*

Proof. Suppose $f : X \to Y$ is continuous on X. Let $E \subset Y$ be open. Let $x_0 \in f^{-1}(E)$, so that $f(x_0) \in E$. Since E is open, there exists $\epsilon > 0$ such that $N^Y_\epsilon(f(x_0)) \subset E$. Choose $\delta > 0$ such that $d_Y(f(x), f(x_0)) < \epsilon$ for all $d_X(x, x_0) < \delta$. It follows that if $x \in N^X_\delta(x_0)$, then $f(x) \in N^Y_\epsilon(f(x_0)) \subset E$, so $x \in f^{-1}(E)$. Therefore, $N^X_\delta(x_0) \subset f^{-1}(E)$, so $f^{-1}(E)$ is open in X.

Conversely, suppose $f^{-1}(E) \subset X$ is open for every open set $E \subset Y$. Fix $x_0 \in X$ and $\epsilon > 0$. Then $N^Y_\epsilon(f(x_0)) \subset Y$ is open, so $f^{-1}(N^Y_\epsilon(x_0)) \subset X$ is open. Note that $f(x_0) \in N^Y_\epsilon(f(x_0))$, so $x_0 \in f^{-1}(N^Y_\epsilon(f(x_0)))$. Hence, there exists $\delta > 0$ such that $N^X_\delta(x_0) \subset f^{-1}(N^Y_\epsilon(f(x_0)))$. Now let $x \in X$ such that $d_X(x, x_0) < \delta$. Then $x \in N^X_\delta(x_0) \subset f^{-1}(N^Y_\epsilon(f(x_0)))$, so $f(x) \in N^Y_\epsilon(f(x_0))$, which means $d_Y(f(x), f(x_0)) < \epsilon$. Therefore, f is continuous at x_0. Since $x_0 \in X$ is arbitrary, f is continuous on X. □

Corollary 4.8.1. *A function $f : X \to Y$ is continuous on X if and only if the pre-image of every closed subset of Y under f is a closed subset of X.*

Proof. Let $f : X \to Y$. We will use the fact that $f^{-1}(E^c) = (f^{-1}(E))^c$ for all $E \subset Y$. Suppose f is continuous on X. Let $E \subset Y$ be closed. Then $E^c \subset Y$ is open, so $f^{-1}(E^c) = (f^{-1}(E))^c \subset X$ is open. Hence, $f^{-1}(E) = ((f^{-1}(E))^c)^c \subset X$ is closed. Conversely, suppose $f^{-1}(E) \subset X$ is closed for every closed $E \subset Y$. Let $S \subset Y$ be open. Then $S^c \subset Y$ is closed, so $f^{-1}(S^c) = (f^{-1}(S))^c \subset X$ is closed. Therefore, $f^{-1}(S) = ((f^{-1}(S))^c)^c \subset X$ is open, so f is continuous because the pre-image of every open set in Y under f is open in X. □

Theorem 4.9. *Let $f : X \to Y$. Suppose $\{x_n\}$ is a sequence in X converging to $x_* \in X$, and suppose f is continuous at x_*. Then $\lim_{n\to\infty} f(x_n) = f(x_*)$.*

Proof. Fix $\epsilon > 0$. Since f is continuous at x_*, there exists $\delta > 0$ such that $d_Y(f(x), f(x_*)) < \epsilon$ whenever $d_X(x, x_*) < \delta$. Since $\lim_{n\to\infty} x_n = x$, there exists $N \in \mathbb{N}$ such that $d_X(x_n, x_*) < \delta$ for all $n \geq N$. If $n \geq N$, then $d_X(x_n, x_*) < \delta$, so $d_Y(f(x_n), f(x_*)) < \epsilon$. Therefore, $\lim_{n\to\infty} f(x_n) = f(x_*)$. $\square$

Theorem 4.10. *Let $E \subset X$, $g : E \to Y$, and $f : Y \to Z$. Suppose f is continuous at $y_0 \in Y$, and suppose $x_0 \in E'$ satisfies $\lim_{x\to x_0} g(x) = y_0$. Then $\lim_{x\to x_0} f(g(x)) = f(y_0)$.*

Proof. Let $\{x_n\}$ be any sequence in E converging to x_0. Then $\{g(x_n)\}$ converges to y_0 by Theorem 4.2 since $\lim_{x\to x_0} g(x) = y_0$. Since f is continuous at y_0, Theorem 4.9 implies that $\lim_{n\to\infty} f(g(x_n)) = f(y_0)$. Since $\{x_n\}$ is an arbitrary sequence satisfying $\lim_{n\to\infty} x_n = x_0$, Theorem 4.2 says that $\lim_{x\to x_0} f(g(x)) = f(y_0)$. $\square$

Theorem 4.11. *If $g : X \to Y$ and $f : Y \to Z$ are continuous, then $f \circ g : X \to Z$ is continuous.*

Proof. Fix $x_0 \in X$. If x_0 is an isolated point of X, then $f \circ g$ is continuous at x_0 by Theorem 4.7. Suppose x_0 is not an isolated point of X. Then $\lim_{x\to x_0} g(x) = g(x_0)$ by Theorem 4.7 since g is continuous at x_0. Since f is continuous at $g(x_0)$, we have that $\lim_{x\to x_0} f(g(x)) = f(g(x_0))$ by Theorem 4.10. Therefore, $f \circ g$ is continuous at x_0 by Theorem 4.7. Since $x_0 \in X$ is arbitrary, $f \circ g$ is continuous on X. $\square$

A simpler proof. Let $E \subset Z$ be open. Then $f^{-1}(E) \subset Y$ is open, so $g^{-1}(f^{-1}(E)) \subset X$ is open. Now note that $g^{-1}(f^{-1}(E)) = (f \circ g)^{-1}(E)$ because $x \in g^{-1}(f^{-1}(E))$ if and only if $f(g(x)) \in E$ if and only if $x \in (f \circ g)^{-1}(E)$. Therefore, $(f \circ g)^{-1}(E) \subset X$ is open for every open $E \subset Z$, so $f \circ g$ is continuous. $\square$

Exercise 4.12. Prove that $f : X \to Y$ is continuous if and only if $f(\overline{E}) \subset \overline{f(E)}$ for all $E \subset X$.

Exercise 4.13. Suppose $f : \mathbb{R}^k \to \mathbb{R}$ is a continuous function such that $\lim_{n\to\infty} f(x_n) = \infty$ for all sequences $\{x_n\}$ in $\mathbb{R}^k$ such that $\lim_{n\to\infty} ||x_n|| = \infty$. Prove that there exists $x^* \in \mathbb{R}^k$ such that $f(x^*) \leq f(x)$ for all $x \in \mathbb{R}^k$. Hint: first construct a sequence $\{x_n\}$ such that $\lim_{n\to\infty} f(x_n) = \inf_x f(x)$.

Exercise 4.14. Let $\mathcal{C}$ be the set of continuous functions $\mathbb{R} \to \mathbb{R}$. In this exercise, we will show that $|\mathcal{C}| = |\mathbb{R}|$.

(a) Suppose $f, g \in \mathcal{C}$ satisfy $f(q) = g(q)$ for all $q \in \mathbb{Q}$. Prove that $f = g$, and hence infer that $|\mathcal{C}| \leq |\mathbb{R}^{\mathbb{Q}}|$ (recall the notation $\mathbb{R}^{\mathbb{Q}}$ from Exercise 2.10).

(b) Prove that $|\mathbb{R}^{\mathbb{Q}}| = |\mathbb{R}|$. Hint: recall that $|\mathbb{R}| = |\mathcal{P}(\mathbb{N})| = |\{0,1\}^{\mathbb{N}}|$ (Theorem 2.8 and Exercise 2.11(a)).

(c) Parts (a) and (b) imply that $|\mathcal{C}| \leq |\mathbb{R}|$. Show that $|\mathbb{R}| \leq |\mathcal{C}|$, and conclude that $|\mathcal{C}| = |\mathbb{R}|$ by the Cantor–Schröder–Bernstein Theorem.

(d) Let $\mathcal{D}$ be the set of functions $\mathbb{R} \to \mathbb{R}$ with at least one discontinuity. Use Exercise 2.11(c) to prove that $|\mathcal{C}| < |\mathcal{D}|$.

Exercise 4.15. Let $\phi : \mathbb{Q} \to \mathbb{R}$ be continuous. Does there necessarily exist a continuous $f : \mathbb{R} \to \mathbb{R}$ such that $f(q) = \phi(q)$ for all $q \in \mathbb{Q}$?

Exercise 4.16. Recall the definitions in Exercise 2.36. Fix $n \in \mathbb{N}$. Prove that the map $T \mapsto T^{-1}$ on $\mathrm{GL}_n(\mathbb{R})$ is continuous.

Hint: one way to proceed is by fixing $A \in \mathrm{GL}_n(\mathbb{R})$ and noting that

$$||A^{-1} - B^{-1}|| = ||A^{-1}(B - A)B^{-1}|| \leq ||A^{-1}|| \cdot ||B^{-1}|| \cdot ||B - A||$$

for any $B \in \mathrm{GL}_n(\mathbb{R})$. We would hope that if $||B - A||$ is sufficiently small, then $||B^{-1}||$ can be bounded above by some expression that only depends on A. You may find the identity

$$\inf_{||x||=1} ||Tx|| = \frac{1}{||T^{-1}||}$$

useful (but you would have to prove this yourself).

Exercise 4.17 (Thomae's Function)**.** Any rational number x can be written uniquely in *simplest form* $\frac{p}{q}$ where $p \in \mathbb{Z}$, $q \in \mathbb{N}$, and $\gcd(p, q) = 1$. Define $f : \mathbb{R} \to \mathbb{R}$ by

$$f(x) = \begin{cases} \frac{1}{q} & \text{if } x \in \mathbb{Q} \text{ with simplest form } \frac{p}{q} \\ 0 & \text{if } x \notin \mathbb{Q}. \end{cases}$$

Prove that f is continuous at every irrational number and discontinuous at every rational number.

Exercise 4.18. In this exercise, we will show that there is no function $f : \mathbb{R} \to \mathbb{R}$ that is continuous on the rationals and discontinuous on the irrationals. First, we need some new definitions. Let $f : \mathbb{R} \to \mathbb{R}$ be a function.

- The *continuity set* of f is

$$C(f) \coloneqq \{x \in \mathbb{R} \mid f \text{ is continuous at } x\}.$$

  For example, the continuity set of Thomae's function from the previous exercise is $\mathbb{R} \setminus \mathbb{Q}$.

- The *oscillation* of f on a non-empty subset $E \subset \mathbb{R}$ is

$$\omega_f(E) \coloneqq \sup_{x \in E} f(x) - \inf_{x \in E} f(x).$$

- The oscillation of f at a fixed $x_0 \in \mathbb{R}$ is

$$\omega_f(x_0) \coloneqq \inf_{\delta > 0} \omega_f((x_0 - \delta, x_0 + \delta)).$$

You will also want to review Exercise 3.48.

(a) Prove that $\omega_f(x_0) = 0$ if and only if f is continuous at x_0.

(b) Fix $\epsilon > 0$. Prove that $C(f, \epsilon) \coloneqq \{x \in \mathbb{R} \mid \omega_f(x) < \epsilon\}$ is an open subset of $\mathbb{R}$.

(c) Find a sequence $\{\mathcal{O}_1, \mathcal{O}_2, \mathcal{O}_3, \dots\}$ of open subsets of $\mathbb{R}$ such that $C(f) = \bigcap_{n=1}^{\infty} \mathcal{O}_n$. Conclude that $C(f)$ cannot be equal to $\mathbb{Q}$.

4.3 Continuity and Compactness

Theorem 4.19. *If $f : X \to Y$ is continuous and $E \subset X$ is compact, then $f(E) \subset Y$ is compact.*

Proof. Let $\{\mathcal{O}_\alpha\}_\alpha$ be an open cover of $f(E)$. For each α, $f^{-1}(\mathcal{O}_\alpha) \subset X$ is open since f is continuous. If $x \in E$, then since $f(x) \in E$, we have that $f(x) \in \mathcal{O}_{\alpha_0}$ for some α_0, and it follows that $x \in f^{-1}(\mathcal{O}_{\alpha_0})$. Therefore, $\{f^{-1}(\mathcal{O}_\alpha)\}_\alpha$ is an open cover of E. Since E is compact, there exists a finite subcover $\{f^{-1}(\mathcal{O}_{\alpha_1}), \dots, f^{-1}(\mathcal{O}_{\alpha_n})\}$ of E. Fix $y \in f(E)$. Then $y = f(t)$ for some $t \in E$. Now $t \in f^{-1}(\mathcal{O}_{\alpha_k})$ for some $1 \le k \le n$ because $\{f^{-1}(\mathcal{O}_{\alpha_j})\}_{j=1}^n$ covers E. Therefore, $y \in \mathcal{O}_{\alpha_k} \subset \bigcup_{j=1}^n \mathcal{O}_{\alpha_j}$. We conclude that $\{\mathcal{O}_{\alpha_1}, \dots, \mathcal{O}_{\alpha_n}\}$ is a finite subcover of $f(E)$, so $f(E)$ is compact. $\square$

Corollary 4.19.1 (The Extreme Value Theorem)**.** *Let $f : [a,b] \to \mathbb{R}$ be continuous where $a \le b$. Then there exist $c_+, c_- \in [a,b]$ such that $\sup_{x\in[a,b]} f(x) = f(c_+)$ and $\inf_{x\in[a,b]} f(x) = f(c_-)$.*

Proof. The Heine–Borel Theorem says that $[a,b]$ is compact, so the image $f([a,b])$ is compact. Hence, $f([a,b])$ is closed and bounded. Clearly, $f([a,b])$ is also non-empty since $f(a) \in f([a,b])$. By Lemma 2.58, $\sup_{x\in[a,b]} f(x) = \sup f([a,b]) \in f([a,b])$ and, similarly, $\inf_{x\in[a,b]} f(x) \in f([a,b])$. The desired result follows. $\square$

Theorem 4.20. *Let X be compact, and let $f : X \to Y$ be a continuous bijection. Then the inverse $f^{-1} : Y \to X$ is continuous.*

The standard proof. Denote $g = f^{-1}$. We want to prove that if $E \subset X$ is open, then $g^{-1}(E) = f(E) \subset Y$ is open. Let $E \subset X$ be open. Then $E^c \subset X$ is closed and hence compact by Theorem 2.43 since X is compact. Since f is continuous, $f(E^c) = f(E)^c \subset Y$ is compact by Theorem 4.19 and hence closed. Therefore, $f(E) = (f(E)^c)^c \subset Y$ is open. $\square$

Alternative proof. Let $y_0 \in Y$, and suppose by way of contradiction that f^{-1} is not continuous at y_0. Then there exists $\epsilon > 0$ such that for all $\delta > 0$, there exists $y \in Y$ such that $d_Y(y, y_0) < \delta$ and $d_X(f^{-1}(y), f^{-1}(y_0)) \ge \epsilon$. For all $n \in \mathbb{N}$, choose $y_n \in Y$ such that $d_Y(y_n, y_0) < \frac{1}{n}$ and $d_X(f^{-1}(y_n), f^{-1}(y_0)) \ge \epsilon$. Then $\{f^{-1}(y_n)\}$ is a sequence in the compact set X, so by the Bolzano–Weierstrass Theorem, there is a convergent subsequence $\{f^{-1}(y_{n_k})\}$. Using the continuity of f and Theorem 4.9, we see that

$$y_0 = \lim_{k\to\infty} y_{n_k} = \lim_{k\to\infty} f(f^{-1}(y_{n_k})) = f(\lim_{k\to\infty} f^{-1}(y_{n_k})).$$

Hence, $\lim_{k\to\infty} f^{-1}(y_{n_k}) = f^{-1}(y_0)$, which contradicts that $d_X(f^{-1}(y_{n_k}), f^{-1}(y_0)) \ge \epsilon$ for all $k \in \mathbb{N}$. Therefore, f^{-1} must be continuous at y_0. Since $y_0 \in Y$ is arbitrary, f^{-1} is continuous on Y. $\square$

Definition 4.21. We say that $f : X \to Y$ is *uniformly continuous* on X if for all $\epsilon > 0$, there exists $\delta > 0$ such that for all $x, y \in X$, $d_Y(f(x), f(y)) < \epsilon$ whenever $d_X(x, y) < \delta$.

The definition of uniform continuity is quite similar to the definition of continuity at a given point $x \in X$. The main difference is that the δ for *uniform* continuity cannot depend on x, whereas the δ for continuity can depend on x.

Proposition 4.22. *If $f : X \to Y$ is uniformly continuous, then f is continuous.*

Proof. Let $f : X \to Y$ be uniformly continuous. Fix $\epsilon > 0$. By uniform continuity, there exists $\delta > 0$ such that $d_Y(f(x), f(y)) < \epsilon$ for all $x, y \in X$ such that $d_X(x, y) < \delta$. Now for any arbitrary $x_0 \in X$, if $x \in X$ satisfies $d_X(x, x_0) < \delta$, then $d_Y(f(x), f(x_0)) < \epsilon$. Hence, f is continuous at every $x_0 \in X$, so f is continuous on X. $\square$

Theorem 4.23. *If X is compact and $f : X \to Y$ is continuous, then f is uniformly continuous.*

Proof. Fix $\epsilon > 0$. For each $x \in X$, there exists $\delta_x > 0$ such that if $y \in X$ and $d_X(y, x) < \delta_x$, then $d_Y(f(y), f(x)) < \epsilon$. Notice that $\{N_{\delta_x/2}(x)\}_{x \in X}$ is an open cover of X. Since X is compact, we can extract a finite subcover $\{N_{\delta_{x_1}/2}(x_1), \ldots, N_{\delta_{x_n}/2}(x_n)\}$. Choose

$$\delta = \min_{1 \le j \le n} \frac{\delta_{x_j}}{2} > 0.$$

Let $x, y \in X$ such that $d_X(x, y) < \delta$. We know that $x \in N_{\delta_{x_k}/2}(x_k)$ for some $1 \le k \le n$ since $\{N_{\delta_{x_j}/2}(x_j)\}_{j=1}^n$ covers X. Hence $d_X(x_k, x) < \frac{\delta_{x_k}}{2} < \delta_{x_k}$, so $d_Y(f(x_k), f(x)) < \epsilon$. Now

$$d_X(x_k, y) \le d_X(x_k, x) + d_X(x, y) < \frac{\delta_{x_k}}{2} + \delta < \frac{\delta_{x_k}}{2} + \frac{\delta_{x_k}}{2} = \delta_{x_k},$$

so $d_Y(f(x_k), f(y)) < \epsilon$ also. Therefore,

$$d_Y(f(x), f(y)) \le d_Y(f(x), f(x_k)) + d_Y(f(x_k), f(y)) < 2\epsilon,$$

proving that f is uniformly continuous. $\square$

Exercise 4.24. Prove the Extreme Value Theorem directly from the Bolzano–Weierstrass Theorem and the definition of continuity.

Exercise 4.25. For any non-empty $S \subset X$, the *diameter* of S is $\operatorname{diam}(S) \coloneqq \sup_{x,y \in S} d(x, y)$.

(a) Prove that if S is bounded, then $\operatorname{diam}(S) < \infty$.

(b) Prove that if S is compact, then there exist $x_0, y_0 \in S$ such that $d(x_0, y_0) = \operatorname{diam}(S)$.

Exercise 4.26. The *distance* between two non-empty sets $S, T \subset X$ is

$$\operatorname{dist}(S, T) \coloneqq \inf_{x \in S, y \in T} d(x, y).$$

Prove that if S and T are compact, then there exist $x \in S$ and $y \in T$ such that $d(x, y) = \operatorname{dist}(S, T)$.

Exercise 4.27. Let K be a non-empty compact set, and let $f : K \to K$ such that $d(f(x), f(y)) < d(x, y)$ for all distinct $x, y \in K$. Prove that there exists a unique $x \in K$ such that $f(x) = x$. Hint: think about $\inf_{x \in K} d(x, f(x))$.

Exercise 4.28. Let K be a compact set and $f : K \to K$ such that $d(f(x), f(y)) \geq d(x, y)$ for all $x, y \in K$.

(a) For all $n \in \mathbb{N}$, let f_n be f composed with itself n times (i.e. $f_n = \underbrace{f \circ \cdots \circ f}_{n \text{ times}}$). Fix $x \in K$, and consider the sequence $\{f_n(x)\}_n$. Prove that there exists a subsequence $\{f_{n_k}(x)\}_k$ that converges to x.

(b) Show that for any $x, y \in K$, there exists an increasing sequence of indices $\{n_k\}_k$ such that $\{f_{n_k}(x)\}_k$ converges to x and $\{f_{n_k}(y)\}_k$ converges to y. (Note: as far as I know, this part requires a slight strengthening of part (a). Look at how you construct $\{n_k\}_k$ in part (a)—the construction should be the same no matter which fixed $x \in K$ we use.)

(c) Hence, prove that $d(x, y) \geq d(f(x), f(y))$, which implies that $d(f(x), f(y)) = d(x, y)$.

(d) Prove that f is bijective.

Exercise 4.29. A metric space (X, d_X) is called *locally compact* if for every $x \in X$, there exists $r > 0$ and a compact $K \subset X$ such that $N_r(x) \subset K$. Suppose X and Y are locally compact, and suppose $f : X \to Y$ is a continuous bijection such that the pre-image of any compact set is compact. Prove that the inverse $f^{-1} : Y \to X$ is continuous.

Exercise 4.30. Give an example of a continuous and bounded function $f : \mathbb{R} \to \mathbb{R}$ that is not uniformly continuous.

Exercise 4.31 (Continuous Extension Theorem)**.** Let X, Y be metric spaces where Y is complete. Let $E \subset X$, and suppose $f : E \to Y$ is uniformly continuous. Prove that there exists a unique continuous $g : \overline{E} \to Y$ such that $g(x) = f(x)$ for all $x \in E$. We call g a *continuous extension* of f.

Exercise 4.32 (Real Exponents)**.** Fix $b > 0$, and let $f : \mathbb{Q} \to \mathbb{R}$ be the map $f(q) = b^q$ (recall Exercise 1.24).

(a) For all $n \in \mathbb{N}$, show that f is uniformly continuous on $\mathbb{Q}_n := \mathbb{Q} \cap [-n, n]$.

(b) Using the Continuous Extension Theorem (Exercise 4.31), prove that there exists a unique continuous function $g : \mathbb{R} \to \mathbb{R}$ such that $g(x) = f(x)$ for all $x \in \mathbb{Q}$. We now *define* $b^x := g(x)$ for all $x \in \mathbb{R}$.

(c) Prove that $b^{x+y} = b^x b^y$ for all $x, y \in \mathbb{R}$. Hint: this identity holds for $x, y \in \mathbb{Q}$; how can we extend it to $x, y \in \mathbb{R}$?

(d) Fix $b, c > 0$. Prove that $(bc)^x = b^x c^x$ for all $x \in \mathbb{R}$.

(e) Show that $b^x > 0$ for all $x \in \mathbb{R}$.

(f) Show that g is strictly increasing if $b > 1$, constant if $b = 1$, and strictly decreasing if $0 < b < 1$.

Exercise 4.33. Suppose $f : \mathbb{R} \to \mathbb{R}$ is continuous and satisfies $f(x + y) = f(x)f(y)$ for all $x, y \in \mathbb{R}$. Let $b = f(1)$. Prove that $f(x) = b^x$ for all $x \in \mathbb{R}$.

4.4 Continuity and Connectedness

Theorem 4.34. *If $f : X \to Y$ is continuous and $E \subset X$ is connected, then $f(E) \subset Y$ is connected.*

Proof. Suppose $f : X \to Y$ is continuous and $E \subset X$ is connected. By Corollary 2.57.1, it suffices to show that $f(E)$ is a connected metric space. To do this, we use Exercise 2.60. Suppose, by way of contradiction, that $f(E)$ is not connected. Then there exists a non-empty proper subset S of $f(E)$ that is both open and closed. Since f is continuous on X, f is also continuous on E, so $f^{-1}(S) \subset E$ is both open and closed. We know that $f^{-1}(S)$ is non-empty since S is non-empty and $S \subset f(E)$. Also, $f^{-1}(S) \neq E$, for if $f^{-1}(S) = E$, then $f(E) = f(f^{-1}(S)) \subset S$, contradicting that $S \neq f(E)$. Therefore, E is not connected because $f^{-1}(S)$ is a non-empty proper subset of E that is both open and closed. This is a contradiction, so $f(E)$ must be connected. □

Corollary 4.34.1 (The Intermediate Value Theorem)**.** *Let $f : [a, b] \to \mathbb{R}$ be continuous where $a \leq b$. If $y \in \mathbb{R}$ is between $f(a)$ and $f(b)$ inclusive, then there exists $x \in [a, b]$ such that $f(x) = y$.*

Proof. Without loss of generality, suppose $f(a) \leq f(b)$. If $x, y \in [a, b]$ and $x < y$, then $a \leq x < z < y \leq b$ for any $z \in (x, y)$, so $(x, y) \subset [a, b]$. Hence, $[a, b]$ is connected by Theorem 2.59. Since f is continuous and $[a, b]$ is connected, $f([a, b])$ must be connected by Theorem 4.34. Let $f(a) \leq y \leq f(b)$. If $y = f(a)$ or $y = f(b)$, then clearly $y \in f([a, b])$. If $f(a) < y < f(b)$, then $(f(a), f(b)) \subset f([a, b])$ by Theorem 2.59, so $y \in (f(a), f(b)) \subset f([a, b])$. Therefore, $y \in f([a, b])$ in all cases, so there exists $x \in [a, b]$ such that $f(x) = y$. □

Exercise 4.35. Prove the Intermediate Value Theorem directly from the supremum property of $\mathbb{R}$ and the definition of continuity.

Exercise 4.36. Let X be a metric space, and let Y be the set $\{0, 1\}$ equipped with the discrete metric

$$d(x, y) = \begin{cases} 1 & \text{if } x \neq y \\ 0 & \text{if } x = y. \end{cases}$$

Prove that X is connected if and only if every continuous function $f : X \to Y$ is constant.

Exercise 4.37. A metric space X is *path-connected* if for all $x, y \in X$, there exists a continuous function $f : [0, 1] \to X$ such that $f(0) = x$ and $f(1) = y$. Prove that every path-connected set is connected.

Exercise 4.38. Let $f : \mathbb{R} \to \mathbb{R}$ be continuous. The *graph* of f is $G_f \coloneqq \{(x, f(x)) \mid x \in \mathbb{R}\}$, which is a subset of $\mathbb{R}^2$. Prove that G_f is path-connected.

Exercise 4.39. Prove that if $f : X \to Y$ is continuous and X is path-connected, then $f(X)$ is path-connected.

4.5 One-Sided Limits and Monotonic Functions

Fix real numbers $a < b$.

Definition 4.40 (One-Sided Limits)**.** Let $f : (a, b) \to \mathbb{R}$; $c_1 \in [a, b)$; $c_2 \in (a, b]$; and $L_+, L_- \in \mathbb{R}$. We say that

$$\lim_{x \to c_1^+} f(x) = L_+$$

if for all $\epsilon > 0$, there exists $\delta > 0$ such that $|f(x) - L_+| < \epsilon$ whenever $0 < x - c_1 < \delta$. Similarly, we say that

$$\lim_{x \to c_2^-} f(x) = L_-$$

if for all $\epsilon > 0$, there exists $\delta > 0$ such that $|f(x) - L_-| < \epsilon$ whenever $0 < c_2 - x < \delta$.

