Real Analysis

Zihan Liang

Monday 15th December, 2025

About these notes. These are English notes for Real Analysis I/II. They follow a Chapter–Lecture–Subsection hierarchy and aim to balance full mathematical rigor with clear intuition and usable proof strategies. Core topics include the structure of the real numbers, sequences and series, limits and continuity, differentiation, the Riemann integral, sequences and series of functions (with uniform convergence), and foundational metric-space ideas when they clarify the real-line theory. The emphasis is on learning how to read and write correct proofs: identifying hypotheses precisely, choosing the right definitions, and organizing arguments around a small set of reliable tools (the completeness of $\mathbb{R}$, compactness, and ε-based estimates). Whenever a theorem is stated, the goal is not only to apply it correctly, but also to understand *why* it is true, *which* assumptions are essential, and *how* to recognize when it is the right tool.

Contents

1 Foundations and Mathematical Preliminaries

Key Idea. Real analysis is built on a small set of precise ideas: (i) *logical form* (especially quantifiers), (ii) *set-theoretic language* for describing collections, (iii) *functions* as rules between sets, and (iv) *size/comparison of infinite sets* (countability vs. uncountability). If you can read and write statements with quantifiers correctly and prove basic claims about sets and functions, the later ε–definitions (limits, continuity, integrability) become routine rather than mysterious.

Notation 1.1. Throughout these notes, we adopt the convention

$$\mathbb{N} = \{1, 2, 3, \dots\}, \qquad \mathbb{N}_0 = \mathbb{N} \cup \{0\}.$$

For sets A, B, we write $A \subseteq B$ for "A is a subset of B," and $A \subsetneq B$ for "$A \subseteq B$ but $A \neq B$." If the universe (ambient set) is not explicitly stated, complements are taken relative to a fixed universe U.

1.1 Logic, Proof Techniques, and Set Theory

1.1.1 Propositions and Logical Connectives

Definition 1.2 (Proposition)**.** A *proposition* is a declarative statement that is either *true* or *false* (but not both).

Common logical connectives for propositions P and Q:

- Negation: $\neg P$ ("not P").
- Conjunction: $P \wedge Q$ ("P and Q").
- Disjunction: $P \vee Q$ ("P or Q").
- Implication: $P \Rightarrow Q$ ("if P, then Q").
- Biconditional: $P \Leftrightarrow Q$ ("P iff Q").

Remark 1.3 (Implication is about *cases*). The statement $P \Rightarrow Q$ is declared false only in the case where P is true and Q is false. If P is false, the implication $P \Rightarrow Q$ is (by definition) true; this reflects the fact that an implication makes a claim only about situations where the hypothesis P actually holds.

Proposition 1.4 (Implication as a disjunction)**.** *For propositions P, Q, one has the logical equivalence*

$$(P \Rightarrow Q) \;\Leftrightarrow\; (\neg P \vee Q).$$

Proof. A truth-table verification is the cleanest method:

P	Q	$P \Rightarrow Q$	$\neg P \vee Q$
T	T	T	T
T	F	F	F
F	T	T	T
F	F	T	T

The last two columns match in all four cases, hence the statements are logically equivalent. ☐

Remark 1.5 (Converse vs. contrapositive). The *converse* of $P \Rightarrow Q$ is $Q \Rightarrow P$ and is generally *not* equivalent to the original statement. The *contrapositive* of $P \Rightarrow Q$ is $\neg Q \Rightarrow \neg P$ and is always logically equivalent to $P \Rightarrow Q$ (apply Theorem 1.4 twice).

1.1.2 Quantifiers and Negation Rules

Definition 1.6 (Quantifiers)**.** Let A be a set and $P(x)$ a predicate (a statement depending on x).

- $\forall x \in A,\ P(x)$ means "for every x in A, $P(x)$ is true."
- $\exists x \in A$ such that $P(x)$ means "there exists at least one x in A for which $P(x)$ is true."

Definition 1.7 (Unique existence)**.** We write $\exists! x \in A$ such that $P(x)$ to mean: there exists *exactly one* $x \in A$ for which $P(x)$ holds. Equivalently,

$$\exists! x \in A,\ P(x) \quad\Longleftrightarrow\quad \Big(\exists x \in A,\ P(x)\Big) \ \wedge\ \Big(\forall y \in A,\ (P(y) \Rightarrow y = x)\Big).$$

Negating quantified statements is a frequent source of errors. The correct rules are:

$$\neg\Big(\forall x \in A,\ P(x)\Big) \iff \exists x \in A \text{ such that } \neg P(x).$$

$$\neg\Big(\exists x \in A \text{ such that } P(x)\Big) \iff \forall x \in A,\ \neg P(x).$$

Remark 1.8. Read Section 1.1.2 literally: "It is not true that *everyone* has property P" means "*someone* fails to have property P." Similarly, Section 1.1.2 says "It is not true that *someone* has property P" means "*everyone* fails to have property P." These are the logical foundations behind many ε–arguments later in analysis.

Example 1.9 (Negating an analysis-style statement)**.** Negate the statement:

$$\forall \varepsilon > 0\ \exists N \in \mathbb{N}\ \forall n \in \mathbb{N},\ (n \geq N \Rightarrow |x_n - L| < \varepsilon).$$

Using Section 1.1.2 repeatedly, the negation is:

$$\exists \varepsilon_0 > 0\ \forall N \in \mathbb{N}\ \exists n \in \mathbb{N} \text{ with } n \geq N \text{ and } |x_n - L| \geq \varepsilon_0.$$

This is precisely the definition of "x_n does not converge to L."

1.1.3 Proof Techniques

Key Idea. Most rigorous proofs in early analysis use one of four patterns: direct proof, contrapositive, contradiction, or induction. Learning to choose the right pattern is half of the work; the other half is writing the quantifiers correctly.

Direct proof. Start from the hypotheses and deduce the conclusion through definitions and previously known facts.

Proposition 1.10. *If $n \in \mathbb{Z}$ is even, then n^2 is even.*

Proof. If n is even, then $n = 2k$ for some $k \in \mathbb{Z}$. Hence $n^2 = 4k^2 = 2(2k^2)$, which is even. □

Contrapositive. To prove $P \Rightarrow Q$, it is often simpler to prove $\neg Q \Rightarrow \neg P$.

Proposition 1.11. *If n^2 is odd for some $n \in \mathbb{Z}$, then n is odd.*

Proof. We prove the contrapositive. If n is even, then by Theorem 1.10 the number n^2 is even. Thus, if n^2 is not even (i.e. is odd), then n cannot be even, hence n is odd. □

Proof by contradiction. Assume the desired statement is false and derive an impossibility (often a violation of a minimality or parity condition).

Theorem 1.12 ($\sqrt{2}$ is irrational)**.** *There do not exist integers $p, q \in \mathbb{Z}$ with $q \neq 0$ such that $\sqrt{2} = p/q$.*

Proof. Assume for contradiction that $\sqrt{2} = p/q$ with $p, q \in \mathbb{Z}$, $q > 0$, and the fraction is in lowest terms (i.e. $\gcd(p, q) = 1$). Then $2 = p^2/q^2$, so

$$p^2 = 2q^2.$$

Hence p^2 is even, so p is even (by Theorem 1.11 contrapositive reasoning). Write $p = 2k$. Then

$$(2k)^2 = 2q^2 \quad \Rightarrow \quad 4k^2 = 2q^2 \quad \Rightarrow \quad q^2 = 2k^2,$$

so q^2 is even and therefore q is even. Thus both p and q are even, contradicting $\gcd(p, q) = 1$. This contradiction shows $\sqrt{2}$ is irrational. □

Remark 1.13 (Counterexamples are proofs). To disprove a universal statement (of the form "for all x ..."), it suffices to give *one* explicit counterexample. This is exactly Section 1.1.2 in action.

1.1.4 Mathematical Induction (Weak and Strong)

Definition 1.14 (Principle of mathematical induction)**.** Let $P(n)$ be a statement for $n \in \mathbb{N}$. If

1. (Base case) $P(1)$ is true, and
2. (Inductive step) for every $n \in \mathbb{N}$, $P(n) \Rightarrow P(n+1)$,

then $P(n)$ is true for all $n \in \mathbb{N}$.

Theorem 1.15 (Sum of the first n integers)**.** *For all* $n \in \mathbb{N}$*,*

$$\sum_{k=1}^{n} k = \frac{n(n+1)}{2}.$$

Proof. Let $P(n)$ be the claim. Base case $n = 1$: $\sum_{k=1}^{1} k = 1 = \frac{1 \cdot 2}{2}$.

Inductive step: Assume $P(n)$ holds. Then

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

Thus $P(n+1)$ holds. By induction, the formula holds for all $n \in \mathbb{N}$. □

Definition 1.16 (Strong induction)**.** Let $P(n)$ be a statement for $n \in \mathbb{N}$. If

1. $P(1)$ is true, and
2. for every $n \geq 1$, the implication

$$\big(P(1) \wedge P(2) \wedge \cdots \wedge P(n)\big) \Rightarrow P(n+1)$$

holds,

then $P(n)$ holds for all $n \in \mathbb{N}$.

Remark 1.17. Weak and strong induction are logically equivalent, but strong induction often matches the structure of proofs about factorizations, recurrences, and constructions (which appear frequently in later examples and counterexamples in analysis).

1.1.5 Basic Set Operations and Identities

Definition 1.18 (Basic operations)**.** Let $A, B \subseteq U$.

- Union: $A \cup B = \{x \in U \ : \ x \in A \text{ or } x \in B\}$.
- Intersection: $A \cap B = \{x \in U \ : \ x \in A \text{ and } x \in B\}$.
- Difference: $A \setminus B = \{x \in U \ : \ x \in A \text{ and } x \notin B\}$.
- Complement: $A^c = U \setminus A$.
- Cartesian product: $A \times B = \{(a, b) \ : \ a \in A, \ b \in B\}$.

Theorem 1.19 (De Morgan's laws)**.** *For $A, B \subseteq U$,*

$$(A \cup B)^c = A^c \cap B^c, \qquad (A \cap B)^c = A^c \cup B^c.$$

Proof. We prove the first identity by element chasing; the second is analogous.

Let $x \in U$. Then

$$x \in (A \cup B)^c \iff x \notin A \cup B \iff \big(x \notin A \text{ and } x \notin B\big) \iff \big(x \in A^c \text{ and } x \in B^c\big) \iff x \in A^c \cap B^c.$$

Since this holds for all $x \in U$, the sets are equal. $\square$

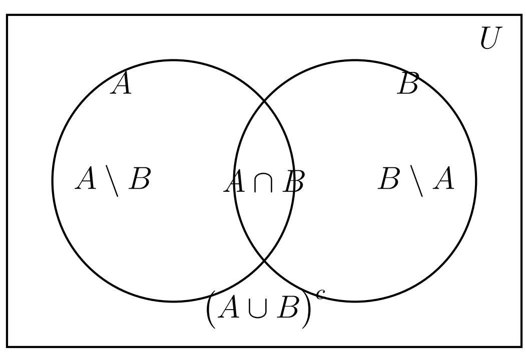


Figure 1: Regions for A, B, and their basic combinations inside a universe U.

Example 1.20 (A Venn diagram intuition)**.**

Proposition 1.21 (Set-builder proof template). *To prove $X = Y$, it suffices to show $X \subseteq Y$ and $Y \subseteq X$, i.e. for every x, $x \in X \Rightarrow x \in Y$ and $x \in Y \Rightarrow x \in X$.*

Exercise 1.22. Prove the distributive laws:

$$A \cap (B \cup C) = (A \cap B) \cup (A \cap C), \qquad A \cup (B \cap C) = (A \cup B) \cap (A \cup C).$$

Exercise 1.23. Let $A, B \subseteq U$. Prove:

$$A \setminus B = A \cap B^c, \qquad A \triangle B := (A \setminus B) \cup (B \setminus A)$$

satisfies $A \triangle B = (A \cup B) \setminus (A \cap B)$.

1.2 Relations, Functions, and Cardinality

1.2.1 Relations and Equivalence Relations

Definition 1.24 (Binary relation). Let A be a set. A *(binary) relation* on A is a subset $R \subseteq A \times A$. We write $a\,R\,b$ to mean $(a, b) \in R$.

Definition 1.25 (Equivalence relation). A relation $\sim$ on A is an *equivalence relation* if it is:

- Reflexive: $a \sim a$ for all $a \in A$.
- Symmetric: $a \sim b \Rightarrow b \sim a$.
- Transitive: $a \sim b$ and $b \sim c \Rightarrow a \sim c$.

Definition 1.26 (Equivalence class). Given $a \in A$, the *equivalence class* of a is

$$[a] := \{x \in A \;:\; x \sim a\}.$$

Theorem 1.27 (Equivalence relations and partitions). *If $\sim$ is an equivalence relation on A, then the set of equivalence classes*

$$A/{\sim} := \{[a] \;:\; a \in A\}$$

is a partition of A: every element of A lies in exactly one class, and distinct classes are disjoint.

Conversely, every partition of A defines an equivalence relation by declaring $x \sim y$ iff x and y lie in the same part.

Proof. Assume $\sim$ is an equivalence relation.

Covering: For any $a \in A$, reflexivity gives $a \sim a$, hence $a \in [a]$. So every a lies in some class.

Disjointness: Suppose $[a] \cap [b] \neq \varnothing$. Choose $x \in [a] \cap [b]$. Then $x \sim a$ and $x \sim b$. By symmetry, $a \sim x$, and by transitivity, $a \sim b$. Now for any $y \in [a]$, we have $y \sim a \sim b$, so $y \in [b]$. Thus $[a] \subseteq [b]$. By symmetry of the argument, $[b] \subseteq [a]$. Hence $[a] = [b]$. So different classes cannot overlap.

Conversely, given a partition $\mathcal{P}$ of A, define $x \sim y$ iff x, y are in the same part of $\mathcal{P}$. Reflexive/symmetric/transitive follow immediately from the fact that parts are disjoint and cover A. □

Example 1.28 (Congruence modulo n)**.** Fix $n \in \mathbb{N}$. Define $a \sim b$ on $\mathbb{Z}$ by $a \sim b$ iff n divides $(a - b)$. This is an equivalence relation, and its equivalence classes are the residue classes modulo n.

1.2.2 Order Relations and Partially Ordered Sets

Definition 1.29 (Partial order)**.** A relation $\preceq$ on a set A is a *partial order* if it is:

- Reflexive: $a \preceq a$,
- Antisymmetric: $a \preceq b$ and $b \preceq a$ imply $a = b$,
- Transitive: $a \preceq b$ and $b \preceq c$ imply $a \preceq c$.

Then $(A, \preceq)$ is called a *partially ordered set* (poset).

Example 1.30 (Subset order)**.** For any set U, the relation $\subseteq$ is a partial order on $\mathcal{P}(U)$ (the power set of U).

Example 1.31 (Divisibility order)**.** On $\mathbb{N}$, define $a \preceq b$ iff a divides b. This is a partial order.

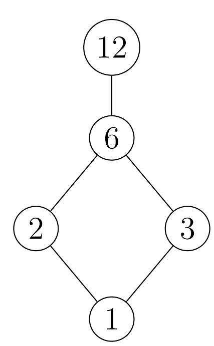


Figure 2: A Hasse diagram for divisibility on $\{1, 2, 3, 6, 12\}$.

Remark 1.32. Later, when we discuss sup and inf, it is useful to remember that "being an upper bound" is an order-theoretic notion: it depends only on the order structure, not on addition/multiplication.

1.2.3 Functions: Definitions, Images, Preimages

Definition 1.33 (Function)**.** A *function* f from a set A to a set B (written $f : A \to B$) assigns to each $a \in A$ a unique element $f(a) \in B$. The set A is the *domain*, B the *codomain*, and $f(A) = \{f(a) \,:\, a \in A\}$ the *image* (or range).

Definition 1.34 (Image and preimage)**.** Let $f : A \to B$ and let $E \subseteq A$, $S \subseteq B$.

$$f(E) \coloneqq \{f(x) \in B \,:\, x \in E\}, \qquad f^{-1}(S) \coloneqq \{x \in A \,:\, f(x) \in S\}.$$

Proposition 1.35 (Basic identities for preimages)**.** *For any $S, T \subseteq B$,*

$$f^{-1}(S \cup T) = f^{-1}(S) \cup f^{-1}(T), \qquad f^{-1}(S \cap T) = f^{-1}(S) \cap f^{-1}(T),$$

and (complements taken in B and A respectively)

$$f^{-1}(S^c) = \big(f^{-1}(S)\big)^c.$$

Proof. We show the first identity; the others follow similarly by chasing membership.

For $x \in A$,

$$\begin{aligned} x \in f^{-1}(S \cup T) &\iff f(x) \in S \cup T \\ &\iff \big(f(x) \in S \text{ or } f(x) \in T\big) \\ &\iff \big(x \in f^{-1}(S) \text{ or } x \in f^{-1}(T)\big) \\ &\iff x \in f^{-1}(S) \cup f^{-1}(T). \end{aligned}$$

□

Proposition 1.36 (Image of unions and intersections)**.** *For $E_1, E_2 \subseteq A$,*

$$f(E_1 \cup E_2) = f(E_1) \cup f(E_2), \qquad f(E_1 \cap E_2) \subseteq f(E_1) \cap f(E_2).$$

If f is injective (defined below), then equality holds in the intersection statement.

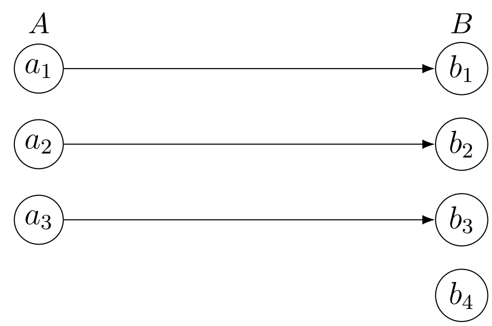
Proof. For unions: $y \in f(E_1 \cup E_2)$ iff $y = f(x)$ for some $x \in E_1 \cup E_2$, i.e. $x \in E_1$ or $x \in E_2$, which is equivalent to $y \in f(E_1) \cup f(E_2)$.

For intersections: if $y \in f(E_1 \cap E_2)$ then $y = f(x)$ with $x \in E_1 \cap E_2$, hence $y \in f(E_1)$ and $y \in f(E_2)$, so $y \in f(E_1) \cap f(E_2)$. This gives $\subseteq$. If f is injective and $y \in f(E_1) \cap f(E_2)$, then $y = f(x_1) = f(x_2)$ with $x_1 \in E_1$ and $x_2 \in E_2$. Injectivity forces $x_1 = x_2 \in E_1 \cap E_2$, hence $y \in f(E_1 \cap E_2)$. $\square$

1.2.4 Injective, Surjective, Bijective

Definition 1.37 (Injective / surjective / bijective)**.** Let $f : A \to B$.

- f is *injective* (one-to-one) if $f(a_1) = f(a_2) \Rightarrow a_1 = a_2$.
- f is *surjective* (onto) if $\forall b \in B\ \exists a \in A$ such that $f(a) = b$.
- f is *bijective* if it is both injective and surjective.



Injective, not surjective (since b_4 has no preimage).

Figure 3: A mapping diagram illustrating injectivity and failure of surjectivity.

Theorem 1.38 (Inverses)**.** *A function $f : A \to B$ is bijective if and only if there exists a function $g : B \to A$ such that*

$$g \circ f = \mathrm{id}_A \quad \textit{and} \quad f \circ g = \mathrm{id}_B,$$

where $\mathrm{id}_A(a) = a$ and $\mathrm{id}_B(b) = b$ are identity maps. In this case g is unique and is denoted f^{-1}.

Proof. If such a g exists, then f is injective: if $f(a_1) = f(a_2)$, applying g gives

$$a_1 = (g \circ f)(a_1) = (g \circ f)(a_2) = a_2.$$

Also f is surjective: for any $b \in B$, take $a = g(b)$; then $(f \circ g)(b) = b$ gives $f(a) = b$.

Conversely, if f is bijective, define $g(b)$ to be the unique $a \in A$ such that $f(a) = b$. Then g is well-defined and satisfies the identities. Uniqueness follows because any g with $g \circ f = \mathrm{id}_A$ must send $b = f(a)$ to a. $\square$

$$\begin{tikzcd} A \arrow[r, "f"] \arrow[dr, "h"'] & B \arrow[d, "g"] \\ & C \end{tikzcd}$$

Figure 4: Composition: $h = g \circ f$. Commutative diagrams help prevent domain/codomain mistakes.

1.2.5 Cardinality: Countable vs. Uncountable

Definition 1.39 (Equipotence and cardinality comparison)**.** Two sets A and B are *equipotent* (same cardinality), written $A \sim B$, if there exists a bijection $f : A \to B$. We write $|A| \leq |B|$ if there exists an injection $A \to B$.

Definition 1.40 (Countable sets)**.** A set A is *countable* if it is finite or if there exists a bijection $A \sim \mathbb{N}$. A set that is not countable is called *uncountable*.

Example 1.41. The sets $\mathbb{N}$, $\mathbb{Z}$, and $\mathbb{Q}$ are countable; the set $\mathbb{R}$ is uncountable (proved below).

Theorem 1.42 ($\mathbb{N} \times \mathbb{N}$ is countable)**.** *The Cartesian product* $\mathbb{N} \times \mathbb{N}$ *is countable.*

Proof. List pairs by increasing sum $m + n$ (a "diagonal" enumeration):

$$(1,1),\ (1,2),(2,1),\ (1,3),(2,2),(3,1),\ \ldots$$

This produces a sequence that hits every pair (m, n) after finitely many steps (namely once the diagonal $m+n$ is reached). Thus there is a surjection $\mathbb{N} \to \mathbb{N} \times \mathbb{N}$. Since $\mathbb{N} \times \mathbb{N}$ is infinite, one can refine this to a bijection (or equivalently, show $\mathbb{N} \times \mathbb{N}$ injects into $\mathbb{N}$ via the Cantor pairing function in Section 1.2.5). $\square$

$$\pi(m, n) = \frac{(m+n)(m+n+1)}{2} + n,$$

which defines an injection $\pi : \mathbb{N}_0 \times \mathbb{N}_0 \to \mathbb{N}_0$ (in fact a bijection).

Theorem 1.43 (Countable unions of countable sets)**.** *If* $\{A_n\}_{n \in \mathbb{N}}$ *is a sequence of countable sets, then* $\bigcup_{n \in \mathbb{N}} A_n$ *is countable.*

Proof. For each n, since A_n is countable, there exists a surjection $f_n : \mathbb{N} \to A_n$. Define $F : \mathbb{N} \times \mathbb{N} \to \bigcup_{n\in\mathbb{N}} A_n$ by $F(n,k) = f_n(k)$. This is surjective by construction. By Theorem 1.42, the domain $\mathbb{N} \times \mathbb{N}$ is countable, so its image under a surjection is also countable. Hence the union is countable. $\square$

Theorem 1.44 ($\mathbb{Q}$ is countable)**.** *The set of rational numbers $\mathbb{Q}$ is countable.*

Proof. For each $q \in \mathbb{N}$, define

$$S_q := \left\{ \frac{p}{q} \in \mathbb{Q} \ : \ p \in \mathbb{Z} \right\}.$$

The map $\mathbb{Z} \to S_q$ given by $p \mapsto p/q$ is surjective, so each S_q is countable. Moreover,

$$\mathbb{Q} = \bigcup_{q\in\mathbb{N}} S_q,$$

because every rational has some (not necessarily reduced) representation with positive denominator. By Theorem 1.43, $\mathbb{Q}$ is countable. $\square$

Remark 1.45 (Why duplicates do not matter). In the union $\bigcup_{q\in\mathbb{N}} S_q$, the same rational number may appear multiple times (e.g. $1/2 = 2/4$). This does not affect countability: a set remains countable even if it is listed with repetitions; removing duplicates only makes the set smaller (and any subset of a countable set is countable).

1.2.6 Cantor's Diagonal Argument

Theorem 1.46 (Uncountability of $(0,1)$)**.** *The interval $(0,1) \subseteq \mathbb{R}$ is uncountable. Consequently, $\mathbb{R}$ is uncountable.*

Proof. Assume for contradiction that $(0,1)$ is countable. Then we can list its elements as a sequence

$$x_1,\ x_2,\ x_3,\ \dots$$

Write each x_n in decimal form:

$$x_n = 0.a_{n1}a_{n2}a_{n3}\dots,$$

where each digit $a_{nk} \in \{0,1,2,\dots,9\}$. (If a number has two decimal expansions, choose the one not ending in repeating 9's; this can always be done.)

Now define a new number $y \in (0,1)$ by specifying its digits:

$$y = 0.b_1 b_2 b_3 \ldots, \qquad b_n := \begin{cases} 1, & \text{if } a_{nn} \neq 1, \\ 2, & \text{if } a_{nn} = 1. \end{cases}$$

Then each $b_n \in \{1, 2\}$, so y does not end with repeating 9's and hence has a unique decimal expansion.

By construction, y differs from x_n in the n-th decimal place (since $b_n \neq a_{nn}$). Therefore $y \neq x_n$ for every $n \in \mathbb{N}$. This contradicts the assumption that $\{x_n\}$ lists all elements of $(0,1)$. Hence $(0,1)$ is uncountable. $\square$

Corollary 1.47. *The set of irrational numbers $\mathbb{R} \setminus \mathbb{Q}$ is uncountable.*

Proof. If $\mathbb{R} \setminus \mathbb{Q}$ were countable, then $\mathbb{R} = (\mathbb{R} \setminus \mathbb{Q}) \cup \mathbb{Q}$ would be a union of two countable sets, hence countable by Theorem 1.43, contradicting Theorem 1.46. $\square$

Remark 1.48 (Cardinality is preserved under bijection). A standard technique in analysis is to compare sets by building explicit bijections. For instance, the map $T : (0,1) \to (a,b)$ given by

$$T(x) = a + (b-a)x$$

is a bijection; hence every open interval has the same cardinality as $(0,1)$.

Also, the map $\Phi : (0,1) \to \mathbb{R}$ defined by

$$\Phi(x) = \tan\bigl(\pi(x - \tfrac{1}{2})\bigr)$$

is a bijection, showing $(0,1) \sim \mathbb{R}$. (This is a good example of how functions encode "same size" even for infinite sets.)

Theorem 1.49 (Cantor's theorem)**.** *For any set A, the power set $\mathcal{P}(A)$ has strictly larger cardinality than A:*

$$|A| < |\mathcal{P}(A)|.$$

Proof. First, there is always an injection $A \to \mathcal{P}(A)$ given by $a \mapsto \{a\}$, so $|A| \leq |\mathcal{P}(A)|$.

To show strict inequality, suppose (for contradiction) that there is a surjection $f : A \to \mathcal{P}(A)$. Define the set

$$D := \{a \in A \ : \ a \notin f(a)\}.$$

Since f is surjective, there exists $d \in A$ with $f(d) = D$. Now ask whether $d \in D$.

- If $d \in D$, then by definition of D we must have $d \notin f(d) = D$, contradiction.

- If $d \notin D$, then by definition of D we must have $d \in f(d) = D$, contradiction.

Both cases contradict, so no surjection $A \to \mathcal{P}(A)$ exists. Hence $|A| < |\mathcal{P}(A)|$. $\square$

Exercise 1.50. Show that the set of all finite subsets of $\mathbb{N}$ is countable.

Exercise 1.51. Let A be countable and $B \subseteq A$. Prove that B is countable.

Exercise 1.52. Give an explicit bijection between $\mathbb{N}$ and $\mathbb{Z}$. (Hint: interleave nonnegative and negative integers.)

Exercise 1.53. Prove that the set of algebraic numbers is countable. Conclude that there exist transcendental real numbers.

2 The Real Number System

> **Key Idea.** The real numbers $\mathbb{R}$ are characterized (up to a unique order-preserving field isomorphism) as the *complete ordered field*. In practice, we do not rebuild $\mathbb{R}$ from scratch every time; instead we work from axioms and prove the familiar properties (algebra, inequalities, bounds, and approximation) *rigorously* from those axioms.