For any $c_1 \in [a, b)$, we can view $[c_1, b)$ as its own metric space. Then the one-sided limit $\lim_{x \to c_1^+} f(x)$ as we have defined it above is the same as the limit $\lim_{x \to c_1} f(x)$ taken in the metric space $[c_1, b)$. Similarly, if $c_2 \in (a, b]$, then $\lim_{x \to c_2^-} f(x)$ is equivalent to the limit $\lim_{x \to c_2} f(x)$ taken in the metric space $(a, c_2]$. As a result, the limit laws that hold for sequences of real numbers also hold analogously for one-sided limits by Theorem 4.2.

Note that the condition "$0 < |x - c| < \delta$" is equivalent to "$0 < x - c < \delta$ or $0 < c - x < \delta$". This observation immediately implies the following result:

Proposition 4.41. *Let $f : (a, b) \to \mathbb{R}$ and $c \in (a, b)$. Then $\lim_{x \to c} f(x)$ exists if and only if $\lim_{x \to c^-} f(x)$ and $\lim_{x \to c^+} f(x)$ both exist and are equal. In that case, we have*

$$\lim_{x \to c} f(x) = \lim_{x \to c^-} f(x) = \lim_{x \to c^+} f(x).$$

Definition 4.42. We call $f : (a, b) \to \mathbb{R}$ *monotonically increasing* if $f(x) \le f(y)$ for all $x, y \in (a, b)$ such that $x \le y$. Similarly, f is *monotonically decreasing* if $f(x) \ge f(y)$ for all $x, y \in (a, b)$ such that $x \le y$. We say that f is *monotone* if f is monotonically increasing or monotonically decreasing.

Remark. Note that the zero function $x \mapsto 0$ on $\mathbb{R}$ is both monotonically increasing and monotonically decreasing.

Theorem 4.43. *Let $f : (a, b) \to \mathbb{R}$ be monotonically increasing. Let $u, v, c \in (a, b)$ such that $u < c < v$. Then*

$$f(u) \le \sup_{x < c} f(x) = \lim_{x \to c^-} f(x) \le f(c) \le \lim_{x \to c^+} f(x) = \inf_{x > c} f(x) \le f(v).$$

Proof. Let

$$A = \{f(x) \mid x \in (a, c)\}$$

and

$$B = \{f(x) \mid x \in (c, b)\}.$$

Then A and B are non-empty since $f(u) \in A$ and $f(v) \in B$. Since f is monotonically increasing, A is bounded above by $f(c)$, and B is bounded below by $f(c)$. Therefore, $\sup_{x<c} f(x) = \sup(A)$ exists and $\inf_{x>c} f(x) = \inf(B)$ exists. Moreover, it is clear that

$$f(u) \le \sup_{x<c} f(x) \le f(c) \le \inf_{x>c} f(x) \le f(v).$$

Now we just need to show that $\lim_{x \to c^-} f(x) = \sup_{x>c} f(x)$ and $\lim_{x \to c^+} f(x) = \inf_{x<c} f(x)$. Let $s = \sup_{x<c} f(x) = \sup(A)$. Fix $\epsilon > 0$. Then there exists $y_0 \in A$ such that $s - \epsilon < y_0 \le s$. Since $y_0 \in A$, there exists $x_0 \in (a, c)$ such that $f(x_0) = y_0$. Now let $\delta = c - x_0 > 0$, and suppose $0 < c - x < \delta$ where $x \in (a, b)$. Then $x_0 = c - \delta < x$, so $f(x_0) \le f(x)$ since f is monotonically increasing. Hence, $s - \epsilon < f(x_0) \le f(x) \le s$, so $|s - f(x)| < \epsilon$. Therefore, $\lim_{x \to c^-} f(x) = s = \sup_{x<c} f(x)$. The proof that $\lim_{x \to c^+} f(x) = \inf_{x>c} f(x)$ proceeds similarly. $\square$

Theorem 4.44. *If $f : (a, b) \to \mathbb{R}$ is monotonically increasing, then the set of discontinuities of f is denumerable.*

Proof. Let $S = \{c \in (a, b) \mid f \text{ is discontinuous at } c\}$. We will associate every $c \in S$ with a unique rational number; this defines an injective map $\phi : S \to \mathbb{Q}$. Since $\mathbb{Q}$ is countable, the result follows.

Suppose f is discontinuous at $c \in (a, b)$. Theorem 4.43 implies that $\lim_{x\to c^-} f(x) \leq \lim_{x\to c^+} f(x)$. By way of contradiction, suppose that $\lim_{x\to c^-} f(x) = \lim_{x\to c^+} f(x)$. Then Theorem 4.43 implies that $\lim_{x\to c^-} f(x) = f(c) = \lim_{x\to c^+} f(x)$. Hence, Proposition 4.41 says that $\lim_{x\to c} f(x) = f(c)$, which means f is continuous at c by Theorem 4.7. This is a contradiction, so it must be that $\lim_{x\to c^-} f(x) < \lim_{x\to c^+} f(x)$. Let $L_- = \lim_{x\to c^-} f(x)$ and $L_+ = \lim_{x\to c^+} f(x)$.

If $L_- < f(c)$, then there exists $q \in \mathbb{Q}$ such that $L_- < q < f(c)$, and we put $\phi(c) = q$.

On the other hand, if $L_- = f(c)$, then $f(c) = L_- < L_+$, so there exists $r \in \mathbb{Q}$ such that $f(c) < r < L_+$. Put $\phi(c) = r$ in this case.

Notice that $L_- < \phi(c) < L_+$ in all cases. Thus, we have defined a map $\phi : S \to \mathbb{Q}$ such that $\lim_{x\to c^-} f(x) < \phi(c) < \lim_{x\to c^+} f(x)$ for all $c \in S$. We just need to show that ϕ is injective. Suppose $c, d \in S$ are distinct. Without loss of generality, we may assume that $c < d$. Pick $t \in (a, b)$ such that $c < t < d$. Then by Theorem 4.43,

$$\phi(c) < \lim_{x\to c^+} f(x) \leq f(t) \leq \lim_{x\to d^-} f(x) < \phi(d).$$

Therefore, ϕ is injective. □

Theorems 4.43 and 4.44 have analogues for monotonically decreasing functions, too—the reader should be able to guess what they are.

5 Differentiation

Fix real numbers $a < b$.

5.1 Definition of the Derivative

Definition 5.1. Let $f : (a, b) \to \mathbb{R}$ and $c \in (a, b)$. Then f is *differentiable at c* if the limit

$$\lim_{x \to c} \frac{f(x) - f(c)}{x - c}$$

exists. The value of this limit is denoted $f'(c)$ and is called the *derivative* of f at c.

If f is differentiable at every $c \in (a, b)$, then we say that f is differentiable (on (a, b)). In this case, the derivative of f is the function $f' : (a, b) \to \mathbb{R}$ defined by $f'(x_0) = \lim_{x \to x_0} \frac{f(x) - f(x_0)}{x - x_0}$ for all $x_0 \in (a, b)$.

Derivatives help us approximate functions locally. Indeed, $f : (a, b) \to \mathbb{R}$ is differentiable at c if and only if there exists $D \in \mathbb{R}$ and $r : (a, b) \to \mathbb{R}$ such that

$$f(x) = f(c) + D(x - c) + r(x)$$

and $\lim_{x \to c} \frac{r(x)}{x - c} = 0$. If D and r exist, then $f'(c) = D$. Hence, if f is differentiable at c, then we have the *linear approximation* $f(x) \approx f(c) + f'(c)(x - c)$ when x is close to c.

If f' itself is differentiable, then the derivative of f' is denoted f'' and is called the *second derivative* of f. In general, we can define repeated derivatives of f recursively as follows:

Definition 5.2. Let $f : (a, b) \to \mathbb{R}$. Define $f^{(0)} = f$. For any integer $n \geq 0$, if $f^{(n)} : (a, b) \to \mathbb{R}$ is differentiable, then we define $f^{(n+1)} = (f^{(n)})'$. We say that f is *n-times differentiable* if $f^{(n)}$ exists. If $f^{(n)}$ exists for all $n \geq 0$, then f is *infinitely differentiable.*

Proposition 5.3. *If $f : (a, b) \to \mathbb{R}$ is differentiable at $c \in (a, b)$, then f is continuous at c.*

Proof. Observe that

$$\begin{aligned}
\lim_{x \to c}[f(x) - f(c)] &= \lim_{x \to c}(x - c)\left(\frac{f(x) - f(c)}{x - c}\right) \\
&= \lim_{x \to c}(x - c) \lim_{x \to c} \frac{f(x) - f(c)}{x - c} \\
&= 0 f'(c) \\
&= 0.
\end{aligned}$$

Hence,

$$\lim_{x \to c} f(x) = \lim_{x \to c}(f(c) + f(x) - f(c)) = \lim_{x \to c} f(c) + \lim_{x \to c}(f(x) - f(c)) = f(c) + 0 = f(c),$$

so f is continuous at c by Theorem 4.7. ☐

Note, however, that differentiability at a point does not imply continuity in a neighbourhood around that point.

Example 5.4. Consider $f : \mathbb{R} \to \mathbb{R}$ defined by

$$f(x) = \begin{cases} x^2 & \text{if } x \in \mathbb{Q} \\ -x^2 & \text{if } x \notin \mathbb{Q}. \end{cases}$$

We will show that f is differentiable at 0 and discontinuous at every non-zero point. First, note that

$$0 \leq \left| \frac{f(x)}{x} \right| \leq \left| \frac{x^2}{x} \right| = |x|$$

for all $x \neq 0$. Since $\lim_{x\to 0} 0 = \lim_{x\to 0} |x| = 0$, the Squeeze Theorem yields that

$$\lim_{x\to 0} \left| \frac{f(x)}{x} \right| = 0.$$

Using Proposition 3.4 and Theorem 4.2, we can show that

$$\lim_{x\to 0} \frac{f(x) - f(0)}{x - 0} = \lim_{x\to 0} \frac{f(x)}{x} = 0.$$

Hence, f is differentiable at 0.

Suppose $c \neq 0$ is rational. Then by Theorem 1.26, c is a limit point of the set of irrational numbers. Hence, there is a sequence $\{x_n\}$ of irrational numbers such that $\lim_{n\to\infty} x_n = c$. Now note that

$$\lim_{n\to\infty} f(x_n) = \lim_{n\to\infty} -x_n^2 = -\lim_{n\to\infty} x_n \lim_{n\to\infty} x_n = -c^2 \neq f(c) = c^2.$$

since c is non-zero. Therefore, f is not continuous at c by Theorems 4.2 and 4.7.

If $c \neq 0$ is irrational, then by Theorem 1.25, c is a limit point of the set of rational numbers. We can then proceed in the same way as the previous paragraph to show that f is discontinuous at c.

Proposition 5.5. *Suppose $f : (a,b) \to \mathbb{R}$ and $g : (a,b) \to \mathbb{R}$ are differentiable at $c \in (a,b)$. Then:*

(a) $(f+g)'(c) = f'(c) + g'(c)$.
(b) $(f-g)'(c) = f'(c) - g'(c)$.
(c) $(\alpha f)'(c) = \alpha f'(c)$ *for any* $\alpha \in \mathbb{R}$.
(d) $(fg)'(c) = f'(c)g(c) + f(c)g'(c)$.
(e) $(\frac{f}{g})'(c) = \frac{1}{[g(c)]^2}(f'(c)g(c) - f(c)g'(c))$ *if* $g(c) \neq 0$.

Proof. These are all direct computations.

(a)

$$\begin{aligned}(f+g)'(c) &= \lim_{x\to c} \frac{f(x) + g(x) - (f(c) + g(c))}{x - c} \\ &= \lim_{x\to c} \left(\frac{f(x) - f(c)}{x - c} + \frac{g(x) - g(c)}{x - c} \right) \\ &= f'(c) + g'(c).\end{aligned}$$

(b)

$$\begin{aligned}(f-g)'(c) &= \lim_{x\to c}\frac{f(x)-g(x)-(f(c)-g(c))}{x-c}\\ &= \lim_{x\to c}\left(\frac{f(x)-f(c)}{x-c}-\frac{g(x)-g(c)}{x-c}\right)\\ &= f'(c)-g'(c).\end{aligned}$$

(c)

$$(\alpha f)'(c) = \lim_{x\to c}\frac{\alpha f(x)-\alpha f(c)}{x-c} = \lim_{x\to c}\alpha\left(\frac{f(x)-f(c)}{x-c}\right) = \alpha f'(c).$$

(d) Since f is differentiable at c, f is continuous at c, so $\lim_{x\to c} f(x) = f(c)$. Hence,

$$\begin{aligned}(fg)'(c) &= \lim_{x\to c}\frac{f(x)g(x)-f(c)g(c)}{x-c}\\ &= \lim_{x\to c}\frac{f(x)g(x)-f(x)g(c)+f(x)g(c)-f(c)g(c)}{x-c}\\ &= \lim_{x\to c}\left(f(x)\frac{g(x)-g(c)}{x-c}+g(c)\frac{f(x)-f(c)}{x-c}\right)\\ &= \lim_{x\to c}f(x)\lim_{x\to c}\frac{g(x)-g(c)}{x-c}+g(c)\lim_{x\to c}\frac{f(x)-f(c)}{x-c}\\ &= f(c)g'(c)+g(c)f'(c).\end{aligned}$$

(e) Note that $\lim_{x\to c} g(x) = g(c)$ since g is differentiable and hence continuous at c. If $g(c) \neq 0$, then

$$\begin{aligned}\left(\frac{f}{g}\right)'(c) &= \lim_{x\to c}\frac{\frac{f(x)}{g(x)}-\frac{f(c)}{g(c)}}{x-c}\\ &= \lim_{x\to c}\frac{f(x)g(c)-f(c)g(x)}{g(x)g(c)(x-c)}\\ &= \lim_{x\to c}\frac{1}{g(x)g(c)}\lim_{x\to c}\frac{f(x)g(c)-f(c)g(x)}{x-c}\\ &= \frac{1}{g(c)^2}\lim_{x\to c}\frac{f(x)g(c)-f(c)g(c)+f(c)g(c)-f(c)g(x)}{x-c}\\ &= \frac{1}{g(c)^2}\lim_{x\to c}\left(g(c)\frac{f(x)-f(c)}{x-c}-f(c)\frac{g(x)-g(c)}{x-c}\right)\\ &= \frac{1}{g(c)^2}[g(c)f'(c)-f(c)g'(c)].\end{aligned}$$

□

Theorem 5.6 (The Chain Rule)**.** *Let $g:(a,b)\to\mathbb{R}$ be differentiable at $c\in(a,b)$. Let $E\subset\mathbb{R}$ be an open interval containing the image of g, and suppose $f:E\to\mathbb{R}$ is differentiable at $g(c)$. Then $f\circ g:(a,b)\to\mathbb{R}$ is differentiable at c, and*

$$(f\circ g)'(c) = f'(g(c))g'(c).$$

Proof. Since g is differentiable at c, there exists $r_1 : (a,b) \to \mathbb{R}$ such that $\lim_{x\to c} \frac{r_1(x)}{x-c} = 0$ and

$$g(x) = g(c) + g'(c)[x-c] + r_1(x)$$

for all $x \in (a,b)$. Similarly, since f is differentiable at $g(c)$, there exists $r_2 : E \to \mathbb{R}$ such that $\lim_{y\to g(c)} \frac{r_2(y)}{y-g(c)} = 0$ and

$$f(y) = f(g(c)) + f'(g(c))[y-g(c)] + r_2(y)$$

for all $y \in E$. Putting $y = g(x)$ for $x \in (a,b)$ gives

$$\begin{aligned} f(g(x)) &= f(g(c)) + f'(g(c))[g(x)-g(c)] + r_2(g(x)) \\ &= f(g(c)) + f'(g(c))g'(c)[x-c] + f'(g(c))r_1(x) + r_2(g(x)). \end{aligned}$$

It suffices to show that

$$\lim_{x\to c} \frac{f'(g(c))r_1(x) + r_2(g(x))}{x-c} = 0.$$

Clearly

$$\lim_{x\to c} \frac{f'(g(c))r_1(x)}{x-c} = f'(g(c)) \lim_{x\to c} \frac{r_1(x)}{x-c} = 0,$$

so now we are left with proving that $\lim_{x\to c} \frac{r_2(g(x))}{x-c} = 0$. Since $\lim_{x\to c} \frac{g(x)-g(c)}{x-c} = g'(c)$, there exists $\delta_1 > 0$ such that

$$\left| \frac{g(x)-g(c)}{x-c} - g'(c) \right| < 1$$

whenever $0 < |x-c| < \delta_1$. Hence,

$$[1-g'(c)](x-c) < g(x) - g(c) < [1+g'(c)](x-c)$$

whenever $0 < |x-c| < \delta_1$. Letting $M = 1 + \max\{|1-g'(c)|, |1+g'(c)|\} > 0$, we see that

$$|g(x) - g(c)| < M|x-c|$$

whenever $0 < |x-c| < \delta_1$. Fix $\epsilon > 0$. Since $\lim_{y\to g(c)} \frac{r_2(y)}{y-g(c)} = 0$, there exists $\delta_2 > 0$ such that

$$\left| \frac{r_2(y)}{y-g(c)} \right| < \frac{\epsilon}{M}$$

whenever $0 < |y-g(c)| < \delta_2$. Therefore, $|r_2(y)| \le \frac{\epsilon}{M}|y-g(c)|$ whenever $|y-g(c)| < \delta_2$. Since g is continuous at c, there exists $\delta_3 > 0$ such that $|g(x)-g(c)| < \delta_2$ whenever $|x-c| < \delta_2$. Let $\delta = \min\{\delta_1, \delta_3\} > 0$, and suppose $0 < |x-c| < \delta$. Then $|g(x)-g(c)| < \delta_2$ since $|x-c| < \delta_3$. Hence, $|r_2(g(x))| \le \frac{\epsilon}{M}|g(x)-g(c)|$. Finally, since $0 < |x-c| < \delta_1$, we have that $|r_2(g(x))| < \frac{\epsilon}{M}M|x-c| = \epsilon|x-c|$. Since $|x-c| > 0$, we can divide by $|x-c|$ on both sides of the inequality to obtain that

$$\left| \frac{r_2(g(x))}{x-c} - 0 \right| = \left| \frac{r_2(g(x))}{x-c} \right| < \epsilon.$$

Therefore, $\lim_{x\to c} \frac{r_2(g(x))}{x-c} = 0$. $\square$

Now let's compute some actual derivatives, so that we have some concrete differentiable functions that we can apply the rules of differentiation to. We will look at the power functions $f(x) = x^n$ where $n \in \mathbb{Z}$. Calculus students will recognize the following results as special cases of the "power rule".

Proposition 5.7. *If $f : \mathbb{R} \to \mathbb{R}$ is a constant function, then $f'(x) = 0$ for all $x \in \mathbb{R}$.*

Proof. Suppose there exists $c \in \mathbb{R}$ such that $f(x) = c$ for all $x \in \mathbb{R}$. Fix $x_0 \in \mathbb{R}$. Then

$$f'(x_0) = \lim_{x \to x_0} \frac{f(x) - f(x_0)}{x - x_0} = \lim_{x \to x_0} \frac{c - c}{x - x_0} = \lim_{x \to x_0} 0 = 0.$$

$\square$

In order for the next proposition to make sense, we need to clarify what 0^0 means. (In general, if $x \neq 0$, then $x^0 = x^{1-1} = \frac{x}{x} = 1$, but this reasoning does not apply if $x = 0$.) By convention, we define $0^0 \coloneqq 1$. There are a few reasons why this definition would make sense:

- For any $x \in \mathbb{R}$ and $n \in \mathbb{N}$, we define x^n as the "repeated multiplication" $x^n = \prod_{k=1}^{n} x$. If $n = 0$, this product becomes the "empty product" which is equal to 1 by convention. This reasoning should apply even if $x = 0$.
- As a result of the previous bullet point, all exponent laws that hold for non-zero bases hold even if the base is zero (as long as no "division by zero" is involved).
- We already know that $x^0 = 1$ for all real numbers x other than 0. It would be the most convenient choice to define 0^0 to equal 1 also, so that we can always write $x^0 = 1$ without any exceptions.
- Defining $0^0 = 1$ makes the function $f(x) = x^0$ continuous on all of $\mathbb{R}$.

Proposition 5.8. *Fix $n \in \mathbb{N}$, and let $f(x) = x^n$ for all $x \in \mathbb{R}$. Then $f'(x) = nx^{n-1}$ for all $x \in \mathbb{R}$.*

Proof. Fix $x_0 \in \mathbb{R}$. For all $x \in \mathbb{R}$, we have that

$$\begin{aligned}
(x - x_0) \sum_{k=0}^{n-1} x^k x_0^{n-1-k} &= \sum_{k=0}^{n-1} x^{k+1} x_0^{n-1-k} - \sum_{k=0}^{n-1} x^k x_0^{n-k} \\
&= x^n - x_0^n + \sum_{k=0}^{n-2} x^{k+1} x_0^{n-1-k} - \sum_{k=1}^{n-1} x^k x_0^{n-k} \\
&\quad \text{(because } y^0 = 1 \text{ for } \textit{all} \text{ real numbers } y\text{, including } y = 0) \\
&= x^n - x_0^n + \sum_{j=1}^{n-1} x^{(j-1)+1} x_0^{n-1-(j-1)} - \sum_{k=1}^{n-1} x^k x_0^{n-k} \\
&= x^n - x_0^n + \sum_{j=1}^{n-1} x^j x_0^{n-j} + \sum_{k=1}^{n-1} x^k x_0^{n-k}
\end{aligned}$$

$$= x^n - x_0^n.$$

Therefore,

$$\begin{aligned}
f'(x_0) &= \lim_{x\to x_0} \frac{f(x)-f(x_0)}{x-x_0} \\
&= \lim_{x\to x_0} \sum_{k=0}^{n-1} x^k x_0^{n-1-k} \\
&= \sum_{k=0}^{n-1} \lim_{x\to x_0} x^k x_0^{n-1-k} \quad \text{since the sum is finite} \\
&= \sum_{k=0}^{n-1} x_0^k x_0^{n-1-k} \quad \text{by Proposition 4.5(c) and induction} \\
&= n x_0^{n-1}.
\end{aligned}$$

□

Proposition 5.9. *Fix $n \in \mathbb{N}$, and let $f(x) = x^{-n}$ for all $x \in \mathbb{R}\setminus\{0\}$. Then $f'(x) = -nx^{-n-1}$ for all $x \in \mathbb{R}\setminus\{0\}$.*

Proof. Let $g(x) = 1$ and $h(x) = x^n$ for all $x \in \mathbb{R}$. Then $f(x) = \frac{g(x)}{h(x)}$ and $h(x) \neq 0$ for all $x \in \mathbb{R}\setminus\{0\}$. Hence,

$$f'(x) = \left(\frac{g}{h}\right)'(x) = \frac{g'(x)h(x) - g(x)h'(x)}{[h(x)]^2} = \frac{0x^n - 1nx^{n-1}}{x^{2n}} = -nx^{n-1-2n} = -nx^{-n-1}$$

for all $x \in \mathbb{R}\setminus\{0\}$. □

Exercise 5.10. Let $f : \mathbb{R} \to \mathbb{R}$ such that $f(x) = x$. If we had defined 0^0 to be equal to 0 instead of 1, would the value of $f'(0)$ change?

Exercise 5.11. Construct a function $g : \mathbb{R} \to \mathbb{R}$ that is differentiable at exactly two points and discontinuous everywhere else. Hint: use the function f from Example 5.4, which is differentiable at 0 and discontinuous everywhere else.

5.2 The Mean Value Theorem

The Mean Value Theorem is probably the most important theorem concerning differentiation. It formalizes the relationship between "derivative" and "rate of change". For example, we can informally deduce that a function whose derivative is zero everywhere must be a constant function, since such a function must have "zero rate of change" at every point. The Mean Value Theorem allows us to formally prove this fact.

Definition 5.12. We say that $f : (a, b) \to \mathbb{R}$ has a *local minimum* (respectively *local maximum*) at $c \in (a, b)$ if there exists $r > 0$ such that $f(x) \geq f(c)$ (respectively $f(x) \leq f(c)$) whenever $|x - c| < r$. A *local extremum* is a local minimum or a local maximum.

Theorem 5.13 (Fermat's Theorem)**.** *If $f : (a, b) \to \mathbb{R}$ has a local extremum at $c \in (a, b)$ and is differentiable at c, then $f'(c) = 0$.*

Proof. Suppose c is a local minimum. Pick $r > 0$ such that $f(x) \geq f(c)$ whenever $|x-c| < r$. Then $\frac{f(x)-f(c)}{x-c} \geq 0$ for all $c < x < c + r$ and $\frac{f(x)-f(c)}{x-c} \leq 0$ for all $c - r < x < c$. Therefore,

$$f'(c) = \lim_{x \to c^+} \frac{f(x) - f(c)}{x - c} \geq 0$$

and

$$f'(c) = \lim_{x \to c^-} \frac{f(x) - f(c)}{x - c} \leq 0,$$

so $f'(c) = 0$. A similar argument works if c is instead a local maximum. □

Theorem 5.14 (Rolle's Theorem)**.** *Let $f : [a, b] \to \mathbb{R}$ be continuous on $[a, b]$ and differentiable on (a, b). Suppose $f(a) = f(b)$. Then there exists $c \in (a, b)$ such that $f'(c) = 0$.*

Proof. Since f is continuous on $[a, b]$, the Extreme Value Theorem says that there exists $c_-, c_+ \in [a, b]$ such that $f(c_-) \leq f(x) \leq f(c_+)$ for all $x \in [a, b]$. If $c_- \in (a, b)$, then f has a local minimum at c_-, so $f'(c_-) = 0$ by Fermat's Theorem. Similarly, if $c_+ \in (a, b)$, then f has a local maximum at c_+, so $f'(c_+) = 0$ by Fermat's Theorem. The only remaining possibility is that $c_-, c_+ \in \{a, b\}$. Since $f(a) = f(b)$, we must have that $f(a) = f(b) = f(c_-) = f(c_+)$. Therefore, $f(c_-) \leq f(x) \leq f(c_+) = f(c_-)$ for all $x \in [a, b]$, so f is constant on $[a, b]$. Therefore, $f'(\frac{a+b}{2}) = 0$ since $\frac{a+b}{2} \in (a, b)$. □

Theorem 5.15 (Cauchy's Mean Value Theorem)**.** *Let $f, g : [a, b] \to \mathbb{R}$ be continuous on $[a, b]$ and differentiable on (a, b). Then there exists $c \in (a, b)$ such that*

$$f'(c)[g(b) - g(a)] = g'(c)[f(b) - f(a)].$$

Proof. Define $h(x) = f(x)[g(b) - g(a)] - g(x)[f(b) - f(a)]$ for all $x \in [a, b]$. Then

$$h(a) = f(a)[g(b) - g(a)] - g(a)[f(b) - f(a)] = f(a)g(b) - g(a)f(b)$$

and

$$h(b) = f(b)[g(b) - g(a)] - g(b)[f(b) - f(a)] = -f(b)g(a) + g(b)f(a) = h(a).$$

Since f and g are continuous on $[a, b]$ and differentiable on (a, b), so is h. Therefore, by Rolle's Theorem, there exists $c \in (a, b)$ such that $h'(c) = 0$. Now

$$h'(x) = f'(x)[g(b) - g(a)] - g'(x)[f(b) - f(a)]$$

for all $x \in (a, b)$, so $0 = h'(c) = f'(c)[g(b) - g(a)] - g'(c)[f(b) - f(a)]$. The desired result follows. □

Corollary 5.15.1 (The Mean Value Theorem)**.** *Let $f : [a, b] \to \mathbb{R}$ be continuous on $[a, b]$ and differentiable on (a, b). Then there exists $c \in (a, b)$ such that*

$$f'(c) = \frac{f(b) - f(a)}{b - a}.$$

Proof. Define $g : [a, b] \to \mathbb{R}$ by $g(x) = x$. Then g is continuous on $[a, b]$ and $g'(c) = 1$ for all $c \in (a, b)$. By Cauchy's Mean Value Theorem, there exists $c \in (a, b)$ such that

$$f'(c)(b-a) = f'(c)[g(b) - g(a)] = g'(c)[f(b) - f(a)] = f(b) - f(a).$$

Hence, $f'(c) = \frac{f(b)-f(a)}{b-a}$. □

Definition 5.16. We call $f : (a, b) \to \mathbb{R}$ *strictly increasing* (resp. *strictly decreasing*) if $f(x) < f(y)$ (resp. $f(x) > f(y)$) for all $x, y \in (a, b)$ such that $x < y$.

Proposition 5.17. *Suppose $f : (a, b) \to \mathbb{R}$ is differentiable.*
(a) If $f'(x) = 0$ for all $x \in (a, b)$, then f is constant.
(b) If $f'(x) > 0$ for all $x \in (a, b)$, then f is strictly increasing.
(c) If $f'(x) < 0$ for all $x \in (a, b)$, then f is strictly decreasing.
(d) If $f'(x) \geq 0$ for all $x \in (a, b)$, then f is monotonically increasing.
(e) If $f'(x) \leq 0$ for all $x \in (a, b)$, then f is monotonically decreasing.

Proof. Let $x, y \in (a, b)$ be distinct. Without loss of generality, suppose $x < y$. By the Mean Value Theorem, there exists $c \in (x, y)$ such that

$$\frac{f(y) - f(x)}{y - x} = f'(c).$$

That is, $f(y) - f(x) = (y - x)f'(c)$. Hence, $f(y) - f(x)$ has the same sign as $f'(c)$ since $y - x > 0$, so all parts of the proposition simultaneously follow. □

Exercise 5.18. Let $f : (a, b) \to \mathbb{R}$ be twice differentiable, and suppose there exists $c \in (a, b)$ such that $f'(c) = 0$ and $f''(c) > 0$. Prove that f has a local minimum at c.

Exercise 5.19. Let $f : (a, b) \to \mathbb{R}$ be continuous on (a, b) and differentiable on $(a, c) \cup (c, b)$. Suppose $\lim_{x \to c} f'(x)$ exists. Prove that f is differentiable at c and that $\lim_{x \to c} f'(x) = f'(c)$.

Exercise 5.20. Let $f : [0, \infty) \to [0, \infty)$ be continuous. Suppose $f(0) = 0$, and suppose f is differentiable on $(0, \infty)$ with $|f'(x)| \leq f(x)$ for all $x > 0$. Prove that $f(x) = 0$ for all $x \geq 0$. (Hint: take $s = \sup_{x \in [0, \frac{1}{2}]} f(x)$, and prove that $s = 0$. Then f is identically zero on the interval $[0, \frac{1}{2}]$. Now repeat this argument for the intervals $[\frac{1}{2}, 1]$, $[1, \frac{3}{2}]$, and so on.)

5.3 Lipschitz Continuity

Definition 5.21. Let (X, d_X) and (Y, d_Y) be metric spaces, and let $f : X \to Y$. Define

$$S = \{c \geq 0 \mid d_Y(f(x), f(y)) \leq c d_X(x, y) \text{ for all } x, y \in X\} \subset \mathbb{R}.$$

We say that $f : X \to Y$ is *Lipschitz continuous* (on X) if S is non-empty. If f is Lipschitz continuous, its *Lipschitz constant* is $\inf(S)$.

Proposition 5.22. *If $f : X \to Y$ is Lipschitz continuous, then f is uniformly continuous.*

Proof. Suppose $f : X \to Y$ is Lipschitz continuous. Then there exists $c \geq 0$ such that $d_Y(f(x), f(y)) \leq c d_X(x, y)$ for all $x, y \in X$. Let $\epsilon > 0$, and choose $\delta = \frac{\epsilon}{c+1} > 0$. If $d_X(x, y) < \delta$, then

$$d_Y(f(x), f(y) \leq c d_X(x, y) \leq c\delta = \frac{c}{c+1}\epsilon < \epsilon.$$

Hence, f is uniformly continuous. □

Proposition 5.23. *For any $x_0 \in X$, the function $f : X \to \mathbb{R}$ defined by $f(x) = d_X(x, x_0)$ is Lipschitz continuous.*

Proof. Fix $x_0 \in X$. Then

$$|f(x) - f(y)| = |d_X(x, x_0) - d_X(y, x_0)| \leq d_X(x, y)$$

for all $x, y \in X$ by the Reverse Triangle Inequality (Proposition 2.13). Hence, f is Lipschitz continuous. □

Theorem 5.24. *Suppose $f : (a, b) \to \mathbb{R}$ is differentiable. Then f is Lipschitz continuous if and only if $\sup_{x\in(a,b)} |f'(x)| < \infty$, in which case the Lipschitz constant of f is $\sup_{x\in(a,b)} |f'(x)|$.*

Proof. Suppose $\sup_{x\in(a,b)} |f'(x)| < \infty$. Let $s = \sup_{x\in(a,b)} |f'(x)|$, and fix $x, y \in (a, b)$. We claim that $|f(x) - f(y)| \leq s|x - y|$. If $x = y$, then clearly $|f(x) - f(y)| = 0 \leq s|x - y|$. Now suppose $x \neq y$. Then by the Mean Value Theorem, there exists z between x and y such that $f'(z) = \frac{f(x)-f(y)}{x-y}$. Since $|\frac{f(x)-f(y)}{x-y}| = |f'(z)| \leq s$, it follows that $|f(x) - f(y)| \leq s|x - y|$. Therefore, f is Lipschitz continuous.

Conversely, suppose f is Lipschitz continuous. Then the set

$$S = \{c \geq 0 : |f(x) - f(y)| \leq c|x - y| \text{ for all } x, y \in (a, b)\}$$

is non-empty. Pick $c \in S$. Then

$$\left|\frac{f(x) - f(y)}{x - y}\right| \leq c$$

for all $x, y \in (a, b)$ such that $x \neq y$. As a consequence of Proposition 3.9(e) and Theorem 4.2, we have that

$$\lim_{x\to x_0} \left|\frac{f(x) - f(x_0)}{x - x_0}\right| = \left|\lim_{x\to x_0} \frac{f(x) - f(x_0)}{x - x_0}\right| = |f'(x_0)|.$$

On the other hand,

$$\lim_{x\to x_0} \left|\frac{f(x) - f(x_0)}{x - x_0}\right| \leq c$$

since $|\frac{f(x)-f(x_0)}{x-x_0}| \leq c$ for all $x \neq x_0$. Therefore, $|f'(x_0)| \leq c$, so $\sup_{x\in(a,b)} |f'(x)| \leq c < \infty$.

Since c is an arbitrary element of S, we know that

$$\sup_{x\in(a,b)} |f'(x)| \leq \inf(S).$$

On the other hand, in the first paragraph, we proved that $\sup_{x\in(a,b)} |f'(x)| \in S$. Therefore, $\sup_{x\in(a,b)} |f'(x)| = \inf(S)$, so $\sup_{x\in(a,b)} |f'(x)|$ is the Lipschitz constant of f. □

Theorem 5.25. *Let $f : [a, b] \to \mathbb{R}$ be continuous. Suppose f is differentiable everywhere on (a, b) except at finitely many points $x_1 < x_2 < \cdots < x_n$. Let $x_0 = a$ and $x_{n+1} = b$, and suppose $s_i = \sup_{x\in(x_{i-1},x_i)} |f'(x)|$ is finite for all $1 \leq i \leq n + 1$. Put $M = \max_{1\leq i\leq n+1} s_i$. Then $|f(x) - f(y)| \leq M|x - y|$ for all $x, y \in [a, b]$.*

Proof. Without loss of generality, suppose $x \leq y$. If there exists $1 \leq i \leq n + 1$ such that x and y are both in $[x_{i-1}, x_i]$, then by Theorem 5.24, $|f(x) - f(y)| \leq s_i(x - y) \leq M(x - y)$. Otherwise, there exists $0 \leq j \leq k \leq n + 1$ such that $x \leq x_j < x_{j+1} < \cdots < x_k \leq y$. Then

$$\begin{aligned}
|f(x) - f(y)| &\leq |f(x) - f(x_j)| + \sum_{i=j+1}^{k} |f(x_i) - f(x_{i-1})| + |f(x_k) - f(y)| \\
&\leq s_j(x_j - x) + \sum_{i=j+1}^{k} s_i(x_i - x_{i-1}) + s_{k+1}(y - x_k) \quad \text{by Theorem 5.24} \\
&\leq M(x_j - x) + \sum_{i=j+1}^{k} M(x_i - x_{i-1}) + M(y - x_k) \\
&= M(y - x).
\end{aligned}$$

□

Exercise 5.26. If $f : X \to Y$ is Lipschitz continuous and $g : Y \to Z$ is Lipschitz continuous, is $g \circ f : X \to Z$ Lipschitz continuous?

5.4 The Intermediate Value Theorem for Derivatives

The following theorem shows that the derivative of a function always satisfies the conclusion of the Intermediate Value Theorem, without assuming that the derivative itself is continuous.

Theorem 5.27 (Darboux)**.** *Let $f : (a, b) \to \mathbb{R}$ be differentiable, and let $c, d \in (a, b)$ such that $c < d$. Then for any $y_0 \in \mathbb{R}$ between $f'(c)$ and $f'(d)$ inclusive, there exists $x_0 \in [c, d]$ such that $y_0 = f'(x_0)$.*

Proof. Let $g(x) = f(x) - y_0x$ for all $x \in (a, b)$. Then $g'(x) = f'(x) - y_0$ for all $x \in (a, b)$. We just need to show that there exists $x_0 \in [c, d]$ such that $g'(x_0) = 0$.

Without loss of generality, suppose $f'(c) \leq f'(d)$. Observe that $g'(c) = f'(c) - y_0 \leq 0$ and $g'(d) = f'(d) - y_0 \geq 0$. If $g'(c) = 0$ or $g'(d) = 0$, then we are done since we can pick $x_0 = c$ or $x_0 = d$ as appropriate. Suppose, then, that $g'(c) < 0 < g'(d)$. By the Extreme Value Theorem, there exists $x_0 \in [c, d]$ such that $g(x_0) \leq g(x)$ for all $x \in [c, d]$. Since $g'(c) = \lim_{x\to c^+} \frac{g(x)-g(c)}{x-c} < 0$, there must exist $x_1 \in (c, d)$ such that $\frac{g(x_1)-g(c)}{x_1-c} < 0$. Then $g(x_0) \leq g(x_1) < g(c)$, so $x_0 \neq c$. A similar argument shows that $x_0 \neq d$. Hence, $x_0 \in (c, d)$, so g has a local minimum at x_0, which means $g'(x_0) = 0$. □

Darboux's Theorem does not imply that derivatives are continuous. In fact, the converse of the Intermediate Value Theorem does not always hold. The following example shows a derivative with a discontinuity.

Example 5.28. Define $f : \mathbb{R} \to \mathbb{R}$ by

$$f(x) = \begin{cases} x^2 \sin\left(\frac{1}{x}\right) & \text{if } x \neq 0 \\ 0 & \text{if } x = 0. \end{cases}$$

For this example, let us recall the derivative of the sine function from introductory calculus. If $x \neq 0$, then by the product and chain rules,

$$f'(x) = 2x \sin\left(\frac{1}{x}\right) + \left[x^2 \cos\left(\frac{1}{x}\right)\right]\left(-\frac{1}{x^2}\right) = 2x \sin\left(\frac{1}{x}\right) - \cos\left(\frac{1}{x}\right).$$

Since $0 \leq |x \sin\left(\frac{1}{x}\right)| \leq |x|$ for all $x \neq 0$, we have that $\lim_{x\to 0} |x \sin\left(\frac{1}{x}\right)| = 0$ by the Squeeze Theorem. Hence, $\lim_{x\to 0} x \sin\left(\frac{1}{x}\right) = 0$, so

$$f'(0) = \lim_{x\to 0} \frac{f(x) - f(0)}{x - 0} = \lim_{x\to 0} \frac{x^2 \sin(\frac{1}{x})}{x} = \lim_{x\to 0} x \sin\left(\frac{1}{x}\right) = 0.$$

Let $x_n = \frac{1}{2n\pi}$ for each $n \in \mathbb{N}$. Then $\lim_{n\to\infty} x_n = 0$, but

$$\lim_{n\to\infty} f'(x_n) = \lim_{n\to\infty} \left[\frac{1}{n\pi} \sin(2n\pi) - \cos(2n\pi)\right] = \lim_{n\to\infty} -1 = -1 \neq f'(0).$$

Therefore, f' is not continuous at 0.

5.5 The One-Dimensional Inverse Function Theorem

Theorem 5.29. *Let $E \subset \mathbb{R}$, $f : (a, b) \to E$, and $c \in (a, b)$. Suppose there exists $r > 0$ such that f is differentiable on $N_r(c)$, and suppose $f'(x) \neq 0$ for all $x \in N_r(c)$. Then f is injective on $N_r(c)$. Let $g : f(N_r(c)) \to N_r(c)$ be the inverse of f on $N_r(c)$. Then g is differentiable and*

$$g'(f(x)) = \frac{1}{f'(x)}$$

for all $x \in N_r(c)$.

Proof. By Darboux's Theorem, $f'(x)$ must have the same sign for all $x \in N_r(c)$. Without loss of generality, suppose $f'(x) > 0$ for all $x \in N_r(c)$. Then f is strictly increasing and hence injective on $N_r(c)$. Let $g : f(N_r(c)) \to N_r(c)$ be the inverse of f on $N_r(c)$. [To be precise, let $h : N_r(c) \to f(N_r(c))$ be the restriction of f to $N_r(c)$. Then h is invertible, and we let $g = h^{-1} : f(N_r(c)) \to N_r(c)$.]

Fix $x_0 \in N_r(c)$ and $\epsilon > 0$. We want to show that

$$\lim_{y\to f(x_0)} \frac{g(y) - g(f(x_0))}{y - f(x_0)} = \frac{1}{f'(x_0)}.$$

It suffices to show that

$$\lim_{y\to f(x_0)} \frac{y - f(x_0)}{g(y) - x_0} = f'(x_0)$$

since this fact (along with the assumption that $f'(x_0) \neq 0$) would imply that

$$\lim_{y \to f(x_0)} \frac{g(y) - g(f(x_0))}{y - f(x_0)} = \lim_{y \to f(x_0)} \frac{g(y) - x_0}{y - f(x_0)} = \frac{1}{\lim_{y \to f(x_0)} \frac{y - f(x_0)}{g(y) - x_0}} = \frac{1}{f'(x_0)}.$$

Since $\lim_{x \to x_0} \frac{f(x) - f(x_0)}{x - x_0} = f'(x_0)$, there exists $\delta_1 > 0$ such that $0 < |x - x_0| < \delta_1$ implies $|\frac{f(x) - f(x_0)}{x - x_0} - f'(x_0)| < \epsilon$. Choose $s > 0$ small enough such that $s < \delta_1$ and $N_s(x_0) \subset N_r(c)$. Now choose $x_-, x_+ \in N_s(x_0)$ such that $x_- < x_0 < x_+$. Then $(x_-, x_+) \subset N_s(x_0) \subset N_r(c)$. Since f is strictly increasing and continuous on $N_r(c)$, the Intermediate Value Theorem implies that $f((x_-, x_+)) = (f(x_-), f(x_+))$. Since $f(x_-) < f(x_0) < f(x_+)$, there exists $\delta_2 > 0$ such that $N_{\delta_2}(f(x_0)) \subset (f(x_-), f(x_+))$. Fix $y \in f(N_r(c))$ such that $0 < |y - f(x_0)| < \delta_2$. Then $f(x_-) < y < f(x_+)$, so $x_- < g(y) < x_+$ since f is strictly increasing on $N_r(c)$. Also, $g(y) \neq x_0$ since $y \neq f(x_0)$, so $0 < |g(y) - x_0| < \delta_1$ since $y \in (x_-, x_+) \subset N_s(x_0)$ and $s < \delta_1$. Therefore,

$$\left| \frac{f(g(y)) - f(x_0)}{g(y) - x_0} - f'(x_0) \right| = \left| \frac{y - f(x_0)}{g(y) - x_0} - f'(x_0) \right| < \epsilon,$$

so $\lim_{y \to f(x_0)} \frac{y - f(x_0)}{g(y) - x_0} = f'(x_0)$ as required. $\square$

Proposition 5.30. *Fix $n \in \mathbb{N}$ and let $f(x) = x^{1/n}$ for all $x \geq 0$. Then $f'(x) = \frac{1}{n} x^{(1/n)-1}$ for all $x > 0$.*

Proof. Let $g : [0, \infty) \to [0, \infty)$ be defined by $g(x) = x^n$. Then $g'(x) = nx^{n-1} > 0$ for all $x > 0$. Also, $f : [0, \infty) \to [0, \infty)$ defined by $f(x) = x^{1/n}$ is the inverse of g since $f(g(x)) = x = g(f(x))$ for all $x \in [0, \infty)$. Fix $x_0 \in (0, \infty)$, and let $y_0 = g(x_0) \in (0, \infty)$. Then

$$f'(x_0) = f'(g(y_0)) = \frac{1}{g'(y_0)} = \frac{1}{n y_0^{n-1}} = \frac{1}{n (x_0^{1/n})^{n-1}} = \frac{x_0^{1/n}}{n x_0} = \frac{1}{n} x_0^{(1/n)-1}.$$

$\square$

From here, we can derive the power rule for rational exponents. Fix $q \in \mathbb{Q}$, and write $q = \frac{m}{n}$ where $m \in \mathbb{Z}$ and $n \in \mathbb{N}$. Let $f(x) = x^q = (x^{1/n})^m$ for all $x > 0$ (see Exercise 1.24). Then by the Chain Rule,

$$f'(x) = [m(x^{1/n})^{m-1}][\frac{1}{n} x^{(1/n)-1}] = \frac{m}{n} (x^{1/n})^m x^{-1} = \frac{m}{n} x^{(m/n)-1} = q x^{q-1}$$

for all $x > 0$. We will use this result to prove the full power rule for real exponents in Exercise 7.24.

5.6 Taylor's Theorem

Taylor's Theorem generalizes the Mean Value Theorem.

Theorem 5.31 (Taylor's Theorem)**.** *Let $n \geq 0$ be an integer, and let $f : (a, b) \to \mathbb{R}$ be $(n+1)$-times differentiable. Let $c \in (a, b)$, and define*

$$T_n(x) = \sum_{k=0}^{n} \frac{f^{(k)}(c)}{k!} (x - c)^k$$

(remember that $0^0 = 1$ by convention). Then for any $x \in (a,b)$, there exists ξ between x and c such that

$$f(x) - T_n(x) = \frac{f^{(n+1)}(\xi)}{(n+1)!}(x-c)^{n+1}.$$

Proof. First, observe that $f^{(k)}(c) = T_n^{(k)}(c)$ for all $0 \le k \le n$ and that $T_n^{(n+1)}(y) = 0$ for all $y \in \mathbb{R}$.

If $x = c$, then we can just pick $\xi = c$ since

$$f(c) - T_n(c) = 0 = \frac{f^{(n+1)}(c)}{(n+1)!}(c-c)^{n+1}.$$

Suppose $x \neq c$. Define $g : (a,b) \to \mathbb{R}$ by

$$g(y) = f(y) - T_n(y) - \frac{A}{(n+1)!}(y-c)^{n+1}$$

where

$$A = \frac{f(x) - T_n(x)}{(x-c)^{n+1}}(n+1)!.$$

Then $g^{(k)}(c) = 0$ for all $0 \le k \le n$. Now note that $g(x) = 0 = g(c)$, so by Rolle's Theorem, there exists x_1 between x and c such that $g'(x_1) = 0$. By induction, suppose that for some integer $1 \le k \le n$, we have found x_k between x and c such that $g^{(k)}(x_k) = 0$. Then $g^{(k)}(x_k) = g^{(k)}(c)$. Since f is $(n+1)$-times diifferentiable, so is g. Hence, $g^{(k)}$ is differentiable, so by Rolle's Theorem, there exists x_{k+1} between x_k and c (hence between x and c) such that $g^{(k+1)}(x_{k+1}) = 0$. By induction, there exists x_{n+1} between x and c such that $g^{(n+1)}(x_{n+1}) = 0$. Let $\xi = x_{n+1}$. Then

$$0 = g^{(n+1)}(\xi) = f^{(n+1)}(\xi) - A = f^{(n+1)}(\xi) - \frac{f(x) - T_n(x)}{(x-c)^{n+1}}(n+1)!,$$

so

$$f(x) - T_n(x) = \frac{f^{(n+1)}(\xi)}{(n+1)!}(x-c)^{n+1}.$$

□

Exercise 5.32.

(a) Let $p : \mathbb{R} \to \mathbb{R}$ be a polynomial of degree n. Show that for any $c \in \mathbb{R}$, there exist constants $a_0, \ldots, a_n \in \mathbb{R}$ (dependent on c) such that

$$p(x) = \sum_{k=0}^{n} a_k (x-c)^k$$

for all $x \in \mathbb{R}$.

(b) Let $n \in \mathbb{N}$ and $c \in \mathbb{R}$. Show that for any polynomial $p : \mathbb{R} \to \mathbb{R}$, there exist a polynomial $q : \mathbb{R} \to \mathbb{R}$ and constants $b_1, \ldots, b_n$ such that

$$\frac{p(x)}{(x-c)^n} = q(x) + \sum_{k=1}^{n} \frac{b_k}{(x-c)^k}.$$

5.7 L'Hôpital's Rule

Theorem 5.33. *Let $f, g : (a, b) \to \mathbb{R}$ be differentiable. Suppose*

$$L = \lim_{x \to a^+} \frac{f'(x)}{g'(x)}$$

exists.

(a) If $\lim_{x \to a^+} f(x) = 0 = \lim_{x \to a^+} g(x)$, then

$$\lim_{x \to a^+} \frac{f(x)}{g(x)} = L.$$

(b) If $\lim_{x \to a^+} |g(x)| = \infty$, then

$$\lim_{x \to a^+} \frac{f(x)}{g(x)} = L.$$

Proof. Fix $\epsilon > 0$. There exists $\delta_1 > 0$ such that if $x \in (a, a+\delta_1)$, then $|\frac{f'(x)}{g'(x)} - L| < \min\{\epsilon, 1\}$. For any $x, y \in (a, a + \delta_1)$ such that $g(x) \neq g(y)$, there exists c between x and y such that

$$\frac{f(y) - f(x)}{g(y) - g(x)} = \frac{f'(c)}{g'(c)}$$

by Cauchy's Mean Value Theorem. Note that $c \in (a, a + \delta_1)$, so

$$\left| \frac{f(y) - f(x)}{g(y) - g(x)} - L \right| < \min\{\epsilon, 1\}. \tag{1}$$

(a) If $\lim_{y \to a^+} f(y) = 0 = \lim_{y \to a^+} g(y)$, we see that

$$\frac{f(x)}{g(x)} = \lim_{y \to a^+} \frac{f(y) - f(x)}{g(y) - g(x)} \in [L - \epsilon, L + \epsilon].$$

Therefore, $|\frac{f(x)}{g(x)} - L| \leq \epsilon$ for any $x \in (a, a + \delta_1)$ in the domain of $\frac{f}{g}$, so $\lim_{x \to a^+} \frac{f(x)}{g(x)} = L$.

(b) Choose $x = \frac{a+\delta_1}{2} \in (a, a + \delta_1)$. Since $\lim_{y \to a^+} |g(y)| = \infty$, we have

$$\lim_{y \to a^+} \frac{f(x)}{g(y)} = \lim_{y \to a^+} \frac{g(x)}{g(y)} = 0.$$

Hence, there exists $0 < \delta_2 < \delta_1$ such that $|\frac{f(x)}{g(y)}| < \epsilon$ and $|\frac{g(x)}{g(y)}| < \min\{\epsilon, 1\}$ whenever $y \in (a, a + \delta_2)$ and $g(y) \neq 0$. Then $g(y) \neq g(x)$ since $|\frac{g(x)}{g(y)}| < 1$, so division by $g(y) - g(x)$ is allowed. Hence,

$$\begin{aligned}
\left| L - \frac{f(y)}{g(y)} \right| &\leq \left| L - \frac{f(y) - f(x)}{g(y) - g(x)} \right| + \left| \frac{f(y) - f(x)}{g(y) - g(x)} - \frac{f(y) - f(x)}{g(y)} \right| + \left| \frac{f(y) - f(x)}{g(y)} - \frac{f(y)}{g(y)} \right| \\
&< \epsilon + \left| \frac{f(y) - f(x)}{g(y) - g(x)} - \frac{f(y) - f(x)}{g(y)} \right| + \left| \frac{-f(x)}{g(y)} \right| \quad \text{by (1)}
\end{aligned}$$

$$< 2\epsilon + \left| \frac{f(y)-f(x)}{g(y)-g(x)} - \frac{f(y)-f(x)}{g(y)} \right|.$$

Focusing on the remaining absolute-value term, we have

$$\begin{aligned}
\left| \frac{f(y)-f(x)}{g(y)-g(x)} - \frac{f(y)-f(x)}{g(y)} \right| &= \left| \frac{f(y)-f(x)}{g(y)-g(x)} \right| \cdot \left| \frac{g(x)}{g(y)} \right| \\
&\le \left(\left| \frac{f(y)-f(x)}{g(y)-g(x)} - L \right| + |L| \right) \left| \frac{g(x)}{g(y)} \right| \\
&< (1+|L|)\epsilon \quad \text{by (1)}.
\end{aligned}$$

Therefore,

$$\left| L - \frac{f(y)}{g(y)} \right| < (3+|L|)\epsilon$$

for all $y \in (a, a+\delta_2)$ in the domain of $\frac{f}{g}$. We conclude that $\lim_{y\to a^+} \frac{f(y)}{g(y)} = L$ since $3+|L|$ is independent of ϵ. $\square$

Remark. The proof can easily be extended to cover the case where $a = -\infty$. By symmetry, L'Hôpital's Rule also applies to the limit $\lim_{x\to b^-} \frac{f(x)}{g(x)}$ (including when $b = \infty$).

Exercise 5.34. Do Exercise 5.19 using L'Hôpital's Rule.

Exercise 5.35. Extend L'Hôpital's Rule to the case where $L = \pm\infty$.

6 The Riemann–Stieltjes Integral

In this chapter, we fix $a \le b$, $f : [a, b] \to \mathbb{R}$ bounded, and $\alpha : [a, b] \to \mathbb{R}$ monotonically increasing.

6.1 Definition of the Integral

Definition 6.1. A *partition* P of an interval $[a, b] \subset \mathbb{R}$ is a list of real numbers $x_0 \le x_1 \le \cdots \le x_n$ where $x_0 = a$ and $x_n = b$. Note that the points x_i need not be distinct. A *refinement* of P is another partition P' such that $P \subset P'$. The *common refinement* of two partitions P_1 and P_2 is $P_1 \cup P_2$.