2.1 Axiomatic Construction of $\mathbb{R}$

2.1.1 Field Axioms

Definition 2.1 (Field)**.** A set F equipped with two operations $+$ and $\cdot$ is a *field* if:

(F1) **Addition:** $(F, +)$ is an abelian group. Concretely:

- (a) $a + b \in F$ for all $a, b \in F$ (closure).
- (b) $(a + b) + c = a + (b + c)$ for all $a, b, c \in F$ (associativity).
- (c) $a + b = b + a$ for all $a, b \in F$ (commutativity).
- (d) There exists $0 \in F$ such that $a + 0 = a$ for all $a \in F$ (additive identity).
- (e) For each $a \in F$, there exists $-a \in F$ such that $a + (-a) = 0$ (additive inverse).

(F2) **Multiplication:** $(F \setminus \{0\}, \cdot)$ is an abelian group:

- (a) $ab \in F$ for all $a, b \in F$ (closure).
- (b) $(ab)c = a(bc)$ for all $a, b, c \in F$ (associativity).
- (c) $ab = ba$ for all $a, b \in F$ (commutativity).
- (d) There exists $1 \in F$ with $1 \neq 0$ such that $a \cdot 1 = a$ for all $a \in F$ (multiplicative identity).
- (e) For each $a \in F \setminus \{0\}$, there exists $a^{-1} \in F$ such that $aa^{-1} = 1$ (multiplicative inverse).

(F3) **Distributivity:** $a(b + c) = ab + ac$ for all $a, b, c \in F$.

Remark 2.2. When $F = \mathbb{R}$, the operations are the usual addition and multiplication. The point of axioms is that every algebraic manipulation must ultimately be justified by these rules (and the consequences derived from them).

Proposition 2.3 (Basic algebraic consequences in a field)**.** *Let F be a field and $a, b, c \in F$. Then:*

*(a) (**Additive cancellation**) If $a + c = b + c$, then $a = b$.*

*(b) (**Multiplicative cancellation**) If $c \neq 0$ and $ac = bc$, then $a = b$.*

(c) $a \cdot 0 = 0$.

*(d) (**Zero-product property**) If $ab = 0$, then $a = 0$ or $b = 0$.*

Proof. (a) Add $(-c)$ to both sides:

$$a + c = b + c \;\Rightarrow\; (a + c) + (-c) = (b + c) + (-c) \;\Rightarrow\; a = b.$$

(b) Multiply both sides by c^{-1}:

$$ac = bc \;\Rightarrow\; ac\,c^{-1} = bc\,c^{-1} \;\Rightarrow\; a = b.$$

(c) Using distributivity and additive cancellation:

$$a \cdot 0 = a(0 + 0) = a \cdot 0 + a \cdot 0 \;\Rightarrow\; a \cdot 0 = 0.$$

(d) If $ab = 0$ and $a \neq 0$, multiply by a^{-1}:

$$ab = 0 \;\Rightarrow\; a^{-1}ab = a^{-1}0 \;\Rightarrow\; b = 0.$$

So either $a = 0$ or $b = 0$.

$\square$

2.1.2 Order Axioms

Definition 2.4 (Ordered field)**.** A field F is an *ordered field* if it is equipped with a relation $<$ such that:

(O1) (**Trichotomy**) For any $a, b \in F$, exactly one of $a < b$, $a = b$, $a > b$ holds.

(O2) (**Transitivity**) If $a < b$ and $b < c$, then $a < c$.

(O3) (**Addition preserves order**) If $a < b$, then $a + c < b + c$ for all $c \in F$.

(O4) (**Multiplication by positives preserves order**) If $a < b$ and $c > 0$, then $ac < bc$.

We also write $a \leq b$ to mean $a < b$ or $a = b$.

Proposition 2.5 (Standard inequality rules)**.** *Let F be an ordered field and $a, b, c, d \in F$.*

(a) If $a < b$ then $-b < -a$.

(b) If $0 < a$ and $0 < b$, then $0 < ab$.

(c) If $a < b$ and $c < d$, then $a + c < b + d$.

(d) If $a < b$ and $0 < c$, then $\frac{a}{c} < \frac{b}{c}$.

(e) If $a < b$ and $c < 0$, then $\frac{a}{c} > \frac{b}{c}$.

Proof. (a) From $a < b$, add $-a - b$ to both sides:

$$a + (-a - b) < b + (-a - b) \;\Rightarrow\; -b < -a.$$

(b) Since $0 < a$, multiplying the inequality $0 < b$ by $a > 0$ gives $0 \cdot a < ba$, i.e. $0 < ab$.

(c) From $a < b$ add c to get $a + c < b + c$. From $c < d$ add b to get $b + c < b + d$. Combine by transitivity:

$$a + c < b + c < b + d.$$

(d) If $0 < c$, then multiplying $a < b$ by $c^{-1} > 0$ preserves the inequality:

$$a < b \;\Rightarrow\; ac^{-1} < bc^{-1} \;\Rightarrow\; \frac{a}{c} < \frac{b}{c}.$$

(e) If $c < 0$, then $-c > 0$. By (d),

$$\frac{a}{-c} < \frac{b}{-c}.$$

Multiplying by -1 reverses order by (a), hence

$$-\frac{a}{-c} > -\frac{b}{-c} \;\Rightarrow\; \frac{a}{c} > \frac{b}{c}.$$

$\square$

2.1.3 Absolute Value as an Order-Theoretic Notion

Definition 2.6 (Absolute value)**.** In an ordered field F, define $|x|$ by

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

Proposition 2.7 (Basic properties of absolute value)**.** *For all x, y in an ordered field:*

(a) $|x| \geq 0$, and $|x| = 0 \iff x = 0$.

(b) $|xy| = |x|\,|y|$.

(c) $|x+y| \le |x| + |y|$ *(triangle inequality).*

(d) $\big||x| - |y|\big| \le |x-y|$ *(reverse triangle inequality).*

Proof. (a) By definition, $|x|$ equals either x (if $x \ge 0$) or $-x$ (if $x < 0$), both of which are ≥ 0. If $|x| = 0$, then either $x = 0$ or $-x = 0$, hence $x = 0$. Conversely, if $x = 0$ then $|x| = 0$.

(b) Consider cases based on the signs of x and y. For example, if $x \ge 0$ and $y \ge 0$, then $|xy| = xy = |x||y|$. If $x < 0$ and $y \ge 0$, then $xy \le 0$ so $|xy| = -(xy) = (-x)y = |x||y|$, etc. All cases give the same identity.

(c) We prove $|x+y| \le |x| + |y|$ by comparing squares (which avoids a long sign split and stays fully rigorous): since $|t| \ge 0$ and $|t|^2 = t^2$, it suffices to show

$$(x+y)^2 \le \big(|x| + |y|\big)^2.$$

Expanding both sides gives

$$x^2 + 2xy + y^2 \le x^2 + 2|x||y| + y^2,$$

which reduces to $xy \le |x||y|$. This inequality holds because:

$$-|x||y| \le xy \le |x||y|,$$

a consequence of $|x||y| \ge 0$ and the definition of $|\cdot|$ applied to xy. Therefore the triangle inequality follows.

(d) Apply the triangle inequality to $x = (x-y) + y$:

$$|x| = |(x-y) + y| \le |x-y| + |y| \;\Rightarrow\; |x| - |y| \le |x-y|.$$

Swap x and y to also get $|y| - |x| \le |x-y|$. Combining,

$$-|x-y| \le |x| - |y| \le |x-y|,$$

which is equivalent to $\big||x| - |y|\big| \le |x-y|$.

□

2.1.4 Completeness Axiom

Definition 2.8 (Upper bounds and least upper bounds)**.** Let $S \subseteq \mathbb{R}$.

- A number $u \in \mathbb{R}$ is an *upper bound* of S if $x \le u$ for all $x \in S$.

- If S has an upper bound, we say S is *bounded above*.
- A number $s \in \mathbb{R}$ is the *least upper bound* (or *supremum*) of S if:
  (i) s is an upper bound of S, and
  (ii) if u is any upper bound of S, then $s \leq u$.

  In this case we write $s = \sup S$.

Theorem 2.9 (Completeness axiom / Least upper bound property)**.** *Every nonempty subset $S \subseteq \mathbb{R}$ that is bounded above has a supremum in $\mathbb{R}$.*

Remark 2.10. This axiom is the key feature that distinguishes $\mathbb{R}$ from $\mathbb{Q}$. Many deep theorems in analysis (convergence, continuity, compactness) ultimately trace back to completeness.

2.1.5 Why $\mathbb{Q}$ Is Not Complete

Example 2.11 (A bounded set in $\mathbb{Q}$ with no supremum in $\mathbb{Q}$)**.** Consider the set

$$S := \{q \in \mathbb{Q} \ : \ q > 0 \text{ and } q^2 < 2\}.$$

- S is nonempty (e.g. $1 \in S$) and bounded above in $\mathbb{Q}$ (e.g. 2 is an upper bound since $q^2 < 2$ implies $q < 2$ for $q > 0$).
- In $\mathbb{R}$, $\sup S = \sqrt{2}$.
- But $\sqrt{2} \notin \mathbb{Q}$, so $\sup S$ does *not* exist in $\mathbb{Q}$.

A more detailed argument shows: if $r \in \mathbb{Q}$ is an upper bound of S, then $r^2 > 2$; if $r^2 > 2$ one can construct a smaller rational upper bound, so no least rational upper bound exists.

Remark 2.12. Completeness is not merely about "having more numbers." It is the precise statement that *bounded monotone approximation processes do not get stuck outside the number system.*

Exercise 2.13. Let $S \subseteq \mathbb{R}$ be nonempty and bounded above, and let $s = \sup S$. Prove that s is unique.

2.2 Bounds and Supremum

2.2.1 Upper/Lower Bounds, Max/Min

Definition 2.14 (Lower bound, infimum, boundedness)**.** Let $S \subseteq \mathbb{R}$.

- A number $\ell \in \mathbb{R}$ is a *lower bound* of S if $\ell \leq x$ for all $x \in S$.
- If S has a lower bound, we say S is *bounded below.*
- If S is bounded above and below, we say S is *bounded.*
- The *greatest lower bound* (or *infimum*) is denoted $\inf S$ and defined analogously to $\sup S$.

Definition 2.15 (Maximum and minimum)**.** Let $S \subseteq \mathbb{R}$.

- $m \in S$ is a *maximum* of S if $x \leq m$ for all $x \in S$. We write $m = \max S$.
- $n \in S$ is a *minimum* of S if $n \leq x$ for all $x \in S$. We write $n = \min S$.

Remark 2.16. A maximum (or minimum) must be an *element* of the set. A supremum (or infimum) need not belong to the set. For example, $\sup(0,1) = 1$ but $(0,1)$ has no maximum.

2.2.2 Characterization of the Supremum

Theorem 2.17 (Approximation property of sup)**.** *Let $S \subseteq \mathbb{R}$ be nonempty and bounded above, and let $s = \sup S$. Then:*

(a) $x \leq s$ for all $x \in S$.

(b) For every $\varepsilon > 0$, there exists $x_\varepsilon \in S$ such that

$$s - \varepsilon < x_\varepsilon \leq s.$$

Equivalently, $s - \varepsilon$ is not *an upper bound of S for any $\varepsilon > 0$.*

Conversely, if $s \in \mathbb{R}$ satisfies (a) and (b), then $s = \sup S$.

Proof. Assume $s = \sup S$. Part (a) is the definition of upper bound.

For (b), fix $\varepsilon > 0$. If no $x \in S$ satisfies $x > s - \varepsilon$, then every $x \in S$ has $x \leq s - \varepsilon$, so $s - \varepsilon$ is an upper bound of S. This contradicts that s is the *least* upper bound (since then $s \leq s - \varepsilon$ would follow). Hence there exists $x_\varepsilon \in S$ with $x_\varepsilon > s - \varepsilon$. Also $x_\varepsilon \leq s$ by (a).

Conversely, suppose (a) and (b) hold. Then s is an upper bound. If u is any upper bound with $u < s$, choose $\varepsilon = s - u > 0$. By (b), there exists $x_\varepsilon \in S$ with $x_\varepsilon > s - \varepsilon = u$, contradicting that u is an upper bound. Therefore no upper bound is $< s$, so s is the least upper bound. $\square$

Corollary 2.18 (When sup is a maximum)**.** *If $S \subseteq \mathbb{R}$ is nonempty and bounded above, then $\sup S \in S$ if and only if S has a maximum, and in that case $\sup S = \max S$.*

Proof. If $\sup S \in S$, then it is an upper bound and belongs to S, hence it is a maximum by definition. Conversely, if $m = \max S$, then m is an upper bound and no smaller number can be an upper bound, so $m = \sup S$. $\square$

2.2.3 Infimum via Supremum

Proposition 2.19 (Infimum and negation)**.** *If $S \subseteq \mathbb{R}$ is nonempty and bounded below, then*

$$\inf S = -\sup(-S), \qquad \text{where } -S := \{-x \,:\, x \in S\}.$$

Proof. Let $T = -S$. A number ℓ is a lower bound of S iff $-\ell$ is an upper bound of T. If $s = \sup T$, then $-s$ is the greatest lower bound of S, i.e. $\inf S = -s$. $\square$

2.2.4 Algebra of Supremum and Infimum

Proposition 2.20 (Shifts and scalings)**.** *Let $S \subseteq \mathbb{R}$ be nonempty and bounded above, and let $c \in \mathbb{R}$.*

(a) $\sup(S + c) = \sup S + c$*, where* $S + c := \{x + c \,:\, x \in S\}$.

(b) If $\lambda > 0$*, then* $\sup(\lambda S) = \lambda \sup S$*, where* $\lambda S := \{\lambda x \,:\, x \in S\}$.

(c) If $\lambda < 0$*, then* $\sup(\lambda S) = \lambda \inf S$ *(assuming S is bounded below so that* $\inf S$ *exists).*

Analogous statements hold for $\inf$*.*

Proof. We prove (a) and (b); (c) follows from (b) and Theorem 2.19.

(a) Let $s = \sup S$. For all $x \in S$, we have $x \leq s$, so $x + c \leq s + c$; hence $s + c$ is an upper bound of $S + c$. If u is any upper bound of $S + c$, then $u - c$ is an upper bound of S, so $s \leq u - c$, i.e. $s + c \leq u$. Thus $s + c$ is the least upper bound, so $\sup(S + c) = s + c$.

(b) Let $\lambda > 0$ and $s = \sup S$. If $x \leq s$ for all $x \in S$, then multiplying by λ preserves order:

$$\lambda x \leq \lambda s \quad \forall x \in S,$$

so λs is an upper bound of λS. If u is any upper bound of λS, then u/λ is an upper bound of S (since $\lambda x \leq u \Rightarrow x \leq u/\lambda$), so $s \leq u/\lambda$ and hence $\lambda s \leq u$. Therefore $\sup(\lambda S) = \lambda s$. $\square$

2.2.5 Examples

Example 2.21. Let $S = (0, 1)$. Then S is bounded above and below. One has

$$\sup S = 1, \qquad \inf S = 0,$$

but S has neither a maximum nor a minimum.

Example 2.22. Let $S = \left\{1 - \frac{1}{n} \ : \ n \in \mathbb{N}\right\}$. Then $S \subset (0, 1)$ and:

$$\sup S = 1 \quad (\text{not in } S), \qquad \inf S = 0 \quad (\text{not in } S).$$

Moreover, S has a minimum: indeed, $1 - \frac{1}{n}$ increases with n, so the smallest occurs at $n = 1$, giving $\min S = 0$? Be careful: at $n = 1$ we get $1 - 1 = 0$, so actually $0 \in S$ and $\min S = 0$ *does* exist, while $\inf S = 0$ as well. Thus $\min S = \inf S = 0$, but $\max S$ does not exist and $\sup S = 1$.

Remark 2.23. This example illustrates a common pattern: $\inf S$ can be attained while $\sup S$ is not (or vice versa). The supremum/infimum concepts are designed to handle both situations cleanly.

2.2.6 A Key Application: Existence of Square Roots

Theorem 2.24 (Existence and uniqueness of $\sqrt{a}$)**.** *For every real number $a > 0$, there exists a unique real number $s > 0$ such that $s^2 = a$.*

Proof. **Uniqueness.** If $s, t > 0$ and $s^2 = t^2$, then

$$0 = s^2 - t^2 = (s - t)(s + t).$$

Since $s + t > 0$, the zero-product property forces $s - t = 0$, hence $s = t$.

Existence. Define

$$S := \left\{x \in \mathbb{R} \ : \ x \geq 0 \text{ and } x^2 < a\right\}.$$

Then S is nonempty (since $0 \in S$) and bounded above (for instance, if $x \geq a + 1$, then $x^2 \geq (a+1)^2 > a$, so such x cannot lie in S; hence $a+1$ is an upper bound). By completeness, S has a supremum; let $s = \sup S$ with $s \geq 0$.

We claim $s^2 = a$. We show s^2 cannot be $< a$ nor $> a$.

Step 1: $s^2 \not< a$. Assume $s^2 < a$. Let

$$\delta := \min\left\{1, \ \frac{a - s^2}{2s + 2}\right\}.$$

Note $\delta > 0$ because $a - s^2 > 0$ and $2s + 2 > 0$. Then

$$\begin{aligned}
(s+\delta)^2 &= s^2 + 2s\delta + \delta^2 \\
&\le s^2 + 2s\delta + \delta \qquad (\text{since } 0 < \delta \le 1 \Rightarrow \delta^2 \le \delta) \\
&= s^2 + (2s+1)\delta \\
&\le s^2 + (2s+2)\delta \\
&\le s^2 + (a - s^2) = a.
\end{aligned}$$

So $(s+\delta)^2 \le a$, and in particular $(s+\delta)^2 < a$ (because equality would force $\delta = 0$ by the inequality chain, impossible). Hence $s + \delta \in S$, contradicting that s is an upper bound of S. Therefore $s^2 \not< a$.

Step 2: $s^2 \not> a$. Assume $s^2 > a$. Choose

$$\delta := \min\left\{1,\ \frac{s^2 - a}{2s}\right\}.$$

Here $s > 0$ (since $a > 0$ and S contains positive numbers; thus $\sup S > 0$), so $\delta > 0$ is well-defined. Then

$$\begin{aligned}
(s-\delta)^2 &= s^2 - 2s\delta + \delta^2 \\
&\ge s^2 - 2s\delta \qquad (\delta^2 \ge 0) \\
&\ge s^2 - (s^2 - a) = a.
\end{aligned}$$

Thus $(s-\delta)^2 \ge a$, so for every $x \in S$ we must have $x \le s - \delta$; otherwise, if $x > s - \delta$, then $x \ge 0$ and monotonicity of $t \mapsto t^2$ on $[0, \infty)$ gives $x^2 > (s-\delta)^2 \ge a$, contradicting $x \in S$. Hence $s - \delta$ is an upper bound of S, contradicting the *least* upper bound property of s (since $s - \delta < s$). Therefore $s^2 \not> a$.

Combining Steps 1 and 2, we conclude $s^2 = a$. This completes the proof. ☐

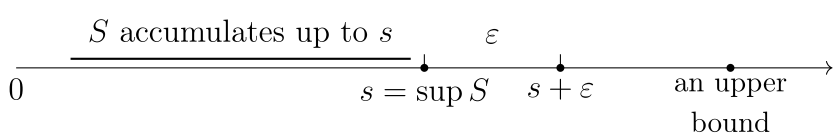


Figure 5: The supremum s is an upper bound that can be approximated from below by points of S.

Exercise 2.25. Let $S \subseteq \mathbb{R}$ be nonempty and bounded above. Show that if $s = \sup S$ and $s \notin S$, then for every $\varepsilon > 0$ there exists $x \in S$ with $s - \varepsilon < x < s$.

2.3 Archimedean Property and Density

2.3.1 The Archimedean Principle

Theorem 2.26 (Archimedean property)**.** *For every real number $x \in \mathbb{R}$, there exists $n \in \mathbb{N}$ such that $n > x$. Equivalently: for every $y > 0$, there exists $n \in \mathbb{N}$ such that $ny > 1$.*

Proof. Assume for contradiction that there exists $x \in \mathbb{R}$ such that $n \leq x$ for all $n \in \mathbb{N}$. Then $\mathbb{N}$ is nonempty and bounded above by x, so by completeness it has a supremum $s = \sup \mathbb{N}$.

By Theorem 2.17 applied to $S = \mathbb{N}$, with $\varepsilon = 1$, there exists $n_0 \in \mathbb{N}$ such that

$$s - 1 < n_0 \leq s.$$

Add 1 to obtain $s < n_0 + 1$. But $n_0 + 1 \in \mathbb{N}$, contradicting that s is an upper bound of $\mathbb{N}$. Therefore no such x exists, and the Archimedean property holds. □

Corollary 2.27. *For every $\varepsilon > 0$, there exists $n \in \mathbb{N}$ such that $\frac{1}{n} < \varepsilon$.*

Proof. Apply Theorem 2.26 to $x = \frac{1}{\varepsilon}$ to obtain $n > \frac{1}{\varepsilon}$, hence $\frac{1}{n} < \varepsilon$. □

2.3.2 Integer Part and the Floor Function

Theorem 2.28 (Existence of integer part)**.** *For every $x \in \mathbb{R}$, there exists a unique integer $m \in \mathbb{Z}$ such that*

$$m \leq x < m + 1.$$

This integer is denoted $\lfloor x \rfloor$ (the floor of x).

Proof. **Existence.** Consider the set

$$A := \{k \in \mathbb{Z} \,:\, k \leq x\}.$$

It is nonempty: by the Archimedean property applied to $-x$, there exists $N \in \mathbb{N}$ with $N > -x$, so $-N < x$, hence $-N \in A$. Also A is bounded above by x (every element is $\leq x$). By completeness, A has a supremum $a = \sup A$.

We claim that $a \in \mathbb{Z}$ and $a \leq x < a + 1$. First, since $a = \sup A$, by Theorem 2.17 with $\varepsilon = 1$ there exists $k \in A$ such that

$$a - 1 < k \leq a.$$

Because $k \in \mathbb{Z}$, the only integers in $(a-1, a]$ are at most one value; in particular, k is the greatest integer $\leq a$, hence $k \leq a < k+1$. But $k \leq a$ and a is an upper bound of A imply k is actually the largest element of A: if there were $\ell \in A$ with $\ell > k$, then $\ell \geq k+1 > a$, contradicting that a is an upper bound. Thus $k = \max A$ and therefore $a = k \in \mathbb{Z}$.

Now $a \in A$ gives $a \leq x$. If $x \geq a+1$, then $a+1 \leq x$ so $a+1 \in A$, contradicting maximality of a in A. Hence $x < a+1$.

Uniqueness. If $m \leq x < m+1$ and $n \leq x < n+1$ with $m, n \in \mathbb{Z}$, then $m \leq x < n+1$ implies $m < n+1$, hence $m \leq n$ (integers). Similarly $n \leq m$. Thus $m = n$. □

Remark 2.29. The floor function provides a precise way to say "choose an integer close to a real number." It is a standard tool in density proofs and in decimal expansions.

2.3.3 Density of $\mathbb{Q}$ in $\mathbb{R}$

Theorem 2.30 (Density of rational numbers)**.** *For any real numbers $a < b$, there exists a rational number $q \in \mathbb{Q}$ such that*

$$a < q < b.$$

Proof. Let $a < b$. By Theorem 2.27, choose $n \in \mathbb{N}$ such that $\frac{1}{n} < b - a$, i.e. $n(b-a) > 1$.

Consider the real number na. By Theorem 2.28, let $m = \lfloor na \rfloor \in \mathbb{Z}$ so that $m \leq na < m+1$. Define

$$q := \frac{m+1}{n} \in \mathbb{Q}.$$

Then $m+1 > na$, hence $q > a$. Also

$$q = \frac{m+1}{n} \leq \frac{na+1}{n} = a + \frac{1}{n} < a + (b-a) = b,$$

because $\frac{1}{n} < b - a$. Therefore $a < q < b$. □

Corollary 2.31. *Every nonempty open interval contains infinitely many rational numbers.*

Proof. Given $a < b$, apply Theorem 2.30 to the smaller interval (a, q) to find another rational strictly between, and iterate to get infinitely many distinct rationals. □

2.3.4 Density of Irrational Numbers

Theorem 2.32 (Density of irrational numbers)**.** *For any real numbers $a < b$, there exists an irrational number $\alpha \in \mathbb{R} \setminus \mathbb{Q}$ such that*

$$a < \alpha < b.$$

Proof. We use $\sqrt{2} \notin \mathbb{Q}$ (proved earlier in these notes). Since $a < b$, divide by $\sqrt{2} > 0$ to get

$$\frac{a}{\sqrt{2}} < \frac{b}{\sqrt{2}}.$$

By Theorem 2.30, choose $q \in \mathbb{Q}$ such that

$$\frac{a}{\sqrt{2}} < q < \frac{b}{\sqrt{2}}.$$

Set $\alpha = q\sqrt{2}$. Then $a < \alpha < b$. Moreover, if α were rational, then $\sqrt{2} = \alpha/q$ would be rational (since $q \neq 0$), a contradiction. Hence α is irrational. $\square$

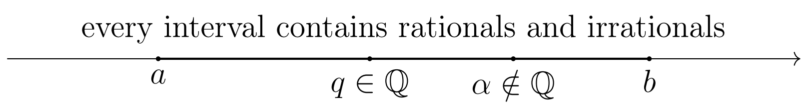


Figure 6: Both $\mathbb{Q}$ and $\mathbb{R} \setminus \mathbb{Q}$ are dense in $\mathbb{R}$.

2.3.5 Decimal Expansions

Theorem 2.33 (Decimal expansion for $x \in [0, 1)$)**.** *For every $x \in [0, 1)$, there exists a sequence of digits $(a_k)_{k\in\mathbb{N}}$ with $a_k \in \{0, 1, \dots, 9\}$ such that*

$$x = \sum_{k=1}^{\infty} \frac{a_k}{10^k}.$$

Moreover, one can choose the digits so that the expansion does not end in repeating 9's.

Proof. Define digits recursively. Set $r_0 = x$. For $k \geq 1$, define

$$a_k := \lfloor 10 r_{k-1} \rfloor \in \{0, 1, \dots, 9\}, \qquad r_k := 10 r_{k-1} - a_k \in [0, 1).$$

(The bounds follow because $a_k \leq 10 r_{k-1} < a_k + 1$ by the floor property.) Unwinding the recursion gives, for each $n \in \mathbb{N}$,

$$x = \sum_{k=1}^{n} \frac{a_k}{10^k} + \frac{r_n}{10^n}, \qquad \text{with } 0 \leq r_n < 1.$$

In particular,

$$0 \leq x - \sum_{k=1}^{n} \frac{a_k}{10^k} = \frac{r_n}{10^n} < \frac{1}{10^n}.$$

Hence the partial sums converge to x, and therefore

$$x = \lim_{n\to\infty} \sum_{k=1}^{n} \frac{a_k}{10^k} = \sum_{k=1}^{\infty} \frac{a_k}{10^k}.$$

If the resulting digits eventually become all 9, then x has an alternative terminating expansion (see Theorem 2.34). We adopt the convention of choosing the non-terminating-9 representation to ensure uniqueness. $\square$

Proposition 2.34 (Why $0.999\ldots = 1$)**.** *In* $\mathbb{R}$,

$$0.999\ldots = 1.$$

More generally, any terminating decimal has a second representation ending in repeating 9*'s.*

Proof. Let $s_n = \sum_{k=1}^{n} 9 \cdot 10^{-k}$. Then $s_n = 1 - 10^{-n}$ (a finite geometric sum), so $\lim_{n\to\infty} s_n = 1$. By definition, $0.999\ldots = \lim_{n\to\infty} s_n$, hence it equals 1. $\square$

Theorem 2.35 (Decimal expansion for general real numbers)**.** *For every real number* $x \in \mathbb{R}$, *there exist an integer* $m \in \mathbb{Z}$ *and digits* $(a_k)_{k\in\mathbb{N}}$ *such that*

$$x = m + \sum_{k=1}^{\infty} \frac{a_k}{10^k}.$$

If we require that the digits do not eventually become all 9*, then this representation is unique.*

Proof. Let $m = \lfloor x \rfloor \in \mathbb{Z}$ so that $m \le x < m + 1$. Then $y = x - m \in [0, 1)$. Apply Theorem 2.33 to y to obtain digits (a_k) with

$$y = \sum_{k=1}^{\infty} \frac{a_k}{10^k}.$$

Thus $x = m + y$ has the desired form. Uniqueness follows from the uniqueness of $\lfloor x \rfloor$ and the convention excluding trailing 9's. $\square$

Exercise 2.36. Show that if a decimal expansion is eventually periodic (i.e. repeats a fixed block of digits from some point onward), then the number is rational.

Exercise 2.37. Give an explicit construction of a rational number strictly between a and b without using the floor function, starting from the Archimedean property.

Exercise 2.38. Prove that for every $x \in \mathbb{R}$ and every $n \in \mathbb{N}$, there exists an integer m such that

$$|x - \frac{m}{10^n}| \le \frac{1}{10^n}.$$

Interpret this as “every real can be approximated by decimals with n digits.”

3 Sequences and Limits

Key Idea. A sequence is a function $x : \mathbb{N} \to \mathbb{R}$, written $(x_n)_{n\in\mathbb{N}}$. The central definition in this chapter is the ε–N definition of convergence. Once that definition is mastered, most theorems become systematic exercises in quantifiers, inequalities, and completeness.

3.1 Sequences of Real Numbers

3.1.1 Definition and Basic Examples

Definition 3.1 (Sequence)**.** A *sequence of real numbers* is a function $x : \mathbb{N} \to \mathbb{R}$. We write $x(n) = x_n$ and denote the sequence by $(x_n)_{n\in\mathbb{N}}$, or simply (x_n).

Example 3.2. (a) $x_n = \frac{1}{n}$ for $n \in \mathbb{N}$.

(b) $x_n = (-1)^n$ (alternating signs).

(c) $x_n = \sqrt{n}$ (unbounded, diverges to $+\infty$).

(d) $x_n = \left(1 + \frac{1}{n}\right)^n$ (approaches e, later proved carefully).

Definition 3.3 (Bounded sequence)**.** A sequence (x_n) is *bounded* if there exists $M > 0$ such that $|x_n| \le M$ for all $n \in \mathbb{N}$. Equivalently, the set $\{x_n : n \in \mathbb{N}\}$ is a bounded subset of $\mathbb{R}$.

3.1.2 Convergence and Divergence

Definition 3.4 (Limit of a sequence)**.** Let (x_n) be a sequence and let $L \in \mathbb{R}$. We say x_n *converges to* L, and write $x_n \to L$ or $\lim_{n\to\infty} x_n = L$, if

$$\forall \varepsilon > 0\ \exists N \in \mathbb{N}\ \forall n \in \mathbb{N}, \quad n \ge N \Rightarrow |x_n - L| < \varepsilon.$$

If no such L exists, we say (x_n) *diverges*.

Remark 3.5 (How to read the definition). The definition says: *no matter how small* a tolerance ε you demand, *eventually* all terms of the sequence lie within the interval $(L - \varepsilon, L + \varepsilon)$. The index N may depend on ε.

Definition 3.6 (Divergence phrased with quantifiers)**.** A sequence (x_n) does *not* converge to L if and only if

$$\exists \varepsilon_0 > 0\ \forall N \in \mathbb{N}\ \exists n \in \mathbb{N}, \quad n \ge N \text{ and } |x_n - L| \ge \varepsilon_0.$$

This is exactly the logical negation of Theorem 3.4.

Example 3.7 ($\frac{1}{n} \to 0$)**.** Let $x_n = \frac{1}{n}$. We prove $x_n \to 0$. Given $\varepsilon > 0$, by the Archimedean property (proved in the previous chapter), choose $N \in \mathbb{N}$ such that $N > \frac{1}{\varepsilon}$. Then for all $n \geq N$,

$$|x_n - 0| = \frac{1}{n} \leq \frac{1}{N} < \varepsilon.$$

Hence $\lim_{n\to\infty} \frac{1}{n} = 0$.

Example 3.8 (A divergent bounded sequence)**.** Let $x_n = (-1)^n$. Suppose $x_n \to L$. Taking $\varepsilon = \frac{1}{2}$, there exists N such that $n \geq N$ implies $|x_n - L| < \frac{1}{2}$. But both 1 and -1 occur infinitely often among the terms, so for $n \geq N$ we must have simultaneously

$$|1 - L| < \tfrac{1}{2} \quad \text{and} \quad |-1 - L| < \tfrac{1}{2},$$

which is impossible because those inequalities imply $L \in (1/2, 3/2)$ and $L \in (-3/2, -1/2)$ at the same time. Therefore $((-1)^n)$ diverges.

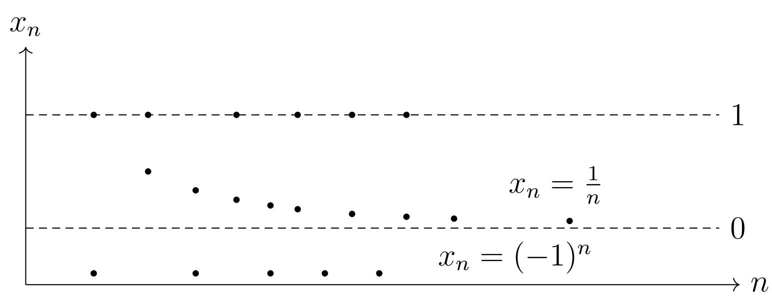


Figure 7: A convergent sequence $\frac{1}{n} \to 0$ versus a bounded divergent sequence $(-1)^n$.

3.1.3 Uniqueness of Limits

Theorem 3.9 (Uniqueness of limits)**.** *If a sequence (x_n) converges, then its limit is unique. That is, if $x_n \to L$ and $x_n \to M$, then $L = M$.*

Proof. Assume $x_n \to L$ and $x_n \to M$. Let $\varepsilon > 0$. Choose N_1 such that $n \geq N_1 \Rightarrow |x_n - L| < \varepsilon/2$, and choose N_2 such that $n \geq N_2 \Rightarrow |x_n - M| < \varepsilon/2$. For $n \geq N = \max\{N_1, N_2\}$, the triangle inequality gives

$$|L - M| \leq |L - x_n| + |x_n - M| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

Since $\varepsilon > 0$ is arbitrary, we must have $|L - M| = 0$, hence $L = M$. $\square$

3.1.4 Convergent Sequences Are Bounded

Proposition 3.10. *If $x_n \to L$, then (x_n) is bounded.*

Proof. Take $\varepsilon = 1$. Then there exists N such that $n \geq N \Rightarrow |x_n - L| < 1$, hence $|x_n| \leq |L| + 1$ for $n \geq N$. Let

$$M := \max\Big\{|x_1|, \ldots, |x_{N-1}|,\ |L| + 1\Big\}.$$

Then $|x_n| \leq M$ for all $n \in \mathbb{N}$. □

3.1.5 Limit Laws

Theorem 3.11 (Algebraic limit laws)**.** *Let (x_n) and (y_n) be sequences with $x_n \to x$ and $y_n \to y$. Then:*

(a) $x_n + y_n \to x + y$.

(b) $x_n y_n \to xy$.

(c) For any $c \in \mathbb{R}$, $cx_n \to cx$.

(d) If $y \neq 0$ and $y_n \neq 0$ for all sufficiently large n, then $\frac{x_n}{y_n} \to \frac{x}{y}$.

Proof. (a) Given $\varepsilon > 0$, choose N_1 so that $|x_n - x| < \varepsilon/2$ for $n \geq N_1$, and N_2 so that $|y_n - y| < \varepsilon/2$ for $n \geq N_2$. Then for $n \geq N = \max\{N_1, N_2\}$,

$$|(x_n + y_n) - (x + y)| \leq |x_n - x| + |y_n - y| < \varepsilon.$$

(b) First note that $y_n \to y$ implies (y_n) is bounded by Theorem 3.10; similarly (x_n) is bounded. Let $M > 0$ be such that $|x_n| \leq M$ and $|y_n| \leq M$ for all n. Then

$$x_n y_n - xy = x_n(y_n - y) + y(x_n - x),$$

so by the triangle inequality

$$|x_n y_n - xy| \leq |x_n|\,|y_n - y| + |y|\,|x_n - x| \leq M|y_n - y| + |y||x_n - x|.$$

Given $\varepsilon > 0$, choose N_1 such that $|y_n - y| < \varepsilon/(2M)$ for $n \geq N_1$, and choose N_2 such that $|x_n - x| < \varepsilon/(2(|y| + 1))$ for $n \geq N_2$. Then for $n \geq N = \max\{N_1, N_2\}$,

$$|x_n y_n - xy| < \frac{\varepsilon}{2} + \frac{|y|}{|y| + 1} \cdot \frac{\varepsilon}{2} \leq \varepsilon.$$

(c) This is immediate from $|cx_n - cx| = |c|\,|x_n - x|$.

(d) Since $y_n \to y \neq 0$, take $\varepsilon_0 = |y|/2 > 0$. Then there exists N_0 such that $n \geq N_0 \Rightarrow |y_n - y| < |y|/2$. Hence for $n \geq N_0$,

$$|y_n| \geq |y| - |y_n - y| > |y| - \frac{|y|}{2} = \frac{|y|}{2},$$

so $y_n \neq 0$ for $n \geq N_0$ and the reciprocals are well-defined.

Now write

$$\frac{x_n}{y_n} - \frac{x}{y} = \frac{xy_n - x_n y}{yy_n} = \frac{x(y_n - y) + y(x - x_n)}{yy_n}.$$

Taking absolute values and using the lower bound on $|y_n|$ for large n,

$$|\frac{x_n}{y_n} - \frac{x}{y}| \leq \frac{|x|\,|y_n - y| + |y|\,|x_n - x|}{|y|\,|y_n|} \leq \frac{2}{|y|^2}\big(|x|\,|y_n - y| + |y|\,|x_n - x|\big)$$

for all $n \geq N_0$. Choose N_1 so that $|y_n - y| < \frac{\varepsilon|y|^2}{4(|x|+1)}$ for $n \geq N_1$ and N_2 so that $|x_n - x| < \frac{\varepsilon|y|}{4}$ for $n \geq N_2$. Then for $n \geq \max\{N_0, N_1, N_2\}$, the right-hand side is $< \varepsilon$.

□

3.1.6 Order Properties and the Squeeze Theorem

Proposition 3.12 (Limit preserves inequalities)**.** *If $x_n \to x$ and $y_n \to y$, and if $x_n \leq y_n$ for all n, then $x \leq y$.*

Proof. Assume for contradiction that $x > y$. Let $\varepsilon = \frac{x-y}{3} > 0$. Choose N_1 such that $n \geq N_1 \Rightarrow |x_n - x| < \varepsilon$, and N_2 such that $n \geq N_2 \Rightarrow |y_n - y| < \varepsilon$. For $n \geq N = \max\{N_1, N_2\}$,

$$x_n > x - \varepsilon = \frac{2x + y}{3} \quad \text{and} \quad y_n < y + \varepsilon = \frac{x + 2y}{3}.$$

Since $\frac{2x+y}{3} > \frac{x+2y}{3}$ (because $x > y$), we get $x_n > y_n$ for all $n \geq N$, contradicting $x_n \leq y_n$. Thus $x \leq y$. □

Theorem 3.13 (Squeeze theorem)**.** *Suppose $a_n \leq x_n \leq b_n$ for all sufficiently large n. If $a_n \to L$ and $b_n \to L$, then $x_n \to L$.*

Proof. Let $\varepsilon > 0$. Choose N_1 such that $n \geq N_1 \Rightarrow |a_n - L| < \varepsilon$, hence $L - \varepsilon < a_n$. Choose N_2 such that $n \geq N_2 \Rightarrow |b_n - L| < \varepsilon$, hence $b_n < L + \varepsilon$. Also choose N_0 such that $n \geq N_0 \Rightarrow a_n \leq x_n \leq b_n$. For $n \geq N = \max\{N_0, N_1, N_2\}$,

$$L - \varepsilon < a_n \leq x_n \leq b_n < L + \varepsilon,$$

so $|x_n - L| < \varepsilon$. Hence $x_n \to L$. $\square$

Example 3.14 ($\frac{\sin n}{n} \to 0$)**.** Since $|\sin n| \leq 1$ for all n, we have

$$-\frac{1}{n} \leq \frac{\sin n}{n} \leq \frac{1}{n}.$$

By Theorem 3.7, both bounds converge to 0, hence by Theorem 3.13 we conclude $\frac{\sin n}{n} \to 0$.

Exercise 3.15. Prove that if $x_n \to 0$ and (y_n) is bounded, then $x_n y_n \to 0$.

Exercise 3.16. Let $x_n \to L$. Show that $|x_n| \to |L|$. (Hint: use the reverse triangle inequality.)

3.2 Monotone and Bounded Sequences

3.2.1 Monotonicity

Definition 3.17 (Monotone sequences)**.** A sequence (x_n) is

- *increasing* if $x_{n+1} \geq x_n$ for all n;
- *strictly increasing* if $x_{n+1} > x_n$ for all n;
- *decreasing* if $x_{n+1} \leq x_n$ for all n;
- *monotone* if it is increasing or decreasing.

Example 3.18. The sequence $x_n = 1 - \frac{1}{n}$ is increasing and bounded above by 1. The sequence $y_n = \frac{1}{n}$ is decreasing and bounded below by 0.

3.2.2 Monotone Convergence Theorem

Theorem 3.19 (Monotone Convergence Theorem)**.** *(a) If (x_n) is increasing and bounded above, then (x_n) converges and*

$$\lim_{n\to\infty} x_n = \sup\{x_n : n \in \mathbb{N}\}.$$

(b) If (x_n) is decreasing and bounded below, then (x_n) converges and

$$\lim_{n\to\infty} x_n = \inf\{x_n : n \in \mathbb{N}\}.$$

Proof. We prove (a); (b) is analogous.

Let $S = \{x_n : n \in \mathbb{N}\}$. Since (x_n) is bounded above, S is bounded above; since the sequence exists, S is nonempty. By completeness, let $s = \sup S$.

We claim $x_n \to s$. Let $\varepsilon > 0$. By the approximation property of the supremum, there exists some index N such that

$$s - \varepsilon < x_N \leq s.$$

Since (x_n) is increasing, for every $n \geq N$ we have $x_n \geq x_N > s - \varepsilon$. Also $x_n \leq s$ for all n because s is an upper bound. Thus for $n \geq N$,

$$s - \varepsilon < x_n \leq s \quad \Rightarrow \quad |x_n - s| < \varepsilon.$$

Therefore $x_n \to s$. $\square$

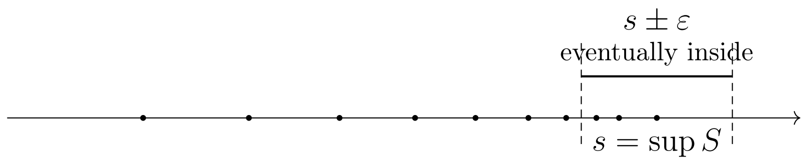


Figure 8: An increasing bounded sequence converges to $\sup S$.

3.2.3 Nested Interval Lemma (Useful Tool)

Theorem 3.20 (Nested Interval Property)**.** *Let (I_n) be a sequence of closed intervals*

$$I_n = [a_n, b_n] \subseteq \mathbb{R}$$

such that $I_{n+1} \subseteq I_n$ for all n (nested), and $a_n \le b_n$. Then

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

If moreover $b_n - a_n \to 0$, then the intersection consists of exactly one point.

Proof. Since the intervals are nested, (a_n) is increasing and bounded above by b_1, while (b_n) is decreasing and bounded below by a_1. Let

$$a = \sup\{a_n : n \in \mathbb{N}\}, \qquad b = \inf\{b_n : n \in \mathbb{N}\}.$$

For each n, we have $a_n \le a$ (as a is an upper bound of $\{a_n\}$) and $b \le b_n$ (as b is a lower bound of $\{b_n\}$). Also $a_n \le b_n$ implies $a \le b$ (one way: take limits using Theorem 3.12 after applying Theorem 3.19 to (a_n) and (b_n)). Hence $a \in [a_n, b_n] = I_n$ for all n (because $a_n \le a \le b_n$), so $a \in \bigcap_{n=1}^{\infty} I_n$ and the intersection is nonempty.

If additionally $b_n - a_n \to 0$, and x, y are in the intersection, then for every n we have $x, y \in [a_n, b_n]$, so $|x - y| \le b_n - a_n$. Letting $n \to \infty$ yields $|x - y| = 0$, hence $x = y$. □

3.2.4 Bolzano–Weierstrass Theorem

Theorem 3.21 (Bolzano–Weierstrass)**.** *Every bounded sequence in $\mathbb{R}$ has a convergent subsequence.*

Proof. Let (x_n) be bounded. Then there exists an interval $I_1 = [\alpha, \beta]$ such that $x_n \in I_1$ for all n (e.g. take $\alpha = -M$, $\beta = M$ where $|x_n| \le M$).

Bisect I_1 into two closed intervals of equal length:

$$I_1^{(L)} = \left[\alpha, \frac{\alpha+\beta}{2}\right], \qquad I_1^{(R)} = \left[\frac{\alpha+\beta}{2}, \beta\right].$$

At least one of these contains infinitely many terms of the sequence (otherwise the union would contain only finitely many terms). Choose I_2 to be a half-interval containing infinitely many terms, and then choose an index n_1 such that $x_{n_1} \in I_2$.

Now repeat: bisect I_2 into two halves, choose $I_3 \subseteq I_2$ that contains infinitely many terms, and pick $n_2 > n_1$ such that $x_{n_2} \in I_3$. Proceed inductively. We obtain:

- a nested sequence of closed intervals $I_k = [a_k, b_k]$ with $I_{k+1} \subseteq I_k$,
- lengths $b_k - a_k = \frac{\beta-\alpha}{2^{k-1}} \to 0$,
- indices $n_1 < n_2 < \cdots$ such that $x_{n_k} \in I_{k+1}$ for each k.

By Theorem 3.20, the intersection $\bigcap_{k=1}^{\infty} I_k$ contains exactly one point; call it L. We claim $x_{n_k} \to L$. Let $\varepsilon > 0$. Choose K such that $b_K - a_K < \varepsilon$ (possible since $b_k - a_k \to 0$). For $k \geq K$, we have $x_{n_k} \in I_{k+1} \subseteq I_K = [a_K, b_K]$ and also $L \in I_K$. Hence

$$|x_{n_k} - L| \leq b_K - a_K < \varepsilon.$$

Therefore the subsequence (x_{n_k}) converges to L. □

Remark 3.22. Bolzano–Weierstrass is a compactness statement in disguise: in $\mathbb{R}$, boundedness is enough to guarantee subsequential convergence. Later, this will be formalized as "closed and bounded sets are compact" (Heine–Borel).

3.2.5 Cauchy Sequences (Optional but Fundamental)

Definition 3.23 (Cauchy sequence)**.** A sequence (x_n) is *Cauchy* if

$$\forall \varepsilon > 0\ \exists N \in \mathbb{N}\ \forall m, n \in \mathbb{N}, \quad m, n \geq N \Rightarrow |x_n - x_m| < \varepsilon.$$

Proposition 3.24. *Every convergent sequence is Cauchy.*

Proof. Assume $x_n \to L$. Let $\varepsilon > 0$ and choose N such that $n \geq N \Rightarrow |x_n - L| < \varepsilon/2$. Then for $m, n \geq N$,

$$|x_n - x_m| \leq |x_n - L| + |L - x_m| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

□

Theorem 3.25 (Cauchy completeness of $\mathbb{R}$)**.** *Every Cauchy sequence in* $\mathbb{R}$ *converges.*

Proof. Let (x_n) be Cauchy. First, (x_n) is bounded: take $\varepsilon = 1$ and choose N so that $m, n \geq N \Rightarrow |x_n - x_m| < 1$. Fix $m = N$. Then for $n \geq N$, $|x_n| \leq |x_N| + 1$; together with finitely many initial terms, this gives boundedness.

By Theorem 3.21, (x_n) has a convergent subsequence $x_{n_k} \to L$. We show the full sequence converges to L.

Let $\varepsilon > 0$. Since (x_n) is Cauchy, choose N_1 such that $m, n \geq N_1 \Rightarrow |x_n - x_m| < \varepsilon/2$. Since $x_{n_k} \to L$, choose K such that $k \geq K \Rightarrow |x_{n_k} - L| < \varepsilon/2$. Also choose $k \geq K$ with $n_k \geq N_1$ (possible because $n_k \to \infty$). Then for any $n \geq N_1$,

$$|x_n - L| \leq |x_n - x_{n_k}| + |x_{n_k} - L| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

Thus $x_n \to L$. ∎

Exercise 3.26. Show that a sequence is Cauchy if and only if it is bounded and has the property that every subsequence has a further subsequence converging to the same limit.

3.3 Subsequences and Limit Points

3.3.1 Subsequences

Definition 3.27 (Subsequence)**.** Let (x_n) be a sequence. A *subsequence* is a sequence of the form $(x_{n_k})_{k\in\mathbb{N}}$ where

$$n_1 < n_2 < n_3 < \cdots$$

is a strictly increasing sequence of indices in $\mathbb{N}$.

Remark 3.28. A subsequence is obtained by *deleting terms without changing the order of the remaining ones.* The map $k \mapsto n_k$ must be strictly increasing; otherwise you could repeat indices, which would destroy key theorems.

Proposition 3.29. *If $x_n \to L$, then every subsequence x_{n_k} also converges to L.*

Proof. Let $\varepsilon > 0$. Choose N such that $n \geq N \Rightarrow |x_n - L| < \varepsilon$. Since $n_k \to \infty$ and is increasing, there exists K such that $k \geq K \Rightarrow n_k \geq N$. Then for $k \geq K$, $|x_{n_k} - L| < \varepsilon$. Hence $x_{n_k} \to L$. $\square$

3.3.2 Limit Points (Cluster Points)

Definition 3.30 (Limit point of a sequence)**.** A number $\ell \in \mathbb{R}$ is called a *limit point* (or *cluster point*) of (x_n) if there exists a subsequence (x_{n_k}) such that $x_{n_k} \to \ell$.

Example 3.31. For $x_n = (-1)^n$, the limit points are 1 and -1 (take the even and odd subsequences). For $x_n = \frac{1}{n}$, the only limit point is 0.

Proposition 3.32 (Characterization via infinitely many terms)**.** *A number $\ell \in \mathbb{R}$ is a limit point of (x_n) if and only if*

$$\forall \varepsilon > 0, \text{ the set } \{n \in \mathbb{N} : |x_n - \ell| < \varepsilon\} \text{ is infinite.}$$

Proof. ($\Rightarrow$) If ℓ is a limit point, there exists a subsequence $x_{n_k} \to \ell$. Fix $\varepsilon > 0$. Then there exists K such that $k \geq K \Rightarrow |x_{n_k} - \ell| < \varepsilon$. Thus infinitely many indices (namely $n_K, n_{K+1}, \ldots$) satisfy $|x_n - \ell| < \varepsilon$.

($\Leftarrow$) Assume the stated infinitude property. Construct a subsequence recursively: choose n_1 with $|x_{n_1} - \ell| < 1$ (possible since infinitely many exist), then choose $n_2 > n_1$ with $|x_{n_2} - \ell| < 1/2$, etc. In general, after choosing n_k, choose $n_{k+1} > n_k$ with $|x_{n_{k+1}} - \ell| < 1/(k+1)$. This is possible because the set of indices with $|x_n - \ell| < 1/(k+1)$ is infinite, hence contains indices larger than n_k. Then $|x_{n_k} - \ell| < 1/k \to 0$, so $x_{n_k} \to \ell$. $\square$

3.3.3 lim sup and lim inf

Definition 3.33 (Tail suprema and tail infima)**.** For a sequence (x_n) define, for each $n \in \mathbb{N}$,

$$s_n := \sup\{x_k : k \geq n\}, \qquad i_n := \inf\{x_k : k \geq n\}.$$

(These are well-defined in $\mathbb{R}$ whenever the sequence is bounded above/below respectively; in particular for bounded sequences both exist.)

Proposition 3.34. *If (x_n) is bounded, then (s_n) is decreasing and (i_n) is increasing. Moreover, for every n,*

$$i_n \leq x_n \leq s_n \quad \textit{and} \quad i_n \leq s_n.$$

Proof. As n increases, the tail set $\{x_k : k \geq n\}$ becomes smaller (fewer elements). Taking sup over a smaller set cannot increase: hence $s_{n+1} \leq s_n$, so (s_n) is decreasing. Similarly, taking inf over a smaller set cannot decrease: hence $i_{n+1} \geq i_n$, so (i_n) is increasing. Finally, x_n is an element of the tail set at index n, hence it lies between the tail infimum and supremum. ☐

Definition 3.35 (Limit superior and limit inferior)**.** For a bounded sequence (x_n), define

$$\limsup_{n\to\infty} x_n := \lim_{n\to\infty} s_n, \qquad \liminf_{n\to\infty} x_n := \lim_{n\to\infty} i_n,$$

where the limits exist by the Monotone Convergence Theorem because (s_n) is decreasing and bounded below, and (i_n) is increasing and bounded above.

Remark 3.36. Even when (x_n) is not convergent, the tails still have a well-defined "eventual ceiling" s_n and "eventual floor" i_n. The quantities $\limsup x_n$ and $\liminf x_n$ capture the long-term upper and lower accumulation behavior of the sequence.

Theorem 3.37 (Basic properties of lim sup and lim inf)**.** *Let (x_n) be bounded and define $L^+ = \limsup x_n$ and $L^- = \liminf x_n$. Then:*

(a) $L^- \leq L^+$.

(b) L^+ *and* L^- *are limit points of* (x_n).

(c) A number ℓ is a limit point of (x_n) if and only if $L^- \leq \ell \leq L^+$ and *for every $\varepsilon > 0$ there are infinitely many n with $|x_n - \ell| < \varepsilon$. (In particular, all limit points lie in $[L^-, L^+]$.)*

(d) The sequence (x_n) converges to L if and only if $\limsup x_n = \liminf x_n = L$.

Proof. (a) By Theorem 3.34, $i_n \leq s_n$ for all n. Taking limits preserves inequalities (apply Theorem 3.12 to the monotone sequences $i_n \to L^-$ and $s_n \to L^+$) to obtain $L^- \leq L^+$.

(b) We prove L^+ is a limit point; the proof for L^- is analogous. Fix $n \in \mathbb{N}$. By definition, $s_n = \sup\{x_k : k \geq n\}$. By the approximation property of the supremum, for each $m \in \mathbb{N}$ there exists an index $k \geq n$ such that

$$s_n - \frac{1}{m} < x_k \leq s_n.$$

Now apply this with $n = 1$ and build a subsequence: choose n_1 such that $s_1 - 1 < x_{n_1} \leq s_1$. Having chosen n_m, choose $n_{m+1} > n_m$ with

$$s_{n_{m+1}} - \frac{1}{m+1} < x_{n_{m+1}} \leq s_{n_{m+1}}.$$

Because $s_n \to L^+$ and $0 \leq s_{n_m} - x_{n_m} < 1/m$, we get

$$|x_{n_m} - L^+| \leq |x_{n_m} - s_{n_m}| + |s_{n_m} - L^+| < \frac{1}{m} + |s_{n_m} - L^+| \to 0,$$

so $x_{n_m} \to L^+$. Hence L^+ is a limit point.

(c) The "only if" direction is immediate from Theorem 3.32 (which is exactly the infinitude condition). Also, every limit point ℓ satisfies $i_n \leq \ell \leq s_n$ for all n because i_n is a lower bound and s_n an upper bound of the tail, and subsequential limits cannot escape those bounds; taking $n \to \infty$ yields $L^- \leq \ell \leq L^+$.

(d) If $x_n \to L$, then by Theorem 3.29 every subsequence converges to L, hence L is the only limit point. By (b), both L^+ and L^- are limit points, so $L^+ = L^- = L$.

Conversely, assume $L^+ = L^- = L$. Then $i_n \to L$ and $s_n \to L$. Since $i_n \leq x_n \leq s_n$ for all n by Theorem 3.34, the Squeeze Theorem (Theorem 3.13) implies $x_n \to L$.