Definition 6.2 (The Riemann–Stieltjes Integral)**.** Let $P = \{x_0, \ldots, x_n\}$ be a partition of $[a, b]$. Denote

$$M_i = \sup_{x \in [x_{i-1}, x_i]} f(x),$$

$$m_i = \inf_{x \in [x_{i-1}, x_i]} f(x),$$

and

$$\Delta\alpha_i = \alpha(x_i) - \alpha(x_{i-1})$$

for all $1 \le i \le n$. Note that M_i and m_i are finite since f is bounded by assumption. Then

$$U(P, f, \alpha) \coloneqq \sum_{i=1}^{n} M_i \Delta\alpha_i$$

and

$$L(P, f, \alpha) \coloneqq \sum_{i=1}^{n} m_i \Delta\alpha_i$$

are the *upper and lower sums* (respectively) of f with respect to P and α. The *upper integral* of f with respect to α is

$$\overline{\int_a^b} f \, d\alpha \coloneqq \inf_P U(P, f, \alpha),$$

and the *lower integral* of f with respect to α is

$$\underline{\int_a^b} f \, d\alpha \coloneqq \sup_P L(P, f, \alpha),$$

where the infimum and supremum are taken over all partitions P of $[a, b]$. If the upper and lower integrals are equal, then f is *Riemann–Stieltjes integrable* with respect to α on $[a, b]$, and we define the *Riemann–Stieltjes integral* of f with respect to α to be

$$\int_a^b f \, d\alpha \coloneqq \overline{\int_a^b} f \, d\alpha = \underline{\int_a^b} f \, d\alpha.$$

If f is integrable with respect to α, we write $f \in \mathcal{R}_\alpha[a, b]$.

Note that removing duplicate points from a partition P does not affect $U(P, f, \alpha)$ or $L(P, f, \alpha)$ since $\Delta\alpha_i = 0$ if $x_{i-1} = x_i$.

If $\alpha(x) = x$, the Riemann–Stieltjes integral $\int_a^b f\,d\alpha$ becomes the *Riemann integral* $\int_a^b f\,dx$. The notations $\mathcal{R}_\alpha[a,b]$, $U(P, f, \alpha)$, and $L(P, f, \alpha)$ become $\mathcal{R}[a,b]$, $U(P, f)$, and $L(P, f)$, respectively. If $f \in \mathcal{R}[a,b]$, we say that f is *Riemann integrable* on $[a,b]$.

6.2 Which Functions are Integrable?

It is now natural to ask which functions are in $\mathcal{R}_\alpha[a,b]$. We will not fully characterize all functions in $\mathcal{R}_\alpha[a,b]$, but we will see that $\mathcal{R}_\alpha[a,b]$ is a fairly general class of functions. For example, every continuous function on $[a,b]$ is in $\mathcal{R}_\alpha[a,b]$. The theory of integration will therefore be applicable to a wide variety of functions.

Theorem 6.3. *Let P be a partition of $[a,b]$. If P' is a refinement of P, then*

$$L(P, f, \alpha) \leq L(P', f, \alpha) \leq U(P', f, \alpha) \leq U(P, f, \alpha).$$

Proof. Write $P' = \{x_0, \ldots, x_n\}$. Since $\inf_{x\in[x_{i-1},x_i]} f(x) \leq \sup_{x\in[x_{i-1},x_i]} f(x)$, we immediately have that $L(P', f, \alpha) \leq U(P', f, \alpha)$. We now show that $L(P, f, \alpha) \leq L(P', f, \alpha)$; the proof that $U(P', f, \alpha) \leq U(P, f, \alpha)$ is similar. Since P' is a refinement of P, we have that $P = \{x_{m_0}, \ldots, x_{m_k}\}$ where $0 = m_0 < m_1 < \cdots < m_k = n$. Now

$$\begin{aligned}
L(P', f, \alpha) &= \sum_{i=1}^{n} \left(\inf_{x\in[x_{i-1},x_i]} f(x) \right) [\alpha(x_i) - \alpha(x_{i-1})] \\
&= \sum_{j=1}^{k} \sum_{i=m_{j-1}+1}^{m_j} \left(\inf_{x\in[x_{i-1},x_i]} f(x) \right) [\alpha(x_i) - \alpha(x_{i-1})] \\
&\geq \sum_{j=1}^{k} \sum_{i=m_{j-1}+1}^{m_j} \left(\inf_{x\in[x_{m_{j-1}},x_{m_j}]} f(x) \right) [\alpha(x_i) - \alpha(x_{i-1})] \\
&\quad \text{since } [x_{i-1}, x_i] \subset [x_{m_{j-1}}, x_{m_j}] \text{ for } m_{j-1} + 1 \leq i \leq m_j \\
&= \sum_{j=1}^{k} \left(\inf_{x\in[x_{m_{j-1}},x_{m_j}]} f(x) \right) [\alpha(m_j) - \alpha(m_{j-1})] \\
&= L(P, f, \alpha).
\end{aligned}$$

□

Theorem 6.4. *For any partition P of $[a,b]$,*

$$L(P, f, \alpha) \leq \underline{\int_a^b} f\,d\alpha \leq \overline{\int_a^b} f\,d\alpha \leq U(P, f, \alpha).$$

Proof. By Theorem 6.3, $L(P, f, \alpha) \leq L(P', f, \alpha) \leq U(P', f, \alpha) \leq U(P, f, \alpha)$ for all refinements $P' \supset P$. Hence,

$$L(P, f, \alpha) \leq \sup_{P' \supset P} L(P', f, \alpha) \leq \underline{\int_a^b} f\,d\alpha$$

and

$$U(P, f, \alpha) \geq \inf_{P' \supset P} U(P', f, \alpha) \geq \overline{\int_a^b} f \, d\alpha.$$

It remains to prove that $\underline{\int_a^b} f \, d\alpha \leq \overline{\int_a^b} f \, d\alpha$. Let Q be a partition of $[a, b]$, and let $P_* = P \cup Q$ be the common refinement. By Theorem 6.3,

$$L(P, f, \alpha) \leq L(P_*, f, \alpha) \leq U(P_*, f, \alpha) \leq U(Q, f, \alpha)$$

since P_* is a refinement of both P and Q. Since Q is arbitrary,

$$L(P, f, \alpha) \leq \inf_Q U(Q, f, \alpha) = \overline{\int_a^b} f \, d\alpha.$$

This inequality holds for any partition P, so

$$\underline{\int_a^b} f \, d\alpha = \sup_P L(P, f, \alpha) \leq \overline{\int_a^b} f \, d\alpha.$$

□

Theorem 6.5. *$f \in \mathcal{R}_\alpha[a, b]$ if and only if for all $\epsilon > 0$, there exists a partition P of $[a, b]$ such that $U(P, f, \alpha) - L(P, f, \alpha) < \epsilon$.*

Proof. Suppose $f \in \mathcal{R}_\alpha[a, b]$. Then $\sup_P L(P, f, \alpha) = \inf_P U(P, f, \alpha)$. Fix $\epsilon > 0$. Then there exist partitions P_1, P_2 of $[a, b]$ such that

$$L(P_1, f, \alpha) > \sup_P L(P, f, \alpha) - \frac{\epsilon}{2}$$

and

$$U(P_2, f, \alpha) < \inf_P U(P, f, \alpha) + \frac{\epsilon}{2}.$$

Let $P = P_1 \cup P_2$ be the common refinement. Then

$$L(P_1, f, \alpha) \leq L(P, f, \alpha) \leq U(P, f, \alpha) \leq U(P_2, f, \alpha),$$

so

$$U(P, f, \alpha) - L(P, f, \alpha) \leq U(P_2, f, \alpha) - L(P_1, f, \alpha) < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon$$

using that $\sup_P L(P, f, \alpha) = \inf_P U(P, f, \alpha)$.

Conversely, suppose that for all $\epsilon > 0$, there exists a partition P of $[a, b]$ such that $U(P, f, \alpha) - L(P, f, \alpha) < \epsilon$. Fix $\epsilon > 0$, and choose such a partition P. Then

$$\overline{\int_a^b} f \, d\alpha - \underline{\int_a^b} f \, d\alpha \leq U(P, f, \alpha) - L(P, f, \alpha) < \epsilon$$

by Theorem 6.4. Since $\epsilon > 0$ is arbitrary,

$$\overline{\int_a^b} f \, d\alpha - \underline{\int_a^b} f \, d\alpha \leq 0.$$

Hence, $\overline{\int_a^b} f \, d\alpha \leq \underline{\int_a^b} f \, d\alpha$. But by Theorem 6.4, $\overline{\int_a^b} f \, d\alpha \geq \underline{\int_a^b} f \, d\alpha$, so in fact, $\overline{\int_a^b} f \, d\alpha = \underline{\int_a^b} f \, d\alpha$, which means $f \in \mathcal{R}_\alpha[a, b]$. □

Theorem 6.6. *Suppose f is continuous. Then $f \in \mathcal{R}_\alpha[a,b]$.*

Proof. Since f is continuous and $[a,b]$ is a compact set, f is uniformly continuous on $[a,b]$. Hence, for any fixed $\epsilon > 0$, we can choose $\delta > 0$ such that if $x, y \in [a,b]$ and $|x-y| < \delta$, then $|f(x) - f(y)| < \epsilon$. Let $n \in \mathbb{N}$ be such that $\frac{1}{n} < \delta$, and choose $P = \{x_0, \ldots, x_n\}$ where $x_i = a + \frac{b-a}{n} i$ for all $0 \le i \le n$. We have that $a = x_0 \le x_1 \le \cdots \le x_n = b$, so P is indeed a partition of $[a,b]$. Note that $x_i - x_{i-1} = \frac{1}{n} < \delta$, so for any $x, y \in [x_{i-1}, x_i]$, we have $|x-y| < \delta$ and consequently $|f(x) - f(y)| < \epsilon$. It follows that $M_i - m_i \le \epsilon$ for all $1 \le i \le n$ (using the notation of Definition 6.2). Therefore,

$$U(P,f,\alpha) - L(P,f,\alpha) = \sum_{i=1}^{n} (M_i - m_i)\Delta\alpha_i \le \epsilon \sum_{i=1}^{n} \Delta\alpha_i = \epsilon[\alpha(b) - \alpha(a)].$$

The last equality follows since $\sum_{i=1}^{n} \Delta\alpha_i = \sum_{i=1}^{n} [\alpha(x_i) - \alpha(x_{i-1})]$ telescopes. By Theorem 6.5, $f \in \mathcal{R}_\alpha[a,b]$ since $\epsilon[\alpha(b) - \alpha(a)]$ becomes arbitrarily small with a suitable choice of ϵ. $\square$

Theorem 6.7. *Suppose α is continuous and f is monotonically increasing. Then $f \in \mathcal{R}_\alpha[a,b]$.*

Proof. The proof is essentially the same as the proof of Theorem 6.6, except that the roles of f and α are switched. For a fixed $\epsilon > 0$, pick $\delta > 0$ such that if $x, y \in [a,b]$ and $|x-y| < \delta$, then $|\alpha(x) - \alpha(y)| < \epsilon$; this is possible due to uniform continuity of α on the compact set $[a,b]$. Let $n \in \mathbb{N}$ be such that $\frac{1}{n} < \delta$, and choose $P = \{x_0, \ldots, x_n\}$ where $x_i = a + \frac{b-a}{n} i$ for each $0 \le i \le n$. Since $|x_i - x_{i-1}| = \frac{1}{n} < \delta$, we have that $\alpha(x_i) - \alpha(x_{i-1}) < \epsilon$. Also, since f is monotonically increasing, we have that $M_i = f(x_i)$ and $m_i = f(x_{i-1})$ for all $1 \le i \le n$. Therefore,

$$\begin{aligned} U(P,f,\alpha) - L(P,f,\alpha) &= \sum_{i=1}^{n} [f(x_i) - f(x_{i-1})][\alpha(x_i) - \alpha(x_{i-1})] \\ &\le \epsilon \sum_{i=1}^{n} [f(x_i) - f(x_{i-1})] \\ &= \epsilon[f(b) - f(a)], \end{aligned}$$

so $f \in \mathcal{R}_\alpha[a,b]$ by Theorem 6.5 since $\epsilon[f(b) - f(a)]$ can be made arbitrarily small. $\square$

Theorem 6.8. *Suppose f is discontinuous at only finitely many points and that α is continuous wherever f is not. Then $f \in \mathcal{R}_\alpha[a,b]$.*

Proof. Let $t_0 < \cdots < t_k$ be the points of $[a,b]$ at which f is discontinuous. The idea is to create partitions P where each t_i is surrounded by a pair of points in P. We will ensure that the space enclosed by each pair of points is very small. As a result, the contribution of the t_i to $U(P,f,\alpha) - L(P,f,\alpha)$ will be very small. The proof of Theorem 6.6 tells us how to handle the portions of $[a,b]$ on which f is continuous. We now give the technical details.

Fix $\epsilon > 0$. Let $\delta_1 = \min_{1 \le i \le k}\{t_i - t_{i-1}\} > 0$. Since α is continuous at each t_i, and there are only finitely many of them, there exists $0 < \delta_2 < \delta_1$ such that for all $0 \le i \le k$, if $x \in [a,b]$ and $|x - t_i| < \delta_2$, then $|\alpha(x) - \alpha(t_i)| < \epsilon$. We construct a partition of $[a,b]$ with the following steps. We start off with an empty list P. First, add a and b to P. Next, for each $0 \leq i \leq k$, if $x_i = t_i - \frac{1}{2}\delta_2 \in [a,b]$, add x_i to P, and do the same with $y_i = t_i + \frac{1}{2}\delta_2$ if $y_i \in [a,b]$. Note that x_i and y_i are not equal to any of the t_j since $\delta_2 \neq 0$ and $|t_j - t_i| \geq \delta_1 > \frac{1}{2}\delta_2$ for any $j \neq i$. For the next step, consider the set

$$E = [a,b] \cap \left(\bigcup_{i=0}^{k} (x_i, y_i) \right)^c,$$

which is closed and bounded in $\mathbb{R}$, hence compact. Note that f is continuous (and hence uniformly continuous) on E since none of the x_i and y_i are points of discontinuity of f. Choose $\delta_3 > 0$ such that if $x, y \in E$ and $|x - y| < \delta_3$, then $|f(x) - f(y)| < \epsilon$. Choose $n \in \mathbb{N}$ such that $\frac{1}{n} < \delta_3$, and for each $0 \leq i \leq n$, if $p_i = a + \frac{b-a}{n} i \in E$, add p_i to P. Finally, ensure that all elements in P are distinct by removing any duplicate elements from P.

Write the resulting list P as $\{s_0, \ldots, s_n\}$. We now observe that if $[s_{i-1}, s_i]$ contains t_j for some $0 \leq j \leq k$, then $[s_{i-1}, s_i] \subset [x_j, y_j]$. Hence, $|s_{i-1} - t_j| < \delta_2$ and $|s_i - t_j| < \delta_2$, so $\Delta\alpha_i = \alpha(s_i) - \alpha(s_{i-1}) = [\alpha(s_i) - \alpha(t_j)] + [\alpha(t_j) - \alpha(s_{i-1})] < 2\epsilon$. Note that there are at most k intervals $[s_{i-1}, s_i]$ that contain one of the t_j. On the other hand, if $[s_{i-1}, s_i]$ does not contain any of the t_j, then $[s_{i-1}, s_i] \subset E$, so $s_i - s_{i-1} \leq \frac{1}{n} < \delta$, which implies that $M_i - m_i \leq \epsilon$. Denote $M = \sup_{x \in [a,b]} |f(x)|$, and note that $M_i - m_i < 2M$ for all $1 \leq i \leq n$. Therefore,

$$\begin{aligned}
U(P,f,\alpha) - L(P,f,\alpha) &= \sum_{t_j \in [s_{i-1}, s_i]} (M_i - m_i)\Delta\alpha_i + \sum_{t_j \notin [s_{i-1}, s_i]} (M_i - m_i)\Delta\alpha_i \\
&\leq 2\epsilon \sum_{t_j \in [s_{i-1}, s_i]} 2M + \epsilon \sum_{t_j \notin [s_{i-1}, s_i]} \Delta\alpha_i \\
&\leq 4Mk\epsilon + \epsilon[\alpha(b) - \alpha(a)] \\
&= [4Mk + \alpha(b) - \alpha(a)]\epsilon.
\end{aligned}$$

Since $4Mk + \alpha(b) - \alpha(a)$ is independent of ϵ, $f \in \mathcal{R}_\alpha[a,b]$ by Theorem 6.5. □

Theorem 6.9. *Suppose $f \in \mathcal{R}_\alpha[a,b]$. Let $m, M \in \mathbb{R}$ such that $m \leq f(x) \leq M$ for all $x \in [a,b]$. Suppose $g : [m,M] \to \mathbb{R}$ is continuous. Then $g \circ f \in \mathcal{R}_\alpha[a,b]$.*

Proof. Fix $\epsilon > 0$. Since $[m,M]$ is compact, g is uniformly continuous, so there exists $\delta > 0$ such that if $x, y \in [m,M]$ and $|x - y| < \delta$, then $|g(x) - g(y)| < \epsilon$. Let $K = \sup_{x \in [a,b]} |g(f(x))|$. For each partition $P = \{x_i\}_{i=0}^n$ of $[a,b]$, let

$$S_P = \{1 \leq i \leq n : M_i - m_i \geq \delta\}$$

and

$$N_P = \sum_{i \in S_P} \Delta\alpha_i,$$

where $M_i = \sup_{x \in [x_{i-1}, x_i]} f(x)$, $m_i = \inf_{x \in [x_{i-1}, x_i]} f(x)$, and $\Delta\alpha_i = \alpha(x_i) - \alpha(x_{i-1})$. We can think of N_P as measuring the portion of the interval $[a,b]$ on which $M_i - m_i$ is uncontrolled.

The key ingredient of this proof is the following observation: there must exist a partition Q of $[a,b]$ such that $N_Q < \epsilon$. If not, then every partition P of $[a,b]$ would satisfy $N_P \geq \epsilon$ and consequently

$$U(P,f,\alpha) - L(P,f,\alpha) \geq \sum_{i \in S_P} (M_i - m_i)\Delta\alpha_i \geq \delta N_P \geq \delta\epsilon.$$

But this means $U(P,f,\alpha) - L(P,f,\alpha)$ cannot get arbitrarily small, contradicting the assumption that $f \in \mathcal{R}_\alpha[a,b]$.

Therefore, we can take a partition $Q = \{x_0, \ldots, x_n\}$ such that $N_Q < \epsilon$. Denote $U_i = \sup_{x \in [x_{i-1}, x_i]} g(f(x))$ and $L_i = \inf_{x \in [x_{i-1}, x_i]} g(f(x))$. Note that if $i \notin S_Q$ and $x, y \in [x_{i-1}, x_i]$, then $|f(x) - f(y)| \leq M_i - m_i < \delta$, so $|g(f(x)) - g(f(y))| < \epsilon$. Hence, if $i \notin S_Q$, then $U_i - L_i \leq \epsilon$. Also, for any $1 \leq i \leq n$, we have $U_i - L_i \leq 2K$. Therefore,

$$\begin{aligned}
U(Q, g \circ f, \alpha) - L(Q, g \circ f, \alpha) &= \sum_{i=1}^{n} (U_i - L_i)\Delta\alpha_i \\
&= \sum_{i \notin S_Q} (U_i - L_i)\Delta\alpha_i + \sum_{i \in S_Q} (U_i - L_i)\Delta\alpha_i \\
&\leq \epsilon[\alpha(b) - \alpha(a)] + 2KN_Q \\
&\leq \epsilon[\alpha(b) - \alpha(a)] + 2K\epsilon \\
&= [\alpha(b) - \alpha(a) + 2K]\epsilon.
\end{aligned}$$

By Theorem 6.5, $g \circ f \in \mathcal{R}_\alpha[a,b]$ since $\alpha(b) - \alpha(a) + 2K$ is independent of ϵ. □

Exercise 6.10 (Right Riemann Sum)**.** Suppose $f \in \mathcal{R}[0,1]$. Prove that

$$\lim_{n \to \infty} \sum_{k=1}^{n} \frac{1}{n} f\left(\frac{k}{n}\right) = \int_0^1 f\, dx.$$

6.3 Properties of the Integral

Theorem 6.11 (Linearity Properties)**.** *Suppose $f, g : [a,b] \to \mathbb{R}$ are bounded, $\alpha, \beta : [a,b] \to \mathbb{R}$ are monotonically increasing, and $c \in \mathbb{R}$.*

(a) If $f \in \mathcal{R}_\alpha[a,b]$, then $cf \in \mathcal{R}_\alpha[a,b]$ and

$$\int_a^b cf\, d\alpha = c \int_a^b f\, d\alpha.$$

(b) If $f, g \in \mathcal{R}_\alpha[a,b]$, then $f + g \in \mathcal{R}_\alpha[a,b]$ and

$$\int_a^b (f+g)\, d\alpha = \int_a^b f\, d\alpha + \int_a^b g\, d\alpha.$$

(c) If $f \in \mathcal{R}_\alpha[a,b]$ and $c \geq 0$, then $f \in \mathcal{R}_{c\alpha}[a,b]$ and

$$\int_a^b f\, d(c\alpha) = c \int_a^b f\, d\alpha.$$

(d) If $f \in \mathcal{R}_\alpha[a,b]$ and $f \in \mathcal{R}_\beta[a,b]$, then $f \in \mathcal{R}_{\alpha+\beta}[a,b]$ and

$$\int_a^b f\, d(\alpha+\beta) = \int_a^b f\, d\alpha + \int_a^b f\, d\beta.$$

Proof. Parts (a) and (b) follow from properties of suprema and infima, and parts (c) and (d) follow from linearity of sums. The proof for part (b) is worth elaborating on. It is not hard to show that

$$\underline{\int_a^b} f\, d\alpha + \underline{\int_a^b} g\, d\alpha \le \underline{\int_a^b} (f+g)\, d\alpha \le \overline{\int_a^b} (f+g)\, d\alpha \le \overline{\int_a^b} f\, d\alpha + \overline{\int_a^b} g\, d\alpha$$

for any bounded $f, g : [a,b] \to \mathbb{R}$. If $f, g \in \mathcal{R}_\alpha[a,b]$, then $\overline{\int_a^b} f\, d\alpha = \underline{\int_a^b} f\, d\alpha = \int_a^b f\, d\alpha$ and $\overline{\int_a^b} g\, d\alpha = \underline{\int_a^b} g\, d\alpha = \int_a^b g\, d\alpha$. This causes the inequalities above to become equalities, and part (b) follows. □

Corollary 6.11.1. *Suppose $f, g \in \mathcal{R}_\alpha[a,b]$. Then $fg \in \mathcal{R}_\alpha[a,b]$.*

Proof. The map $x \mapsto x^2$ is continuous on $\mathbb{R}$, so $f^2, g^2, (f+g)^2 \in \mathcal{R}_\alpha[a,b]$ by Theorems 6.9 and 6.11. Consequently, $fg = \frac{1}{2}[(f+g)^2 - f^2 - g^2] \in \mathcal{R}_\alpha[a,b]$ by Theorem 6.11. □

Remark. Theorem 6.11 and its corollary show that $\mathcal{R}_\alpha[a,b]$ is an "$\mathbb{R}$-algebra" (i.e. $\mathcal{R}_\alpha[a,b]$ is a vector space over $\mathbb{R}$ together with a multiplication operation that interacts nicely with the vector-space operations).

Theorem 6.12. *Let $c \in [a,b]$. Then $f \in \mathcal{R}_\alpha[a,b]$ if and only if $f \in \mathcal{R}_\alpha[a,c]$ and $f \in \mathcal{R}_\alpha[c,b]$, in which case we have that*

$$\int_a^b f\, d\alpha = \int_a^c f\, d\alpha + \int_c^b f\, d\alpha.$$

Proof. Fix $\epsilon > 0$. Suppose $f \in \mathcal{R}_\alpha[a,b]$. By Theorem 6.5, there exists a partition $P = \{x_0, \ldots, x_n\}$ of $[a,b]$ such that $U(P,f,\alpha) - L(P,f,\alpha) < \epsilon$. Construct a refinement $P_* = \{y_0, \ldots, y_{n+1}\}$ of P by inserting c into P. Since P_* is a refinement, we have that

$$U(P_*,f,\alpha) - L(P_*,f,\alpha) \le U(P,f,\alpha) - L(P,f,\alpha) < \epsilon$$

by Theorem 6.3. Let $P_1 = \{y_0, \ldots, c\}$ and $P_2 = \{c, \ldots, y_{n+1}\}$. Then

$$U(P_1,f,\alpha) - L(P_1,f,\alpha) \le U(P_*,f,\alpha) - L(P_*,f,\alpha) < \epsilon$$

and

$$U(P_2,f,\alpha) - L(P_2,f,\alpha) \le U(P_*,f,\alpha) - L(P_*,f,\alpha) < \epsilon.$$

Since P_1 and P_2 are parititons of $[a,c]$ and $[c,b]$, respectively, Theorem 6.5 implies that $f \in \mathcal{R}_\alpha[a,c]$ and $f \in \mathcal{R}_\alpha[c,b]$. Notice that

$$U(P_1,f,\alpha) + U(P_2,f,\alpha) = U(P_*,f,\alpha)$$

and

$$L(P_1, f, \alpha) + L(P_2, f, \alpha) = L(P_*, f, \alpha).$$

Using Theorem 6.4, we see that

$$\begin{aligned}\left(\int_a^c f\,d\alpha + \int_c^b f\,d\alpha\right) - \int_a^b f\,d\alpha &\leq U(P_1, f, \alpha) + U(P_2, f, \alpha) - L(P_*, f, \alpha)\\ &= U(P_*, f, \alpha) - L(P_*, f, \alpha)\\ &< \epsilon\end{aligned}$$

and

$$\begin{aligned}\int_a^b f\,d\alpha - \left(\int_a^c f\,d\alpha + \int_c^b f\,d\alpha\right) &\leq U(P_*, f, \alpha) - L(P_1, f, \alpha) - L(P_2, f, \alpha)\\ &= U(P_*, f, \alpha) - L(P_*, f, \alpha)\\ &< \epsilon,\end{aligned}$$

so taking $\epsilon \to 0^+$ yields the formula $\int_a^b f\,d\alpha = \int_a^c f\,d\alpha + \int_c^b f\,d\alpha$.

Conversely, suppose $f \in \mathcal{R}_\alpha[a, c]$ and $f \in \mathcal{R}_\alpha[c, b]$. Let $P_1 = \{x_0, \ldots, x_n\}$ be a partition of $[a, c]$ and $P_2 = \{y_0, \ldots, y_m\}$ be a partition $[c, b]$ such that $U(P_1, f, \alpha) - L(P_1, f, \alpha) < \frac{\epsilon}{2}$ and $U(P_2, f, \alpha) - L(P_2, f, \alpha) < \frac{\epsilon}{2}$. Note that $x_n = c = y_0$. Consider the partition $P = \{x_0, \ldots, x_n, y_0, \ldots, y_m\}$ of $[a, b]$. Since $x_n = y_0$, it follows that

$$U(P, f, \alpha) - L(P, f, \alpha) = U(P_1, f, \alpha) + U(P_2, f, \alpha) - L(P_1, f, \alpha) - L(P_2, f, \alpha) < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon.$$

Hence, $f \in \mathcal{R}_\alpha[a, b]$. □

As a corollary of Theorem 6.12, we note that if $[c, d] \subset [a, b]$ and $f \in \mathcal{R}_\alpha[a, b]$, then $f \in \mathcal{R}_\alpha[c, d]$. Indeed, if $f \in \mathcal{R}_\alpha[a, b]$, then $f \in \mathcal{R}_\alpha[c, b]$ since $c \in [a, b]$, and hence $f \in \mathcal{R}_\alpha[c, d]$ since $d \in [c, b]$. We now use this fact to prove a theorem about the continuity of Riemann-integrable functions.