□

Example 3.38 (Computing lim sup and lim inf)**.** Let $x_n = (-1)^n$. For each n, the tail $\{x_k : k \geq n\}$ contains both 1 and -1, hence $s_n = 1$ and $i_n = -1$ for all n. Therefore

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

Example 3.39 (A sequence with three limit points)**.** Define

$$x_n = \begin{cases} 0, & n \equiv 0 \pmod 3, \\ 1, & n \equiv 1 \pmod 3, \\ 2, & n \equiv 2 \pmod 3. \end{cases}$$

Then the limit points are $\{0, 1, 2\}$, and

$$\limsup x_n = 2, \qquad \liminf x_n = 0.$$

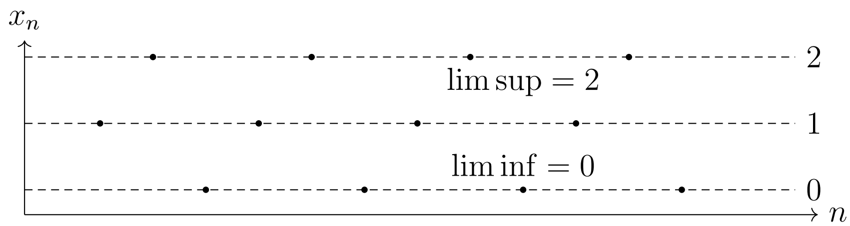


Figure 9: A bounded nonconvergent sequence with multiple limit points; lim inf and lim sup detect the eventual extremes.

3.3.4 Useful Criteria and Counterexamples

Proposition 3.40 (Eventual closeness implies convergence). *Let (x_n) and (y_n) be sequences. If $x_n - y_n \to 0$ and $y_n \to L$, then $x_n \to L$.*

Proof. We have

$$x_n - L = (x_n - y_n) + (y_n - L).$$

Take limits and use the limit laws (Theorem 3.11): the right-hand side tends to $0 + 0 = 0$, hence $x_n \to L$. $\square$

Example 3.41 (Bounded does not imply convergent). The sequence $(-1)^n$ is bounded but not convergent. This is the simplest example showing that boundedness alone is insufficient; one needs additional structure (e.g. monotonicity) or weaker conclusions (e.g. subsequential convergence).

Example 3.42 (A sequence can have infinitely many limit points). The sequence $x_n = \sin n$ is bounded. It has infinitely many limit points (in fact, the set of limit points is an interval). A full proof requires more tools than we have introduced so far, but this example is a warning: *Bolzano–Weierstrass guarantees at least one limit point, not just finitely many.*

Exercise 3.43. Let (x_n) be bounded. Prove that $s_n = \sup\{x_k : k \geq n\}$ satisfies

$$\limsup_{n\to\infty} x_n = \inf_{n\in\mathbb{N}} s_n,$$

and similarly

$$\liminf_{n\to\infty} x_n = \sup_{n\in\mathbb{N}} i_n.$$

(Hint: apply the Monotone Convergence Theorem to (s_n) and (i_n).)

Exercise 3.44. Construct a bounded sequence whose set of limit points is exactly the interval $[0, 1]$. (Hint: list rationals in $[0, 1]$ with repetitions and use density.)

Exercise 3.45. Suppose $x_n \geq 0$ for all n and $\liminf x_n = 0$. Must there exist a subsequence $x_{n_k} \to 0$? Prove your answer.

Exercise 3.46. Show that if $\limsup x_n < a$, then there exists N such that $x_n < a$ for all $n \geq N$. Similarly, if $\liminf x_n > a$, then $x_n > a$ eventually.

4 Series of Real Numbers

Key Idea. A series $\sum_{n=1}^{\infty} a_n$ is defined through its *sequence of partial sums*

$$s_N := \sum_{n=1}^{N} a_n.$$

The series converges if and only if (s_N) converges, and its value is $\lim_{N\to\infty} s_N$. Almost every convergence test is a way to control partial sums using positivity, comparison, or cancellation.

4.1 Definition and Basic Properties

4.1.1 Series and Partial Sums

Definition 4.1 (Infinite series)**.** Let $(a_n)_{n\in\mathbb{N}}$ be a sequence in $\mathbb{R}$. Define the *partial sums*

$$s_N := \sum_{n=1}^{N} a_n.$$

We say the *series* $\sum_{n=1}^{\infty} a_n$ *converges* if (s_N) converges in $\mathbb{R}$. In that case we define

$$\sum_{n=1}^{\infty} a_n := \lim_{N\to\infty} s_N.$$

If (s_N) diverges, we say the series diverges.

Remark 4.2. A series is not an infinite "sum" in the elementary sense; it is the *limit* of a sequence of finite sums. All rigor comes from working with partial sums and applying the convergence theory of sequences.

4.1.2 Necessary Condition for Convergence

Proposition 4.3 (Term test)**.** *If $\sum_{n=1}^{\infty} a_n$ converges, then $a_n \to 0$.*

Proof. Let $s_N = \sum_{n=1}^{N} a_n$. If $s_N \to s$, then

$$a_N = s_N - s_{N-1} \quad (N \geq 2).$$

Since $s_N \to s$ and $s_{N-1} \to s$ as well, the difference tends to 0 by the limit laws for sequences. Hence $a_N \to 0$. $\square$

Remark 4.4. The converse is false: $a_n \to 0$ does *not* imply $\sum a_n$ converges. The harmonic series is the standard counterexample.

4.1.3 Linearity

Theorem 4.5 (Linearity of convergent series). *If $\sum a_n$ and $\sum b_n$ converge and $\alpha, \beta \in \mathbb{R}$, then*

$$\sum_{n=1}^{\infty}(\alpha a_n + \beta b_n)$$

converges and

$$\sum_{n=1}^{\infty}(\alpha a_n + \beta b_n) = \alpha \sum_{n=1}^{\infty} a_n + \beta \sum_{n=1}^{\infty} b_n.$$

Proof. Let $s_N = \sum_{n=1}^{N} a_n$ and $t_N = \sum_{n=1}^{N} b_n$ with limits s and t. Then the partial sums of $\sum(\alpha a_n + \beta b_n)$ are

$$u_N = \sum_{n=1}^{N}(\alpha a_n + \beta b_n) = \alpha s_N + \beta t_N.$$

By sequence limit laws, $u_N \to \alpha s + \beta t$. Hence the series converges with the stated sum. $\square$

4.1.4 Geometric Series

Theorem 4.6 (Geometric series). *Let $r \in \mathbb{R}$.*

(a) If $|r| < 1$, then

$$\sum_{n=0}^{\infty} r^n = \frac{1}{1-r}.$$

(b) If $|r| \geq 1$, then $\sum_{n=0}^{\infty} r^n$ diverges.

Proof. For $N \geq 0$, the partial sum is

$$s_N = \sum_{n=0}^{N} r^n.$$

If $r \neq 1$, multiply by $(1-r)$:

$$(1-r)s_N = (1-r)\sum_{n=0}^{N} r^n = \sum_{n=0}^{N} r^n - \sum_{n=0}^{N} r^{n+1} = 1 - r^{N+1}.$$

Thus

$$s_N = \frac{1-r^{N+1}}{1-r}.$$

If $|r| < 1$, then $r^{N+1} \to 0$, so $s_N \to \frac{1}{1-r}$.

If $|r| \geq 1$, then r^N does not converge to 0 (indeed, if $r = 1$ then $s_N = N + 1 \to \infty$; if $r = -1$ then s_N oscillates; if $|r| > 1$ then $|s_N| \to \infty$), so the series diverges. $\square$

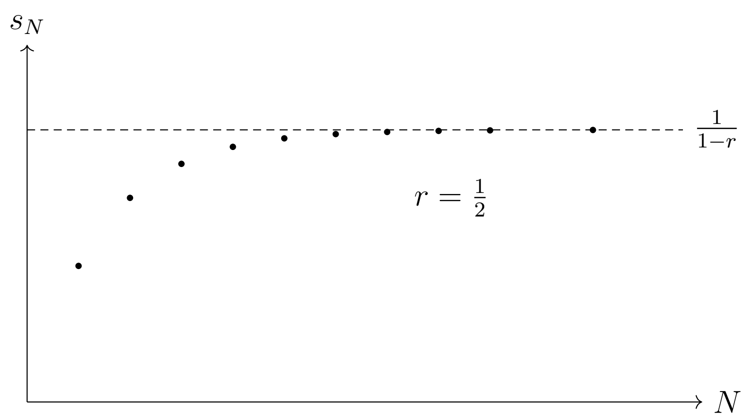


Figure 10: Partial sums of a geometric series approach a finite limit when $|r| < 1$.

Exercise 4.7. Show that $\sum_{n=1}^{\infty} 2^{-n} = 1$ and interpret this as a binary expansion of 1.

4.2 Positive Term Series and Comparison Tests

4.2.1 Monotonicity of Partial Sums

Proposition 4.8. *If $a_n \geq 0$ for all n, then the sequence of partial sums (s_N) is increasing. Hence $\sum_{n=1}^{\infty} a_n$ converges if and only if (s_N) is bounded above.*

Proof. If $a_n \geq 0$, then

$$s_{N+1} = s_N + a_{N+1} \geq s_N,$$

so (s_N) is increasing. By the Monotone Convergence Theorem for sequences, an increasing sequence converges if and only if it is bounded above. □

4.2.2 Comparison Test

Theorem 4.9 (Direct comparison test)**.** *Let $a_n, b_n \geq 0$ for all n and suppose $a_n \leq b_n$ for all sufficiently large n.*

(a) If $\sum b_n$ converges, then $\sum a_n$ converges.

(b) If $\sum a_n$ diverges, then $\sum b_n$ diverges.

Proof. Assume $a_n \leq b_n$ for $n \geq N_0$ and define partial sums

$$s_N = \sum_{n=1}^{N} a_n, \qquad t_N = \sum_{n=1}^{N} b_n.$$

For $N \geq N_0$,

$$s_N = \sum_{n=1}^{N_0-1} a_n + \sum_{n=N_0}^{N} a_n \leq \sum_{n=1}^{N_0-1} a_n + \sum_{n=N_0}^{N} b_n \leq \sum_{n=1}^{N_0-1} a_n + t_N.$$

If $\sum b_n$ converges, then (t_N) is bounded above, hence the right-hand side is bounded above, so (s_N) is bounded above. Since (s_N) is increasing, $\sum a_n$ converges by Theorem 4.8. This proves (a). Statement (b) is the contrapositive of (a). □

4.2.3 Limit Comparison

Theorem 4.10 (Limit comparison test)**.** *Let $a_n, b_n > 0$ for all n and suppose*

$$\lim_{n\to\infty} \frac{a_n}{b_n} = c$$

for some finite c with $c > 0$. Then $\sum a_n$ converges if and only if $\sum b_n$ converges.

Proof. Choose $\varepsilon = \frac{c}{2} > 0$. Then there exists N such that for $n \geq N$,

$$\left|\frac{a_n}{b_n} - c\right| < \frac{c}{2} \quad \Rightarrow \quad \frac{c}{2} < \frac{a_n}{b_n} < \frac{3c}{2}.$$

Thus for $n \geq N$,

$$\frac{c}{2} b_n < a_n < \frac{3c}{2} b_n.$$

By the direct comparison test applied twice, $\sum a_n$ converges iff $\sum b_n$ converges. □

4.2.4 p-Series

Theorem 4.11 (p-series)**.** *For $p \in \mathbb{R}$, the series*

$$\sum_{n=1}^{\infty} \frac{1}{n^p}$$

converges if and only if $p > 1$.

Proof. **Case 1: $p \leq 1$ (divergence).** If $p = 1$, we obtain the harmonic series $\sum 1/n$, which diverges (proved below via condensation). If $p < 1$, then $n^{-p} \geq n^{-1}$ for all $n \geq 1$ because $-p > -1$ implies $n^{-p}/n^{-1} = n^{1-p} \geq 1$. Since $\sum 1/n$ diverges and $n^{-p} \geq 1/n$, the comparison test implies $\sum 1/n^p$ diverges.

Case 2: $p > 1$ (convergence). We use Cauchy condensation (proved below) on the decreasing sequence $a_n = 1/n^p$. Then $\sum_{n=1}^{\infty} a_n$ converges iff $\sum_{k=0}^{\infty} 2^k a_{2^k}$ converges:

$$2^k a_{2^k} = 2^k \cdot \frac{1}{(2^k)^p} = 2^{k(1-p)}.$$

Since $p > 1$, we have $1 - p < 0$, so this is a geometric series with ratio $2^{1-p} \in (0, 1)$ and thus converges. Therefore $\sum 1/n^p$ converges. □

Remark 4.12. The fact that the threshold is exactly $p = 1$ is a fundamental "scale" in analysis: sums like $\sum 1/n$ are just barely too large, while $\sum 1/n^{1+\delta}$ becomes summable for any $\delta > 0$.

4.2.5 Cauchy Condensation Test

Theorem 4.13 (Cauchy condensation)**.** *Let (a_n) be a nonincreasing sequence of nonnegative real numbers. Then*

$$\sum_{n=1}^{\infty} a_n \text{ converges} \quad \Longleftrightarrow \quad \sum_{k=0}^{\infty} 2^k a_{2^k} \text{ converges.}$$

Proof. Let $s_N = \sum_{n=1}^{N} a_n$. Group terms in dyadic blocks:

$$\sum_{n=1}^{2^{m+1}-1} a_n = a_1 + \sum_{k=0}^{m} \sum_{n=2^k}^{2^{k+1}-1} a_n.$$

Since (a_n) is nonincreasing, for $n \in [2^k, 2^{k+1}-1]$ we have

$$a_{2^{k+1}-1} \le a_n \le a_{2^k}.$$

There are exactly 2^k integers in the block $[2^k, 2^{k+1}-1]$, hence

$$2^k\, a_{2^{k+1}-1} \le \sum_{n=2^k}^{2^{k+1}-1} a_n \le 2^k\, a_{2^k}.$$

Also $a_{2^{k+1}-1} \ge a_{2^{k+1}}$ (monotonicity), so

$$2^k\, a_{2^{k+1}} \le \sum_{n=2^k}^{2^{k+1}-1} a_n \le 2^k\, a_{2^k}.$$

Summing over $k = 0, \dots, m$ gives two-sided comparisons between partial sums of $\sum a_n$ and partial sums of $\sum 2^k a_{2^k}$ (up to harmless finite shifts and constants). Concretely,

$$\sum_{k=0}^{m} 2^k a_{2^{k+1}} \le \sum_{n=1}^{2^{m+1}-1} a_n \le a_1 + \sum_{k=0}^{m} 2^k a_{2^k}.$$

Since these are positive-term sums, boundedness (hence convergence) of one family of partial sums is equivalent to boundedness of the other. Therefore the two series converge or diverge together. ∎

Corollary 4.14 (Harmonic series diverges)**.** *The harmonic series $\sum_{n=1}^{\infty} \frac{1}{n}$ diverges.*

Proof. Apply Theorem 4.13 with $a_n = 1/n$ (positive and decreasing):

$$2^k a_{2^k} = 2^k \cdot \frac{1}{2^k} = 1.$$

Hence the condensed series is $\sum_{k=0}^{\infty} 1$, which diverges, so the harmonic series diverges. $\square$

Exercise 4.15. Use condensation to determine convergence of $\sum_{n=2}^{\infty} \frac{1}{n(\log n)^p}$ for $p > 0$. (You may assume standard logarithm identities.)

4.3 Absolute and Conditional Convergence

4.3.1 Absolute Convergence

Definition 4.16 (Absolute convergence)**.** A series $\sum a_n$ *converges absolutely* if $\sum |a_n|$ converges. If $\sum a_n$ converges but $\sum |a_n|$ diverges, the series is *conditionally convergent.*

Theorem 4.17 (Absolute convergence implies convergence)**.** *If $\sum |a_n|$ converges, then $\sum a_n$ converges.*

Proof. Let $s_N = \sum_{n=1}^{N} a_n$ be the partial sums. For $m > n$,

$$|s_m - s_n| = |\sum_{k=n+1}^{m} a_k| \le \sum_{k=n+1}^{m} |a_k|.$$

Since $\sum |a_k|$ converges, its partial sums are Cauchy, so the tail sums $\sum_{k=n+1}^{m} |a_k|$ can be made arbitrarily small uniformly for all $m > n$ once n is large. Hence (s_N) is Cauchy, and therefore convergent in $\mathbb{R}$ (Cauchy completeness of $\mathbb{R}$). Thus $\sum a_n$ converges. □

Remark 4.18. The key idea is that absolute convergence controls *all* cancellation patterns by dominating the tail differences with a positive-term tail.

4.3.2 Alternating Series

Theorem 4.19 (Leibniz alternating series test)**.** *Let (b_n) be a decreasing sequence of nonnegative numbers with $b_n \to 0$. Then the alternating series*

$$\sum_{n=1}^{\infty} (-1)^{n+1} b_n$$

converges. Moreover, if S is its sum and s_N its N-th partial sum, then the remainder satisfies the error bound

$$|S - s_N| \le b_{N+1}.$$

Proof. Let $s_N = \sum_{n=1}^{N} (-1)^{n+1} b_n$. Consider the even and odd partial sums:

$$s_{2m} = (b_1 - b_2) + (b_3 - b_4) + \cdots + (b_{2m-1} - b_{2m}),$$

$$s_{2m+1} = s_{2m} + b_{2m+1}.$$

Because b_n is decreasing, each difference $(b_{2k-1} - b_{2k}) \ge 0$, so (s_{2m}) is increasing. Also $s_{2m} \le b_1$ since each pair contributes at most b_{2k-1} and the alternating pattern never exceeds the first term; more directly,

$$s_{2m} = b_1 - (b_2 - b_3) - (b_4 - b_5) - \cdots - b_{2m} \le b_1.$$

Hence (s_{2m}) is increasing and bounded above, so it converges; let

$$\ell = \lim_{m\to\infty} s_{2m}.$$

Now note that $s_{2m+1} = s_{2m} + b_{2m+1}$, and $b_{2m+1} \to 0$, so $s_{2m+1} \to \ell$ as well. Thus the full sequence (s_N) converges to ℓ, proving convergence.

For the error bound: the partial sums alternate around the limit. Indeed $s_{2m} \le \ell \le s_{2m+1}$ for all m (because s_{2m} increases to ℓ and $s_{2m+1} = s_{2m} + b_{2m+1} \ge s_{2m}$ decreases to ℓ). Therefore

$$|\ell - s_N| \le |s_{N+1} - s_N| = b_{N+1},$$

which is exactly the stated bound. □

Example 4.20 (Alternating harmonic series)**.** The series

$$\sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n}$$

converges by Theorem 4.19 since $b_n = 1/n$ decreases to 0. It is not absolutely convergent because $\sum |(-1)^{n+1}/n| = \sum 1/n$ diverges (Theorem 4.14). Thus it is conditionally convergent.

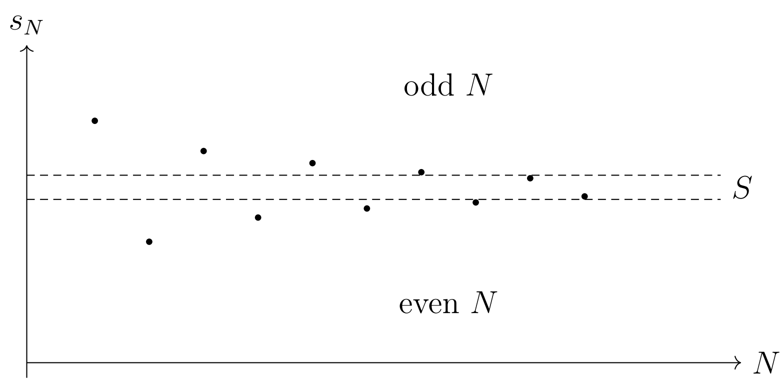


Figure 11: Alternating series partial sums oscillate and squeeze toward the limit when $b_n \searrow 0$.

4.3.3 Rearrangements (Statement Only for Now)

Remark 4.21 (Why absolute vs. conditional matters). For absolutely convergent series, rearranging terms does not change the sum. For conditionally convergent series, rearrangements can change the sum (Riemann rearrangement theorem). A rigorous treatment requires more machinery; we will return to this later.

Exercise 4.22. Use the alternating series error bound to estimate how many terms of $\sum_{n=1}^{\infty}(-1)^{n+1}/n$ are needed to approximate the sum within 10^{-3}.

4.4 Ratio and Root Tests; Power Series

4.4.1 Ratio Test

Theorem 4.23 (Ratio test). *Let $\sum a_n$ be a series with $a_n \neq 0$ for all sufficiently large n and define*

$$L := \limsup_{n\to\infty} \left|\frac{a_{n+1}}{a_n}\right| \in [0, \infty].$$

(a) If $L < 1$, then $\sum a_n$ converges absolutely.

(b) If $L > 1$ (including $L = \infty$), then $\sum a_n$ diverges.

(c) If $L = 1$, the test is inconclusive.

Proof. (a) Choose a number r with $L < r < 1$. By the definition of lim sup, there exists N such that for all $n \geq N$,

$$\left|\frac{a_{n+1}}{a_n}\right| \leq r, \quad \text{hence} \quad |a_{n+1}| \leq r|a_n|.$$

Iterating gives for $k \geq 0$,

$$|a_{N+k}| \leq |a_N| r^k.$$

Therefore

$$\sum_{n=N}^{\infty} |a_n| \leq |a_N| \sum_{k=0}^{\infty} r^k = \frac{|a_N|}{1-r} < \infty,$$

so $\sum |a_n|$ converges, and hence $\sum a_n$ converges absolutely.

(b) If $L > 1$, pick r with $1 < r < L$. Then by the definition of lim sup, there are infinitely many n such that

$$\left|\frac{a_{n+1}}{a_n}\right| \geq r, \quad \text{so} \quad |a_{n+1}| \geq r|a_n|.$$

In particular, $|a_n|$ cannot tend to 0; otherwise the inequality would force a subsequence to grow away from 0. Thus $a_n \not\to 0$, and by the term test (Theorem 4.3) the series diverges. □

4.4.2 Root Test

Theorem 4.24 (Root test). *Let $\sum a_n$ be a series and define*

$$L := \limsup_{n\to\infty} \sqrt[n]{|a_n|} \in [0, \infty].$$

(a) If $L < 1$, then $\sum a_n$ converges absolutely.

(b) If $L > 1$, then $\sum a_n$ diverges.

(c) If $L = 1$, the test is inconclusive.

Proof. (a) Choose r with $L < r < 1$. By definition of lim sup, there exists N such that for all $n \geq N$,

$$\sqrt[n]{|a_n|} \leq r \quad \Rightarrow \quad |a_n| \leq r^n.$$

Then $\sum_{n=N}^{\infty} |a_n| \leq \sum_{n=N}^{\infty} r^n$, a convergent geometric series. Hence $\sum a_n$ converges absolutely.

(b) If $L > 1$, choose r with $1 < r < L$. Then infinitely many n satisfy $\sqrt[n]{|a_n|} \geq r$, hence $|a_n| \geq r^n$ along a subsequence. In particular $a_n \not\to 0$, so the series diverges by Theorem 4.3. □

4.4.3 Power Series and Radius of Convergence

Definition 4.25 (Power series)**.** A *power series centered at x_0* is a series of the form

$$\sum_{n=0}^{\infty} c_n (x - x_0)^n,$$

where (c_n) is a sequence of real coefficients and x is a real variable.

Theorem 4.26 (Radius of convergence)**.** *For a power series $\sum_{n=0}^{\infty} c_n (x - x_0)^n$, define*

$$L := \limsup_{n \to \infty} \sqrt[n]{|c_n|} \in [0, \infty].$$

Then the series converges absolutely for all x satisfying $|x - x_0| < R$ and diverges for all x satisfying $|x - x_0| > R$, where

$$R := \begin{cases} \frac{1}{L}, & L > 0, \\ \infty, & L = 0, \\ 0, & L = \infty. \end{cases}$$

The number R is called the radius of convergence*.*

Proof. Fix $x \in \mathbb{R}$ and consider $a_n = c_n (x - x_0)^n$. Then

$$\sqrt[n]{|a_n|} = \sqrt[n]{|c_n|}\, |x - x_0|.$$

Taking lim sup and using basic properties of lim sup for nonnegative sequences yields

$$\limsup_{n \to \infty} \sqrt[n]{|a_n|} = \left(\limsup_{n \to \infty} \sqrt[n]{|c_n|} \right) |x - x_0| = L |x - x_0|.$$

By the root test (Theorem 4.24), the series $\sum a_n$ converges absolutely if $L|x - x_0| < 1$, i.e. $|x - x_0| < 1/L$ when $L > 0$, and diverges if $L|x - x_0| > 1$, i.e. $|x - x_0| > 1/L$. The cases $L = 0$ and $L = \infty$ are interpreted as $R = \infty$ and $R = 0$ respectively and follow from the same inequality. $\square$

Remark 4.27. The theorem is deliberately silent about the boundary $|x - x_0| = R$: convergence there must be checked case-by-case.

Exercise 4.28. Determine the radius of convergence of $\sum_{n=0}^{\infty} \frac{(x-2)^n}{n!}$ using the ratio test.

Exercise 4.29. Determine the radius of convergence of $\sum_{n=1}^{\infty} n(x+1)^n$ using the root test or ratio test.

Exercise 4.30. Give an example of a power series that converges at $x = x_0 + R$ but diverges at $x = x_0 - R$.

5 Topology of the Real Line

Key Idea. Topology provides the language for "closeness" without relying on coordinates. On $\mathbb{R}$ (with its usual distance), the key topological notions are: *open/closed sets*, *compactness*, and *connectedness*. These ideas will later become the natural setting for limits, continuity, and uniform convergence.

Notation 5.1. We view $\mathbb{R}$ as a metric space with the usual distance

$$d(x, y) := |x - y|.$$

For $x \in \mathbb{R}$ and $r > 0$, the (open) ball centered at x of radius r is

$$B(x, r) := \{y \in \mathbb{R} : |y - x| < r\} = (x - r, x + r).$$

5.1 Open and Closed Sets

5.1.1 Open Intervals, Neighborhoods, and Open Sets

Definition 5.2 (Neighborhood)**.** A set $N \subseteq \mathbb{R}$ is a *neighborhood* of $x \in \mathbb{R}$ if there exists $r > 0$ such that

$$B(x, r) \subseteq N.$$

Equivalently, x is an interior point of N.

Definition 5.3 (Open set)**.** A set $U \subseteq \mathbb{R}$ is *open* if for every $x \in U$ there exists $r > 0$ such that

$$B(x, r) \subseteq U.$$

Example 5.4. (a) Every open interval (a, b) is open.

(b) The set $[a, b)$ is not open (the point a has no ball contained in $[a, b)$).

(c) $\varnothing$ and $\mathbb{R}$ are open (vacuously / by definition).

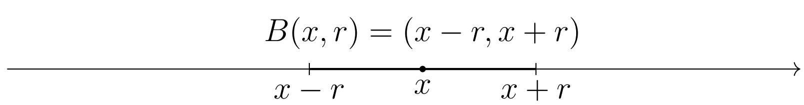


Figure 12: In $\mathbb{R}$, an open ball is exactly an open interval centered at x.

Proposition 5.5 (Basic closure properties of open sets)**.** *(a) The union of any family of open sets is open.*

(b) The intersection of finitely many open sets is open.

Proof. (a) Let $\{U_\alpha\}_{\alpha\in A}$ be open and set $U = \bigcup_{\alpha\in A} U_\alpha$. Take $x \in U$. Then $x \in U_{\alpha_0}$ for some α_0, and since U_{α_0} is open there exists $r > 0$ with $B(x, r) \subseteq U_{\alpha_0} \subseteq U$. Hence U is open.

(b) Let $U_1, \dots, U_m$ be open and set $U = \bigcap_{j=1}^m U_j$. Take $x \in U$. For each j there exists $r_j > 0$ such that $B(x, r_j) \subseteq U_j$. Let $r = \min\{r_1, \dots, r_m\} > 0$. Then $B(x, r) \subseteq U_j$ for all j, hence $B(x, r) \subseteq U$. So U is open. □

5.1.2 Closed Sets and Equivalent Characterizations

Definition 5.6 (Closed set)**.** A set $F \subseteq \mathbb{R}$ is *closed* if its complement $\mathbb{R} \setminus F$ is open.

Proposition 5.7 (Basic closure properties of closed sets)**.** *(a) The intersection of any family of closed sets is closed.*

(b) The union of finitely many closed sets is closed.