Lemma 6.13. *Let $a < b$. If $f \in \mathcal{R}[a, b]$, then f is continuous at some point in $[a, b]$.*

Proof. Suppose $f \in \mathcal{R}[a, b]$. For any partition $P = \{x_0, \ldots, x_n\}$ and any integer $1 \leq i \leq n - 1$, denote $M_{i,P} = \sup_{x \in [x_{i-1}, x_i]} f(x)$ and $m_{i,P} = \inf_{x \in [x_{i-1}, x_i]} f(x)$. There must exist a partition $P_1 = \{x_{0,1}, \ldots, x_{n_1,1}\}$ of distinct points of $[a, b]$ such that $M_{i_1,P_1} - m_{i_1,P_1} < 1$ for some $1 \leq i_1 \leq n_1$. Otherwise, $U(P, f) - L(P, f) \geq 1(b - a)$ for all partitions P of $[a, b]$, contradicting that $f \in \mathcal{R}[a, b]$. Without loss of generality, we can assume that $2 \leq i_1 \leq n_1 - 1$ because if needed, we can split $[x_0, x_1]$ or $[x_{n_1-1}, x_{n_1}]$ into three equispaced subintervals and take the endpoints of the middle subinterval to be x_{i_1-1} and x_{i_1}. We know that $f \in \mathcal{R}[x_{i_1-1}, x_{i_1}]$ since $[x_{i_1-1}, x_{i_1}] \subset [a, b]$. By similar reasoning as before, there must exist a partition $P_2 = \{x_{0,2}, \ldots, x_{n_2,2}\}$ of distinct points of $[x_{i_1-1}, x_{i_1}]$ such that $M_{i_2,P_2} - m_{i_2,P_2} < \frac{1}{2}$ for some $2 \leq i_2 \leq n_2 - 1$. Continuing inductively, we see that for all $j \geq 2$, there must exist a partition $P_j = \{x_{0,j}, \ldots, x_{n_j,j}\}$ of distinct points of $[x_{i_{j-1}-1}, x_{i_{j-1}}]$ such that $M_{i_j,P_j} - m_{i_j,P_j} < \frac{1}{j}$ for some $2 \leq i_j \leq n_j - 1$. Denote $K_j = [x_{i_j-1}, x_{i_j}]$ for each $j \in \mathbb{N}$. Then K_j is compact and

$K_{j+1} \subset K_j$ for all $j \in \mathbb{N}$, so $K = \bigcap_{j=1}^{\infty} K_j$ is non-empty. Pick $x \in K$, and fix $\epsilon > 0$. Pick $j \in \mathbb{N}$ such that $\frac{1}{j} < \epsilon$. Then $x \in K_{j+1} = [x_{i_{j+1}-1}, x_{i_{j+1}}] \subset (x_{i_j-1}, x_{i_j})$ by construction since $2 \leq i_{j+1} \leq n_{j+1} - 1$. There exists $\delta > 0$ such that $N_\delta(x) \subset (x_{i_j-1}, x_{i_j})$ since (x_{i_j-1}, x_{i_j}) is open. Note that

$$\sup_{x \in [x_{i_j-1}, x_{i_j}]} f(x) - \inf_{x \in [x_{i_j-1}, x_{i_j}]} f(x) = M_{i_j, P_j} - m_{i_j, P_j} < \frac{1}{j} < \epsilon,$$

so if $|y - x| < \delta$, then $|f(y) - f(x)| < \epsilon$ since $y \in (x_{i_j-1}, x_{i_j})$. Thus, f is continuous at x. $\square$

Theorem 6.14. *Suppose $f \in \mathcal{R}[a,b]$ where $a < b$. Let $S \subset [a,b]$ be the set of points at which f is continuous. Then S is dense in $[a,b]$.*

Proof. Suppose $x \in [a,b]$. Fix $r > 0$, and define $s = \max\{a, x - \frac{r}{2}\}$ and $t = \min\{b, x + \frac{r}{2}\}$. A case-by-case analysis shows that $s < t$. We have that $[s,t] \subset [a,b]$ and $[s,t] \subset N_r(x)$. Hence, $f \in \mathcal{R}[s,t]$. By Lemma 6.13, there exists $x_0 \in [s,t] \subset N_r(x)$ such that $x_0 \in S$. Therefore, x is a limit point of S. It follows that $\overline{S} = [a,b]$. $\square$

Now we resume our discussion of some basic properties of the integral.

Theorem 6.15 (Basic Inequalities)**.**
(a) Suppose $f, g \in \mathcal{R}_\alpha[a,b]$ and $f(x) \leq g(x)$ for all $x \in [a,b]$. Then

$$\int_a^b f\,d\alpha \leq \int_a^b g\,d\alpha.$$

(b) Suppose $f \in \mathcal{R}_\alpha[a,b]$. Then

$$\left|\int_a^b f\,d\alpha\right| \leq \int_a^b |f|\,d\alpha.$$

(c) Suppose $f \in \mathcal{R}_\alpha[a,b]$, and suppose $|f(x)| \leq M$ for all $x \in [a,b]$. Then

$$\left|\int_a^b f\,d\alpha\right| \leq M[\alpha(b) - \alpha(a)].$$

Proof. (a) Observe that $U(P,f,\alpha) \leq U(P,g,\alpha)$ for any partition P. Thus, $\overline{\int_a^b} f\,d\alpha \leq \overline{\int_a^b} g\,d\alpha$. Conclude by using that $f, g \in \mathcal{R}_\alpha[a,b]$.

(b) Since $x \mapsto |x|$ is continuous on $\mathbb{R}$, Theorem 6.9 implies that $|f| \in \mathcal{R}_\alpha[a,b]$. For any partition $P = \{x_0, \ldots, x_n\}$, we have that

$$\left|\sup_{x \in [x_{i-1}, x_i]} f(x)\right| \leq \sup_{x \in [x_{i-1}, x_i]} |f(x)|$$

for all $1 \leq i \leq n$ (if not, a contradiction is easily obtained). The Triangle Inequality now implies that $|U(P,f,\alpha)| \leq U(P,|f|,\alpha)$. Hence, $\left|\overline{\int_a^b} f\,d\alpha\right| \leq \overline{\int_a^b} |f|\,d\alpha$, and the desired result follows since $f, |f| \in \mathcal{R}_\alpha[a,b]$.

(c) Apply part (b) and then part (a) with $g(x) = M$. A full proof would show that $\int_a^b M\,d\alpha = M[\alpha(b) - \alpha(a)]$, but this is an easy exercise. $\square$

Theorem 6.16 (Mean Value Theorem for Integrals)**.** *Let f be continuous on $[a, b]$. Then there exists $c \in [a, b]$ such that*

$$[\alpha(b) - \alpha(a)]f(c) = \int_a^b f \, d\alpha.$$

Proof. By the Extreme Value Theorem, there exists $x_+, x_- \in [a, b]$ such that $f(x_+)$ is a global maximum and $f(x_-)$ is a global minimum. Let $g(x) = [\alpha(b) - \alpha(a)]f(x)$ for $x \in [a, b]$. Then $g(x_-) \leq \int_a^b f \, d\alpha \leq g(x_+)$. By the Intermediate Value Theorem, there exists c between x_+ and x_- such that $g(c) = \int_a^b f \, d\alpha$. Since x_+ and x_- are in $[a, b]$, so is c. □

Theorem 6.17. *Suppose α is differentiable with $\alpha' \in \mathcal{R}[a, b]$. Then $f \in \mathcal{R}_\alpha[a, b]$ if and only if $f\alpha' \in \mathcal{R}[a, b]$. If these conditions are satisfied, then*

$$\int_a^b f \, d\alpha = \int_a^b f\alpha' \, dx.$$

Proof. We claim that

$$\overline{\int_a^b} f \, d\alpha = \overline{\int_a^b} f\alpha' \, dx$$

and

$$\underline{\int_a^b} f \, d\alpha = \underline{\int_a^b} f\alpha' \, dx$$

for any bounded $f : [a, b] \to \mathbb{R}$. The theorem will immediately follow from these equations. Therefore, fix $\epsilon > 0$. Since $\alpha' \in \mathcal{R}[a, b]$, there is a partition $P = \{x_0, \ldots, x_n\}$ such that

$$U(P, \alpha') - L(P, \alpha') < \epsilon.$$

For each $1 \leq i \leq n$, the Mean Value Theorem implies that there exists $t_i \in (x_{i-1}, x_i)$ such that $\alpha'(t_i)\Delta x_i = \Delta\alpha_i$. For any choice of $s_i \in [x_{i-1}, x_i]$, we have that

$$\sum_{i=1}^n |\alpha'(s_i) - \alpha'(t_i)|\Delta x_i \leq U(P, \alpha') - L(P, \alpha') < \epsilon.$$

Letting $M = \sup_{x \in [a,b]} |f(x)|$, we see that

$$\left| \sum_{i=1}^n f(s_i)\alpha'(s_i)\Delta x_i - \sum_{i=1}^n f(s_i)\alpha'(t_i)\Delta x_i \right| \leq M \sum_{i=1}^n |\alpha'(s_i) - \alpha'(t_i)|\Delta x_i \leq M\epsilon.$$

Hence,

$$\sum_{i=1}^n f(s_i)\alpha'(s_i)\Delta x_i \leq \sum_{i=1}^n f(s_i)\alpha'(t_i)\Delta x_i + M\epsilon = \sum_{i=1}^n f(s_i)\Delta\alpha_i + M\epsilon \leq U(P, f, \alpha) + M\epsilon.$$

Since the $s_i \in [x_{i-1}, x_i]$ are arbitrary, the inequality holds even when we take the supremum of $\sum_{i=1}^n f(s_i)\alpha'(s_i)\Delta x_i$ over all choices of the s_i. Hence,

$$U(P, f\alpha') = \sum_{i=1}^n \sup_{s_i \in [x_{i-1}, x_i]} f(s_i)\alpha'(s_i)\Delta x_i \leq U(P, f, \alpha) + M\epsilon.$$

Thus, we have proved that if P_* is any partition such that $U(P_*, \alpha') - L(P_*, \alpha') < \epsilon$, then $U(P_*, f\alpha') \leq U(P_*, f, \alpha) + M\epsilon$. Hence, for any refinement P_* of P, we have that

$$U(P_*, f\alpha') \leq U(P_*, f, \alpha) + M\epsilon.$$

Taking infima on both sides yields

$$\inf_{P_* \supset P} U(P_*, f\alpha') \leq \inf_{P_* \supset P} U(P_*, f, \alpha) + M\epsilon.$$

But in fact,

$$\inf_{P_* \supset P} U(P_*, f\alpha') = \inf_Q U(Q, f\alpha') = \overline{\int_a^b} f\alpha'\,dx$$

because every partition Q of $[a, b]$ has a common refinement with P. Similarly,

$$\inf_{P_* \supset P} U(P_*, f, \alpha) = \inf_Q U(Q, f, \alpha) = \overline{\int_a^b} f\,d\alpha.$$

We conclude that

$$\overline{\int_a^b} f\alpha'\,dx \leq \overline{\int_a^b} f\,d\alpha + M\epsilon.$$

A similar argument shows that

$$\overline{\int_a^b} f\alpha'\,dx \geq \overline{\int_a^b} f\,d\alpha - M\epsilon,$$

so taking $\epsilon \to 0^+$ yields that

$$\overline{\int_a^b} f\alpha'\,dx = \overline{\int_a^b} f\,d\alpha.$$

The argument for showing that

$$\underline{\int_a^b} f\alpha'\,dx = \underline{\int_a^b} f\,d\alpha$$

is similar. □

Theorem 6.18. *Suppose $\phi : [A, B] \to [a, b]$ is a strictly increasing bijection. Let $g = f \circ \phi : [A, B] \to \mathbb{R}$ and $\beta = \alpha \circ \phi : [A, B] \to \mathbb{R}$. Then $g \in \mathcal{R}_\beta[A, B]$ if and only if $f \in \mathcal{R}_\alpha[a, b]$, and in this case,*

$$\int_A^B g\,d\beta = \int_a^b f\,d\alpha.$$

Proof. First note that β is monotonically increasing since α and ϕ are monotonically increasing, so we can integrate with respect to β. Let $P = \{x_0, \ldots, x_n\}$ be a partition of $[A, B]$. Let $y_i = \phi(x_i)$ for each $0 \leq i \leq n$. Then $Q = \{y_0, \ldots, y_n\}$ is a partition of $[a, b]$ since $\phi(x_0) = a$, $\phi(x_n) = b$, and ϕ is strictly increasing. For each $1 \leq i \leq n$, $\phi : [x_{i-1}, x_i] \to [y_{i-1}, y_i]$ is a bijection, so

$$\sup_{x \in [x_{i-1}, x_i]} f(\phi(x)) = \sup_{y \in [y_{i-1}, y_i]} f(y).$$

Therefore,

$$\begin{aligned}
U(P, g, \beta) &= \sum_{i=1}^{n} \left(\sup_{x \in [x_{i-1}, x_i]} g(x) \right) [\beta(x_i) - \beta(x_{i-1})] \\
&= \sum_{i=1}^{n} \left(\sup_{x \in [x_{i-1}, x_i]} f(\phi(x)) \right) [\alpha(\phi(x_i)) - \alpha(\phi(x_{i-1}))] \\
&= \sum_{i=1}^{n} \left(\sup_{y \in [y_{i-1}, y_i]} f(y) \right) [\alpha(y_i) - \alpha(y_{i-1})] \\
&= U(P, f, \alpha).
\end{aligned}$$

This equation holds for any partition P, so taking infima over all partitions yields that

$$\overline{\int_a^b} g \, d\beta = \overline{\int_a^b} f \, d\alpha.$$

A similar argument shows that $L(P, g, \beta) = L(P, f, \alpha)$ for all partitions P and hence that

$$\underline{\int_a^b} g \, d\beta = \underline{\int_a^b} f \, d\alpha.$$

$\square$

Corollary 6.18.1 (Substitution Rule)**.** *Suppose $\phi : [A, B] \to \mathbb{R}$ is differentiable and strictly increasing with $\phi' \in \mathcal{R}[A, B]$, and suppose $(f \circ \phi)\phi' \in \mathcal{R}[A, B]$. Then $f \in \mathcal{R}[\phi(A), \phi(B)]$ gives*

$$\int_A^B (f \circ \phi)\phi' \, dx = \int_{\phi(A)}^{\phi(B)} f \, dx.$$

Proof. Theorem 6.17 says that $f \circ \phi \in \mathcal{R}_\phi[A, B]$ and

$$\int_A^B (f \circ \phi)\phi' \, dx = \int_A^B f \circ \phi \, d\phi.$$

Now Theorem 6.18 with $\alpha(x) = x$ says that $f \in \mathcal{R}[\phi(A), \phi(B)]$ and

$$\int_A^B f \circ \phi \, d\phi = \int_{\phi(A)}^{\phi(B)} f \, dx.$$

$\square$

Exercise 6.19 (Hölder's Inequality)**.** Let $f, g \in \mathcal{R}_\alpha[a,b]$ and $p, q > 1$ such that $\frac{1}{p} + \frac{1}{q} = 1$. Our goal is to show *Hölder's Inequality*, which says that

$$\int_a^b |fg|\, d\alpha \leq \left(\int_a^b |f|^p\, d\alpha\right)^{1/p} \left(\int_a^b |g|^q\, d\alpha\right)^{1/q}.$$

(a) First, show that if $x, y > 0$, then

$$xy \leq \frac{x^p}{p} + \frac{y^q}{q}.$$

Hint: this is a differential calculus problem. You can assume that the power rule holds for all real exponents.

(b) Suppose $\int_a^b |f|^p\, d\alpha = 1 = \int_a^b |g|^q\, d\alpha$. Use part (a) to prove that $\int_a^b |fg|\, d\alpha \leq 1$.

(c) Prove Hölder's Inequality assuming that $\int_a^b |f|^p\, d\alpha > 0$ and $\int_a^b |g|^q\, d\alpha > 0$. Hint: use part (b) with a carefully-crafted choice of f and g.

(d) Argue that if $\int_a^b |f|^p\, d\alpha = 0$ or $\int_a^b |g|^q\, d\alpha = 0$, then $\int_a^b |fg|\, d\alpha = 0$. Hint: this basically comes down to showing that if $\int_a^b |f|^p\, d\alpha = 0$, then $\int_a^b |f|\, d\alpha = 0$. You can start by imitating the "key ingredient" of the proof of Theorem 6.9. That is, argue that $|f|^p$ must be "small" throughout $[a,b]$ except possibly on some intervals whose lengths (with respect to α) sum to be less than ϵ. You can then bound $|f|$ by a small number almost everywhere on $[a,b]$.

Exercise 6.20 (Minkowski's Inequality)**.** This exercise gives a very important application of Hölder's Inequality (Exercise 6.19). Let $f, g \in \mathcal{R}_\alpha[a,b]$ and $p > 1$.

(a) Prove that

$$\int_a^b |f| \cdot |f+g|^{p-1}\, d\alpha \leq \left(\int_a^b |f|^p\, d\alpha\right)^{1/p} \left(\int_a^b |f+g|^p\, d\alpha\right)^{1-1/p}$$

and similarly

$$\int_a^b |g| \cdot |f+g|^{p-1}\, d\alpha \leq \left(\int_a^b |g|^p\, d\alpha\right)^{1/p} \left(\int_a^b |f+g|^p\, d\alpha\right)^{1-1/p}.$$

(b) Hence, prove that

$$\left(\int_a^b |f+g|^p\, d\alpha\right)^{1/p} \leq \left(\int_a^b |f|^p\, d\alpha\right)^{1/p} + \left(\int_a^b |g|^p\, d\alpha\right)^{1/p}.$$

This is *Minkowski's Inequality*; it is a triangle inequality for integrals.

6.4 Step Functions

Definition 6.21. The *unit step function* is $I : \mathbb{R} \to \{0,1\}$ defined by

$$I(x) = \begin{cases} 0 & \text{if } x \leq 0 \\ 1 & \text{if } x > 0. \end{cases}$$

Theorem 6.22. *Let $\alpha(x) = I(x - s)$ for some $s \in (a, b)$. If f is continuous at s, then $f \in \mathcal{R}_\alpha[a, b]$ and*

$$\int_a^b f\,d\alpha = f(s).$$

Proof. Suppose f is continuous at s. Fix $\epsilon > 0$, and choose $\delta > 0$ such that $N_\delta(s) \subset [a, b]$ and $|f(s) - f(t)| < \epsilon$ whenever $t \in [a, b]$ and $|s - t| < \delta$. Consider the partition $P = \{a, s - \frac{\delta}{2}, s + \frac{\delta}{2}, b\} = \{x_0, x_1, x_2, x_3\}$. Note that $\Delta\alpha_1 = \Delta\alpha_3 = 0$ and $\Delta\alpha_2 = 1$. If $t \in [x_1, x_2] = [s - \frac{\delta}{2}, s + \frac{\delta}{2}]$, then $|f(t) - f(s)| < \epsilon$, or equivalently, $f(s) - \epsilon < f(t) < f(s) + \epsilon$. Hence,

$$U(P, f, \alpha) = M_2 \leq f(s) + \epsilon$$

and

$$L(P, f, \alpha) = m_2 \geq f(s) - \epsilon,$$

so

$$f(s) - \epsilon \leq L(P, f, \alpha) \leq \underline{\int_a^b} f\,d\alpha \leq \overline{\int_a^b} f\,d\alpha \leq U(P, f, \alpha) \leq f(s) + \epsilon.$$

It follows that

$$\left|\overline{\int_a^b} f\,d\alpha - f(s)\right| \leq \epsilon$$

and

$$\left|\underline{\int_a^b} f\,d\alpha - f(s)\right| \leq \epsilon.$$

But $\epsilon > 0$ is arbitrary, so

$$\overline{\int_a^b} f\,d\alpha = f(s) = \underline{\int_a^b} f\,d\alpha.$$

□

Theorem 6.23. *Let $\{c_n\}$ be a sequence of nonnegative real numbers such that $\sum_{n=0}^\infty c_n$ converges. Let $\{s_n\}$ be a sequence of real numbers in (a, b). Let $\alpha(x) = \sum_{n=0}^\infty c_n I(x - s_n)$. If f is continuous on $[a, b]$, then*

$$\int_a^b f\,d\alpha = \sum_{n=0}^\infty c_n f(s_n).$$

Proof. For any $x \in [a, b]$, $\sum_{n=0}^\infty c_n I(x - s_n)$ converges absolutely by the Comparison Test since $\sum_{n=0}^\infty |c_n| = \sum_{n=0}^\infty c_n$ converges and $|I(x - s_n)| \leq 1$ for all $n \geq 0$. Thus, α is well-defined. Similarly, $\sum_{n=0}^\infty c_n f(s_n)$ converges by the Comparison Test since f is continuous on $[a, b]$ and hence bounded. Let

$$A_x = \{n \geq 0 \mid s_n < x\}$$

for each $x \in [a, b]$. It is clear that if $x_1 \leq x_2$, then $A_{x_1} \subset A_{x_2}$. Hence, α is monotonically increasing, so it makes sense to integrate with respect to α. Since f is continuous on $[a, b]$, Theorem 6.6 says that $f \in \mathcal{R}_\alpha[a, b]$.

Fix $\epsilon > 0$. Since $\sum_{n=0}^{\infty} c_n$ converges, there exists $N \in \mathbb{N}$ such that $\sum_{n=N+1}^{\infty} c_n < \epsilon$. Define

$$\alpha_1(x) = \sum_{n=0}^{N} c_n I(x - s_n)$$

and

$$\alpha_2(x) = \sum_{n=N+1}^{\infty} c_n I(x - s_n).$$

Then α_1 and α_2 are both monotonically increasing, and $\alpha = \alpha_1 + \alpha_2$. Since f is continuous, we have that $f \in \mathcal{R}_{\alpha_1}[a, b]$ and $f \in \mathcal{R}_{\alpha_2}[a, b]$. Notice that $\alpha_2(a) = 0$ and $\alpha_2(b) = \sum_{n=N+1}^{\infty} c_n$ since $a < s_n < b$ for all n. Hence,

$$\begin{aligned}
\left| \int_a^b f\, d\alpha - \sum_{n=0}^{\infty} c_n f(s_n) \right| &= \left| \int_a^b f\, d\alpha_1 + \int_a^b f\, d\alpha_2 - \sum_{n=0}^{\infty} c_n f(s_n) \right| \quad \text{by Theorem 6.11} \\
&= \left| \int_a^b f\, d\alpha_2 - \sum_{n=N+1}^{\infty} c_n f(s_n) \right| \quad \text{by Theorem 6.22} \\
&\leq M[\alpha_2(b) - \alpha_2(a)] + \sum_{n=N+1}^{\infty} c_n |f(s_n)| \quad \text{where } M = \sup_{x \in [a,b]} |f(x)| \\
&\leq M \sum_{n=N+1}^{\infty} c_n + M \sum_{n=N+1}^{\infty} c_n \\
&= 2M\epsilon.
\end{aligned}$$

The result follows by taking ϵ arbitrarily small. □

Exercise 6.24. Let $p > 1$; $x_1, \dots, x_n \in \mathbb{R}$; and $y_1, \dots, y_n \in \mathbb{R}$.

(a) Use Minkowski's Inequality (Exercise 6.20) to prove that

$$\left(\sum_{i=1}^{n} |x_i + y_i|^p \right)^{1/p} \leq \left(\sum_{i=1}^{n} |x_i|^p \right)^{1/p} + \left(\sum_{i=1}^{n} |y_i|^p \right)^{1/p}.$$

(b) Now prove the same inequality assuming that the x_i and y_i are complex numbers.

6.5 The Fundamental Theorem of Calculus

Definition 6.25. If $x > y$ and $f \in \mathcal{R}_\alpha[y, x]$, we denote

$$\int_x^y f\, d\alpha := - \int_y^x f\, d\alpha.$$

Theorem 6.26 (Fundamental Theorem of Calculus, Part 1)**.** *Let $f \in \mathcal{R}[a, b]$. Define $F : [a, b] \to \mathbb{R}$ by*

$$F(x) = \int_a^x f(t)\, dt.$$

Then F is Lipschitz continuous. Moreover, if f is continuous at some $x_0 \in [a, b]$, then F is differentiable at x_0 and $F'(x_0) = f(x_0)$.

Proof. Let $M = \sup_{x\in[a,b]} |f(x)|$. Then

$$|F(x) - F(y)| = \left|\int_x^y f\,dx\right| \leq M|x-y|$$

for all $x, y \in [a,b]$. Hence, F is Lipschitz continuous on $[a,b]$.

Suppose f is continuous at $x_0 \in [a,b]$. Fix $\epsilon > 0$, and choose $\delta > 0$ such that if $x \in [a,b]$ and $|x - x_0| < \delta$, then $|f(x) - f(x_0)| < \epsilon$. If $0 < |h| < \delta$ and $x_0 + h \in [a,b]$, then

$$\begin{aligned}
\left|\frac{F(x_0+h) - F(x_0)}{h} - f(x_0)\right| &= \left|\frac{1}{h}\int_{x_0}^{x_0+h} f(t)\,dt - \frac{1}{h}\int_{x_0}^{x_0+h} f(x_0)\,dt\right| \\
&= \left|\frac{1}{h}\int_{x_0}^{x_0+h} [f(t) - f(x_0)]\,dt\right| \\
&\leq \frac{1}{|h|}|(x_0+h) - x_0|\epsilon \quad \text{since } |t - x_0| \leq h \text{ implies } |t - x_0| < \delta \\
&= \epsilon.
\end{aligned}$$

Therefore, $F'(x_0) = f(x_0)$ by the limit definition of the derivative. □

Theorem 6.27 (Fundamental Theorem of Calculus, Part 2)**.** *Let $f \in \mathcal{R}[a,b]$ and $F : [a,b] \to \mathbb{R}$ such that $F' = f$. Then*

$$\int_a^b f(x)\,dx = F(b) - F(a).$$

Proof. Fix $\epsilon > 0$, and let $P = \{x_0, \ldots, x_n\}$ be a partition of $[a,b]$ such that $U(P,f) - L(P,f) < \epsilon$. By the Mean Value Theorem, for each $1 \leq i \leq n$, there exists $t_i \in (x_{i-1}, x_i)$ such that $F(x_i) - F(x_{i-1}) = f(t_i)\Delta x_i$. Notice that

$$F(b) - F(a) = \sum_{i=1}^n [F(x_i) - F(x_{i-1})] = \sum_{i=1}^n f(t_i)\Delta x_i.$$

Hence,

$$L(P,f) \leq F(b) - F(a) \leq U(P,f),$$

so

$$\left|F(b) - F(a) - \int_a^b f\,dx\right| \leq U(P,f) - L(P,f) < \epsilon.$$

Since $\epsilon > 0$ is arbitrary, it must be that $|F(b) - F(a) - \int_a^b f\,dx| = 0$, so $F(b) - F(a) = \int_a^b f\,dx$. □

Corollary 6.27.1 (Integration by Parts)**.** *Let $f, g \in \mathcal{R}[a,b]$, and let $F, G : [a,b] \to \mathbb{R}$ such that $F' = f$ and $G' = g$. Then*

$$\int_a^b F(x)g(x)\,dx = F(b)G(b) - F(a)G(a) - \int_a^b f(x)G(x)\,dx.$$

Proof. Let $H(x) = F(x)G(x)$. Then $H'(x) = f(x)G(x) + F(x)g(x)$ by the product rule. By Theorems 6.11 and 6.27,

$$\begin{aligned}\int_a^b f(x)G(x)\,dx + \int_a^b F(x)g(x)\,dx &= \int_a^b H'(x)\,dx \\ &= H(b) - H(a) \\ &= F(b)G(b) - F(a)G(a).\end{aligned}$$

Subtracting $\int_a^b f(x)G(x)\,dx$ from both sides yields the result. $\square$

6.6 Integrating Vector-Valued Functions

Definition 6.28. Let $f : [a, b] \to \mathbb{R}^n$, and write $f = (f_1, \ldots, f_n)$ where $f_i : [a, b] \to \mathbb{R}$. Let $\alpha : [a, b] \to \mathbb{R}$ be monotonically increasing. We say that $f \in \mathcal{R}_\alpha[a, b]$ if each f_i is in $\mathcal{R}_\alpha[a, b]$. If $f \in \mathcal{R}_\alpha[a, b]$, then

$$\int_a^b f\,d\alpha \coloneqq \left(\int_a^b f_1\,d\alpha, \ldots, \int_a^b f_n\,d\alpha\right) \in \mathbb{R}^n.$$

In particular, for complex-valued functions $f : [a, b] \to \mathbb{C}$, we say that $f \in \mathcal{R}_\alpha[a, b]$ if $\mathrm{Re}(f) \in \mathcal{R}_\alpha[a, b]$ and $\mathrm{Im}(f) \in \mathcal{R}_\alpha[a, b]$.

Since integration of vector-valued functions is just component-wise integration of real-valued functions, many of the theorems we have seen in this chapter concerning integration of real-valued functions extend naturally to vector-valued functions. For example, we can formulate the following analogue of the Fundamental Theorem of Calculus for vector-valued functions.

Theorem 6.29.

(a) Let $f : [a, b] \to \mathbb{R}^n$ be Riemann integrable. Define $F : [a, b] \to \mathbb{R}^n$ by $F(x) = \int_a^x f(t)\,dt$. Then F is Lipschitz continuous. If f is continuous at $x_0 \in [a, b]$, then F is differentiable at x_0 with $F'(x_0) = f(x_0)$.

(b) If $g : [a, b] \to \mathbb{R}^n$ is Riemann integrable and $G : [a, b] \to \mathbb{R}^n$ satisfies $G' = g$, then $\int_a^b g(x)\,dx = G(b) - G(a)$.

The proof consists of applying the Fundamental Theorem of Calculus for real-valued functions to each component of the vector-valued integral.

The vector-valued analogue of part (b) of Theorem 6.15 is interesting because its proof is not merely a reduction to the real-valued case.

Theorem 6.30. *Let $f : [a, b] \to \mathbb{R}^n$ be in $\mathcal{R}_\alpha[a, b]$. Then $|\int_a^b f\,d\alpha| \le \int_a^b |f|\,d\alpha$.*

Proof. Write $f = (f_1, \ldots, f_n)$. We should first check that $|f| \in \mathcal{R}_\alpha[a, b]$. By assumption, $f_j \in \mathcal{R}_\alpha[a, b]$ for all $1 \le j \le n$. By Theorems 6.9 and 6.11, $|f| = \sqrt{f_1^2 + \cdots + f_n^2} \in \mathcal{R}_\alpha[a, b]$ because the maps $x \mapsto x^2$ (for $x \in \mathbb{R}$) and $x \mapsto \sqrt{x}$ (for $x \ge 0$) are continuous.

Note that $\int_a^b |f|\,d\alpha \geq 0$ since $|f(x)| \geq 0$ for all $x \in [a,b]$. Hence, if $|\int_a^b f\,d\alpha| = 0$, then the desired inequality holds. Now suppose $|\int_a^b f\,d\alpha| > 0$. We will want to exploit the Cauchy–Schwarz Inequality. For any $c_1, \ldots, c_n \in \mathbb{R}$, we have that

$$\begin{aligned}
\int_a^b (c_1 f_1 + \cdots + c_n f_n)\,d\alpha &\leq \int_a^b |c_1 f_1 + \cdots + c_n f_n|\,d\alpha \quad \text{by Theorem 6.15(a)} \\
&\leq \int_a^b \sqrt{c_1^2 + \cdots + c_n^2}\sqrt{f_1^2 + \cdots + f_n^2}\,d\alpha \quad \text{by Cauchy–Schwarz} \\
&= \sqrt{c_1^2 + \cdots + c_n^2} \int_a^b |f|\,d\alpha.
\end{aligned}$$

Our goal is to select good values of $c_1, \ldots, c_n$. We observe that

$$\left|\int_a^b f\,d\alpha\right| = \sqrt{\left(\int_a^b f_1\,d\alpha\right)^2 + \cdots + \left(\int_a^b f_n\,d\alpha\right)^2},$$

so this motivates us to let $c_j = \int_a^b f_j\,d\alpha$ for each $1 \leq j \leq n$. It follows that

$$\begin{aligned}
\left|\int_a^b f\,d\alpha\right|^2 &= \left(\int_a^b f_1\,d\alpha\right)^2 + \cdots + \left(\int_a^b f_n\,d\alpha\right)^2 \\
&= c_1 \int_a^b f_1\,d\alpha + \cdots + c_n \int_a^b f_n\,d\alpha \\
&= \int_a^b (c_1 f_1 + \cdots + c_n f_n)\,d\alpha \\
&\leq \sqrt{c_1^2 + \cdots + c_n^2} \int_a^b |f|\,d\alpha \\
&= \left|\int_a^b f\,d\alpha\right| \int_a^b |f|\,d\alpha.
\end{aligned}$$

Since $|\int_a^b f\,d\alpha| > 0$ by assumption, we can divide both sides of the above inequality by $|\int_a^b f\,d\alpha|$. This gives the desired inequality. $\square$

7 Uniform Convergence

Define $f : [0,1] \to \mathbb{R}$ by

$$f(x) = \sum_{n=1}^{\infty} \frac{1}{n^2 + x}.$$

Is f continuous? Consider the following proof: f is continuous because for every $c \in [0,1]$, we have

$$\lim_{x\to c} f(x) = \lim_{x\to c} \sum_{n=1}^{\infty} \frac{1}{n^2+x} = \sum_{n=1}^{\infty} \lim_{x\to c} \frac{1}{n^2+x} = \sum_{n=1}^{\infty} \frac{1}{n^2+c} = f(c) \tag{1}$$

using the fact that limits distribute across sums (Proposition 4.5(a)). Unfortunately, this proof is flawed because Proposition 4.5(a) only applies to *finite* sums, and infinite sums (which are not really sums, but rather limits) are fundamentally different from finite sums.

Even though the "proof" in the previous paragraph is invalid, we do often want to pass limits through operators such as infinite sums, derivatives, and integrals. This is especially true when studying functions defined using power series (e.g. the exponential function $\exp(z) = \sum_{n=0}^{\infty} \frac{z^n}{n!}$). In this chapter, we study *uniform convergence* of sequences of functions, which is the main condition that allows us to move limits around. In fact, the chain of equations in (1) turns out to be correct because the sequence of functions $f_N(x) = \sum_{n=1}^{N} \frac{1}{n^2+x}$ (for $N = 1,2,3,\ldots$ and $x \in [0,1]$) converges uniformly to $f(x) = \sum_{n=1}^{\infty} \frac{1}{n^2+x}$. This chapter will give us the tools to rigorously justify (1).

7.1 Examples of Limit Interchange Failures

Example 7.1. For each $n \in \mathbb{N}$, let

$$f_n(x) = \begin{cases} 1 + nx & \text{if } -\frac{1}{n} < x \le 0 \\ 1 - nx & \text{if } 0 < x < \frac{1}{n} \\ 0 & \text{otherwise.} \end{cases}$$

Then f_n is continuous for each n, but

$$f(x) \coloneqq \lim_{n\to\infty} f_n(x) = \begin{cases} 1 & \text{if } x = 0 \\ 0 & \text{otherwise} \end{cases}$$

is discontinuous at $x = 0$. Symbolically,

$$\lim_{n\to\infty} \lim_{x\to 0} f_n(x) = 1 \ne 0 = \lim_{x\to 0} \lim_{n\to\infty} f_n(x).$$

Example 7.2. For each $n \in \mathbb{N}$, define $f_n : [0,2] \to \mathbb{R}$ by

$$f_n(x) = \begin{cases} x^n & \text{if } 0 \le x \le 1 \\ 2 - (2-x)^n & \text{if } 1 < x \le 2. \end{cases}$$

Then each f_n is differentiable at $x = 1$ (by Exercise 5.19), but the limiting function

$$f(x) = \lim_{n\to\infty} f_n(x) = \begin{cases} 0 & \text{if } 0 \le x < 1 \\ 1 & \text{if } x = 1 \\ 2 & \text{if } 1 < x \le 2 \end{cases}$$

is not.

Example 7.3. For each $n \in \mathbb{N}$, define $f_n : [0,1] \to \mathbb{R}$ by

$$f_n(x) = \begin{cases} 1 & \text{if } x = \frac{p}{q} \text{ for some integers } 1 \le p \le q \le n \\ 0 & \text{otherwise.} \end{cases}$$

Each f_n has only finitely many discontinuities, so by Theorem 6.8, each f_n is in $\mathcal{R}[0,1]$. But

$$f(x) \coloneqq \lim_{n\to\infty} f_n(x) = \begin{cases} 1 & \text{if } x \in \mathbb{Q} \cap [0,1] \\ 0 & \text{if } x \in [0,1] \setminus \mathbb{Q} \end{cases}$$

is discontinuous everywhere on $[0,1]$ and is therefore not in $\mathcal{R}[0,1]$ by Lemma 6.13.

Example 7.4. For each $n \in \mathbb{N}$, define $f_n : [0,1] \to \mathbb{R}$ by

$$f_n(x) = \begin{cases} n & \text{if } 0 < x \le \frac{1}{n} \\ 0 & \text{otherwise.} \end{cases}$$

Then $\int_0^1 f_n(x)\,dx = 1$ for all n, but $\int_0^1 \lim_{n\to\infty} f_n(x)\,dx = 0$ since $\lim_{n\to\infty} f_n(x) = 0$ for all $x \in [0,1]$. Hence, $\lim_{n\to\infty} \int_0^1 f_n(x)\,dx \ne \int_0^1 \lim_{n\to\infty} f_n(x)\,dx$.

7.2 Introduction to Uniform Convergence

Definition 7.5. Let E be a set. For each $n \in \mathbb{N}$, let $f_n : E \to \mathbb{C}$. We say that the sequence of functions $\{f_n\}$ *converges uniformly to* $f : E \to \mathbb{C}$ *on* E if for all $\epsilon > 0$, there exists $N \in \mathbb{N}$ such that if $n \ge N$, then $|f_n(x) - f(x)| < \epsilon$ for all $x \in E$. In particular, note that N is independent of any $x \in E$.

We say that $\{f_n\}$ *converges pointwise to* f *on* E if for all $x \in E$ and $\epsilon > 0$, there exists $N \in \mathbb{N}$ (possibly dependent on x) such that $|f_n(x) - f(x)| < \epsilon$ for all $n \ge N$. In other words, $f_n \to f$ pointwise if $f(x) = \lim_{n\to\infty} f_n(x)$ for all $x \in E$.

The reader should quickly verify that uniform convergence implies pointwise convergence. On occasion, we will encounter sequences of *real-valued* functions $f_n : E \to \mathbb{R}$. Note that Definition 7.5 still makes sense for such sequences because we can view $\mathbb{R}$ as a subset of $\mathbb{C}$.

The definition of uniform convergence may be difficult to parse or remember, so the following proposition gives an equivalent definition.

Proposition 7.6. *A sequence of functions* $f_n : E \to \mathbb{C}$ *converges uniformly to* $f : E \to \mathbb{C}$ *on* E *(in the sense of Definition 7.5) if and only if*

$$\lim_{n\to\infty} \sup_{x\in E} |f_n(x) - f(x)| = 0.$$

Proof. Suppose $f_n \to f$ uniformly on E in the sense of Definition 7.5. Fix $\epsilon > 0$, and choose $N \in \mathbb{N}$ such that $|f_n(x) - f(x)| < \epsilon$ for all $n \geq N$ and $x \in E$. Then $\sup_{x\in E} |f_n(x) - f(x)| \leq \epsilon$ for all $n \geq N$. By definition of limit, $\lim_{n\to\infty} \sup_{x\in E} |f_n(x) - f(x)| = 0$.

Conversely, suppose $\lim_{n\to\infty} \sup_{x\in E} |f_n(x) - f(x)| = 0$. Then for all $\epsilon > 0$, there exists $N \in \mathbb{N}$ such that $\sup_{x\in E} |f_n(x) - f(x)| < \epsilon$ for all $n \geq N$. If $n \geq N$, then by definition of supremum, $|f_n(x) - f(x)| < \epsilon$ for all $x \in E$. Thus, $f_n \to f$ uniformly on E in the sense of Definition 7.5. □

Proposition 7.7. *Let $f_n : E \to \mathbb{C}$ be bounded functions that converge uniformly to f on E. Then f is bounded.*

Proof. Since $f_n \to f$ uniformly, there exists $N \in \mathbb{N}$ such that $|f_N(x) - f(x)| < 1$ for all $x \in E$. By assumption, $|f_N(x)|$ is bounded above by some $M \geq 0$ for all x. Then $|f(x)| \leq |f_N(x)| + |f(x) - f_N(x)| < M + 1$ for all x, so f is bounded. □

Theorem 7.8 (Cauchy Criterion for Uniform Convergence)**.** *A sequence of functions $f_n : E \to \mathbb{C}$ converges uniformly to some $f : E \to \mathbb{C}$ if and only if for all $\epsilon > 0$, there exists $N \in \mathbb{N}$ such that $|f_n(x) - f_m(x)| < \epsilon$ whenever $n, m \geq N$ and $x \in E$.*

Proof. Suppose $f_n \to f$ uniformly on E. Then for all $\epsilon > 0$, there exists $N \in \mathbb{N}$ such that $|f_n(x) - f(x)| < \frac{\epsilon}{2}$ for all $n \geq N$ and $x \in E$. If $n, m \geq N$, then

$$|f_n(x) - f_m(x)| \leq |f_n(x) - f(x)| + |f(x) - f_m(x)| < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon$$

for all $x \in E$.

Conversely, suppose that for all $\epsilon > 0$, there exists $N \in \mathbb{N}$ such that $|f_n(x) - f_m(x)| < \epsilon$ whenever $n, m \geq N$ and $x \in E$. It follows that for any fixed $x \in E$, $\{f_n(x)\}$ is a Cauchy sequence in $\mathbb{C}$, so $\lim_{n\to\infty} f_n(x)$ exists since $\mathbb{C}$ is complete. Define $f : E \to \mathbb{C}$ by $f(x) = \lim_{n\to\infty} f_n(x)$. Fix $\epsilon > 0$, and pick $N_1 \in \mathbb{N}$ such that $|f_n(x) - f_m(x)| < \frac{\epsilon}{2}$ for all $n, m \geq N_1$ and $x \in E$. Let $n \geq N_1$ and $x \in E$. Since $f(x) = \lim_{m\to\infty} f_m(x)$, there exists $N_2 \geq N_1$ such that $|f(x) - f_{N_2}(x)| < \frac{\epsilon}{2}$. Hence,

$$|f_n(x) - f(x)| \leq |f_n(x) - f_{N_2}(x)| + |f_{N_2}(x) - f(x)| < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon,$$

so $f_n \to f$ uniformly on E since N_1 is independent of x. □

Theorem 7.9. *Suppose $f_n : E \to \mathbb{C}$ and $g_n : E \to \mathbb{C}$ converge uniformly to f and g on E, respectively. Let $c \in \mathbb{C}$. Then:*

(a) $f_n + g_n \to f + g$ uniformly on E.

(b) $cf_n \to cf$ uniformly on E.

(c) If f_n and g_n are bounded for each n, then $f_n g_n \to fg$ uniformly on E.

Proof. (a) Fix $\epsilon > 0$. Let $N_f, N_g \in \mathbb{N}$ such that if $n \geq N_f$ and $m \geq N_g$, then $|f_n(x) - f(x)| < \frac{\epsilon}{2}$ and $|g_m(x) - g(x)| < \frac{\epsilon}{2}$ for all $x \in E$. Then for all $n \geq \max\{N_f, N_g\}$ and $x \in E$,

$$|(f_n + g_n)(x) - (f + g)(x)| \leq |f_n(x) - f(x)| + |g_n(x) - g(x)| < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon.$$

Hence, $f_n + g_n \to f + g$ uniformly on E.

(b) If $c = 0$, then $cf_n = cf = 0$ for all n, which immediately implies that $cf_n \to cf$ uniformly. Suppose $c \neq 0$. Fix $\epsilon > 0$, and choose $N \in \mathbb{N}$ such that if $n \geq N$, then $|f_n(x) - f(x)| < \frac{\epsilon}{|c|}$ for all $x \in E$. If $n \geq N$, then

$$|(cf_n)(x) - (cf)(x)| = |c| \cdot |f_n(x) - f(x)| < |c| \left(\frac{\epsilon}{|c|} \right) = \epsilon$$

for all $x \in E$. Hence, $cf_n \to cf$ uniformly on E.

(c) Choose $N_1, N_2 \in \mathbb{N}$ such that if $n \geq N_1$ and $m \geq N_2$, then $|f_n(x) - f(x)| < 1$ and $|g_m(x) - g(x)| < 1$ for all $x \in E$. By assumption, f_{N_1} and g_{N_2} are bounded, so there exists $M_1, M_2 \in \mathbb{R}$ such that $|f_{N_1}(x)| \leq M_1$ and $|g_{N_2}(x)| \leq M_2$ for all $x \in E$. Hence,

$$|f(x)| \leq |f_{N_1}(x)| + |f(x) - f_{N_1}(x)| < M_1 + 1$$

and

$$|g(x)| \leq |g_{N_2}(x)| + |g(x) - g_{N_2}(x)| < M_2 + 1$$

for all $x \in E$. If $n \geq N_1$, then

$$|f_n(x)| \leq |f(x)| + |f_n(x) - f(x)| < (M_1 + 1) + 1 = M_1 + 2$$

for all $x \in E$. Fix $\epsilon > 0$, and choose $N_f, N_g \in \mathbb{N}$ such that if $n \geq N_f$ and $m \geq N_g$, then $|f_n(x) - f(x)| < \epsilon$ and $|g_m(x) - g(x)| < \epsilon$ for all $x \in E$. Suppose $n \geq \max\{N_1, N_f, N_g\}$. Then for all $x \in E$,

$$\begin{aligned}
|(f_n g_n)(x) - (fg)(x)| &= |f_n(x)g_n(x) - f_n(x)g(x) + f_n(x)g(x) - f(x)g(x)| \\
&\leq |f_n(x)g_n(x) - f_n(x)g(x)| + |f_n(x)g(x) - f(x)g(x)| \\
&= |f_n(x)| \cdot |g_n(x) - g(x)| + |g(x)| \cdot |f_n(x) - f(x)| \\
&< (M_1 + 2)\epsilon + (M_2 + 1)\epsilon \\
&= (M_1 + M_2 + 3)\epsilon.
\end{aligned}$$

Hence, $f_n g_n \to fg$ uniformly on E since $M_1 + M_2 + 3$ is independent of ϵ. □

Theorem 7.10 (Dini)**.** *Suppose K is compact and that $f_n : K \to \mathbb{R}$ is continuous for each $n \in \mathbb{N}$. Suppose that $f_n(x) \leq f_{n+1}(x)$ for all $n \in \mathbb{N}$ and $x \in K$, and suppose $f_n \to f$ pointwise on K where $f : K \to \mathbb{R}$ is continuous. Then $f_n \to f$ uniformly on K.*

Proof. Since $f_n \leq f_{n+1}$ for all $n \in \mathbb{N}$, we have that $f_n \leq f$ for all $n \in \mathbb{N}$. Fix $\epsilon > 0$. For each $n \in \mathbb{N}$, let $g_n : K \to \mathbb{R}$ be defined by $g_n(x) = f(x) - f_n(x)$, and let $K_n = g_n^{-1}([\epsilon, \infty))$. We want to show that K_n is empty for n large enough.

We claim that $\{K_n\}_n$ is a decreasing family of compact sets. Fix $n \in \mathbb{N}$. First, note that g_n is continuous since f and f_n is continuous. Next, K_n is closed because $[\epsilon, \infty)$ is closed in $\mathbb{R}$ and the pre-image of a closed set under a continuous function is closed. A closed subset of a compact set is compact, so $K_n \subset K$ is compact. If $x \in K_{n+1}$, then $\epsilon \leq f(x) - f_{n+1}(x) \leq f(x) - f_n(x)$ since $f_n(x) \leq f_{n+1}(x)$, so $x \in K_n$. Therefore, $K_{n+1} \subset K_n$ for all $n \in \mathbb{N}$, proving our claim.

Suppose, by way of contradiction, that K_n is non-empty for all n. Then $\bigcap_{n\in\mathbb{N}} K_n$ is non-empty by Corollary 2.45.1. Choose $x_0 \in \bigcap_{n\in\mathbb{N}} K_n$. Then $g_n(x_0) = f(x_0) - f_n(x_0) \geq \epsilon$ for all $n \in \mathbb{N}$. But $\lim_{n\to\infty} f_n(x_0) = f(x_0)$ by assumption, so there exists $N \in \mathbb{N}$ such that $f(x_0) - f_N(x_0) < \epsilon$. We have obtained a contradiction, so there must exist $N_0 \in \mathbb{N}$ such that K_{N_0} is empty.

Let $n \geq N_0$ and $x \in K$. Then

$$|f(x) - f_n(x)| = f(x) - f_n(x) = g_n(x) < \epsilon$$

because $g_n^{-1}([\epsilon, \infty)) = K_n \subset K_{N_0}$ is empty. Therefore, $f_n \to f$ uniformly. □

Definition 7.11. Let $f_n : E \to \mathbb{C}$ be a sequence of functions such that $f(x) = \sum_{n=1}^{\infty} f_n(x)$ exists for each $x \in E$. We say that the series $\sum_{n=1}^{\infty} f_n(x)$ converges uniformly to f on E if the sequence of partial sums $s_N = \sum_{n=1}^{N} f_n$ converges uniformly to f on E in the sense of Definition 7.5.

Theorem 7.12 (Weierstrass M-test)**.** *Let $f_n : E \to \mathbb{C}$ for each $n \in \mathbb{N}$. Suppose there exists a sequence of nonnegative real numbers $\{M_n\}$ such that $|f_n(x)| \leq M_n$ for all $n \in \mathbb{N}$ and $x \in E$. If $\sum_{n=1}^{\infty} M_n$ converges, then $\sum_{n=1}^{\infty} f_n(x)$ converges uniformly on E.*

Proof. Let $s_N(x) = \sum_{n=1}^{N} f_n(x)$ for all $N \in \mathbb{N}$. Fix $\epsilon > 0$. Since $\sum_{n=1}^{\infty} M_n$ converges, there exists $T \in \mathbb{N}$ such that $\sum_{n=i+1}^{j} M_n < \epsilon$ for all $j \geq i \geq T$ (see Proposition 3.32). Then

$$|s_j(x) - s_i(x)| = \left| \sum_{n=i+1}^{j} f_n(x) \right| \leq \sum_{n=i+1}^{j} |f_n(x)| \leq \sum_{n=i+1}^{j} M_n < \epsilon$$

for all $j \geq i \geq T$ and $x \in E$. Therefore, s_N converges uniformly by the Cauchy Criterion. □

Exercise 7.13. Give an example of a sequence of bounded functions $f_n : \mathbb{R} \to \mathbb{R}$ that converges pointwise on $\mathbb{R}$ to an unbounded function f. (Compare with Proposition 7.7.)

7.3 Limit Interchange Under Uniform Convergence

The following limit interchange theorem justifies our efforts in studying uniform convergence.

Theorem 7.14. *Let X be a metric space and $E \subset X$. Suppose a sequence of functions $f_n : E \to \mathbb{C}$ converges uniformly to f on E. Let $x \in E'$, and suppose $A_n = \lim_{t\to x} f_n(t)$ exists for each n. Then $A = \lim_{n\to\infty} A_n$ exists and $\lim_{t\to x} f(t) = A$.*

More concisely, if $\lim_{t\to x} f_n(t)$ exists for each n (or at least for all n large enough), then the equation

$$\lim_{n\to\infty} \lim_{t\to x} f_n(t) = \lim_{t\to x} \lim_{n\to\infty} f_n(t)$$

is valid as long as $\{f_n\}$ uniformly converges.

Proof. Fix $\epsilon > 0$. By the Cauchy Criterion, there exists $N_1 \in \mathbb{N}$ such that $|f_n(t) - f_m(t)| < \epsilon$ for all $n, m \geq N$ and $t \in E$. Also, for each n, there exists $\delta_n > 0$ such that $|f_n(t) - A_n| < \epsilon$ whenever $d(t, x) < \delta_n$. Since $x \in E'$, we can choose $t_n \in E$ for each n to satisfy $d(t_n, x) < \delta_n$. For all $n, m \geq N_1$,

$$\begin{aligned} |A_n - A_m| &\leq |A_n - f_n(t_n)| + |f_n(t_n) - f_m(t_m)| + |f_m(t_m) - A_m| \\ &< \epsilon + \epsilon + \epsilon \\ &= 3\epsilon. \end{aligned}$$

Hence, $\{A_n\}$ is a Cauchy sequence, so $A = \lim_{n\to\infty} A_n$ exists.

Let $N_2 \in \mathbb{N}$ be such that $|A - A_n| < \epsilon$ for all $n \geq N_2$. Since the f_n converge uniformly to f, there exists $N_3 \in \mathbb{N}$ such that $|f(t) - f_n(t)| < \epsilon$ for all $n \geq N_3$ and $t \in E$. Let $n = \max\{N_2, N_3\}$ and $\delta = \delta_n$. Then for all $t \in E$ such that $d(t, x) < \delta$,

$$\begin{aligned} |f(t) - A| &\leq |f(t) - f_n(t)| + |f_n(t) - A_n| + |A_n - A| \\ &< \epsilon + \epsilon + \epsilon \\ &= 3\epsilon, \end{aligned}$$

so $\lim_{t\to x} f(t) = A$. □

As a result, uniform convergence preserves continuity.

Theorem 7.15. *If $f_n : X \to \mathbb{C}$ are continuous functions that converge uniformly to $f : X \to \mathbb{C}$ on X, then f is continuous.*

Proof. Let $x \in X$. First, suppose $x \in X'$. Since each f_n is continuous at x, we have that $\lim_{t\to x} f_n(t) = f_n(x)$ for each n by Theorem 4.7. By Theorem 7.14, $\lim_{t\to x} f(t)$ exists and

$$\lim_{t\to x} f(t) = \lim_{n\to\infty} \lim_{t\to x} f_n(t) = \lim_{n\to\infty} f_n(x) = f(x).$$

Therefore, f is continuous at x.