Proof. This is immediate from Theorem 5.5 by taking complements and using De Morgan's laws. □

Definition 5.8 (Limit point / accumulation point)**.** Let $A \subseteq \mathbb{R}$. A point $x \in \mathbb{R}$ is a *limit point* of A if for every $r > 0$,

$$\big(B(x, r) \setminus \{x\}\big) \cap A \neq \varnothing.$$

Equivalently: every punctured neighborhood of x contains a point of A.

Remark 5.9. A limit point of A need not belong to A. For example, 0 is a limit point of $(0, 1)$, but $0 \notin (0, 1)$.

Theorem 5.10 (Closed sets contain their limit points)**.** *A set $F \subseteq \mathbb{R}$ is closed if and only if it contains all of its limit points.*

Proof. ($\Rightarrow$) Assume F is closed and let x be a limit point of F. If $x \notin F$, then $x \in \mathbb{R} \setminus F$, and since $\mathbb{R} \setminus F$ is open, there exists $r > 0$ with $B(x, r) \subseteq \mathbb{R} \setminus F$. This implies $B(x, r) \cap F = \varnothing$, contradicting that x is a limit point of F. Hence $x \in F$.

($\Leftarrow$) Assume F contains all of its limit points. We show $\mathbb{R} \setminus F$ is open. Take $x \in \mathbb{R} \setminus F$. If every ball $B(x, r)$ met F, then x would be a limit point of F (indeed, $B(x, r) \setminus \{x\} = B(x, r)$ since $x \notin F$), hence $x \in F$ by assumption, a contradiction. Therefore, there exists $r > 0$ with $B(x, r) \cap F = \varnothing$, i.e. $B(x, r) \subseteq \mathbb{R} \setminus F$. Thus $\mathbb{R} \setminus F$ is open, and F is closed. □

5.1.3 Closure, Interior, and Boundary

Definition 5.11 (Closure)**.** For $A \subseteq \mathbb{R}$, the *closure* of A, denoted $\overline{A}$, is the intersection of all closed sets containing A:

$$\overline{A} \coloneqq \bigcap \{F \subseteq \mathbb{R} : \ F \text{ closed and } A \subseteq F\}.$$

Proposition 5.12 (Basic facts about closure)**.** *For any $A \subseteq \mathbb{R}$:*

(a) $A \subseteq \overline{A}$ and $\overline{A}$ is closed.

(b) $\overline{A}$ is the smallest *closed set containing A: if F is closed and $A \subseteq F$, then $\overline{A} \subseteq F$.*

(c) $\overline{A} = A$ if and only if A is closed.

Proof. (a) Each closed superset F of A contains A, so the intersection contains A. Intersections of closed sets are closed (Theorem 5.7).

(b) If F is closed with $A \subseteq F$, then F appears in the intersection defining $\overline{A}$, so $\overline{A} \subseteq F$.

(c) If A is closed, then by (a) and (b) we have $A \subseteq \overline{A} \subseteq A$, so $\overline{A} = A$. Conversely, if $\overline{A} = A$, then A is closed because $\overline{A}$ is closed by (a). □

Theorem 5.13 (Characterizations of closure)**.** *Let $A \subseteq \mathbb{R}$ and $x \in \mathbb{R}$. The following are equivalent:*

(a) $x \in \overline{A}$.

(b) For every $r > 0$, $B(x, r) \cap A \neq \varnothing$.

(c) There exists a sequence (a_n) with $a_n \in A$ for all n and $a_n \to x$.

Proof. (a)⇒(b): If (b) fails, then for some $r > 0$ we have $B(x, r) \cap A = \varnothing$. Then $U \coloneqq B(x, r)$ is open and disjoint from A, so $\mathbb{R} \setminus U$ is closed and contains A. Hence $\overline{A} \subseteq \mathbb{R} \setminus U$, so $x \notin \overline{A}$, contradicting (a).

(b)⇒(c): For each $n \in \mathbb{N}$, apply (b) with $r = 1/n$ to choose $a_n \in A \cap B(x, 1/n)$. Then $|a_n - x| < 1/n \to 0$, so $a_n \to x$.

(c)⇒(a): Let F be any closed set with $A \subseteq F$. Since $a_n \in A \subseteq F$ and F is closed, Theorem 5.10 implies the limit x of (a_n) lies in F. This holds for every such F, hence $x \in \bigcap F = \overline{A}$. □

Definition 5.14 (Interior)**.** For $A \subseteq \mathbb{R}$, the *interior* of A, denoted $\operatorname{int}(A)$, is the set of points $x \in A$ for which there exists $r > 0$ such that

$$B(x, r) \subseteq A.$$

Equivalently, $\operatorname{int}(A)$ is the union of all open sets contained in A.

Proposition 5.15 (Basic facts about interior)**.** *For any $A \subseteq \mathbb{R}$:*

(a) $\operatorname{int}(A)$ *is open and* $\operatorname{int}(A) \subseteq A$.

(b) $\operatorname{int}(A)$ *is the* largest *open set contained in A: if U is open and $U \subseteq A$, then $U \subseteq \operatorname{int}(A)$.*

(c) $\operatorname{int}(A) = A$ *if and only if A is open.*

Proof. (a) If $x \in \operatorname{int}(A)$, there exists $r > 0$ with $B(x, r) \subseteq A$. But then for any $y \in B(x, r/2)$ we have $B(y, r/2) \subseteq B(x, r) \subseteq A$, so $y \in \operatorname{int}(A)$. Hence $B(x, r/2) \subseteq \operatorname{int}(A)$ and $\operatorname{int}(A)$ is open. Also $\operatorname{int}(A) \subseteq A$ by definition.

(b) If U is open and $U \subseteq A$, then for any $x \in U$ there exists $r > 0$ with $B(x, r) \subseteq U \subseteq A$, so $x \in \operatorname{int}(A)$.

(c) If A is open, then every $x \in A$ has a ball contained in A, so $A \subseteq \operatorname{int}(A)$; combined with (a), $\operatorname{int}(A) = A$. Conversely, if $\operatorname{int}(A) = A$, then A is open by (a). $\square$

Definition 5.16 (Boundary)**.** For $A \subseteq \mathbb{R}$, the *boundary* of A is

$$\partial A := \overline{A} \setminus \operatorname{int}(A).$$

Proposition 5.17 (Boundary characterizations)**.** *Let $A \subseteq \mathbb{R}$ and $x \in \mathbb{R}$. The following are equivalent:*

(a) $x \in \partial A$.

(b) For every $r > 0$, both $B(x, r) \cap A \neq \varnothing$ and $B(x, r) \cap (\mathbb{R} \setminus A) \neq \varnothing$.

Moreover, ∂A is closed.

Proof. (a)$\Leftrightarrow$(b): $x \in \overline{A}$ is equivalent to $B(x, r) \cap A \neq \varnothing$ for all $r > 0$ by Theorem 5.13. Also $x \notin \operatorname{int}(A)$ means: for every $r > 0$, $B(x, r) \nsubseteq A$, i.e. $B(x, r) \cap (\mathbb{R} \setminus A) \neq \varnothing$. Combining yields (b).

Closedness: since $\partial A = \overline{A} \cap (\operatorname{int}(A))^c$ and $\overline{A}$ is closed while $\operatorname{int}(A)$ is open, $(\operatorname{int}(A))^c$ is closed. Hence ∂A is an intersection of closed sets, therefore closed. $\square$

Example 5.18. Let $A = (0, 1) \cup (2, 3)$. Then

$$\operatorname{int}(A) = A, \qquad \overline{A} = [0, 1] \cup [2, 3], \qquad \partial A = \{0, 1, 2, 3\}.$$

Figure 13: For $A = (0, 1) \cup (2, 3)$, the boundary is $\{0, 1, 2, 3\}$ and the closure adds endpoints.

5.1.4 Subspace Topology (Relative Openness/Closedness)

Definition 5.19 (Open in a subset / relative topology)**.** Let $E \subseteq \mathbb{R}$ and $U \subseteq E$. We say U is *open in* E (or *relatively open*) if there exists an open set $V \subseteq \mathbb{R}$ such that

$$U = E \cap V.$$

Similarly, $F \subseteq E$ is *closed in* E if there exists a closed set $C \subseteq \mathbb{R}$ such that $F = E \cap C$, equivalently if $E \setminus F$ is open in E.

Example 5.20. In $E = [0, 1]$, the set $[0, \frac{1}{2})$ is open in E because

$$[0, \tfrac{1}{2}) = [0, 1] \cap (-1, \tfrac{1}{2}).$$

But $[0, \frac{1}{2})$ is not open in $\mathbb{R}$.

Exercise 5.21. Prove that for $A \subseteq \mathbb{R}$,

$$\overline{A} = \mathbb{R} \setminus \operatorname{int}(\mathbb{R} \setminus A).$$

(Hint: use the ball characterizations of closure and interior.)

Exercise 5.22. Let $A \subseteq \mathbb{R}$. Show that $\overline{A}$ is the set of all limits of sequences from A (allowing eventually constant sequences). Show that x is a limit point of A if and only if there exists a sequence (a_n) with $a_n \in A \setminus \{x\}$ and $a_n \to x$.

5.2 Compactness

5.2.1 Open Covers and Compact Sets

Definition 5.23 (Open cover)**.** Let $K \subseteq \mathbb{R}$. A family $\{U_\alpha\}_{\alpha\in A}$ of open sets in $\mathbb{R}$ is an *open cover* of K if

$$K \subseteq \bigcup_{\alpha\in A} U_\alpha.$$

A *finite subcover* is a finite subcollection $U_{\alpha_1}, \dots, U_{\alpha_m}$ such that

$$K \subseteq \bigcup_{j=1}^{m} U_{\alpha_j}.$$

Definition 5.24 (Compactness)**.** A set $K \subseteq \mathbb{R}$ is *compact* if every open cover of K has a finite subcover.

Remark 5.25. Compactness is a finiteness principle: it says that even though K may be infinite, any attempt to cover it by open sets can be reduced to finitely many of them. On $\mathbb{R}$, compactness will turn out to be exactly the same as being *closed and bounded* (Heine–Borel).

5.2.2 Compactness of Closed Intervals

Theorem 5.26 (Heine–Borel for intervals)**.** *Every closed interval* $[a, b] \subseteq \mathbb{R}$ *is compact.*

Proof. Let $\{U_\alpha\}_{\alpha\in A}$ be an open cover of $[a, b]$. Define

$$S := \{x \in [a, b] : [a, x] \text{ can be covered by finitely many } U_\alpha\}.$$

We prove $b \in S$, which is exactly the statement that $[a, b]$ has a finite subcover.

Step 1: S is nonempty. Since $\{U_\alpha\}$ covers $[a, b]$, there exists α_0 with $a \in U_{\alpha_0}$. Because U_{α_0} is open, there exists $\delta > 0$ such that $B(a, \delta) \subseteq U_{\alpha_0}$. Then $[a, \min\{a + \delta/2, b\}] \subseteq U_{\alpha_0}$, so $\min\{a + \delta/2, b\} \in S$. Hence $S \neq \varnothing$.

Step 2: Let $c = \sup S$. Then $c = b$. Since $S \subseteq [a, b]$, it is bounded above, hence $c = \sup S$ exists. Assume for contradiction that $c < b$.

Because $\{U_\alpha\}$ covers $[a, b]$, pick β such that $c \in U_\beta$. Since U_β is open, choose $\eta > 0$ with $B(c, \eta) \subseteq U_\beta$.

By the supremum approximation property, there exists $x \in S$ such that

$$c - \frac{\eta}{2} < x \leq c.$$

By definition of S, the interval $[a, x]$ is covered by finitely many sets from the family, say

$$[a, x] \subseteq U_{\alpha_1} \cup \cdots \cup U_{\alpha_m}.$$

Also, since $x > c - \eta/2$, we have

$$[x,\, c + \eta/2] \subseteq [c - \eta/2,\, c + \eta/2] \subseteq B(c, \eta) \subseteq U_\beta.$$

Therefore,

$$[a,\, c + \eta/2] \subseteq U_{\alpha_1} \cup \cdots \cup U_{\alpha_m} \cup U_\beta,$$

which is a finite subcover of $[a, c + \eta/2]$. Hence $c + \eta/2 \in S$, contradicting that c is an upper bound of S. Thus $c = b$.

Step 3: Conclude $b \in S$. Since $c = \sup S = b$ and $b \in [a, b]$, the definition of S implies $[a, b]$ has a finite subcover. Therefore $[a, b]$ is compact. □

5.2.3 Closed Subsets of Compact Sets

Theorem 5.27 (Closed subset of compact is compact)**.** *If $K \subseteq \mathbb{R}$ is compact and $F \subseteq K$ is closed in $\mathbb{R}$, then F is compact.*

Proof. Let $\{U_\alpha\}_{\alpha \in A}$ be an open cover of F (open in $\mathbb{R}$). Since F is closed, $\mathbb{R} \setminus F$ is open. Then

$$K \subseteq \Big(\bigcup_{\alpha \in A} U_\alpha \Big) \cup (\mathbb{R} \setminus F),$$

so $\{U_\alpha\}_{\alpha \in A} \cup \{\mathbb{R} \setminus F\}$ is an open cover of K. By compactness of K, there exists a finite subcover:

$$K \subseteq U_{\alpha_1} \cup \cdots \cup U_{\alpha_m} \cup (\mathbb{R} \setminus F).$$

Intersecting both sides with F eliminates $\mathbb{R} \setminus F$ and yields

$$F \subseteq U_{\alpha_1} \cup \cdots \cup U_{\alpha_m}.$$

Thus F has a finite subcover and is compact. □

5.2.4 Compact Sets Are Closed and Bounded

Theorem 5.28 (Compact $\Rightarrow$ bounded)**.** *If $K \subseteq \mathbb{R}$ is compact, then K is bounded.*

Proof. For each $n \in \mathbb{N}$, the open interval $(-n, n)$ is open and

$$\mathbb{R} = \bigcup_{n\in\mathbb{N}} (-n, n).$$

Hence $\{(-n, n)\}_{n\in\mathbb{N}}$ is an open cover of K. By compactness, there exists $n_1, \ldots, n_m$ such that

$$K \subseteq \bigcup_{j=1}^{m} (-n_j, n_j).$$

Let $N = \max\{n_1, \ldots, n_m\}$. Then $\bigcup_{j=1}^{m}(-n_j, n_j) \subseteq (-N, N)$, so $K \subseteq (-N, N)$ and K is bounded. $\square$

Theorem 5.29 (Compact $\Rightarrow$ closed)**.** *If $K \subseteq \mathbb{R}$ is compact, then K is closed.*

Proof. We show that $\mathbb{R} \setminus K$ is open. Let $x \in \mathbb{R} \setminus K$. For each $y \in K$, define

$$r_y := \frac{|x-y|}{2} > 0, \qquad V_y := B(y, r_y).$$

Then V_y is open, $y \in V_y$, and importantly $x \notin V_y$ (since $|x - y| = 2r_y$). The family $\{V_y\}_{y\in K}$ is an open cover of K, so by compactness there exist $y_1, \ldots, y_m \in K$ such that

$$K \subseteq V_{y_1} \cup \cdots \cup V_{y_m}.$$

Let

$$r := \min\{r_{y_1}, \ldots, r_{y_m}\} > 0.$$

We claim that $B(x, r) \cap K = \varnothing$.

Indeed, take any $z \in B(x, r)$. For each j,

$$|z - y_j| \geq |x - y_j| - |z - x| > |x - y_j| - r \geq |x - y_j| - r_{y_j} = r_{y_j},$$

so $z \notin V_{y_j}$. Thus $z \notin \bigcup_{j=1}^{m} V_{y_j}$, and since $K \subseteq \bigcup_{j=1}^{m} V_{y_j}$, we must have $z \notin K$. Therefore $B(x, r) \subseteq \mathbb{R} \setminus K$, showing $\mathbb{R} \setminus K$ is open. Hence K is closed. $\square$

5.2.5 Heine–Borel Theorem

Theorem 5.30 (Heine–Borel on $\mathbb{R}$)**.** *A subset $K \subseteq \mathbb{R}$ is compact if and only if it is closed and bounded.*

Proof. ($\Rightarrow$) If K is compact, then it is bounded by Theorem 5.28 and closed by Theorem 5.29.

($\Leftarrow$) If K is closed and bounded, then $K \subseteq [a, b]$ for some $a < b$. By Theorem 5.26, $[a, b]$ is compact. Since K is closed in $\mathbb{R}$, it is a closed subset of $[a, b]$ as well, hence K is compact by Theorem 5.27. $\square$

5.2.6 Sequential Compactness

Definition 5.31 (Sequential compactness)**.** A set $K \subseteq \mathbb{R}$ is *sequentially compact* if every sequence in K has a subsequence that converges to a point in K.

Theorem 5.32 (Compactness and sequential compactness on $\mathbb{R}$)**.** *For $K \subseteq \mathbb{R}$, the following are equivalent:*

(a) K is compact.

(b) K is sequentially compact.

(c) K is closed and bounded.

Proof. (a)$\Rightarrow$(c) is Theorems 5.28 and 5.29, and (c)$\Rightarrow$(a) is Theorem 5.30.

(c)$\Rightarrow$(b): If K is bounded, then any sequence in K is bounded, so by Bolzano–Weierstrass it has a convergent subsequence in $\mathbb{R}$. If K is also closed, the limit of that subsequence must lie in K (closed sets contain limits of convergent sequences). Thus K is sequentially compact.

(b)$\Rightarrow$(c): If K were unbounded, we could choose $x_n \in K$ with $|x_n| \geq n$. Then (x_n) has no convergent subsequence (any convergent subsequence must be bounded), contradicting sequential compactness. So K is bounded.

To see K is closed, let (x_n) be a sequence in K with $x_n \to x \in \mathbb{R}$. By sequential compactness, (x_n) has a subsequence $x_{n_k} \to y$ with $y \in K$. But a subsequence of a convergent sequence converges to the same limit, so $y = x$. Hence $x \in K$, showing K is closed. $\square$

Exercise 5.33. Show that a set $K \subseteq \mathbb{R}$ is compact if and only if every sequence in K has a Cauchy subsequence. Then explain why completeness of $\mathbb{R}$ upgrades this to an actually convergent subsequence.

Exercise 5.34. Give an example of a closed set in $\mathbb{R}$ that is not compact, and explain precisely which hypothesis in Heine–Borel fails.

5.3 Connectedness

5.3.1 Connected Sets and Separations

Definition 5.35 (Separation and connectedness). Let $E \subseteq \mathbb{R}$. A *separation* of E is a pair of sets $A, B \subseteq E$ such that:

(i) $A \neq \varnothing$ and $B \neq \varnothing$,

(ii) $A \cap B = \varnothing$ and $E = A \cup B$,

(iii) A and B are open in the relative topology on E (equivalently, open subsets of E).

We say E is *connected* if it admits no separation.

Remark 5.36. Connectedness is the formal version of "E has no gaps." On $\mathbb{R}$, it will be equivalent to E being an interval (possibly a single point).

Definition 5.37 (Clopen sets in a subspace). A subset $U \subseteq E$ is *clopen in E* if it is both open in E and closed in E.

Proposition 5.38 (Connectedness via clopen sets). *A set $E \subseteq \mathbb{R}$ is connected if and only if the only clopen subsets of E are $\varnothing$ and E itself.*

Proof. If E has a separation $E = A \cup B$ with A, B nonempty disjoint and open in E, then $A = E \setminus B$ so A is also closed in E. Thus A is a nontrivial clopen subset.

Conversely, if $U \subseteq E$ is nonempty and not all of E and is clopen in E, then $E = U \cup (E \setminus U)$ is a separation because U and $E \setminus U$ are disjoint, nonempty, and open in E (since U is closed in E implies $E \setminus U$ is open in E). □

5.3.2 Intervals as Connected Sets

Definition 5.39 (Interval). A set $I \subseteq \mathbb{R}$ is an *interval* if for all $a, b \in I$ with $a < b$, every c with $a < c < b$ also lies in I. Equivalently,

$$a, b \in I,\ a < b \quad \Rightarrow \quad (a, b) \subseteq I.$$

Theorem 5.40 (Intervals are connected). *Every interval $I \subseteq \mathbb{R}$ is connected.*

Proof. Assume for contradiction that I is an interval and $I = A \cup B$ is a separation with A, B nonempty, disjoint, and open in I. Choose $a \in A$ and $b \in B$. WLOG assume $a < b$ (otherwise swap A, B).

Define

$$S := A \cap [a, b] \subseteq [a, b].$$

Then S is nonempty (since $a \in S$) and bounded above by b, so $c = \sup S$ exists with $a \le c \le b$.

We show $c \notin A$ and $c \notin B$, a contradiction because $I = A \cup B$.

Claim 1: $c \notin A$**.** If $c \in A$, since A is open in I, there exists $\varepsilon > 0$ such that

$$I \cap (c - \varepsilon, c + \varepsilon) \subseteq A.$$

Because $c \le b$, the point $c' = \min\{c + \varepsilon/2, b\}$ belongs to $I \cap (c - \varepsilon, c + \varepsilon)$ and hence to A. Moreover, $c' \in [a, b]$ and $c' > c$, so $c' \in S$ with $c' > c$, contradicting that c is an upper bound of S. Thus $c \notin A$.

Claim 2: $c \notin B$**.** If $c \in B$, since B is open in I, there exists $\varepsilon > 0$ such that

$$I \cap (c - \varepsilon, c + \varepsilon) \subseteq B.$$

By definition of supremum, there exists $s \in S$ with $c - \varepsilon/2 < s \le c$. Then $s \in A$ and also $s \in I \cap (c - \varepsilon, c + \varepsilon)$, so $s \in B$ as well. This contradicts $A \cap B = \varnothing$. Hence $c \notin B$.

Therefore $c \notin A \cup B = I$, impossible since $c \in [a, b] \subseteq I$ (because I is an interval containing a and b). So no separation exists and I is connected. □

5.3.3 Connected Subsets of $\mathbb{R}$ Are Intervals

Theorem 5.41 (Connected $\Leftrightarrow$ interval in $\mathbb{R}$)**.** *A set $E \subseteq \mathbb{R}$ is connected if and only if it is an interval (possibly a single point).*

Proof. ($\Leftarrow$) If E is an interval, it is connected by Theorem 5.40.

($\Rightarrow$) Assume E is connected. Take $a, b \in E$ with $a < b$ and let c satisfy $a < c < b$. If $c \notin E$, then

$$E = \big(E \cap (-\infty, c)\big) \;\cup\; \big(E \cap (c, \infty)\big),$$

a disjoint union. Both pieces are open in E (they are intersections of E with open sets of $\mathbb{R}$), both are nonempty (they contain a and b respectively), and they cover E. This is a separation of E, contradicting connectedness. Hence $c \in E$. Therefore $(a, b) \subseteq E$, so E is an interval. □

5.3.4 Intermediate Value Property for Sets

Proposition 5.42 (Intermediate value property for connected sets)**.** *Let $E \subseteq \mathbb{R}$. The following are equivalent:*

(a) E is connected.

(b) Whenever $a, b \in E$ with $a < b$, then every c with $a \leq c \leq b$ lies in E.

Proof. This is exactly Theorem 5.41, noting that including endpoints corresponds to the same interval property stated with $\leq$. $\square$

Remark 5.43. Later, we will prove that continuous functions map connected sets to connected sets. Combining that with Theorem 5.42 yields the Intermediate Value Theorem for continuous functions on intervals.

5.3.5 Connected Components

Definition 5.44 (Connected component)**.** Let $E \subseteq \mathbb{R}$ and $x \in E$. The *connected component of x in E* is the union of all connected subsets of E that contain x.

Theorem 5.45 (Components partition the set)**.** *The connected components of E form a partition of E:*

(a) Every $x \in E$ lies in exactly one connected component.

(b) Distinct connected components are disjoint.

Moreover, each component is connected and is maximal among connected subsets of E (by inclusion).

Proof. Fix $x \in E$ and let C_x be the union of all connected subsets of E containing x. Then C_x is connected: if it had a separation, restricting to any connected subset containing x would produce a separation there, impossible. (Equivalently: unions of connected sets with nonempty intersection are connected; here all sets intersect at x.)

If $y \in C_x$, then $C_y = C_x$ because any connected set containing y and contained in E can be joined (via union) with a connected set containing x and y inside C_x, forcing both maximal unions to coincide. Hence components either coincide or are disjoint, so they partition E. Maximality follows directly from the definition: any connected subset of E containing x is contained in the union C_x. $\square$

Corollary 5.46 (Components in $\mathbb{R}$ are intervals)**.** *If $E \subseteq \mathbb{R}$, then each connected component of E is an interval (possibly a single point).*

Proof. Each component is a connected subset of $\mathbb{R}$, hence an interval by Theorem 5.41. $\square$

Example 5.47. Let $E = (-1, 0) \cup (1, 3] \cup \{5\}$. Then E has three connected components: $(-1, 0)$, $(1, 3]$, and $\{5\}$.

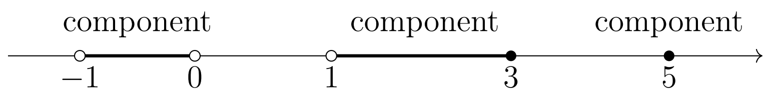


Figure 14: Connected components in $\mathbb{R}$ are precisely the "maximal pieces with no gaps" (intervals or points).

Exercise 5.48. Prove directly from the definition that if $E \subseteq \mathbb{R}$ is connected and $f : E \to \mathbb{R}$ is continuous (definition later), then $f(E)$ is connected. Explain how this implies the Intermediate Value Theorem.

Exercise 5.49. Let $A \subseteq \mathbb{R}$. Prove that $\overline{A}$ is connected whenever A is connected. (Hint: use Theorem 5.41.)

Exercise 5.50. Let $E \subseteq \mathbb{R}$ and let $\{C_i\}_{i \in I}$ be its connected components. Show that each C_i is closed in E (relative topology). Is each C_i necessarily open in E? Give a criterion or counterexample.

6 Continuity

Key Idea. Continuity is the rigorous formulation of "small changes in input produce small changes in output." There are multiple equivalent ways to express continuity:

- ε–δ definition (local control),
- sequential characterization (continuity preserves limits of sequences),
- topological characterization (preimages of open/closed sets).

On $\mathbb{R}$, these viewpoints are equivalent and each is useful in different proofs.

6.1 The ε–δ Definition

6.1.1 Continuity at a Point

Definition 6.1 (Continuity at a point)**.** Let $f : D \to \mathbb{R}$ with $D \subseteq \mathbb{R}$, and let $a \in D$. We say f is *continuous at* a if

$$\forall \varepsilon > 0 \ \exists \delta > 0 \ \forall x \in D, \quad |x - a| < \delta \Rightarrow |f(x) - f(a)| < \varepsilon.$$

If f is continuous at every $a \in D$, we say f is *continuous on* D.

Remark 6.2 (Domain matters). The quantifier "for all x" is restricted to $x \in D$. Continuity at a boundary point of D is therefore a one-sided phenomenon automatically built into the definition. For example, $f(x) = \sqrt{x}$ is continuous at 0 as a function on $D = [0, \infty)$.

6.1.2 Equivalent "Neighborhood" Formulation

Proposition 6.3 (Ball formulation)**.** *Let $f : D \to \mathbb{R}$ and $a \in D$. Then f is continuous at a if and only if: for every $r > 0$ there exists $\delta > 0$ such that*

$$f\big(B(a, \delta) \cap D\big) \subseteq B(f(a), r).$$

Proof. This is a direct rewriting of Theorem 6.1 with $r = \varepsilon$ and the notation $B(\cdot, \cdot)$ for open balls in $\mathbb{R}$. $\square$

6.1.3 Examples and Non-Examples

Example 6.4 (Constant and identity functions)**.** If $f(x) = c$, then $|f(x) - f(a)| = 0 < \varepsilon$ for all x, so any $\delta > 0$ works; hence f is continuous everywhere. If $f(x) = x$, then

$|f(x) - f(a)| = |x - a|$, so choosing $\delta = \varepsilon$ proves continuity.

Example 6.5 (Linear functions)**.** Let $f(x) = mx + b$. Fix $a \in \mathbb{R}$ and $\varepsilon > 0$. Then

$$|f(x) - f(a)| = |m(x - a)| = |m||x - a|.$$

If $m = 0$, f is constant. If $m \neq 0$, choosing $\delta = \varepsilon/|m|$ yields $|f(x) - f(a)| < \varepsilon$ whenever $|x - a| < \delta$. Thus every linear function is continuous on $\mathbb{R}$.

Example 6.6 (Quadratic function)**.** Let $f(x) = x^2$. Fix $a \in \mathbb{R}$. We estimate

$$|x^2 - a^2| = |x - a||x + a|.$$

To control $|x + a|$, impose $|x - a| < 1$ so that $|x| \leq |a| + 1$ and hence

$$|x + a| \leq |x| + |a| \leq 2|a| + 1.$$

Given $\varepsilon > 0$, choose

$$\delta = \min\left\{1,\ \frac{\varepsilon}{2|a| + 1}\right\}.$$

Then $|x - a| < \delta$ implies

$$|x^2 - a^2| = |x - a||x + a| < \delta(2|a| + 1) \leq \varepsilon,$$

so x^2 is continuous at a.

Example 6.7 (A discontinuous function: step function)**.** Define $f : \mathbb{R} \to \mathbb{R}$ by

$$f(x) = \begin{cases} 0, & x < 0, \\ 1, & x \geq 0. \end{cases}$$

Then f is discontinuous at 0. Indeed, take $\varepsilon = \frac{1}{2}$. For any $\delta > 0$, choose $x = -\delta/2$ (so $|x - 0| < \delta$) but then

$$|f(x) - f(0)| = |0 - 1| = 1 \geq \tfrac{1}{2},$$

so the ε–δ condition fails. However, f is continuous at every $a \neq 0$.

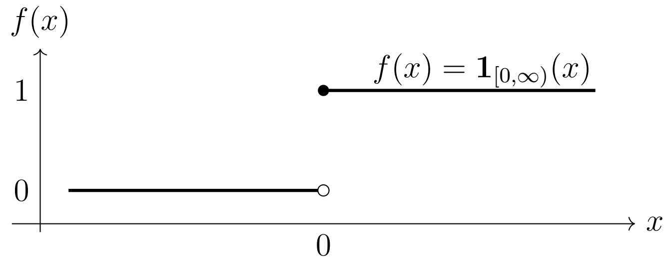


Figure 15: A jump discontinuity at $x = 0$: no matter how small δ is, values on both sides appear.

6.1.4 Algebra of Continuous Functions

Theorem 6.8 (Algebraic operations preserve continuity)**.** *Let $f, g : D \to \mathbb{R}$ be continuous at $a \in D$. Then:*

(a) $f + g$ is continuous at a.

(b) fg is continuous at a.

(c) If $g(a) \neq 0$, then $\frac{f}{g}$ is continuous at a (on points where $g \neq 0$).

(d) If $c \in \mathbb{R}$, then cf is continuous at a.

Proof. Fix $\varepsilon > 0$.

(a) Since f and g are continuous at a, there exist $\delta_1, \delta_2 > 0$ such that $|x - a| < \delta_1 \Rightarrow |f(x) - f(a)| < \varepsilon/2$ and $|x - a| < \delta_2 \Rightarrow |g(x) - g(a)| < \varepsilon/2$. Let $\delta = \min\{\delta_1, \delta_2\}$. Then for $|x - a| < \delta$,

$$|(f + g)(x) - (f + g)(a)| \leq |f(x) - f(a)| + |g(x) - g(a)| < \varepsilon.$$

(b) First note that continuity of f at a implies f is bounded in a neighborhood of a: take $\eta = 1$ and choose $\delta_0 > 0$ such that $|x - a| < \delta_0 \Rightarrow |f(x) - f(a)| < 1$. Then $|f(x)| \leq |f(a)| + 1$ whenever $|x - a| < \delta_0$.

Now write

$$f(x)g(x) - f(a)g(a) = f(x)\big(g(x) - g(a)\big) + g(a)\big(f(x) - f(a)\big).$$

Let $M = |f(a)| + 1$. Choose $\delta_1 > 0$ such that $|x - a| < \delta_1 \Rightarrow |f(x) - f(a)| < \varepsilon/(2(|g(a)| + 1))$, and choose $\delta_2 > 0$ such that $|x - a| < \delta_2 \Rightarrow |g(x) - g(a)| < \varepsilon/(2M)$.

Let $\delta = \min\{\delta_0, \delta_1, \delta_2\}$. Then for $|x-a| < \delta$ we have $|f(x)| \le M$ and

$$\begin{aligned} |f(x)g(x) - f(a)g(a)| &\le |f(x)||g(x)-g(a)| + |g(a)||f(x)-f(a)| \\ &< M \cdot \frac{\varepsilon}{2M} + |g(a)| \cdot \frac{\varepsilon}{2(|g(a)|+1)} \le \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon. \end{aligned}$$

(c) Since $g(a) \neq 0$, let $\eta = |g(a)|/2 > 0$. By continuity of g, choose $\delta_0 > 0$ such that $|x-a| < \delta_0 \Rightarrow |g(x) - g(a)| < \eta$. Then $|g(x)| \ge |g(a)| - \eta = |g(a)|/2$ for $|x-a| < \delta_0$, so $1/g(x)$ is well-defined and bounded there.

We show $1/g$ is continuous at a:

$$\left| \frac{1}{g(x)} - \frac{1}{g(a)} \right| = \frac{|g(x)-g(a)|}{|g(x)||g(a)|} \le \frac{2}{|g(a)|^2} |g(x) - g(a)| \quad (|x-a| < \delta_0).$$

Given $\varepsilon > 0$, choose $\delta_1 > 0$ so that $|x-a| < \delta_1 \Rightarrow |g(x)-g(a)| < \frac{\varepsilon |g(a)|^2}{2}$. Then for $|x-a| < \min\{\delta_0, \delta_1\}$ we have $|\frac{1}{g(x)} - \frac{1}{g(a)}| < \varepsilon$. Hence $1/g$ is continuous at a.

Finally, $\frac{f}{g} = f \cdot (1/g)$, and products of continuous functions are continuous by (b).

(d) Immediate from $|cf(x) - cf(a)| = |c|\,|f(x) - f(a)|$.

□

6.1.5 Composition

Theorem 6.9 (Continuity of composition)**.** *Let $f : D \to \mathbb{R}$ and $g : E \to \mathbb{R}$ with $f(D) \subseteq E$. If f is continuous at $a \in D$ and g is continuous at $f(a) \in E$, then $g \circ f$ is continuous at a.*

Proof. Let $\varepsilon > 0$. Since g is continuous at $f(a)$, there exists $\eta > 0$ such that

$$|y - f(a)| < \eta \Rightarrow |g(y) - g(f(a))| < \varepsilon.$$

Since f is continuous at a, there exists $\delta > 0$ such that

$$|x - a| < \delta \Rightarrow |f(x) - f(a)| < \eta.$$

Therefore, if $|x-a| < \delta$, then with $y = f(x)$ we have $|y - f(a)| < \eta$ and hence

$$|g(f(x)) - g(f(a))| < \varepsilon.$$

Thus $g \circ f$ is continuous at a. □

Exercise 6.10. Prove that polynomials are continuous on $\mathbb{R}$ and rational functions are continuous on their domains. (Hint: use Theorem 6.8 inductively.)

6.2 Sequential Characterization of Continuity

6.2.1 Continuity Preserves Limits of Sequences

Theorem 6.11 (Sequential criterion)**.** *Let $f : D \to \mathbb{R}$ and $a \in D$. Then f is continuous at a if and only if for every sequence (x_n) in D with $x_n \to a$, we have*

$$f(x_n) \to f(a).$$

Proof. ($\Rightarrow$) Assume f is continuous at a and let $x_n \to a$ with $x_n \in D$. Let $\varepsilon > 0$. By continuity, choose $\delta > 0$ such that $|x - a| < \delta \Rightarrow |f(x) - f(a)| < \varepsilon$. Since $x_n \to a$, there exists N such that $n \geq N \Rightarrow |x_n - a| < \delta$. Then $n \geq N$ implies $|f(x_n) - f(a)| < \varepsilon$, so $f(x_n) \to f(a)$.

($\Leftarrow$) Assume the sequential property holds, and suppose for contradiction that f is not continuous at a. Then there exists $\varepsilon_0 > 0$ such that for every $\delta > 0$ there exists $x \in D$ with

$$|x - a| < \delta \quad \text{and} \quad |f(x) - f(a)| \geq \varepsilon_0.$$

For each $n \in \mathbb{N}$, apply this with $\delta = 1/n$ to choose $x_n \in D$ such that

$$|x_n - a| < \frac{1}{n} \quad \text{and} \quad |f(x_n) - f(a)| \geq \varepsilon_0.$$

Then $x_n \to a$, but $f(x_n) \not\to f(a)$ because the distance to $f(a)$ is always at least ε_0. This contradicts the assumed sequential property. Hence f must be continuous at a. □

6.2.2 Closed Sets and Continuity

Theorem 6.12 (Preimages of closed sets)**.** *Let $f : D \to \mathbb{R}$ be continuous (on D). If $C \subseteq \mathbb{R}$ is closed, then $f^{-1}(C) := \{x \in D : f(x) \in C\}$ is closed in the relative topology on D.*

Proof. Let (x_n) be a sequence in $f^{-1}(C)$ that converges to some $x \in D$. We must show $x \in f^{-1}(C)$. Since $x_n \in f^{-1}(C)$, we have $f(x_n) \in C$ for all n. By continuity and Theorem 6.11, $f(x_n) \to f(x)$. Because C is closed, it contains limits of convergent sequences in C, so $f(x) \in C$. Thus $x \in f^{-1}(C)$, and $f^{-1}(C)$ is closed in D. □

Theorem 6.13 (Preimages of open sets)**.** *Let $f : D \to \mathbb{R}$ be continuous (on D). If $U \subseteq \mathbb{R}$ is open, then $f^{-1}(U)$ is open in the relative topology on D.*

Proof. Since U is open, $\mathbb{R} \setminus U$ is closed. By Theorem 6.12,

$$f^{-1}(\mathbb{R} \setminus U) \text{ is closed in } D.$$

But

$$f^{-1}(\mathbb{R} \setminus U) = D \setminus f^{-1}(U),$$

so $f^{-1}(U)$ is open in D. $\square$

Remark 6.14. Theorems Theorems 6.12 and 6.13 are the topological formulation of continuity. On metric spaces, this is equivalent to the ε–δ definition.

Exercise 6.15. Use Theorem 6.13 to prove that if f is continuous and $a \in D$, then for every $\varepsilon > 0$ the set

$$\{x \in D : |f(x) - f(a)| < \varepsilon\}$$

is open in D.

6.3 Continuous Functions on Compact Sets

6.3.1 Extreme Value Theorem

Theorem 6.16 (Extreme Value Theorem). *Let $K \subseteq \mathbb{R}$ be compact and let $f : K \to \mathbb{R}$ be continuous. Then f attains a maximum and a minimum on K: there exist $x_{\max}, x_{\min} \in K$ such that*

$$f(x_{\min}) = \min_{x \in K} f(x), \qquad f(x_{\max}) = \max_{x \in K} f(x).$$

Proof. We prove existence of the maximum; the minimum follows by applying the same argument to $-f$.

Let $M = \sup f(K)$, where $f(K) = \{f(x) : x \in K\}$. This set is nonempty, so M exists provided we show $f(K)$ is bounded above. We will show boundedness first, then attainability.

Step 1: $f(K)$ is bounded. Assume for contradiction that $f(K)$ is unbounded above. Then we can choose a sequence (x_n) in K such that $f(x_n) \geq n$. By sequential compactness of K (equivalent to compactness in $\mathbb{R}$), there exists a subsequence $x_{n_k} \to x^*$ with $x^* \in K$. By continuity, $f(x_{n_k}) \to f(x^*)$, so $(f(x_{n_k}))$ is bounded. But $f(x_{n_k}) \geq n_k \to \infty$, a contradiction. Hence $f(K)$ is bounded above; similarly it is bounded below.

Step 2: M is attained. By definition of supremum, for each $n \in \mathbb{N}$ there exists $x_n \in K$ such that

$$M - \frac{1}{n} < f(x_n) \leq M.$$

Again, by compactness, choose a convergent subsequence $x_{n_k} \to x_{\max} \in K$. By continuity,

$$f(x_{n_k}) \to f(x_{\max}).$$

But also $M - \frac{1}{n_k} < f(x_{n_k}) \leq M$, so by the squeeze theorem $f(x_{n_k}) \to M$. Therefore $f(x_{\max}) = M$, so the maximum is attained. □

6.3.2 Uniform Continuity

Definition 6.17 (Uniform continuity). Let $f : D \to \mathbb{R}$. We say f is *uniformly continuous* on D if

$$\forall \varepsilon > 0 \; \exists \delta > 0 \; \forall x, y \in D, \quad |x - y| < \delta \Rightarrow |f(x) - f(y)| < \varepsilon.$$

The key distinction from ordinary continuity is that δ depends only on ε, not on a base point.

Remark 6.18. Uniform continuity is a global property. A function can be continuous everywhere but fail to be uniformly continuous on an unbounded domain (e.g. $f(x) = x^2$ on $\mathbb{R}$).

Theorem 6.19 (Heine–Cantor)**.** *If $K \subseteq \mathbb{R}$ is compact and $f : K \to \mathbb{R}$ is continuous, then f is uniformly continuous on K.*

Proof. Assume for contradiction that f is not uniformly continuous on K. Then there exists $\varepsilon_0 > 0$ such that for every $n \in \mathbb{N}$ (taking $\delta = 1/n$) there exist $x_n, y_n \in K$ with

$$|x_n - y_n| < \frac{1}{n} \quad \text{but} \quad |f(x_n) - f(y_n)| \geq \varepsilon_0.$$

By compactness, extract a subsequence $x_{n_k} \to x^* \in K$. Then

$$|y_{n_k} - x^*| \leq |y_{n_k} - x_{n_k}| + |x_{n_k} - x^*| < \frac{1}{n_k} + |x_{n_k} - x^*| \to 0,$$

so $y_{n_k} \to x^*$ as well.

By continuity and the sequential criterion (Theorem 6.11),

$$f(x_{n_k}) \to f(x^*), \qquad f(y_{n_k}) \to f(x^*),$$

so $|f(x_{n_k}) - f(y_{n_k})| \to 0$. This contradicts $|f(x_{n_k}) - f(y_{n_k})| \geq \varepsilon_0$ for all k. Therefore f is uniformly continuous on K. □

6.3.3 Continuous Images of Compact Sets

Theorem 6.20 (Continuous image of a compact set is compact)**.** *If $K \subseteq \mathbb{R}$ is compact and $f : K \to \mathbb{R}$ is continuous, then $f(K)$ is compact.*

Proof. Let $\{U_\alpha\}_{\alpha \in A}$ be an open cover of $f(K)$. Then $\{f^{-1}(U_\alpha)\}_{\alpha \in A}$ is an open cover of K (open in K) by Theorem 6.13. By compactness of K, select $\alpha_1, \ldots, \alpha_m$ such that

$$K \subseteq \bigcup_{j=1}^{m} f^{-1}(U_{\alpha_j}).$$

Apply f to get

$$f(K) \subseteq \bigcup_{j=1}^{m} U_{\alpha_j}.$$

Thus $f(K)$ has a finite subcover and is compact. □

Corollary 6.21. *If K is compact and $f : K \to \mathbb{R}$ is continuous, then $f(K)$ is closed and bounded. In particular, f is bounded on K and attains its bounds (Extreme Value Theorem).*

Proof. In $\mathbb{R}$, compact sets are exactly closed and bounded sets (Heine–Borel). Apply Theorem 6.20. $\square$

Exercise 6.22. Show that if f is uniformly continuous on D and (x_n) is Cauchy in D, then $(f(x_n))$ is Cauchy in $\mathbb{R}$. Explain why this is useful for extending continuous functions.

6.4 Intermediate Value and Related Theorems

6.4.1 Intermediate Value Theorem

Theorem 6.23 (Intermediate Value Theorem)**.** *Let $f : [a, b] \to \mathbb{R}$ be continuous and let y be a real number between $f(a)$ and $f(b)$. Then there exists $c \in [a, b]$ such that $f(c) = y$.*

Proof. WLOG assume $f(a) < y < f(b)$ (the other cases are analogous; if $y = f(a)$ or $y = f(b)$ the claim is trivial).

Define
$$S \coloneqq \{x \in [a, b] : f(x) \leq y\}.$$
Then S is nonempty because $a \in S$ (since $f(a) \leq y$), and S is bounded above by b. Let $c = \sup S \in [a, b]$.

We claim $f(c) = y$. First show $f(c) \leq y$: if $f(c) > y$, then by continuity at c there exists $\delta > 0$ such that $|x - c| < \delta \Rightarrow f(x) > y$. In particular, for $x \in (c - \delta, c) \cap [a, b]$ we would have $f(x) > y$, so no such x lies in S. But then $c - \delta$ would be an upper bound of S, contradicting that $c = \sup S$.

Next show $f(c) \geq y$: if $f(c) < y$, then by continuity there exists $\delta > 0$ such that $|x - c| < \delta \Rightarrow f(x) < y$. Then for $x \in (c, c + \delta) \cap [a, b]$, we have $f(x) < y$, hence $x \in S$. This produces points of S strictly larger than c, contradicting that c is an upper bound of S.

Therefore $f(c) \leq y$ and $f(c) \geq y$, so $f(c) = y$. □

Remark 6.24. This proof is a standard "supremum argument": define a set of points on one side of the target value and use completeness of $\mathbb{R}$ plus continuity to show the boundary point hits the target.

6.4.2 Bisection Method (Existence of Roots)

Corollary 6.25 (Existence of a root)**.** *Let $f : [a, b] \to \mathbb{R}$ be continuous. If $f(a)f(b) < 0$, then there exists $c \in (a, b)$ such that $f(c) = 0$.*

Proof. If $f(a)f(b) < 0$, then $f(a)$ and $f(b)$ have opposite signs, so 0 lies between them. Apply Theorem 6.23 with $y = 0$. □

Proposition 6.26 (Bisection intervals converge)**.** *Assume $f : [a, b] \to \mathbb{R}$ is continuous and $f(a)f(b) < 0$. Construct nested intervals $[a_n, b_n]$ by repeatedly bisecting:*

$$[a_1, b_1] = [a, b], \qquad m_n = \frac{a_n + b_n}{2},$$

and set

$$[a_{n+1}, b_{n+1}] = \begin{cases} [a_n, m_n], & f(a_n)f(m_n) \le 0, \\ [m_n, b_n], & f(m_n)f(b_n) \le 0. \end{cases}$$

Then:

(a) $[a_{n+1}, b_{n+1}] \subseteq [a_n, b_n]$ *and* $b_n - a_n = \frac{b-a}{2^{n-1}}$.

(b) There exists a unique $c \in \bigcap_{n=1}^{\infty} [a_n, b_n]$.

(c) $f(c) = 0$.

Proof. (a) Each step chooses a subinterval of half-length, so nesting is immediate and the length formula follows by induction.

(b) By the Nested Interval Property, $\bigcap [a_n, b_n] \neq \varnothing$. Since lengths tend to 0, the intersection contains exactly one point.

(c) Let c be that point. Then $a_n \to c$ and $b_n \to c$ (because $c \in [a_n, b_n]$ and the interval length goes to 0). Also, by construction, $f(a_n)f(b_n) \le 0$ for all n (the sign-change property is preserved). By continuity, $f(a_n) \to f(c)$ and $f(b_n) \to f(c)$. Taking limits in the inequality $f(a_n)f(b_n) \le 0$ yields $f(c)^2 \le 0$, hence $f(c) = 0$. □

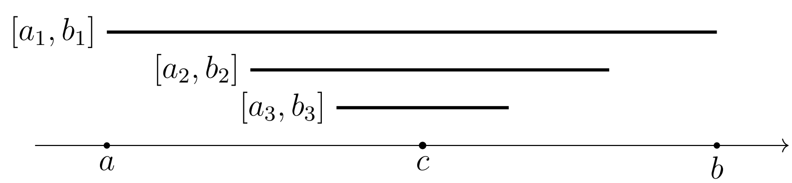


Figure 16: Bisection produces nested intervals shrinking to a unique root location.

Exercise 6.27. Apply the bisection method to $f(x) = x^3 - 2$ on $[1, 2]$ to produce an explicit n such that the interval length is $< 10^{-4}$. Explain why this yields an approximation to $\sqrt[3]{2}$.

7 Differentiation

Key Idea. Differentiation formalizes the best linear approximation of a function near a point. The derivative $f'(a)$, when it exists, is the unique real number such that

$$f(a+h) = f(a) + f'(a)\,h + o(h) \quad (h \to 0).$$

In this chapter we develop:

- the definition and basic rules of derivatives,
- the Mean Value Theorem and its consequences,
- Taylor's theorem (with a rigorous remainder),
- qualitative behavior of functions via derivatives.

7.1 Derivative and Basic Rules

7.1.1 Definition of the Derivative

Definition 7.1 (Derivative at a point)**.** Let $f : D \to \mathbb{R}$ with $D \subseteq \mathbb{R}$, and let $a \in D$ be a point such that a is a limit point of D (i.e. there exist points of D arbitrarily close to a). We say that f is *differentiable at* a if the limit

$$f'(a) := \lim_{x \to a,\ x \in D,\ x \neq a} \frac{f(x) - f(a)}{x - a}$$

exists in $\mathbb{R}$. Equivalently, using $h = x - a$,

$$f'(a) = \lim_{h \to 0,\ a+h \in D} \frac{f(a+h) - f(a)}{h}.$$

Remark 7.2. Differentiability is strictly stronger than continuity. Indeed, if $f'(a)$ exists then f must be continuous at a (proved below).

Definition 7.3 (Differentiability on an interval)**.** If $I \subseteq \mathbb{R}$ is an interval and $f : I \to \mathbb{R}$, we say f is *differentiable on* I if it is differentiable at every interior point of I. At endpoints, one can define one-sided derivatives (useful later), but the classical Mean Value Theorem uses interior derivatives.

7.1.2 Geometric Meaning

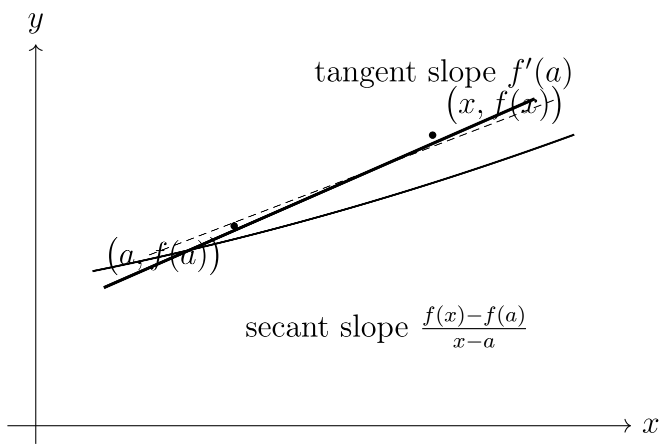


Figure 17: The derivative is the limiting slope of secant lines as $x \to a$.

7.1.3 Differentiability Implies Continuity

Theorem 7.4 (Differentiability $\Rightarrow$ continuity)**.** *If f is differentiable at a, then f is continuous at a.*

Proof. Assume $f'(a)$ exists. For $x \neq a$ (with $x \in D$),

$$f(x) - f(a) = \frac{f(x) - f(a)}{x - a}\,(x - a).$$

Take the limit as $x \to a$. The difference quotient tends to $f'(a)$ by differentiability, and $(x - a) \to 0$. By limit laws,

$$\lim_{x \to a}\big(f(x) - f(a)\big) = f'(a) \cdot 0 = 0,$$

so $f(x) \to f(a)$, i.e. f is continuous at a. $\square$

7.1.4 Linearity and Product/Quotient Rules

Theorem 7.5 (Basic derivative rules)**.** *Let $f, g : D \to \mathbb{R}$ be differentiable at $a \in D$. Then:*

(a) $(f + g)'(a) = f'(a) + g'(a)$.

(b) *For* $c \in \mathbb{R}$, $(cf)'(a) = cf'(a)$.

(c) $(fg)'(a) = f'(a)g(a) + f(a)g'(a)$.

(d) If $g(a) \neq 0$, then $(f/g)'(a) = \dfrac{f'(a)g(a) - f(a)g'(a)}{g(a)^2}$.

Proof. (a) Using the definition,

$$\frac{(f+g)(a+h) - (f+g)(a)}{h} = \frac{f(a+h) - f(a)}{h} + \frac{g(a+h) - g(a)}{h}.$$

Taking $h \to 0$ gives $f'(a) + g'(a)$.

(b) Immediate from

$$\frac{(cf)(a+h) - (cf)(a)}{h} = c\frac{f(a+h) - f(a)}{h}.$$

(c) Compute

$$\begin{aligned}\frac{f(a+h)g(a+h) - f(a)g(a)}{h} &= \frac{f(a+h)g(a+h) - f(a)g(a+h)}{h} + \frac{f(a)g(a+h) - f(a)g(a)}{h} \\ &= g(a+h)\frac{f(a+h) - f(a)}{h} + f(a)\frac{g(a+h) - g(a)}{h}.\end{aligned}$$

As $h \to 0$, $g(a+h) \to g(a)$ by continuity of g (from Theorem 7.4), and the difference quotients tend to $f'(a)$ and $g'(a)$. Thus the limit is $g(a)f'(a) + f(a)g'(a)$.

(d) First note that $g(a) \neq 0$ and continuity of g imply $g(a+h) \neq 0$ for all sufficiently small h. Then apply (c) to $f \cdot (1/g)$ once we show $(1/g)'(a) = -g'(a)/g(a)^2$. Compute:

$$\frac{\frac{1}{g(a+h)} - \frac{1}{g(a)}}{h} = \frac{g(a) - g(a+h)}{h\, g(a+h)\, g(a)} = -\frac{1}{g(a+h)g(a)} \cdot \frac{g(a+h) - g(a)}{h}.$$

Let $h \to 0$. The factor $1/(g(a+h)g(a)) \to 1/g(a)^2$, and the last quotient $\to g'(a)$, giving $(1/g)'(a) = -g'(a)/g(a)^2$. Then

$$\left(\frac{f}{g}\right)'(a) = (f \cdot (1/g))'(a) = f'(a)\frac{1}{g(a)} + f(a)\left(-\frac{g'(a)}{g(a)^2}\right) = \frac{f'(a)g(a) - f(a)g'(a)}{g(a)^2}.$$

□

7.1.5 Chain Rule

Theorem 7.6 (Chain rule)**.** *Let $f : D \to \mathbb{R}$ be differentiable at $a \in D$, and let $g : E \to \mathbb{R}$ be differentiable at $f(a)$, where $f(D) \subseteq E$. Then $g \circ f$ is differentiable at a and*

$$(g \circ f)'(a) = g'(f(a))\, f'(a).$$

Proof. Let $h \to 0$ with $a + h \in D$ and define $\Delta f = f(a+h) - f(a)$. Then

$$\frac{g(f(a+h)) - g(f(a))}{h} = \frac{g(f(a) + \Delta f) - g(f(a))}{\Delta f} \cdot \frac{\Delta f}{h},$$

provided $\Delta f \neq 0$. If $\Delta f = 0$ along some sequence $h_k \to 0$, then the expression is 0 along that subsequence, and the desired limit still holds because $\Delta f / h \to f'(a)$ and $g'(f(a))$ is finite; we may proceed with the standard argument below which covers all cases.