If $x \notin X'$, then x is an isolated point of f, so f is automatically continuous at x. □

We can prove Theorem 3.53 using Theorem 7.14. Let $\{a_{n,m}\}$ be a double sequence of complex numbers such that $\sum_{n=1}^\infty \sum_{m=1}^\infty |a_{n,m}|$ converges. We recognize that

$$\sum_{n=1}^\infty \sum_{m=1}^\infty a_{n,m} = \lim_{N\to\infty} \lim_{M\to\infty} \sum_{n=1}^N \sum_{m=1}^M a_{n,m} = \lim_{M\to\infty} \lim_{N\to\infty} \sum_{n=1}^N \sum_{m=1}^M a_{n,m},$$

so if we could interchange the two limits, we would obtain the desired result. However, Theorem 7.14 only allows for the interchanging of a "continuous" limit (of the form "$\lim_{t\to x}$") with a "sequential" limit (of the form "$\lim_{n\to\infty}$"). The limits we want to interchange are both sequential limits, so we cannot apply the theorem directly. We work around this by defining $E = \{\frac{1}{n} \mid n \in \mathbb{N}\} \subset \mathbb{R}$ and a sequence of functions $f_N : E \to \mathbb{C}$ such that

$$f_N(x) = \sum_{m=1}^{1/x} \sum_{n=1}^N a_{n,m} = \sum_{n=1}^N \sum_{m=1}^{1/x} a_{n,m}.$$

Note that $0 \in E'$ and that $\lim_{x \to 0} f_N(x) = \sum_{n=1}^{N} \sum_{m=1}^{\infty} a_{n,m}$ exists for each N. Since

$$\sum_{n=1}^{\infty} \sum_{m=1}^{\infty} a_{n,m} = \lim_{N \to \infty} \lim_{x \to 0} f_N(x)$$

and

$$\sum_{m=1}^{\infty} \sum_{n=1}^{\infty} a_{n,m} = \lim_{x \to 0} \lim_{N \to \infty} f_N(x),$$

we can use Theorem 7.14 if the f_N converge uniformly on E. Let

$$U_n = \sum_{m=1}^{\infty} |a_{n,m}|$$

for all $n \geq 1$. Notice that $U_n \geq |\sum_{m=1}^{1/x} a_{n,m}|$ for all $x \in E$. Since $\sum_{n=1}^{\infty} U_n$ converges by assumption, the f_N converge uniformly on E by the Weierstrass M-test. Hence,

$$\sum_{n=1}^{\infty} \sum_{m=1}^{\infty} a_{n,m} = \lim_{N \to \infty} \lim_{x \to 0} f_N(x) = \lim_{x \to 0} \lim_{N \to \infty} f_N(x) = \sum_{m=1}^{\infty} \sum_{n=1}^{\infty} a_{n,m}$$

by Theorem 7.14.

Definition 7.16. Let X be a metric space, and let F be $\mathbb{R}$ or $\mathbb{C}$. We define

$$\mathcal{C}(X; F) \coloneqq \{f : X \to F \mid f \text{ is continuous and bounded}\}.$$

Proposition 7.17. *Let E be a set, and let B be the set of bounded functions $f : E \to \mathbb{C}$. For $f \in B$, define*

$$||f||_\infty \coloneqq \sup_{x \in E} |f(x)|,$$

which is finite since f is bounded. Then $d(f, g) = ||f - g||_\infty$ is a metric on B.

Proof. For all $f \in B$, we have $d(f, f) = ||f - f||_\infty = 0$ since $f - f$ is the zero function. If $f, g \in B$ are distinct, then there exists $x \in E$ such that $f(x) \neq g(x)$, so $d(f, g) \geq |f(x) - g(x)| > 0$.

Let $f, g \in B$. Then $|f(x) - g(x)| = |g(x) - f(x)|$ for all $x \in E$, so $d(f, g) = d(g, f)$. Finally, if $h \in B$, then

$$|f(x) - g(x)| \leq |f(x) - h(x)| + |h(x) - g(x)| \leq d(f, h) + d(h, g)$$

for all $x \in E$, so $d(f, g) \leq d(f, h) + d(h, g)$ since x is arbitrary. $\square$

Definition 7.18. The metric from Proposition 7.17 is called the *supremum metric.*

By Proposition 7.6, a sequence of bounded functions $f_n : E \to \mathbb{C}$ converges uniformly to f if and only if the f_n converge to f with respect to the supremum metric. Hence, the supremum metric is also known as the *uniform metric.*

When we apply the Cauchy Criterion for uniform convergence to a set of continuous functions, we get the following theorem.

Theorem 7.19. $\mathcal{C}(X;F)$ *with the supremum metric is a complete metric space.*

Proof. Let $\{f_n\}$ be a Cauchy sequence in $\mathcal{C}(X;F)$. Then for all $\epsilon > 0$, there exists $N \in \mathbb{N}$ such that $||f_n - f_m||_\infty < \epsilon$ whenever $n, m \geq N$. Hence, $|f_n(x) - f_m(x)| < \epsilon$ for all $n, m \geq N$ and $x \in E$, so the f_n converge uniformly to some function f by the Cauchy Criterion. Moreover, f is continuous by Theorem 7.15. Since the f_n are bounded, f is bounded by Proposition 7.7. Hence, $f \in \mathcal{C}(X;F)$. $\square$

Exercise 7.20. Let $\{c_{n,m}\}$ be a double complex sequence. Create your own theorem that gives sufficient conditions for making the limit interchange

$$\lim_{n\to\infty} \lim_{m\to\infty} c_{n,m} = \lim_{m\to\infty} \lim_{n\to\infty} c_{n,m}.$$

(You will probably obtain the Moore–Osgood Theorem.)

Exercise 7.21. Does uniform convergence preserve uniform continuity?

7.4 Differentiation and Integration

Uniform convergence also interacts nicely with differentiation and integration.

Theorem 7.22. *Suppose* $f_n : (a,b) \to \mathbb{R}$ *are differentiable,* $f_n' \to g$ *uniformly on* (a,b)*, and* $\lim_{n\to\infty} f_n(x_0)$ *exists for some* $x_0 \in (a,b)$*. Then* f_n *converges uniformly to some function* f *on* (a,b)*, and* f *is differentiable with* $f' = g$*.*

The main idea of the proof is to exploit Lipschitz continuity of $\phi_{n,m}(x) = f_n(x) - f_m(x)$ when n, m are large enough.

Proof. For each $n, m \in \mathbb{N}$, let $\phi_{n,m}(x) = f_n(x) - f_m(x)$. Fix $\epsilon > 0$. Since f_n' converges uniformly, there exists $N_1 \in \mathbb{N}$ such that $|f_n'(x) - f_m'(x)| < \epsilon$ for all $n, m \geq N_1$ and $x \in (a,b)$. Since $\lim_{n\to\infty} f_n(x_0)$ exists, there exists $N_2 \in \mathbb{N}$ such that $|f_n(x_0) - f_m(x_0)| < \epsilon$ for all $n, m \geq N_2$. Suppose $n, m \geq \max\{N_1, N_2\}$. Then

$$|\phi_{n,m}'(t)| = |f_n'(t) - f_m'(t)| < \epsilon$$

for all $t \in (a,b)$. Hence, for all $x \in (a,b)$,

$$\begin{aligned}
|f_n(x) - f_m(x)| &= |\phi_{n,m}(x)| \\
&\leq |\phi_{n,m}(x) - \phi_{n,m}(x_0)| + |\phi_{n,m}(x_0)| \\
&\leq \epsilon|x - x_0| + |\phi_{n,m}(x_0)| \quad \text{by the Mean Value Theorem} \\
&= \epsilon|x - x_0| + |f_n(x_0) - f_m(x_0)| \\
&< \epsilon|x - x_0| + \epsilon \quad \text{since } n, m \geq N_2 \\
&\leq \epsilon(b - a) + \epsilon \\
&= (1 + b - a)\epsilon.
\end{aligned}$$

By the Cauchy Criterion, f_n converges uniformly on (a,b) to some function $f : (a,b) \to \mathbb{R}$. Fix $x \in (a,b)$, and let $h \neq 0$ such that $x + h \in (a,b)$. If $n, m \geq N_1$, then

$$|f(x+h) - f(x) - [f_m(x+h) - f_m(x)]|$$

$$
\begin{aligned}
&\leq |f_n(x+h) - f_n(x) - [f_m(x+h) - f_m(x)]| + |f(x+h) - f(x) - [f_n(x+h) - f_n(x)]| \\
&= |\phi_{n,m}(x+h) - \phi_{n,m}(x)| + |f(x+h) - f(x) - [f_n(x+h) - f_n(x)]| \\
&< \epsilon|h| + |f(x+h) - f(x) - [f_n(x+h) - f_n(x)]| \quad \text{by the Mean Value Theorem} \\
&\leq \epsilon|h| + |f(x+h) - f_n(x+h)| + |f(x) - f_n(x)|.
\end{aligned}
$$

Since $f_n \to f$ uniformly, we can choose $n \geq N_1$ large enough such that $|f(t) - f_n(t)| < \epsilon|h|$ for all $t \in (a, b)$. As a result,

$$
|f(x+h) - f(x) - [f_m(x+h) - f_m(x)]| < \epsilon|h| + \epsilon|h| + \epsilon|h| = 3\epsilon|h|.
$$

Therefore,

$$
\begin{aligned}
\left|\frac{f(x+h) - f(x)}{h} - g(x)\right| &\leq \left|\frac{f(x+h) - f(x) - [f_m(x+h) - f_m(x)]}{h}\right| \\
&\quad + \left|\frac{f_m(x+h) - f_m(x)}{h} - f_m'(x)\right| + |f_m'(x) - g(x)| \\
&< 3\epsilon + \left|\frac{f_m(x+h) - f_m(x)}{h} - f_m'(x)\right| + |f_m'(x) - g(x)|.
\end{aligned}
$$

Recall that $m \geq N_1$ is arbitrary. Since $f_m' \to g$ uniformly, there exists $M \geq N_1$ such that $|f_M'(x) - g(x)| < \epsilon$. Choose $\delta > 0$ such that

$$
\left|\frac{f_M(x+h) - f_M(x)}{h} - f_M'(x)\right| < \epsilon
$$

if $0 < |h| < \delta$ and $x + h \in (a, b)$. Then

$$
\left|\frac{f(x+h) - f(x)}{h} - g(x)\right| < 3\epsilon + \epsilon + \epsilon = 5\epsilon
$$

if $0 < |h| < \delta$ and $x + h \in (a, b)$. Therefore, $\lim_{h\to 0} \frac{f(x+h)-f(x)}{h} = g(x)$, so $f'(x)$ exists and equals $g(x)$. $\square$

Theorem 7.23. *Let $\alpha : [a, b] \to \mathbb{R}$ be monotonically increasing. Suppose $f_n \in \mathcal{R}_\alpha[a, b]$ for each n, and suppose $f_n \to f$ uniformly on $[a, b]$. Then $f \in \mathcal{R}_\alpha[a, b]$ and*

$$
\int_a^b f\, d\alpha = \lim_{n\to\infty} \int_a^b f_n\, d\alpha.
$$

Proof. Fix $\epsilon > 0$ and choose $N \in \mathbb{N}$ such that $|f_n(x) - f(x)| < \epsilon$ for all $x \in [a, b]$ and $n \geq N$. Let $P = \{x_0, \ldots, x_m\}$ be a partition of $[a, b]$ such that

$$
U(P, f_N, \alpha) - L(P, f_N, \alpha) = \sum_{i=1}^{m} \left(\sup_{x\in[x_{i-1},x_i]} f_N(x) - \inf_{x\in[x_{i-1},x_i]} f_N(x) \right) \Delta\alpha_i < \epsilon.
$$

Since $f_N(x) - \epsilon < f(x) < f_N(x) + \epsilon$ for all $x \in [a, b]$, it follows that

$$
\inf_{x\in[x_{i-1},x_i]} f_N(x) - \epsilon \leq \inf_{x\in[x_{i-1},x_i]} f(x) \leq \sup_{x\in[x_{i-1},x_i]} f(x) \leq \sup_{x\in[x_{i-1},x_i]} f_N(x) + \epsilon
$$

for all $1 \le i \le m$. Therefore,

$$\sup_{x\in[x_{i-1},x_i]} f(x) - \inf_{x\in[x_{i-1},x_i]} f(x) \le \sup_{x\in[x_{i-1},x_i]} f_N(x) - \inf_{x\in[x_{i-1},x_i]} f_N(x) + 2\epsilon$$

for all $1 \le i \le m$, so

$$\begin{aligned} U(P, f, \alpha) - L(P, f, \alpha) &\le U(P, f_N, \alpha) - L(P, f_N, \alpha) + 2\epsilon[\alpha(b) - \alpha(a)] \\ &< [1 + 2\alpha(b) - 2\alpha(a)]\epsilon. \end{aligned}$$

Hence, $f \in \mathcal{R}_\alpha[a, b]$ since $1 + 2\alpha(b) - 2\alpha(a)$ is independent of ϵ. Finally,

$$\left|\int_a^b f\,d\alpha - \int_a^b f_n\,d\alpha\right| \le \int_a^b |f - f_n|\,d\alpha \le \int_a^b \epsilon\,d\alpha = [\alpha(b) - \alpha(a)]\epsilon$$

for all $n \ge N$, so $\lim_{n\to\infty} \int_a^b f_n\,d\alpha = \int_a^b f\,d\alpha$ since $\alpha(b) - \alpha(a)$ is independent of ϵ. □

Exercise 7.24 (The Power Rule)**.** Fix $c \in \mathbb{R}$. Let $f(x) = x^c$ for all $x > 0$ (recall that we defined what x^c means in Exercise 4.32). We proved in Section 5.5 that if $c \in \mathbb{Q}$, then $f'(x) = cx^{c-1}$. Now we will drop the assumption that $c \in \mathbb{Q}$ and (finally!) prove the full power rule: $f'(x) = cx^{c-1}$ for all $x > 0$.

(a) Fix $x_0 > 0$. Let $\{q_n\}$ be an increasing sequence of rational numbers that converges to c, and for each n, let $f_n(x) = x^{q_n}$. Let $g(x) = cx^{c-1}$ and $0 < a < x_0 < b$. Prove that the derivatives $f_n'(x) = q_n x^{q_n-1}$ converge uniformly to g on (a, b).

(b) Conclude that the f_n converge uniformly to f on (a, b) and that $f' = g$ on this interval. Therefore, $f'(x_0) = cx_0^{c-1}$.

Exercise 7.25. Prove that $(b^x)^y = b^{xy}$ for all $b > 0$ and $x, y \in \mathbb{R}$.

Exercise 7.26. Let

$$f(x) = \sum_{n=0}^{\infty} \frac{x^n}{n!}$$

for all $x \in \mathbb{R}$.

(a) Use Exercise 4.33 to prove that $f(x) = e^x$.

(b) Prove that f is differentiable and that $f'(x) = e^x$. That is, $f' = f$.

(c) Suppose $g : \mathbb{R} \to \mathbb{R}$ satisfies $g' = g$. Prove that there exists a constant $C \in \mathbb{R}$ such that $g = Cf$.

Exercise 7.27. Let $f_n : (0, \infty) \to \mathbb{R}$ such that $f_n(t) = \frac{1}{1+tn^2}$. For each $t \in (0, \infty)$, $\sum_{n=1}^{\infty} f_n(t)$ converges by comparison with $\sum_{n=1}^{\infty} \frac{1}{n^2}$. Let $f(t) = \sum_{n=1}^{\infty} f_n(t)$.

(a) For any fixed $\epsilon > 0$, show that $\sum_{n=1}^{\infty} f_n'(t)$ converges uniformly on (ϵ, ∞). Hence, conclude that f is differentiable on $(0, \infty)$ with $f'(t) = \sum_{n=1}^{\infty} f_n'(t)$.

(b) For $x \in (0, 1]$, let $F(x) = \int_x^1 f(t)\,dt$. Is F bounded?

7.5 A Continuous but Nowhere-Differentiable Function

Using the fact that uniform convergence preserves continuity, we can prove the following surprising theorem.

Theorem 7.28. *There exists a continuous $f : \mathbb{R} \to \mathbb{R}$ that is nowhere differentiable.*

Proof. Let $\phi : \mathbb{R} \to \mathbb{R}$ be such that $\phi(x) = |x|$ for $-1 \le x < 1$ and $\phi(x+2) = \phi(x)$ for all $x \in \mathbb{R}$. Let n be an odd integer. Then ϕ is continuous on $(n, n+2)$. Since $n-1$ and $n+1$ are even, and ϕ has a period of 2, we have that

$$\lim_{x\to n^-} \phi(x) = \lim_{x\to 1^-} \phi(x + (n-1)) = \lim_{x\to 1^-} \phi(x) = 1$$

and

$$\lim_{x\to n^+} \phi(x) = \lim_{x\to -1^+} \phi(x + (n+1)) = \lim_{x\to -1^+} \phi(x) = 1.$$

Hence, $\lim_{x\to n} \phi(x) = 1 = \phi(n)$, so ϕ is continuous at n. Therefore, ϕ is continuous on $\mathbb{R}$ because the union of the intervals $[n, n+2)$ over all odd n is $\mathbb{R}$.

Define

$$f(x) = \sum_{n=1}^{\infty} \left(\frac{3}{4}\right)^n \phi(4^n x). \tag{2}$$

Note that $|(\frac{3}{4})^n \phi(4^n x)| \le (\frac{3}{4})^n$ for all $n \in \mathbb{N}$ and $x \in \mathbb{R}$. Since $\sum_{n=1}^{\infty} (\frac{3}{4})^n$ converges, the series (2) converges uniformly on $\mathbb{R}$ by the Weierstrass M-test. By Theorem 7.15, f is continuous on $\mathbb{R}$.

Fix $x \in \mathbb{R}$. We will show that $f'(x)$ does not exist by constructing a sequence $\{y_m\}$ converging to x such that $\lim_{m\to\infty} \left|\frac{f(y_m)-f(x)}{y_m - x}\right| = \infty$. Note that if $z \in \mathbb{R}$ and $\lfloor z - \frac{1}{2} \rfloor \ne \lfloor z \rfloor$, then

$$\lfloor z \rfloor \le \left\lfloor z + \frac{1}{2} \right\rfloor = \left\lfloor z - \frac{1}{2} \right\rfloor + 1 < \lfloor z \rfloor + 1,$$

so $\lfloor z \rfloor = \lfloor z + \frac{1}{2} \rfloor$. Hence, for all $m \in \mathbb{N}$, we can let

$$y_m = x \pm \frac{1}{2(4^m)}$$

where the sign (plus or minus) is chosen such that $\lfloor 4^m y_m \rfloor = \lfloor 4^m x \rfloor$. It is immediate that $\lim_{m\to\infty} y_m = x$. If $n > m$, then $4^n y_m = 4^n x \pm 2(4^{n-m-1})$, so $\phi(4^n y_m) = \phi(4^n x)$ because 4^{n-m-1} is an integer and ϕ has a period of 2. On the other hand, if $1 \le n \le m$, then $\lfloor 4^n y_m \rfloor = \lfloor 4^n x \rfloor$, so $|\phi(4^n y_m) - \phi(4^n x)| = |4^n y_m - 4^n x|$. Hence,

$$\begin{aligned}
|f(y_m) - f(x)| &= \left| \sum_{n=1}^{\infty} \left(\frac{3}{4}\right)^n [\phi(4^n y_m) - \phi(4^n x)] \right| \\
&= \left| \sum_{n=1}^{m} \left(\frac{3}{4}\right)^n [\phi(4^n y_m) - \phi(4^n x)] \right| \\
&\ge \left(\frac{3}{4}\right)^m |\phi(4^m y_m) - \phi(4^m x)| - \sum_{n=1}^{m-1} \left(\frac{3}{4}\right)^n |\phi(4^n y_m) - \phi(4^n x)|
\end{aligned}$$

$$
\begin{aligned}
&= \left(\frac{3}{4}\right)^m |4^m y_m - 4^m x| - \sum_{n=1}^{m-1} \left(\frac{3}{4}\right)^n |4^n y_m - 4^n x| \\
&= 3^m |y_m - x| - \sum_{n=1}^{m-1} 3^n |y_m - x| \\
&= \left(3^m - \sum_{n=1}^{m-1} 3^n\right) |y_m - x| \\
&= \frac{3^m + 3}{2} |y_m - x|,
\end{aligned}
$$

so

$$\left|\frac{f(y_m) - f(x)}{y_m - x}\right| \geq \frac{3^m + 3}{2}.$$

Since $\lim_{m\to\infty} \frac{3^m+3}{2} = \infty$, we have that $\lim_{m\to\infty} \left|\frac{f(y_m)-f(x)}{y_m-x}\right| = \infty$. □

7.6 The Arzelà–Ascoli Theorem

This section is dedicated to proving the Arzelà–Ascoli Theorem, which gives sufficient conditions for a sequence of functions to have a uniformly-convergent subsequence. The Arzelà–Ascoli Theorem is much like the Bolzano–Weierstrass Theorem (which says that any sequence in a compact set has a convergent subsequence).

Definition 7.29. Let X be a metric space. A set of functions $\mathcal{F}$ from X to $\mathbb{C}$ is *equicontinuous* (on X) if for all $\epsilon > 0$, there exists $\delta > 0$ such that for all $f \in \mathcal{F}$ and $x, y \in X$, if $d(x, y) < \delta$, then $|f(x) - f(y)| < \epsilon$.

Remark. Equicontinuity is like uniform continuity for sets of functions. Every function in an equicontinuous set is uniformly continuous, and there is a $\delta > 0$ that satisfies the definition of uniform continuity for *every* function in the set.

Definition 7.30. Let E be a set and $\mathcal{F}$ be a set of functions from E to $\mathbb{C}$.

(a) $\mathcal{F}$ is *pointwise bounded* if for all $x \in E$, there exists $M_x \geq 0$ such that $|f(x)| \leq M_x$ for all $f \in \mathcal{F}$.

(b) $\mathcal{F}$ is *uniformly bounded* if there exists $M \geq 0$ such that $|f(x)| \leq M$ for all $f \in \mathcal{F}$ and $x \in E$.

Theorem 7.31. *If $\{f_n\}$ is a uniformly-convergent sequence of continuous functions $K \to \mathbb{C}$ where K is compact, then $\{f_n\}$ is equicontinuous.*

Proof. Fix $\epsilon > 0$. Since $\{f_n\}$ converges uniformly, there exists $N \in \mathbb{N}$ such that if $n, m \geq N$ and $x \in K$, then $|f_n(x) - f_m(x)| < \frac{\epsilon}{3}$. Since $f_1, \ldots, f_N$ are continuous on the compact set K, these functions are uniformly continuous. For each $1 \leq j \leq N$, we can choose $\delta_j > 0$ such that if $d(x, y) < \delta_j$, then $|f_j(x) - f_j(y)| < \frac{\epsilon}{3}$. Choose $\delta = \min\{\delta_1, \ldots, \delta_N\} > 0$. Suppose $n \in \mathbb{N}$ and $d(x, y) < \delta$. If $1 \leq n \leq N$, then $|f_n(x) - f_n(y)| < \frac{\epsilon}{3} < \epsilon$ since $d(x, y) < \delta_n$. If $n > N$, then

$$|f_n(x) - f_n(y)| \leq |f_n(x) - f_N(x)| + |f_N(x) - f_N(y)| + |f_N(y) - f_n(y)|$$

$$
\begin{aligned}
&< \frac{\epsilon}{3} + \frac{\epsilon}{3} + \frac{\epsilon}{3} \quad \text{since } n \geq N \text{ and } d(x,y) < \delta_N \\
&= \epsilon.
\end{aligned}
$$

Therefore, $|f_n(x) - f_n(y)| < \epsilon$ for all $n \in \mathbb{N}$ and $d(x,y) < \delta$, so $\{f_n\}$ is equicontinuous. $\square$

Lemma 7.32. *Let K be compact. Then K has a dense subset $E \subset K$ that is at most countable. (We say that K is "separable".)*

Proof. For each $n \in \mathbb{N}$, the open cover $\{N_{1/n}(x)\}_{x \in K}$ of K has a finite subcover

$$\{N_{1/n}(x_{1,n}), \ldots, N_{1/n}(x_{r_n,n})\}$$

where $r_n \in \mathbb{N}$. Consider the set

$$E = \bigcup_{n=1}^{\infty} \{x_{1,n}, \ldots, x_{r_n,n}\},$$

which is a countable union of finite sets and is therefore at most countable. Let $y \in K \setminus E$ and $\delta > 0$. Choose $n \in \mathbb{N}$ such that $\frac{1}{n} < \delta$. Then there exists $1 \leq j \leq r_n$ such that $y \in N_{1/n}(x_{j,n})$ since $\{N_{1/n}(x_{i,n})\}_{i=1}^{r_n}$ covers K. Hence, $x_{j,n} \in N_{1/n}(y) \subset N_\delta(y)$, so y is a limit point of E. Therefore, $\overline{E} = K$. $\square$

Lemma 7.33. *Let $\{f_n\}$ be a pointwise-bounded sequence of functions $E \to \mathbb{C}$ where E is at most countable. Then there exists a subsequence $\{f_{n_k}\}$ that converges pointwise on E.*

Proof. Suppose E is infinite. Write $E = \{x_1, x_2, x_3, \ldots\}$. Since $\{f_n(x_1)\}_n$ is a bounded sequence in $\mathbb{C}$, there exists a subsequence $S_1 = \{f_{n_{1,j}}\}$ such that $\{f_{n_{1,j}}(x_1)\}_j$ converges (in $\mathbb{C}$). Recursively, suppose we have a subsequence $S_i = \{f_{n_{i,j}}\}_j$ such that $\{f_{n_{i,j}}(x_i)\}_j$ converges. Then $\{f_{n_{i,j}}(x_{i+1})\}_j$ is a bounded sequence in $\mathbb{C}$, so there is a subsequence $S_{i+1} = \{f_{n_{i+1,j}}\}_j$ of S_i such that $\{f_{n_{i+1,j}}(x_{i+1})\}_j$ converges. This recursive procedure constructs subsequences $S_i = \{f_{n_{i,j}}\}_j$ such that for each i, S_{i+1} is a subsequence of S_i and $\{f_{n_{i,j}}(x_i)\}_j$ converges. Now form the subsequence $S = \{f_{n_{k,k}}\}_k$. Then for each i, we have that $f_{n_{k,k}} \in S_i$ for all $k \geq i$. In other words, S is eventually a subsequence of S_i for all i. Hence, $\{f_{n_{k,k}}(x_i)\}_k$ converges for each i, so the subsequence S converges pointwise on E.

If E is finite, the recursive procedure for constructing the subsequences S_i terminates after a finite number of steps. The final subsequence $S_{|E|}$ is a subsequence of all of $S_1, \ldots, S_{|E|-1}$, so $S_{|E|}$ is pointwise-convergent on E. $\square$

Theorem 7.34 (Arzelà–Ascoli)**.** *Let K be compact, and suppose $\{f_n\}$ is an equicontinuous and pointwise-bounded sequence in $\mathcal{C}(K;\mathbb{C})$. Then:*

(a) $\{f_n\}$ is uniformly bounded.

(b) There exists a subsequence $\{f_{n_k}\}$ that converges uniformly on K.

Proof. (a) The pointwise-bounded assumption means that for all $x \in K$, there exists $M_x \geq 0$ such that $|f_n(x)| \leq M_x$ for all $n \in \mathbb{N}$. Since $\{f_n\}$ is equicontinuous, there exists $\delta > 0$ such that $|f_n(x) - f_n(y)| < 1$ for all $n \in \mathbb{N}$ and $x, y \in K$ such that $d(x,y) < \delta$. Note that $\{N_\delta(x)\}_{x \in K}$ is an open cover of K. By compactness of K, there exists a finite subcover

$\{N_\delta(x_1), \ldots, N_\delta(x_r)\}$. Let $M = \max\{M_{x_1}, \ldots, M_{x_r}\}$. Now fix $n \in \mathbb{N}$ and $x \in K$. Then there exists $1 \le j \le r$ such that $x \in N_\delta(x_j)$ since $\{N_\delta(x_1), \ldots, N_\delta(x_r)\}$ covers K. Hence,

$$\begin{aligned} |f_n(x)| &= |f_n(x) - f_n(x_j) + f_n(x_j)| \\ &\le |f_n(x) - f_n(x_j)| + |f_n(x_j)| \\ &< 1 + M_{x_j} \quad \text{since } d(x, x_j) < \delta \\ &\le 1 + M. \end{aligned}$$

Therefore, $\{f_n\}$ is uniformly bounded.