Since f is differentiable at a, we have $\Delta f \to 0$ and $\frac{\Delta f}{h} \to f'(a)$. Since g is differentiable at $f(a)$, we have

$$\lim_{\Delta f \to 0} \frac{g(f(a) + \Delta f) - g(f(a))}{\Delta f} = g'(f(a)).$$

Thus the product tends to $g'(f(a))\, f'(a)$ by limit laws. ☐

7.1.6 Standard Derivatives (From First Principles)

Proposition 7.7 (Derivative of x^n for $n \in \mathbb{N}$)**.** *For $n \in \mathbb{N}$, the function $f(x) = x^n$ is differentiable on $\mathbb{R}$ and*

$$\frac{d}{dx} x^n = n x^{n-1}.$$

Proof. Fix $a \in \mathbb{R}$. Using the factorization $x^n - a^n = (x-a) \sum_{k=0}^{n-1} x^{n-1-k} a^k$ (valid for all real x, a), for $x \neq a$ we have

$$\frac{x^n - a^n}{x - a} = \sum_{k=0}^{n-1} x^{n-1-k} a^k.$$

Let $x \to a$. Each term $x^{n-1-k} a^k \to a^{n-1-k} a^k = a^{n-1}$, and there are n terms. Hence the limit is $n a^{n-1}$, i.e. $f'(a) = n a^{n-1}$. ☐

Corollary 7.8. *Every polynomial is differentiable on $\mathbb{R}$ and its derivative is obtained by term-by-term differentiation. Every rational function is differentiable on its domain.*

Exercise 7.9. Prove from the definition that $\frac{d}{dx}|x|$ does not exist at $x = 0$. (Hint: compute the difference quotient along $h > 0$ and $h < 0$.)

7.2 Mean Value Theorem and Consequences

7.2.1 Rolle's Theorem

Theorem 7.10 (Rolle's theorem)**.** *Let $f : [a, b] \to \mathbb{R}$ be continuous on $[a, b]$ and differentiable on (a, b). If $f(a) = f(b)$, then there exists $c \in (a, b)$ such that $f'(c) = 0$.*

Proof. By the Extreme Value Theorem, f attains a maximum and a minimum on $[a, b]$. Let $x_{\max}$ be a point where the maximum is attained. If f is constant on $[a, b]$, then $f'(x) = 0$ for all $x \in (a, b)$ and we are done. Otherwise, the maximum value is strictly larger than the minimum value. Since $f(a) = f(b)$, at least one of the extrema must occur at an interior point (otherwise both max and min would be at endpoints, forcing constancy). Thus we can choose an interior extremum point $c \in (a, b)$ where either a maximum or a minimum occurs.

At such an interior extremum, we claim $f'(c) = 0$. For h small with $c + h \in (a, b)$,

$$\frac{f(c+h) - f(c)}{h} \le 0 \quad \text{for } h > 0 \qquad \text{and} \qquad \frac{f(c+h) - f(c)}{h} \ge 0 \quad \text{for } h < 0$$

if c is a maximum (the inequalities reverse for a minimum, but the conclusion is the same). Taking limits $h \to 0^{\pm}$ yields $f'(c) \le 0$ and $f'(c) \ge 0$, hence $f'(c) = 0$. □

7.2.2 Mean Value Theorem

Theorem 7.11 (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 the secant line

$$\ell(x) \coloneqq f(a) + \frac{f(b) - f(a)}{b - a}(x - a),$$

and consider $g(x) = f(x) - \ell(x)$. Then g is continuous on $[a, b]$ and differentiable on (a, b), and

$$g(a) = f(a) - \ell(a) = 0, \qquad g(b) = f(b) - \ell(b) = 0.$$

By Rolle's theorem (Theorem 7.10), there exists $c \in (a, b)$ such that $g'(c) = 0$. But

$$g'(x) = f'(x) - \frac{f(b) - f(a)}{b - a},$$

so $g'(c) = 0$ implies the desired identity. □

7.2.3 Consequences: Monotonicity and Constant Functions

Theorem 7.12 (Derivative sign and monotonicity). *Let $f : (a, b) \to \mathbb{R}$ be differentiable.*

(a) If $f'(x) \geq 0$ for all $x \in (a, b)$, then f is nondecreasing on (a, b).

(b) If $f'(x) > 0$ for all $x \in (a, b)$, then f is strictly increasing on (a, b).

(c) If $f'(x) \leq 0$ for all $x \in (a, b)$, then f is nonincreasing on (a, b).

(d) If $f'(x) = 0$ for all $x \in (a, b)$, then f is constant on (a, b).

Proof. We prove (a) and (d); (b) and (c) are analogous.

(a) Take $x_1 < x_2$ in (a, b). Apply the Mean Value Theorem to f on $[x_1, x_2]$ to find $c \in (x_1, x_2)$ such that

$$f(x_2) - f(x_1) = f'(c)(x_2 - x_1).$$

Since $x_2 - x_1 > 0$ and $f'(c) \geq 0$, we have $f(x_2) \geq f(x_1)$, so f is nondecreasing.

(d) If $f'(x) = 0$ for all x, then by (a) f is nondecreasing and by (c) f is nonincreasing, hence constant. Equivalently, apply MVT: for any $x_1 < x_2$,

$$f(x_2) - f(x_1) = f'(c)(x_2 - x_1) = 0,$$

so $f(x_2) = f(x_1)$. □

7.2.4 Uniqueness of Antiderivatives (Preview)

Corollary 7.13 (Antiderivatives differ by constants). *Let $F, G : (a, b) \to \mathbb{R}$ be differentiable and suppose $F'(x) = G'(x)$ for all $x \in (a, b)$. Then $F - G$ is constant on (a, b).*

Proof. $(F - G)' = F' - G' = 0$ on (a, b). Apply Theorem 7.12(d). □

7.2.5 Lipschitz Estimates

Theorem 7.14 (Bounded derivative implies Lipschitz). *Let $f : [a, b] \to \mathbb{R}$ be continuous on $[a, b]$ and differentiable on (a, b). If $|f'(x)| \leq M$ for all $x \in (a, b)$, then for all $x, y \in [a, b]$,*

$$|f(x) - f(y)| \leq M|x - y|.$$

Proof. If $x = y$ there is nothing to show. Assume $x < y$. Apply the Mean Value Theorem on $[x, y]$ to get $c \in (x, y)$ with

$$f(y) - f(x) = f'(c)(y - x).$$

Taking absolute values and using $|f'(c)| \leq M$ yields $|f(y) - f(x)| \leq M(y - x) = M|x - y|$. $\square$

Remark 7.15. This is stronger than uniform continuity: a Lipschitz function has a global linear modulus of continuity.

Exercise 7.16. Let $f : [0, 1] \to \mathbb{R}$ be differentiable on $(0, 1)$ and continuous on $[0, 1]$. If $f'(x) = 0$ for all $x \in (0, 1)$, show f is constant on $[0, 1]$. Be explicit about how endpoints are handled.

7.3 Inverse Function Theorem on the Real Line

7.3.1 Local Invertibility from a Nonzero Derivative

Theorem 7.17 (Inverse function theorem in $\mathbb{R}$)**.** *Let $f : (a, b) \to \mathbb{R}$ be differentiable and let $x_0 \in (a, b)$ with $f'(x_0) \neq 0$. Then there exists an open interval I containing x_0 such that:*

(a) f is strictly monotone on I, hence one-to-one on I;

(b) $f(I)$ is an open interval;

(c) the inverse function $f^{-1} : f(I) \to I$ is differentiable at $y_0 = f(x_0)$ and

$$(f^{-1})'(y_0) = \frac{1}{f'(x_0)}.$$

Proof. Assume $f'(x_0) > 0$ (the case $f'(x_0) < 0$ is similar with inequalities reversed). By the definition of derivative,

$$\lim_{x \to x_0} \frac{f(x) - f(x_0)}{x - x_0} = f'(x_0) > 0.$$

Hence there exists $\delta > 0$ such that for all $x \in (x_0 - \delta, x_0 + \delta)$ with $x \neq x_0$,

$$\frac{f(x) - f(x_0)}{x - x_0} > \frac{f'(x_0)}{2} > 0.$$

This implies:

- if $x > x_0$ then $f(x) > f(x_0)$,
- if $x < x_0$ then $f(x) < f(x_0)$.

So f is strictly increasing on $I = (x_0 - \delta, x_0 + \delta)$, hence one-to-one there. Since f is continuous on I, the image $f(I)$ is an interval, and because f is strictly monotone it is in fact open (a standard argument using intermediate values; details omitted here but can be proved carefully). Let $y_0 = f(x_0)$.

Now consider the inverse $g = f^{-1}$ defined on $f(I)$. To compute $g'(y_0)$, take $y \to y_0$ with $y \in f(I)$ and set $x = g(y) \in I$. Then $x \to x_0$ and

$$\frac{g(y) - g(y_0)}{y - y_0} = \frac{x - x_0}{f(x) - f(x_0)}.$$

Taking $x \to x_0$ gives

$$\lim_{y \to y_0} \frac{g(y) - g(y_0)}{y - y_0} = \left(\lim_{x \to x_0} \frac{f(x) - f(x_0)}{x - x_0} \right)^{-1} = \frac{1}{f'(x_0)}.$$

Thus g is differentiable at y_0 with derivative $1/f'(x_0)$. $\square$

Remark 7.18. This is the one-dimensional version of the Inverse Function Theorem. In higher dimensions, the condition is invertibility of the Jacobian matrix.

Exercise 7.19. Let $f(x) = x^3$. Use Theorem 7.17 to justify that $f^{-1}(y) = \sqrt[3]{y}$ is differentiable for all y and compute $(f^{-1})'(y)$.

7.4 Taylor's Theorem (One Variable)

7.4.1 Statement and Remainder

Definition 7.20 (Taylor polynomial)**.** Let $n \in \mathbb{N}_0$ and suppose f has derivatives up to order n at a. The *Taylor polynomial of order n at a* is

$$T_n(x) = \sum_{k=0}^{n} \frac{f^{(k)}(a)}{k!}(x-a)^k.$$

Theorem 7.21 (Taylor's theorem with Lagrange remainder)**.** *Let $f : [a,b] \to \mathbb{R}$ be such that $f^{(n+1)}$ exists on (a,b) and extends continuously to $[a,b]$. Then for each $x \in [a,b]$ there exists ξ between a and x such that*

$$f(x) = T_n(x) + R_{n+1}(x), \qquad R_{n+1}(x) = \frac{f^{(n+1)}(\xi)}{(n+1)!}(x-a)^{n+1}.$$

Proof. Fix $x \in [a,b]$, $x \neq a$ (the case $x = a$ is trivial). Define the auxiliary function

$$\Phi(t) \coloneqq f(t) - T_n(t) - \lambda(t-a)^{n+1},$$

where λ is chosen so that $\Phi(x) = 0$. Explicitly,

$$\lambda \coloneqq \frac{f(x) - T_n(x)}{(x-a)^{n+1}}.$$

By construction, $\Phi(a) = 0$ (because $T_n(a) = f(a)$ and $(a-a)^{n+1} = 0$) and $\Phi(x) = 0$.

Moreover, Φ has derivatives up to order $n+1$ on (a,b). A direct computation shows that for $k = 0, 1, \ldots, n$,

$$\Phi^{(k)}(a) = 0$$

because T_n matches the first n derivatives of f at a and $(t-a)^{n+1}$ has a zero of order $n+1$ at $t = a$.

Now apply Rolle's theorem repeatedly: Since $\Phi(a) = \Phi(x) = 0$, there exists c_1 between a and x with $\Phi'(c_1) = 0$. Also $\Phi'(a) = 0$ and $\Phi'(c_1) = 0$, so there exists c_2 between a and c_1 with $\Phi''(c_2) = 0$. Continuing in this way, we obtain a point ξ between a and x such that

$$\Phi^{(n+1)}(\xi) = 0.$$

But

$$\Phi^{(n+1)}(t) = f^{(n+1)}(t) - \lambda(n+1)!,$$

since T_n has degree n and its $(n+1)$st derivative is 0, while $(t-a)^{n+1}$ differentiates to $(n+1)!$.

Thus $\Phi^{(n+1)}(\xi) = 0$ implies

$$\lambda = \frac{f^{(n+1)}(\xi)}{(n+1)!}.$$

Substitute this into the definition of λ:

$$\frac{f(x) - T_n(x)}{(x-a)^{n+1}} = \frac{f^{(n+1)}(\xi)}{(n+1)!},$$

which rearranges to the stated remainder formula. □

7.4.2 Error Bounds

Corollary 7.22 (Uniform bound on the remainder). *Under the hypotheses of Theorem 7.21, if $|f^{(n+1)}(t)| \le M$ for all $t \in [a,b]$, then for all $x \in [a,b]$,*

$$|R_{n+1}(x)| \le \frac{M}{(n+1)!}|x-a|^{n+1}.$$

Proof. Use the remainder formula

$$R_{n+1}(x) = \frac{f^{(n+1)}(\xi)}{(n+1)!}(x-a)^{n+1}$$

for some ξ between a and x, then take absolute values and apply $|f^{(n+1)}(\xi)| \le M$. □

Exercise 7.23. Assume $|f''(t)| \le 10$ on $[-1,1]$ and $f(0) = 0$, $f'(0) = 2$. Use Taylor's theorem with $n = 1$ at $a = 0$ to give a quantitative bound on $|f(x) - 2x|$ for $x \in [-1,1]$.

7.5 Second Derivative and Shape

7.5.1 Local Extrema and the Second Derivative Test

Theorem 7.24 (Second derivative test)**.** *Let f be twice differentiable in a neighborhood of a and assume $f'(a) = 0$.*

(a) If $f''(a) > 0$, then f has a strict local minimum at a.

(b) If $f''(a) < 0$, then f has a strict local maximum at a.

(c) If $f''(a) = 0$, no conclusion is possible in general.

Proof. We prove (a). By Taylor's theorem of order 2 at a, for x near a there exists ξ between a and x such that

$$f(x) = f(a) + f'(a)(x - a) + \frac{f''(\xi)}{2}(x - a)^2.$$

Since $f'(a) = 0$,

$$f(x) - f(a) = \frac{f''(\xi)}{2}(x - a)^2.$$

If $f''(a) > 0$, continuity of f'' implies there exists a neighborhood of a in which $f''(\xi) > 0$. Then for $x \neq a$ sufficiently close to a, we have $(x-a)^2 > 0$ and $f''(\xi) > 0$, so $f(x)-f(a) > 0$. Hence $f(a)$ is a strict local minimum. The proof of (b) is analogous with signs reversed. □

7.5.2 Convexity (Preview)

Definition 7.25 (Convex function)**.** A function $f : I \to \mathbb{R}$ on an interval I is *convex* if for all $x, y \in I$ and all $\lambda \in [0, 1]$,

$$f(\lambda x + (1 - \lambda)y) \leq \lambda f(x) + (1 - \lambda)f(y).$$

Proposition 7.26 (Second derivative and convexity)**.** *If f is twice differentiable on an interval I and $f''(x) \geq 0$ for all $x \in I$, then f is convex on I.*

Proof. A standard proof uses the Mean Value Theorem on the derivative: If $f'' \geq 0$, then f' is nondecreasing by Theorem 7.12. This implies that secant slopes are increasing in their endpoints, which is equivalent to convexity. (A complete derivation can be included once secant-slope characterization of convexity is established.) □

Exercise 7.27. Prove rigorously that if $f'' \geq 0$ on I, then for $x < y < z$ in I,

$$\frac{f(y) - f(x)}{y - x} \leq \frac{f(z) - f(y)}{z - y}.$$

Then deduce convexity from this inequality.

8 The Riemann Integral

Key Idea. The Riemann integral is a precise way to define "area under a curve" for bounded functions on an interval. Its core idea is approximation by rectangles:

$$\int_a^b f(x)\,\mathrm{d}x \approx \sum_{i=1}^{n} f(\xi_i)\,(x_i - x_{i-1}),$$

where $a = x_0 < \cdots < x_n = b$ is a partition and $\xi_i \in [x_{i-1}, x_i]$ are sample points. A function is Riemann integrable exactly when these approximations can be made arbitrarily accurate by refining the partition.

8.1 Riemann Sums and Integrability

8.1.1 Partitions, Mesh, and Tagged Partitions

Definition 8.1 (Partition)**.** Let $a < b$. A *partition* P of $[a, b]$ is a finite set of points

$$P = \{x_0, x_1, \ldots, x_n\} \quad \text{with} \quad a = x_0 < x_1 < \cdots < x_n = b.$$

We write $\Delta x_i \coloneqq x_i - x_{i-1}$ for the ith subinterval length. The *mesh* (or norm) of P is

$$\|P\| \coloneqq \max_{1\le i\le n} \Delta x_i.$$

Definition 8.2 (Tagged partition and Riemann sum)**.** A *tagged partition* is a partition $P = \{x_0, \ldots, x_n\}$ together with a choice of tags

$$\xi_i \in [x_{i-1}, x_i] \qquad (i = 1, \ldots, n).$$

Given a bounded function $f : [a, b] \to \mathbb{R}$, the *Riemann sum* of f with respect to (P, ξ) is

$$S(f, P, \xi) \coloneqq \sum_{i=1}^{n} f(\xi_i)\,\Delta x_i.$$

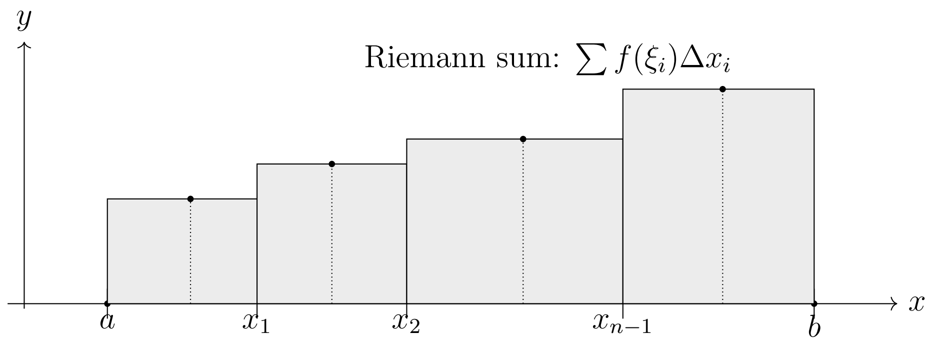


Figure 18: A partition with sample points (tags) produces a rectangle approximation to the area.

8.1.2 Upper and Lower Sums (Darboux Sums)

Definition 8.3 (Upper and lower sums)**.** Let $f : [a, b] \to \mathbb{R}$ be bounded and let $P = \{x_0, \dots, x_n\}$ be a partition. For each subinterval $[x_{i-1}, x_i]$, define

$$M_i(f, P) \coloneqq \sup_{x \in [x_{i-1}, x_i]} f(x), \qquad m_i(f, P) \coloneqq \inf_{x \in [x_{i-1}, x_i]} f(x).$$

The *upper sum* and *lower sum* are

$$U(f, P) \coloneqq \sum_{i=1}^{n} M_i(f, P)\, \Delta x_i, \qquad L(f, P) \coloneqq \sum_{i=1}^{n} m_i(f, P)\, \Delta x_i.$$

Proposition 8.4 (Refinement inequalities)**.** *If Q is a refinement of P (i.e. $P \subseteq Q$), then*

$$L(f, P) \le L(f, Q) \le U(f, Q) \le U(f, P).$$

Proof. Refining a partition splits some subintervals into smaller ones. On a smaller interval the infimum can only increase and the supremum can only decrease, so the lower-sum terms weakly increase and the upper-sum terms weakly decrease. Summing over subintervals gives the chain of inequalities. □

Proposition 8.5 (Riemann sums sit between Darboux sums)**.** *For any tagged partition (P, ξ),*

$$L(f, P) \le S(f, P, \xi) \le U(f, P).$$

Proof. Fix i. Since $\xi_i \in [x_{i-1}, x_i]$, we have

$$m_i(f, P) \le f(\xi_i) \le M_i(f, P).$$

Multiply by $\Delta x_i \geq 0$ and sum from $i = 1$ to n. $\square$

8.1.3 Darboux Integrability and the Definition of the Integral

Definition 8.6 (Upper/lower integrals)**.** Let $f : [a, b] \to \mathbb{R}$ be bounded. Define the *upper integral* and *lower integral* by

$$\overline{\int_a^b} f := \inf_P U(f, P), \qquad \underline{\int_a^b} f := \sup_P L(f, P),$$

where the infimum/supremum range over all partitions of $[a, b]$.

Proposition 8.7. *For every bounded $f : [a, b] \to \mathbb{R}$,*

$$\underline{\int_a^b} f \;\leq\; \overline{\int_a^b} f.$$

Proof. For any partition P, we have $L(f, P) \leq U(f, P)$. Taking the supremum over P on the left and the infimum over P on the right preserves the inequality. $\square$

Definition 8.8 (Riemann (Darboux) integrability)**.** A bounded function $f : [a, b] \to \mathbb{R}$ is *Riemann integrable* if

$$\underline{\int_a^b} f \;=\; \overline{\int_a^b} f.$$

In this case the common value is denoted

$$\int_a^b f(x)\, \mathrm{d}x$$

and is called the *Riemann integral* of f over $[a, b]$.

Theorem 8.9 (Characterization by upper–lower sums)**.** *A bounded function $f : [a, b] \to \mathbb{R}$ is Riemann integrable if and only if*

$$\forall \varepsilon > 0 \;\exists \text{ a partition } P \text{ such that } U(f, P) - L(f, P) < \varepsilon.$$

Proof. ($\Rightarrow$) If f is integrable, then

$$\inf_P U(f, P) = \sup_P L(f, P) = I.$$

Choose P_1 with $U(f, P_1) < I + \varepsilon/2$ and P_2 with $L(f, P_2) > I - \varepsilon/2$. Let P be a common refinement of P_1 and P_2 (take the union of their points). Then by Theorem 8.4,

$$U(f,P) \le U(f,P_1) < I + \frac{\varepsilon}{2}, \qquad L(f,P) \ge L(f,P_2) > I - \frac{\varepsilon}{2},$$

hence $U(f,P) - L(f,P) < \varepsilon$.

($\Leftarrow$) Assume the ε-partition condition holds. Let $I^- = \sup_P L(f,P)$ and $I^+ = \inf_P U(f,P)$. For any $\varepsilon > 0$ choose P with $U(f,P) - L(f,P) < \varepsilon$. Then

$$I^+ \le U(f,P) < L(f,P) + \varepsilon \le I^- + \varepsilon.$$

Since always $I^- \le I^+$ by Theorem 8.7, we get

$$0 \le I^+ - I^- < \varepsilon.$$

As $\varepsilon > 0$ is arbitrary, $I^+ = I^-$, so f is integrable. □

8.1.4 Equivalence with the "Riemann Sum" Definition

Theorem 8.10 (Riemann sums characterize the integral). *Let $f : [a,b] \to \mathbb{R}$ be bounded and let $I \in \mathbb{R}$. Then the following are equivalent:*

(a) f is Riemann integrable and $\displaystyle\int_a^b f(x)\,\mathrm{d}x = I$.

(b) For every $\varepsilon > 0$ there exists $\delta > 0$ such that for every tagged partition (P,ξ) with $\|P\| < \delta$,

$$|S(f,P,\xi) - I| < \varepsilon.$$

Proof. We use the Darboux framework.

(a)⇒(b): Let $I = \int_a^b f$ and fix $\varepsilon > 0$. By Theorem 8.9, choose a partition P_0 with

$$U(f,P_0) - L(f,P_0) < \varepsilon.$$

Now take any tagged partition (P,ξ) (no restriction on mesh is needed, but we proceed systematically) and form the common refinement

$$Q := P \cup P_0.$$

By Theorem 8.4 and Theorem 8.5,

$$L(f,Q) \le S(f,P,\xi) \le U(f,Q), \qquad L(f,P_0) \le L(f,Q) \le U(f,Q) \le U(f,P_0).$$

Also, by definition of I as both a supremum of lower sums and an infimum of upper sums,

$$L(f,Q) \le I \le U(f,Q).$$

Therefore,

$$|S(f,P,\xi) - I| \le U(f,Q) - L(f,Q) \le U(f,P_0) - L(f,P_0) < \varepsilon.$$

Thus every Riemann sum is within ε of I; in particular the statement holds for any $\delta > 0$.

(b)$\Rightarrow$(a): Assume (b). Fix $\varepsilon > 0$ and choose $\delta > 0$ such that $\|P\| < \delta$ implies $|S(f,P,\xi) - I| < \varepsilon/3$ for all tags. Let P be any partition with $\|P\| < \delta$.

For each subinterval $[x_{i-1}, x_i]$, choose points $\xi_i, \eta_i \in [x_{i-1}, x_i]$ such that

$$f(\xi_i) > M_i(f,P) - \gamma, \qquad f(\eta_i) < m_i(f,P) + \gamma,$$

where $\gamma > 0$ will be specified momentarily. (Such points exist because suprema/infima can be approximated arbitrarily well.)

Define the corresponding Riemann sums

$$S^+ := \sum_{i=1}^{n} f(\xi_i)\,\Delta x_i, \qquad S^- := \sum_{i=1}^{n} f(\eta_i)\,\Delta x_i.$$

Then

$$U(f,P) - \gamma(b-a) \ \le\ S^+ \ \le\ U(f,P), \qquad L(f,P) \ \le\ S^- \ \le\ L(f,P) + \gamma(b-a).$$

Choose $\gamma = \varepsilon\big/3(b-a)$ (any positive choice with $\gamma(b-a) < \varepsilon/3$ works). Now (b) gives $|S^+ - I| < \varepsilon/3$ and $|S^- - I| < \varepsilon/3$. Hence

$$\begin{aligned} U(f,P) - L(f,P) &\le \big(U(f,P) - S^+\big) + \big(S^+ - I\big) + \big(I - S^-\big) + \big(S^- - L(f,P)\big) \\ &< \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \frac{4\varepsilon}{3}. \end{aligned}$$

In particular, there exists a partition (namely this P) with $U(f,P) - L(f,P) < 2\varepsilon$. Since $\varepsilon > 0$ was arbitrary, Theorem 8.9 implies f is Riemann integrable.

Finally, to identify the integral with I, note that for integrable f every Riemann sum can be made arbitrarily close to the integral, so the only number satisfying (b) is the integral itself. Thus $\int_a^b f = I$. $\square$

8.1.5 Examples and Non-Examples

Example 8.11 (Constant and step functions)**.** If $f(x) = c$ on $[a, b]$, then every upper and lower sum equals $c(b-a)$, hence

$$\int_a^b c \, \mathrm{d}x = c(b-a).$$

More generally, any step function (finite linear combination of indicator functions of intervals) is Riemann integrable.

Example 8.12 (Dirichlet function is not Riemann integrable)**.** Define $f : [0,1] \to \mathbb{R}$ by

$$f(x) = \begin{cases} 1, & x \in \mathbb{Q}, \\ 0, & x \in \mathbb{R} \setminus \mathbb{Q}. \end{cases}$$

On every nontrivial interval, both rationals and irrationals occur, so on each subinterval $[x_{i-1}, x_i]$ we have

$$m_i(f, P) = 0, \qquad M_i(f, P) = 1.$$

Hence for every partition P,

$$L(f, P) = 0, \qquad U(f, P) = 1,$$

so $\underline{\int_0^1} f = 0$ and $\overline{\int_0^1} f = 1$. Therefore f is not Riemann integrable.

Example 8.13 (Thomae's function is Riemann integrable with integral 0)**.** Define $t : [0,1] \to \mathbb{R}$ by

$$t(0) = 1, \qquad t(x) = \begin{cases} \frac{1}{q}, & x = \frac{p}{q} \in (0,1] \text{ in lowest terms}, \\ 0, & x \notin \mathbb{Q}. \end{cases}$$

Then t is Riemann integrable on $[0,1]$ and

$$\int_0^1 t(x) \, \mathrm{d}x = 0.$$

Sketch of a fully rigorous estimate. Fix $\varepsilon > 0$. Choose $N \in \mathbb{N}$ so that $1/N < \varepsilon/2$. Let

$$A_N := \left\{ \frac{p}{q} \in [0,1] : 1 \le q \le N, \ \gcd(p,q) = 1 \right\}.$$

This set is finite. For each $r \in A_N$, choose an open interval I_r around r so small that the total length satisfies

$$\sum_{r \in A_N} |I_r| < \frac{\varepsilon}{2}.$$

(We can do this because A_N is finite.) Now build a partition P that includes all endpoints of the intervals I_r.

On subintervals lying outside $\bigcup_{r\in A_N} I_r$, any rational point has denominator $q > N$, hence $t(x) \le 1/N < \varepsilon/2$, and irrationals have value 0. Thus the supremum on those subintervals is at most $\varepsilon/2$.

On subintervals contained in $\bigcup_{r\in A_N} I_r$, the supremum is at most 1 (since $t(x) \le 1$ always).