(b) Fix $\epsilon > 0$. By equicontinuity, there exists $\delta > 0$ such that $|f_n(x) - f_n(y)| < \frac{\epsilon}{3}$ for all $n \in \mathbb{N}$ and $x, y \in K$ such that $d(x, y) < \delta$. By Lemma 7.32, K has a dense subset E that is at most countable. The density of E ensures that for all $x \in K$, there exists $y \in E$ such that $x \in N_\delta(y)$. Hence, $\{N_\delta(x)\}_{x \in E}$ is an open cover of K, and by compactness of K, there exists a finite subcover $\{N_\delta(x_1), \ldots, N_\delta(x_r)\}$. Lemma 7.33 says that there is a subsequence $\{f_{n_k}\}$ of $\{f_n\}$ that converges pointwise on E. For all integers $1 \le j \le r$, since $x_j \in E$, there exists $N_j \in \mathbb{N}$ such that $|f_{n_k}(x_j) - f_{n_\ell}(x_j)| < \frac{\epsilon}{3}$ whenever $k, \ell \ge N_j$. Put $N = \max\{N_1, \ldots, N_r\} \in \mathbb{N}$. Fix $x \in K$ and $k, \ell \ge N$. Then there exists $1 \le j \le r$ such that $x \in N_\delta(x_j)$ since $\{N_\delta(x_1), \ldots, N_\delta(x_r)\}$ covers K. Hence,

$$\begin{aligned} |f_{n_k}(x) - f_{n_\ell}(x)| &\le |f_{n_k}(x) - f_{n_k}(x_j)| + |f_{n_k}(x_j) - f_{n_\ell}(x_j)| + |f_{n_\ell}(x_j) - f_{n_\ell}(x)| \\ &< \frac{\epsilon}{3} + \frac{\epsilon}{3} + \frac{\epsilon}{3} \quad \text{since } d(x, x_j) < \delta \text{ and } k, \ell \ge N \ge N_j \\ &= \epsilon, \end{aligned}$$

so $\{f_{n_k}\}$ converges uniformly by the Cauchy Criterion. □

Exercise 7.35. Let K be compact. Prove that a subset E of $\mathcal{C}(K; \mathbb{C})$ (equipped with the supremum metric) is compact if and only if E is closed, bounded, and equicontinuous. (Recall Theorem 3.64, which says that compactness is equivalent to sequential compactness.)

7.7 The Stone–Weierstrass Theorem

Approximating functions using simpler functions is a common theme in analysis. Taylor series use polynomials to approximate infinitely-differentiable functions. Fourier series use periodic functions (e^{inx}) to approximate square-integrable functions. The big theorem that this section builds up to is called the Stone–Weierstrass Theorem and looks like this:

Let K be a compact set and $\mathcal{A}$ be a set of continuous functions $K \to \mathbb{R}$ that satisfy certain hypotheses. Then for any continuous function $f : K \to \mathbb{R}$, there exists a sequence of functions $\phi_n \in \mathcal{A}$ that converge uniformly to f on K.

We start with a special case: the Weierstrass Approximation Theorem.

Theorem 7.36 (Weierstrass Approximation Theorem)**.** *Let $f : [a, b] \to \mathbb{R}$ be continuous. Then there exists a sequence of real polynomials p_n that converges uniformly to f on $[a, b]$.*

The proof we present is motivated by probability theory. After giving the proof, we will explain the key steps from a probabilistic perspective.

Proof. It suffices to prove the theorem for $a = 0$ and $b = 1$ because any function on $[a, b]$ can be linearly transformed into a function on $[0, 1]$ and vice versa. We claim that the polynomials

$$p_n(x) = \sum_{k=0}^{n} f\left(\frac{k}{n}\right)\binom{n}{k}x^k(1-x)^{n-k} \tag{3}$$

converge to f uniformly on $[0, 1]$. We first note that

$$\sum_{k=0}^{n}\binom{n}{k}x^k(1-x)^{n-k} = 1, \tag{4}$$

$$\sum_{k=0}^{n}k\binom{n}{k}x^k(1-x)^{n-k} = nx, \tag{5}$$

and

$$\sum_{k=0}^{n}k^2\binom{n}{k}x^k(1-x)^{n-k} = n(n-1)x^2 + nx \tag{6}$$

for all $n \geq 0$ and $x \in \mathbb{R}$; the proofs mainly rely on the Binomial Theorem. These three equations together imply that

$$\sum_{k=0}^{n}(k-nx)^2\binom{n}{k}x^k(1-x)^{n-k} = nx(1-x). \tag{7}$$

The key ingredient of the proof is the observation that for all $n \geq 1$, $0 < x < 1$, and $m > 0$,

$$\begin{aligned}
nx(1-x) &= \sum_{k=0}^{n}(k-nx)^2\binom{n}{k}x^k(1-x)^{n-k} \\
&\geq \sum_{|k-nx|\geq m\sqrt{nx(1-x)}}(k-nx)^2\binom{n}{k}x^k(1-x)^{n-k} \\
&\geq \sum_{|k-nx|\geq m\sqrt{nx(1-x)}}m^2nx(1-x)\binom{n}{k}x^k(1-x)^{n-k} \\
&= m^2nx(1-x)\sum_{|k-nx|\geq m\sqrt{nx(1-x)}}\binom{n}{k}x^k(1-x)^{n-k},
\end{aligned}$$

which implies that

$$\sum_{|k-nx|\geq m\sqrt{nx(1-x)}}\binom{n}{k}x^k(1-x)^k \leq \frac{1}{m^2}. \tag{8}$$

It is clear that (8) also holds when $x = 0$ or $x = 1$, since the left-hand side equals 0 in these cases. Note that the inequality $|k-nx| \geq m\sqrt{nx(1-x)}$ is equivalent to $|\frac{k}{n}-x| \geq m\sqrt{\frac{x(1-x)}{n}}$. Therefore, by applying (8) with $m = n^{1/3}$, we obtain that

$$\sum_{|\frac{k}{n}-x|\geq n^{-1/6}\sqrt{x(1-x)}}\binom{n}{k}x^k(1-x)^{n-k} \leq \frac{1}{n^{2/3}}. \tag{9}$$

Fix $\epsilon > 0$. Let $M = \sup_{t\in[0,1]} |f(t)|$, which is finite since f is continuous and $[0,1]$ is compact. Note that f is *uniformly* continuous on $[0,1]$ since $[0,1]$ is compact. Hence, there exists $\delta > 0$ such that if $s, t \in [0,1]$ and $|s-t| < \delta$, then $|f(s) - f(t)| < \epsilon$. Pick $N \in \mathbb{N}$ such that $\frac{1}{n^{2/3}} < \epsilon$ and $\frac{1}{2n^{1/6}} < \delta$ for all $n \geq N$. For a fixed $x \in [0,1]$ and $n \geq N$,

$$\begin{aligned}
|p_n(x) - f(x)| &= \left|\sum_{k=0}^{n} f\left(\frac{k}{n}\right)\binom{n}{k}x^k(1-x)^{n-k} - f(x)\right| \\
&= \left|\sum_{k=0}^{n}\left[f\left(\frac{k}{n}\right) - f(x)\right]\binom{n}{k}x^k(1-x)^{n-k}\right| \quad \text{by (4)} \\
&\leq \sum_{k=0}^{n}\left|f\left(\frac{k}{n}\right) - f(x)\right|\binom{n}{k}x^k(1-x)^{n-k} \\
&= \sum_{|\frac{k}{n}-x|\leq n^{-1/6}\sqrt{x(1-x)}}\left|f\left(\frac{k}{n}\right) - f(x)\right|\binom{n}{k}x^k(1-x)^{n-k} \\
&\quad + \sum_{|\frac{k}{n}-x|\geq n^{-1/6}\sqrt{x(1-x)}}\left|f\left(\frac{k}{n}\right) - f(x)\right|\binom{n}{k}x^k(1-x)^{n-k}.
\end{aligned}$$

Now note that $\sqrt{t(1-t)} \leq \frac{1}{2}$ for all $t \in [0,1]$. Hence, $n^{-1/6}\sqrt{x(1-x)} \leq \frac{1}{2}n^{-1/6} < \delta$, so if $|\frac{k}{n} - x| \leq n^{-1/6}\sqrt{x(1-x)}$, then $|f(\frac{k}{n}) - f(x)| < \epsilon$. We therefore obtain the bound

$$\sum_{|\frac{k}{n}-x|\leq n^{-1/6}\sqrt{x(1-x)}}\left|f\left(\frac{k}{n}\right) - f(x)\right|\binom{n}{k}x^k(1-x)^{n-k} \leq \epsilon\sum_{k=0}^{n}\binom{n}{k}x^k(1-x)^{n-k} = \epsilon.$$

We also have that

$$\begin{aligned}
&\sum_{|\frac{k}{n}-x|\geq n^{-1/6}\sqrt{x(1-x)}}\left|f\left(\frac{k}{n}\right) - f(x)\right|\binom{n}{k}x^k(1-x)^{n-k} \\
&\leq 2M\sum_{|\frac{k}{n}-x|\geq n^{-1/6}\sqrt{x(1-x)}}\binom{n}{k}x^k(1-x)^{n-k} \\
&\leq 2M\epsilon
\end{aligned}$$

by (9). Therefore,

$$|p_n(x) - f(x)| \leq \epsilon + 2M\epsilon = (1+2M)\epsilon$$

for all $x \in [0,1]$ and $n \geq N$, so $p_n \to f$ uniformly on $[0,1]$. □

We now give the probabilistic motivation behind the proof. Suppose we have a coin which, when flipped, lands on heads with a fixed probability $x \in [0,1]$. For any $n \in \mathbb{N}$, define the discrete random variable X_n to be the number of times the coin lands on heads when flipped n times in succession. Then for any integer $0 \leq k \leq n$, we have that

$$\Pr(X_n = k) = \binom{n}{k}x^k(1-x)^{n-k}.$$

Equation (4) encodes the fact that the sum of these probabilities for $0 \leq k \leq n$ should be 1. The *expectation* of a discrete random variable Y is

$$\mathbb{E}(X_n) := \sum_{y \in S} y \Pr(Y = y),$$

where S is the (at most countable) set of possible values that Y can take. The *variance* of Y is

$$\operatorname{Var}(Y) := \mathbb{E}([Y - \mathbb{E}(Y)]^2) = \mathbb{E}(Y^2) - \mathbb{E}(Y)^2.$$

Equations (5) and (7) encode the facts that $\mathbb{E}(X_n) = nx$ and $\operatorname{Var}(X_n) = nx(1-x)$. The key inequality (8) is an instance of *Chebyshev's inequality*: for any random variable Y, if $m > 0$, then

$$\Pr(|Y - \mu| \geq m\sigma) \leq \frac{1}{m^2},$$

where $\mu = \mathbb{E}(Y)$ and $\sigma = \sqrt{\operatorname{Var}(Y)}$. Applying Chebyshev's inequality with $Y = X_n$ gives (8).

The polynomials defined in (3) may be more easily remembered as

$$p_n(x) = \mathbb{E}(f(V_n))$$

where $V_n = \frac{X_n}{n}$. These polynomials are called *Bernstein polynomials*, named after the mathematician who found the probabilistic proof of the Weierstrass Approximation Theorem. Note that V_n represents the proportion of times that the coin lands on heads. By linearity of expectation, $\mathbb{E}(V_n) = x$ and $\operatorname{Var}(V_n) = \frac{x(1-x)}{n}$, so when n is large, the value of V_n should be close to x most of the time since the variance tends to 0. Since f is continuous, $f(V_n)$ should be close to $f(x)$ most of the time. Thus, $\mathbb{E}(f(V_n))$ should be close to $f(x)$ when n is large, and the *uniform* convergence results from the *uniform* continuity of f.

Definition 7.37. Let E be a set, and let F be $\mathbb{R}$ or $\mathbb{C}$. Let $\mathcal{A}$ be a set of functions $f : E \to F$. We say that $\mathcal{A}$ is an F-*algebra* if for all $f, g \in \mathcal{A}$ and $c \in F$,

(i) $f + g \in \mathcal{A}$,
(ii) $fg \in \mathcal{A}$, and
(iii) $cf \in \mathcal{A}$.

The *uniform closure* of $\mathcal{A}$, denoted $\overline{\mathcal{A}}$, is the set of functions $\phi : E \to F$ for which there exists a sequence of functions $\phi_n \in \mathcal{A}$ that converge uniformly to ϕ. It is clear that $\mathcal{A} \subset \overline{\mathcal{A}}$ since constant sequences of functions converge uniformly. We say that $\mathcal{A}$ is *uniformly closed* if $\mathcal{A} = \overline{\mathcal{A}}$.

Theorem 7.38. *If $\mathcal{A}$ is an F-algebra of bounded functions $f : E \to F$, then $\overline{\mathcal{A}}$ is a uniformly-closed F-algebra.*

Proof. Theorem 7.9 implies that $\overline{\mathcal{A}}$ is an F-algebra since $\mathcal{A}$ is an F-algebra of bounded functions. (It may seem that Theorem 7.9 only applies if $\mathcal{A}$ is a $\mathbb{C}$-algebra, but since we can view $\mathbb{R}$ as a subset of $\mathbb{C}$, Theorem 7.9 also applies if $\mathcal{A}$ is an $\mathbb{R}$-algebra.) We just need to prove that $\overline{\mathcal{A}}$ is uniformly closed. Suppose a sequence of functions $f_n \in \overline{\mathcal{A}}$ converges uniformly on E to some function $f : E \to F$. Fix $\epsilon > 0$. For each n, since $f_n \in \overline{\mathcal{A}}$, there exists $g_n \in \mathcal{A}$

such that $|f_n(x) - g_n(x)| < \epsilon$ for all $x \in E$. Choose $N \in \mathbb{N}$ such that $|f_n(x) - f(x)| < \epsilon$ for all $n \geq N$ and $x \in E$. If $n \geq N$, then

$$|g_n(x) - f(x)| \leq |g_n(x) - f_n(x)| + |f_n(x) - f(x)| < \epsilon + \epsilon = 2\epsilon$$

for all $x \in E$. Therefore, $g_n \to f$ uniformly on E, so $f \in \overline{\mathcal{A}}$, which means $\overline{\mathcal{A}}$ is uniformly closed. □

Definition 7.39. Let $\mathcal{A}$ be an F-algebra of functions $f : E \to F$. We say that $\mathcal{A}$ *separates points of* E if for all distinct $x_1, x_2 \in E$, there exists $f \in \mathcal{A}$ such that $f(x_1) \neq f(x_2)$. We say that $\mathcal{A}$ *vanishes at no point of* E if for all $x \in E$, there exists $f \in \mathcal{A}$ such that $f(x) \neq 0$.

Lemma 7.40. *Let $\mathcal{A}$ be an $\mathbb{R}$-algebra of bounded functions $f : E \to \mathbb{R}$. If $f, g \in \mathcal{A}$, then*
(a) $|f| \in \overline{\mathcal{A}}$,
(b) $\max\{f, g\} \in \overline{\mathcal{A}}$, *and*
(c) $\min\{f, g\} \in \overline{\mathcal{A}}$.

Proof. Let $M \in \mathbb{R}$ such that $|f(x)| \leq M$ for all $x \in E$. Fix $\epsilon > 0$. Since the map $x \mapsto |x|$ is continuous on $[-M, M]$, Theorem 7.36 says that there exists a polynomial $p(x) = c_0 + c_1 x + \cdots + c_n x^n$ such that $|p(x) - |x|| < \frac{\epsilon}{2}$ for all $x \in [-M, M]$. Notice that $|c_0| = |p(0)| = |p(0) - |0|| < \frac{\epsilon}{2}$. Let $q(x) = p(x) - c_0 = c_1 x + \cdots + c_n x^n$. Then

$$|q(x) - |x|| \leq |p(x) - |x|| + |-c_0| < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon$$

for all $x \in [-M, M]$. Hence, if $x \in E$, then

$$|q(f(x)) - |f(x)|| < \epsilon$$

because $f(x) \in [-M, M]$. Now notice that $q \circ f = c_1 f + \cdots + c_n f^n \in \mathcal{A}$. Therefore, $|f|$ can be uniformly approximated by functions in $\mathcal{A}$, so $|f| \in \overline{\mathcal{A}}$. This proves part (a).

Parts (b) and (c) now follow from the identities

$$\max\{f, g\} = \frac{1}{2}(f + g) + \frac{1}{2}|f - g|$$

and

$$\min\{f, g\} = \frac{1}{2}(f + g) - \frac{1}{2}|f - g|,$$

together with the fact that $\overline{\mathcal{A}}$ is an $\mathbb{R}$-algebra (Theorem 7.38). □

Lemma 7.41. *Let $\mathcal{A}$ be an F-algebra of functions $f : E \to F$ that separates points and vanishes at no point. For any $x_1, x_2 \in E$ and $c_1, c_2 \in F$ with $x_1 \neq x_2$, there exists $f \in \mathcal{A}$ such that $f(x_1) = c_1$ and $f(x_2) = c_2$.*

Proof. Since $\mathcal{A}$ separates points of E, there exists $g \in \mathcal{A}$ such that $g(x_1) \neq g(x_2)$. Since $\mathcal{A}$ vanishes at no point, there exist $h_1, h_2 \in \mathcal{A}$ such that $h_1(x_1) \neq 0$ and $h_2(x_2) \neq 0$. Define

$$\phi_1(x) = \frac{[g(x) - g(x_2)]h_1(x)}{[g(x_1) - g(x_2)]h_1(x_1)} = \frac{1}{[g(x_1) - g(x_2)]h_1(x_1)}[g(x)h_1(x) - g(x_2)h_1(x)]$$

and

$$\phi_2(x) = \frac{[g(x) - g(x_1)]h_2(x)}{[g(x_2) - g(x_1)]h_2(x_2)} = \frac{1}{[g(x_2) - g(x_1)]h_2(x_2)}[g(x)h_2(x) - g(x_1)h_2(x)].$$

Then $\phi_1, \phi_2 \in \mathcal{A}$. Also, $\phi_1(x_1) = 1$, $\phi_1(x_2) = 0$, $\phi_2(x_1) = 0$, and $\phi_2(x_2) = 1$. Let $f = c_1\phi_1 + c_2\phi_2 \in \mathcal{A}$. Then $f(x_1) = c_1$ and $f(x_2) = c_2$. □

We are now ready to state and prove the Stone–Weierstrass Theorem. The proof is an excellent demonstration of the power of compactness.

Theorem 7.42 (Real Stone–Weierstrass Theorem)**.** *Let K be compact and $\mathcal{A} \subset \mathcal{C}(K;\mathbb{R})$ be an $\mathbb{R}$-algebra that separates points and vanishes at no point. Then $\overline{\mathcal{A}} = \mathcal{C}(K;\mathbb{R})$.*

Proof. Theorem 7.15 implies that $\overline{\mathcal{A}} \subset \mathcal{C}(K;\mathbb{R})$. All functions in $\mathcal{C}(K;\mathbb{R})$ and hence in $\overline{\mathcal{A}}$ are bounded. Therefore, $\overline{\mathcal{A}}$ is uniformly closed by Theorem 7.38.

Fix $\phi \in \mathcal{C}(K;\mathbb{R})$ and $\epsilon > 0$. We claim that for all $x \in K$, there exists $f_x \in \overline{\mathcal{A}}$ such that $f_x(x) = \phi(x)$ and $f_x(y) < \phi(y) + \epsilon$ for all $y \in K$. For all $t \in K \setminus \{x\}$, Lemma 7.41 says that there exists a function $g_{x,t} \in \overline{\mathcal{A}}$ such that $g_{x,t}(x) = \phi(x)$ and $g_{x,t}(t) = \phi(t)$. Let $S_{x,t} = \{y \in K \mid g_{x,t}(y) < \phi(y) + \epsilon\}$, and notice that

$$S_{x,t} = (g_{x,t} - \phi)^{-1}((-\infty, \epsilon)).$$

Hence, $S_{x,t}$ is an open set because $g_{x,t} - \phi$ is continuous and $(-\infty, \epsilon)$ is an open set. Moreover, $t \in S_{x,t}$ because $g_{x,t}(t) = \phi(t) < \phi(t) + \epsilon$. The collection $\{S_{x,t}\}_{t \in K}$ is therefore an open cover of K. Since K is compact, there exists a finite subcover $\{S_{x,t_1}, \ldots, S_{x,t_n}\}$. Put $f_x = \min_{1 \le i \le n} g_{x,t_i}$. Repeated application of Lemma 7.40 yields that $f_x \in \overline{\mathcal{A}}$. We immediately have that $f_x(x) = \phi(x)$ since $g_{x,t_i}(x) = \phi(x)$ for all i. Fix $y \in K$, and let $1 \le j \le n$ such that $y \in S_{x,t_j}$. Then

$$f_x(y) \le g_{x,t_j}(y) < \phi(y) + \epsilon,$$

proving our claim.

Now we construct $f \in \overline{\mathcal{A}}$ such that $|f(y) - \phi(y)| < \epsilon$ for all $y \in K$. For each $x \in K$, let $T_x = \{y \in K \mid f_x(y) > \phi(y) - \epsilon\}$. Then

$$T_x = (f_x - \phi)^{-1}((-\epsilon, \infty))$$

is an open set, and $x \in T_x$ since $f_x(x) = \phi(x) > \phi(x) - \epsilon$. Hence, $\{T_x\}_{x \in K}$ is an open cover of K, so there exists a finite subcover $\{T_{x_1}, \ldots, T_{x_m}\}$. Put $f = \max_{1 \le i \le m} f_{x_i}$, and note that $f \in \overline{\mathcal{A}}$ by Lemma 7.40. Fix $y \in K$, and let $1 \le j \le n$ such that $y \in T_{x_j}$. Then

$$f(y) \ge f_{x_j}(y) > \phi(y) - \epsilon.$$

On the other hand, $f_{x_i}(y) < \phi(y) + \epsilon$ for all i by construction of the functions f_x, so $f(y) < \phi(y) + \epsilon$. Altogether, we have that $\phi(y) - \epsilon < f(y) < \phi(y) + \epsilon$, so $|f(y) - \phi(y)| < \epsilon$.

Since $\epsilon > 0$ is arbitrary, we can obtain a sequence $\{\phi_n\}_{n \in \mathbb{N}}$ of functions in $\overline{\mathcal{A}}$ such that $|\phi_n(y) - \phi(y)| < \frac{1}{n}$ for all $n \in \mathbb{N}$ and $y \in K$. The ϕ_n uniformly converge to ϕ since $\lim_{n \to \infty} \frac{1}{n} = 0$. Therefore, $\phi \in \overline{\mathcal{A}}$ since $\overline{\mathcal{A}}$ is uniformly closed. □

Definition 7.43. A $\mathbb{C}$-algebra $\mathcal{A}$ of functions $f : E \to \mathbb{C}$ is *self-adjoint* if $f \in \mathcal{A}$ implies that $\overline{f} \in \mathcal{A}$, where $\overline{f}(x) \coloneqq \overline{f(x)}$.

Theorem 7.44 (Complex Stone–Weierstrass Theorem)**.** *Let K be compact and $\mathcal{A} \subset \mathcal{C}(K;\mathbb{C})$ be a self-adjoint $\mathbb{C}$-algebra that separates points and vanishes at no point. Then $\overline{\mathcal{A}} = \mathcal{C}(K;\mathbb{C})$.*

Proof. Just as in the proof of Theorem 7.42, it is immediate that $\overline{\mathcal{A}} \subset \mathcal{C}(K;\mathbb{C})$ and that $\overline{\mathcal{A}}$ is uniformly closed. Fix $\phi \in \mathcal{C}(K;\mathbb{C})$. We can easily verify that $|\overline{\phi}(x) - \overline{\phi}(y)| = |\phi(x) - \phi(y)|$ for all $x, y \in K$. Therefore, $\overline{\phi}$ is continuous (write out the $\epsilon - \delta$ definition of continuity on ϕ and realize that it also applies to $\overline{\phi}$). Also, $\overline{\phi}$ is bounded because $|\phi(x)| = |\overline{\phi}(x)|$ for all $x \in K$. Consequently,

$$\mathrm{Re}(\phi) = \frac{\phi + \overline{\phi}}{2} \in \mathcal{C}(K;\mathbb{R}) \tag{10}$$

and

$$\mathrm{Im}(\phi) = \frac{\phi - \overline{\phi}}{2i} \in \mathcal{C}(K;\mathbb{R}).$$

Let $\mathcal{B} = \{\mathrm{Re}(f) \mid f \in \mathcal{A}\}$. Then $\mathcal{B} \subset \mathcal{C}(K;\mathbb{R})$. Also, since $\mathcal{A}$ is self-adjoint, equation (10) implies that $\mathcal{B} \subset \mathcal{A}$. It now suffices to show that $\mathcal{B}$ is an $\mathbb{R}$-algebra that separates points and vanishes at no point. Indeed, if $\mathcal{B}$ satisfies the hypotheses of Theorem 7.42, then $\mathrm{Re}(\phi) \in \overline{\mathcal{B}} \subset \overline{\mathcal{A}}$ and $\mathrm{Im}(\phi) \in \overline{\mathcal{B}} \subset \overline{\mathcal{A}}$. Since $\overline{\mathcal{A}}$ is uniformly closed, we would conclude that $\phi = \mathrm{Re}(\phi) + i\mathrm{Im}(\phi) \in \overline{\mathcal{A}}$.

Let $\alpha, \beta \in \mathcal{B}$ and $c \in \mathbb{R}$. Then there exist $f, g \in \mathcal{A}$ such that $\alpha = \mathrm{Re}(f)$ and $\beta = \mathrm{Re}(g)$. Note that

$$c\alpha = \mathrm{Re}(cf) \in \mathcal{B},$$

$$\alpha + \beta = \mathrm{Re}(f + g) \in \mathcal{B},$$

and

$$\alpha\beta = \left(\frac{f + \overline{f}}{2}\right)\left(\frac{g + \overline{g}}{2}\right) = \frac{fg + f\overline{g} + \overline{f}g + \overline{fg}}{4} = \mathrm{Re}\left(\frac{fg + f\overline{g}}{2}\right) \in \mathcal{B}.$$

Therefore, $\mathcal{B}$ is an $\mathbb{R}$-algebra.

Let $x, y \in K$ such that $x \neq y$. Since $\mathcal{A}$ vanishes at no point, there exists $f \in \mathcal{A}$ such that $f(x) \neq 0$. Let $g = \mathrm{Re}(\frac{1}{f(x)}f) \in \mathcal{B}$. Then $g(x) = 1 \neq 0$, so $\mathcal{B}$ vanishes at no point. Also, by Lemma 7.41, there exists $h \in \mathcal{A}$ such that $h(x) = 1$ and $h(y) = 2$. Letting $\ell = \mathrm{Re}(h) \in \mathcal{B}$, we have that $\ell(x) = 1 \neq 2 = \ell(y)$, so $\mathcal{B}$ separates points of K. $\square$