Therefore the upper sum satisfies

$$U(t,P) \le 1 \cdot \Big|\bigcup_{r\in A_N} I_r\Big| \;+\; \frac{\varepsilon}{2}\cdot\Big(1 - \Big|\bigcup_{r\in A_N} I_r\Big|\Big) \le \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

Since $t \ge 0$, all lower sums are ≥ 0, so $\underline{\int_0^1} t \ge 0$. Letting $\varepsilon \downarrow 0$ forces $\overline{\int_0^1} t = 0$ and hence the integral is 0.

8.1.6 Continuous and Monotone Functions Are Integrable

Theorem 8.14 (Continuous functions are Riemann integrable)**.** *If $f : [a,b] \to \mathbb{R}$ is continuous, then f is Riemann integrable on $[a,b]$.*

Proof. Since $[a,b]$ is compact, continuity implies uniform continuity. Fix $\varepsilon > 0$. Choose $\delta > 0$ such that $|x-y| < \delta$ implies

$$|f(x) - f(y)| < \frac{\varepsilon}{b-a}.$$

Let P be any partition with $\|P\| < \delta$. For each subinterval $[x_{i-1}, x_i]$, the oscillation satisfies

$$M_i(f,P) - m_i(f,P) \le \frac{\varepsilon}{b-a}$$

because any two points in the subinterval are within δ of each other. Hence

$$U(f,P) - L(f,P) = \sum_{i=1}^{n} \big(M_i(f,P) - m_i(f,P)\big)\Delta x_i \le \frac{\varepsilon}{b-a}\sum_{i=1}^{n} \Delta x_i = \frac{\varepsilon}{b-a}(b-a) = \varepsilon.$$

By Theorem 8.9, f is Riemann integrable. □

Theorem 8.15 (Monotone functions are Riemann integrable)**.** *If $f : [a,b] \to \mathbb{R}$ is monotone (nondecreasing or nonincreasing), then f is Riemann integrable on $[a,b]$.*

Proof. Assume f is nondecreasing (the other case is analogous). Take an equal partition P_n with n subintervals:

$$x_i = a + i\frac{b-a}{n}.$$

On $[x_{i-1}, x_i]$, we have

$$m_i(f, P_n) = f(x_{i-1}), \qquad M_i(f, P_n) = f(x_i).$$

Thus

$$\begin{aligned} U(f, P_n) - L(f, P_n) &= \sum_{i=1}^{n} \big(f(x_i) - f(x_{i-1})\big)\Delta x_i \\ &= \frac{b-a}{n} \sum_{i=1}^{n} \big(f(x_i) - f(x_{i-1})\big) = \frac{b-a}{n}\big(f(b) - f(a)\big). \end{aligned}$$

As $n \to \infty$ this tends to 0, so for large n it is $< \varepsilon$. By Theorem 8.9, f is integrable. $\square$

Exercise 8.16. Let $f : [a, b] \to \mathbb{R}$ be bounded. Prove that if f is Riemann integrable and $g = f$ except at finitely many points, then g is Riemann integrable and

$$\int_a^b g(x)\ \mathrm{d}x = \int_a^b f(x)\ \mathrm{d}x.$$

(Hint: isolate each exceptional point inside a tiny subinterval of total length $< \varepsilon$.)

8.2 Properties of the Riemann Integral

8.2.1 Linearity, Order, and Bounds

Theorem 8.17 (Linearity). *If f, g are Riemann integrable on $[a, b]$ and $\alpha, \beta \in \mathbb{R}$, then $\alpha f + \beta g$ is Riemann integrable and*

$$\int_a^b (\alpha f(x) + \beta g(x)) \, \mathrm{d}x = \alpha \int_a^b f(x) \, \mathrm{d}x + \beta \int_a^b g(x) \, \mathrm{d}x.$$

Proof. Using upper/lower sums, one checks that for any partition P,

$$U(\alpha f + \beta g, P) \leq \alpha U(f, P) + \beta U(g, P) \quad \text{and} \quad L(\alpha f + \beta g, P) \geq \alpha L(f, P) + \beta L(g, P)$$

when $\alpha, \beta \geq 0$, and extends to general scalars by decomposing into positive and negative parts. Equivalently (and often more directly), use the Riemann sum characterization Theorem 8.10 and limit laws for sequences of sums. □

Theorem 8.18 (Monotonicity). *If f, g are integrable on $[a, b]$ and $f(x) \leq g(x)$ for all $x \in [a, b]$, then*

$$\int_a^b f(x) \, \mathrm{d}x \leq \int_a^b g(x) \, \mathrm{d}x.$$

Proof. For any partition P, we have $m_i(f, P) \leq m_i(g, P)$ and $M_i(f, P) \leq M_i(g, P)$, hence

$$L(f, P) \leq L(g, P), \qquad U(f, P) \leq U(g, P).$$

Taking suprema/infima and using integrability yields the desired inequality. □

Corollary 8.19 (Integral bounds from pointwise bounds). *If f is integrable and $m \leq f(x) \leq M$ for all $x \in [a, b]$, then*

$$m(b - a) \leq \int_a^b f(x) \, \mathrm{d}x \leq M(b - a).$$

Proof. Apply Theorem 8.18 to the constant functions m and M and use $\int_a^b c \, \mathrm{d}x = c(b-a)$. □

8.2.2 Additivity Over Intervals

Theorem 8.20 (Additivity). *Let f be Riemann integrable on $[a, b]$ and let $c \in (a, b)$. Then f is integrable on $[a, c]$ and on $[c, b]$, and*

$$\int_a^b f(x) \, \mathrm{d}x = \int_a^c f(x) \, \mathrm{d}x + \int_c^b f(x) \, \mathrm{d}x.$$

Proof. Given a partition P of $[a, b]$, refine it (if needed) so that c is a partition point. Then the upper and lower sums split exactly into sums over $[a, c]$ and $[c, b]$. Using Theorem 8.9 on each subinterval (and conversely combining partitions) shows integrability and the identity of integrals. □

8.2.3 Absolute Values and Continuous Transformations

Theorem 8.21 (Integrability is stable under continuous transforms)**.** *Let f be Riemann integrable on $[a, b]$. Let $\varphi : \mathbb{R} \to \mathbb{R}$ be continuous. Then $\varphi \circ f$ is Riemann integrable on $[a, b]$.*

Proof. Since f is integrable, it is bounded: $|f(x)| \le M$ on $[a, b]$ for some M. The restriction of φ to the compact interval $[-M, M]$ is uniformly continuous. Hence, for every $\varepsilon > 0$ there exists $\eta > 0$ such that for all $u, v \in [-M, M]$,

$$|u - v| < \eta \;\Rightarrow\; |\varphi(u) - \varphi(v)| < \frac{\varepsilon}{b - a}.$$

Because f is integrable, there exists a partition P such that on each subinterval the oscillation of f is $< \eta$ in the sense that

$$M_i(f, P) - m_i(f, P) < \eta \quad \text{for all } i$$

(after refining enough; a standard consequence of Theorem 8.9). Then on each subinterval, the oscillation of $\varphi \circ f$ is $< \varepsilon/(b - a)$. Therefore,

$$U(\varphi \circ f, P) - L(\varphi \circ f, P) \le \frac{\varepsilon}{b - a} \sum_{i=1}^{n} \Delta x_i = \varepsilon.$$

By Theorem 8.9, $\varphi \circ f$ is integrable. □

Corollary 8.22 (Absolute value and triangle inequality)**.** *If f is Riemann integrable on $[a, b]$, then $|f|$ is Riemann integrable and*

$$\left|\int_a^b f(x)\,\mathrm{d}x\right| \le \int_a^b |f(x)|\,\mathrm{d}x.$$

Proof. Integrability of $|f|$ follows from Theorem 8.21 with $\varphi(u) = |u|$. For the inequality, note that $-|f| \le f \le |f|$ pointwise, so by Theorem 8.18,

$$-\int_a^b |f|\,\mathrm{d}x \le \int_a^b f\,\mathrm{d}x \le \int_a^b |f|\,\mathrm{d}x,$$

which is equivalent to the stated estimate. □

Corollary 8.23 (Products and quotients)**.** *If f and g are Riemann integrable on $[a, b]$, then fg is Riemann integrable. If additionally $g(x) \neq 0$ for all x and $1/g$ is bounded on $[a, b]$, then f/g is Riemann integrable.*

Proof. The map $(u, v) \mapsto uv$ is continuous on $\mathbb{R}^2$; on bounded ranges one can reduce to a one-variable continuous transform using algebraic identities (or prove directly via oscillation estimates). For the quotient, use $f/g = f \cdot (1/g)$ and continuity of $u \mapsto 1/u$ on any set bounded away from 0. □

8.2.4 Mean Value Theorem for Integrals

Theorem 8.24 (Mean value theorem for integrals)**.** *If $f : [a, b] \to \mathbb{R}$ is continuous, then there exists $c \in [a, b]$ such that*

$$\int_a^b f(x)\, \mathrm{d}x = f(c)\,(b - a).$$

Proof. By the Extreme Value Theorem, f attains a minimum m and maximum M on $[a, b]$. Thus $m \leq f(x) \leq M$ for all x, so by Theorem 8.19,

$$m(b - a) \leq \int_a^b f(x)\, \mathrm{d}x \leq M(b - a).$$

Divide by $(b - a) > 0$:

$$m \leq \frac{1}{b - a} \int_a^b f(x)\, \mathrm{d}x \leq M.$$

Because $f([a, b])$ is an interval (continuous image of a connected set), every value between m and M is achieved. Hence there exists c with $f(c) = \frac{1}{b-a} \int_a^b f$. □

Exercise 8.25. Let f be integrable on $[a, b]$ and define $F(x) = \int_a^x f(t)\, \mathrm{d}t$. Prove that F is uniformly continuous on $[a, b]$. (Hint: use Theorem 8.19 with a bound on $|f|$.)

8.3 Fundamental Theorem of Calculus

8.3.1 The Integral Defines a Continuous Accumulation Function

Theorem 8.26 (Continuity of the accumulation function)**.** *Let $f : [a, b] \to \mathbb{R}$ be Riemann integrable, and define for $x \in [a, b]$,*

$$F(x) := \int_a^x f(t)\ \mathrm{d}t.$$

Then F is continuous on $[a, b]$. Moreover, if $|f(t)| \le M$ on $[a, b]$, then for all $x, y \in [a, b]$,

$$|F(x) - F(y)| \le M|x - y|.$$

Proof. Assume $x < y$ (the other case is symmetric). By additivity (Theorem 8.20),

$$F(y) - F(x) = \int_x^y f(t)\ \mathrm{d}t.$$

If $|f(t)| \le M$, then $-M \le f(t) \le M$, so by Theorem 8.19,

$$-M(y - x) \le \int_x^y f(t)\ \mathrm{d}t \le M(y - x).$$

Thus $|F(y) - F(x)| \le M|y - x|$, which implies continuity (indeed Lipschitz continuity). □

8.3.2 Fundamental Theorem of Calculus, Part I

Theorem 8.27 (FTC I: differentiation of the integral)**.** *Let $f : [a, b] \to \mathbb{R}$ be Riemann integrable and define $F(x) = \int_a^x f(t)\ \mathrm{d}t$. If f is continuous at a point $x_0 \in (a, b)$, then F is differentiable at x_0 and*

$$F'(x_0) = f(x_0).$$

In particular, if f is continuous on $[a, b]$, then F is differentiable on (a, b) with $F' = f$.

Proof. Fix $x_0 \in (a, b)$ where f is continuous. For h small enough that $x_0 + h \in [a, b]$, we have

$$\frac{F(x_0 + h) - F(x_0)}{h} = \frac{1}{h} \int_{x_0}^{x_0+h} f(t)\ \mathrm{d}t.$$

Assume $h > 0$ first. Let

$$m(h) := \inf_{t \in [x_0, x_0+h]} f(t), \qquad M(h) := \sup_{t \in [x_0, x_0+h]} f(t).$$

Then

$$m(h)\,h \le \int_{x_0}^{x_0+h} f(t)\,\mathrm{d}t \le M(h)\,h,$$

so dividing by $h > 0$ gives

$$m(h) \le \frac{F(x_0+h) - F(x_0)}{h} \le M(h).$$

Because f is continuous at x_0, the oscillation on $[x_0, x_0+h]$ shrinks as $h \to 0^+$, and specifically

$$m(h) \to f(x_0), \qquad M(h) \to f(x_0) \qquad (h \to 0^+).$$

By the squeeze theorem,

$$\lim_{h\to 0^+} \frac{F(x_0+h) - F(x_0)}{h} = f(x_0).$$

For $h < 0$, apply the same argument on $[x_0 + h, x_0]$ and note that

$$\frac{F(x_0+h) - F(x_0)}{h} = \frac{1}{h}\int_{x_0}^{x_0+h} f(t)\,\mathrm{d}t = \frac{1}{h}\left(-\int_{x_0+h}^{x_0} f(t)\,\mathrm{d}t\right),$$

which again is squeezed between the infimum and supremum on the interval, leading to the same limit $f(x_0)$ as $h \to 0^-$. Therefore the two-sided derivative exists and equals $f(x_0)$. □

8.3.3 Fundamental Theorem of Calculus, Part II

Theorem 8.28 (FTC II: integrating a derivative)**.** *Let $F : [a, b] \to \mathbb{R}$ be continuous on $[a, b]$ and differentiable on (a, b). Assume F' is Riemann integrable on $[a, b]$. Then*

$$\int_a^b F'(x)\,\mathrm{d}x = F(b) - F(a).$$

Proof. Let $P = \{x_0, \ldots, x_n\}$ be any partition of $[a, b]$. For each i, apply the Mean Value Theorem to F on $[x_{i-1}, x_i]$ to obtain $c_i \in (x_{i-1}, x_i)$ such that

$$F(x_i) - F(x_{i-1}) = F'(c_i)\,(x_i - x_{i-1}) = F'(c_i)\,\Delta x_i.$$

Summing over i gives a telescoping sum:

$$F(b) - F(a) = \sum_{i=1}^{n} F'(c_i)\,\Delta x_i.$$

The right-hand side is precisely a Riemann sum $S(F', P, \xi)$ with tags $\xi_i = c_i$. Since F'

is integrable, by Theorem 8.10 these Riemann sums converge to $\int_a^b F'(x)\,\mathrm{d}x$ as $\|P\| \to 0$. Therefore $F(b) - F(a)$ equals the integral of F'. □

Remark 8.29. The hypothesis "F' is integrable" is automatic if F' is continuous. In many calculus settings, F' is assumed continuous, and then FTC II follows immediately.

8.3.4 Change of Variables (Substitution)

Theorem 8.30 (Change of variables)**.** *Let $\varphi : [\alpha, \beta] \to [a, b]$ be continuously differentiable and strictly monotone, with $\varphi(\alpha) = a$ and $\varphi(\beta) = b$. Let $f : [a, b] \to \mathbb{R}$ be continuous. Then*

$$\int_a^b f(x)\,\mathrm{d}x = \int_\alpha^\beta f(\varphi(t))\,\varphi'(t)\,\mathrm{d}t.$$

Proof. Let $F(x) := \int_a^x f(u)\,\mathrm{d}u$. By Theorem 8.27, F is differentiable with $F'(x) = f(x)$. Consider the composition $H(t) := F(\varphi(t))$. By the chain rule,

$$H'(t) = F'(\varphi(t))\,\varphi'(t) = f(\varphi(t))\,\varphi'(t).$$

Integrate H' over $[\alpha, \beta]$ and apply Theorem 8.28:

$$\int_\alpha^\beta f(\varphi(t))\,\varphi'(t)\,\mathrm{d}t = \int_\alpha^\beta H'(t)\,\mathrm{d}t = H(\beta) - H(\alpha) = F(b) - F(a) = \int_a^b f(x)\,\mathrm{d}x.$$

□

8.3.5 Integration by Parts

Theorem 8.31 (Integration by parts)**.** *Let $u, v : [a, b] \to \mathbb{R}$ be continuously differentiable. Then*

$$\int_a^b u(x)\,v'(x)\,\mathrm{d}x = u(b)v(b) - u(a)v(a) - \int_a^b u'(x)\,v(x)\,\mathrm{d}x.$$

Proof. By the product rule, $(uv)' = u'v + uv'$. Integrate both sides and apply Theorem 8.28:

$$\int_a^b (u(x)v(x))'\,\mathrm{d}x = u(b)v(b) - u(a)v(a) = \int_a^b u'(x)v(x)\,\mathrm{d}x + \int_a^b u(x)v'(x)\,\mathrm{d}x.$$

Rearrange to obtain the formula. □

8.3.6 A Visual: Upper/Lower Sums Converging to Area

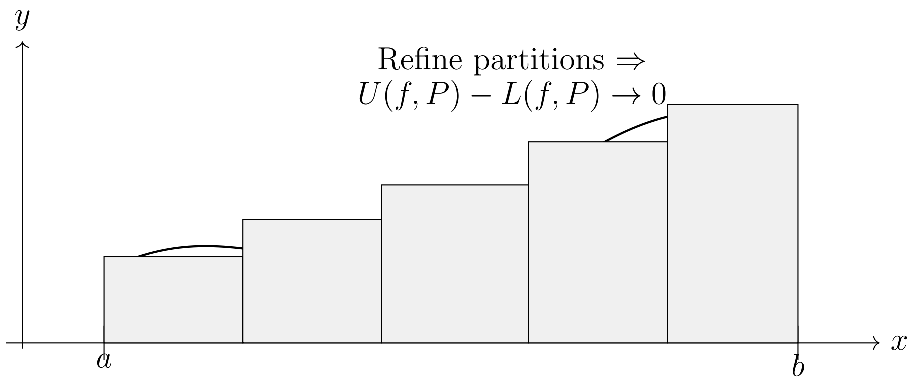


Figure 19: As partitions are refined, lower and upper rectangle approximations squeeze toward the true integral.

8.3.7 Applications and Quick Computations (with Full Justification)

Example 8.32 (A classic consequence: $\int_a^b x^n \, \mathrm{d}x$)**.** For $n \in \mathbb{N}_0$, define $F(x) = \frac{x^{n+1}}{n+1}$. Then $F'(x) = x^n$ is continuous, hence integrable, and by FTC II,

$$\int_a^b x^n \, \mathrm{d}x = F(b) - F(a) = \frac{b^{n+1} - a^{n+1}}{n+1}.$$

Example 8.33 (Substitution)**.** Let f be continuous and compute $\int_0^1 2t\, f(t^2) \, \mathrm{d}t$. Set $\varphi(t) = t^2$, so $\varphi'(t) = 2t$ and $\varphi([0,1]) = [0,1]$. By Theorem 8.30,

$$\int_0^1 2t\, f(t^2) \, \mathrm{d}t = \int_0^1 f(x) \, \mathrm{d}x.$$

Example 8.34 (Integration by parts)**.** If u, v are C^1, then

$$\int_a^b x\, f'(x) \, \mathrm{d}x = \big[x f(x)\big]_a^b - \int_a^b 1 \cdot f(x) \, \mathrm{d}x = b f(b) - a f(a) - \int_a^b f(x) \, \mathrm{d}x.$$

This is Theorem 8.31 with $u(x) = x$ and $v(x) = f(x)$.

Exercise 8.35. Let $f : [a,b] \to \mathbb{R}$ be continuous and define $F(x) = \int_a^x f$. Show that F is differentiable on (a,b) and that F is constant on $[a,b]$ if and only if $f \equiv 0$.

Exercise 8.36. Assume f is continuous on $[a,b]$ and $\int_a^b f(x) \, \mathrm{d}x = 0$. Is it necessarily true that $f(x) = 0$ for all x? If not, give a counterexample; if yes, state the additional hypotheses needed.

Exercise 8.37. Prove that if f is integrable on $[a, b]$ and $f(x) \geq 0$ for all x, then $\int_a^b f = 0$ implies: for every $\varepsilon > 0$, the set $\{x \in [a, b] : f(x) \geq \varepsilon\}$ can be covered by finitely many intervals of arbitrarily small total length. (Interpretation: a nonnegative integrable function with zero integral must be "small almost everywhere" in a Riemann sense.)

9 Sequences of Functions and Uniform Convergence

Key Idea. For sequences of real numbers, convergence is about controlling a single tail. For sequences of functions, there are *two* variables: the index n and the point x. Uniform convergence is the correct notion when we want to pass limits through *continuous operations* such as continuity, integration, and differentiation.

9.1 Pointwise vs. Uniform Convergence

9.1.1 Two Notions of Convergence

Definition 9.1 (Pointwise convergence)**.** Let $D \subseteq \mathbb{R}$ and let (f_n) be a sequence of functions $f_n : D \to \mathbb{R}$. We say $f_n \to f$ *pointwise* on D if for every $x \in D$,

$$\lim_{n\to\infty} f_n(x) = f(x).$$

Equivalently: for every $x \in D$ and every $\varepsilon > 0$, there exists $N = N(\varepsilon, x)$ such that $n \geq N \Rightarrow |f_n(x) - f(x)| < \varepsilon$.

Definition 9.2 (Uniform convergence)**.** Let $D \subseteq \mathbb{R}$ and $f_n, f : D \to \mathbb{R}$. We say $f_n \to f$ *uniformly* on D if

$$\forall \varepsilon > 0 \ \exists N = N(\varepsilon) \ \forall n \geq N \ \forall x \in D, \quad |f_n(x) - f(x)| < \varepsilon.$$

The key difference from pointwise convergence is that N depends on ε but *not* on x.

Remark 9.3. Uniform convergence is strictly stronger than pointwise convergence:

$$f_n \to f \text{ uniformly} \quad \Rightarrow \quad f_n \to f \text{ pointwise}.$$

The converse is false in general.

9.1.2 The Sup Norm Viewpoint

Definition 9.4 (Sup norm)**.** If $f : D \to \mathbb{R}$ is bounded, define

$$\|f\|_{\infty,D} \coloneqq \sup_{x\in D} |f(x)| \in [0, \infty).$$

When D is clear, we write $\|f\|_\infty$.

Proposition 9.5 (Uniform convergence $\Leftrightarrow$ sup norm convergence)**.** *Assume $f_n - f$ is bounded on D for all n. Then*

$$f_n \to f \text{ uniformly on } D \quad \Longleftrightarrow \quad \|f_n - f\|_{\infty,D} \to 0.$$

Proof. By definition,

$$\|f_n - f\|_{\infty,D} = \sup_{x\in D}|f_n(x) - f(x)|.$$

The statement $\|f_n - f\|_{\infty,D} \to 0$ means: for every $\varepsilon > 0$ there exists N such that $n \geq N$ implies $\sup_{x\in D}|f_n(x) - f(x)| < \varepsilon$, which is equivalent to $|f_n(x) - f(x)| < \varepsilon$ for all $x \in D$. This is exactly uniform convergence. $\square$

9.1.3 Uniform Cauchy Criterion

Definition 9.6 (Uniformly Cauchy)**.** A sequence (f_n) on D is *uniformly Cauchy* if

$$\forall\varepsilon > 0\ \exists N\ \forall m, n \geq N\ \forall x \in D, \quad |f_n(x) - f_m(x)| < \varepsilon.$$

Equivalently, $\|f_n - f_m\|_{\infty,D} \to 0$ as $m, n \to \infty$.

Theorem 9.7 (Completeness under uniform convergence)**.** *Let $D \subseteq \mathbb{R}$. If (f_n) is uniformly Cauchy on D, then there exists a function $f : D \to \mathbb{R}$ such that $f_n \to f$ uniformly on D.*

Proof. Fix $x \in D$. The condition "uniformly Cauchy" implies in particular that $(f_n(x))$ is a Cauchy sequence in $\mathbb{R}$: for every $\varepsilon > 0$ there exists N such that for $m, n \geq N$,

$$|f_n(x) - f_m(x)| < \varepsilon.$$

Since $\mathbb{R}$ is complete, $(f_n(x))$ converges to some real number; define

$$f(x) := \lim_{n\to\infty} f_n(x).$$

We claim $f_n \to f$ uniformly. Let $\varepsilon > 0$ and choose N from the uniform Cauchy property so that $|f_n(x) - f_m(x)| < \varepsilon$ for all $m, n \geq N$ and all $x \in D$. Fix $n \geq N$ and $x \in D$. Let $m \to \infty$; then $f_m(x) \to f(x)$, and by taking limits in $|f_n(x) - f_m(x)| < \varepsilon$, we obtain

$$|f_n(x) - f(x)| \leq \varepsilon.$$

This holds for all $x \in D$, hence $\|f_n - f\|_{\infty,D} \leq \varepsilon$ for all $n \geq N$. Therefore $f_n \to f$ uniformly on D. $\square$

9.1.4 A Canonical Example: Pointwise but Not Uniform

Example 9.8 (The sequence x^n on $[0,1]$)**.** Let $f_n : [0,1] \to \mathbb{R}$ be $f_n(x) = x^n$. Then $f_n \to f$ pointwise on $[0,1]$, where

$$f(x) = \begin{cases} 0, & 0 \le x < 1, \\ 1, & x = 1. \end{cases}$$

However, f_n does *not* converge uniformly to f on $[0,1]$.

Proof. **Pointwise convergence.** Fix $x \in [0,1)$. Then $0 \le x < 1$, so $x^n \to 0$ as $n \to \infty$ (geometric decay). At $x = 1$, $f_n(1) = 1$ for all n, so the limit is 1.

Failure of uniform convergence. Take $\varepsilon = \frac{1}{2}$. For each n, choose

$$x_n := \left(\tfrac{1}{2}\right)^{1/n} \in (0,1).$$

Then $f(x_n) = 0$ (since $x_n < 1$) and

$$|f_n(x_n) - f(x_n)| = |x_n^n - 0| = \tfrac{1}{2}.$$

Thus for every N we can take $n \ge N$ and find an $x = x_n$ with $|f_n(x) - f(x)| = \frac{1}{2} \not< \frac{1}{2}$. So no single N works for all x, and uniform convergence fails. $\square$

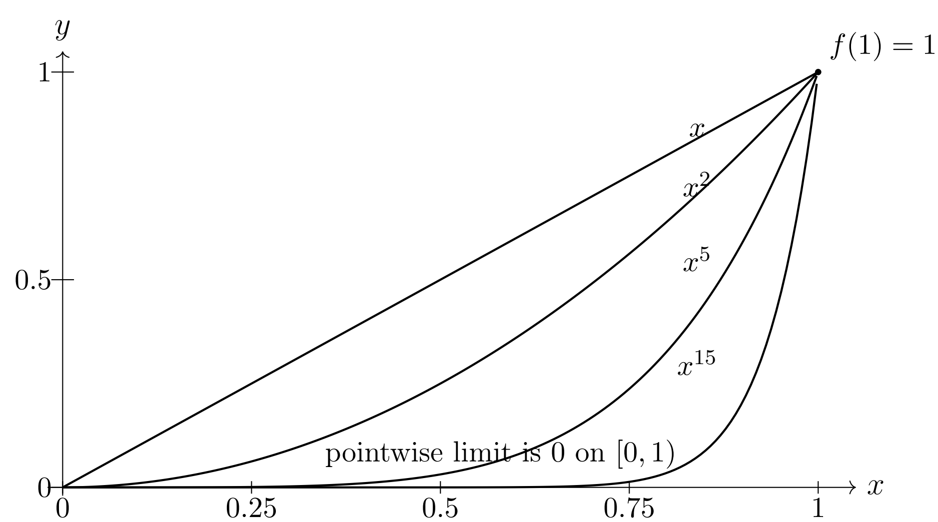


Figure 20: The functions x^n collapse toward 0 on $[0,1)$, but near $x = 1$ the convergence cannot be made uniform.

9.1.5 A Simple Uniform Convergence Example

Example 9.9 (A uniformly convergent sequence)**.** Let $f_n : [0,1] \to \mathbb{R}$ be

$$f_n(x) = \frac{x}{1+nx}.$$

Then $f_n \to 0$ uniformly on $[0,1]$.

Proof. For $x \in [0,1]$, we have $0 \leq f_n(x) \leq \frac{x}{nx} = \frac{1}{n}$ for $x > 0$, and also $f_n(0) = 0$. Hence

$$\|f_n - 0\|_{\infty,[0,1]} = \sup_{x\in[0,1]} \frac{x}{1+nx} = \frac{1}{1+n} \leq \frac{1}{n} \to 0.$$

So $f_n \to 0$ uniformly. $\square$

Exercise 9.10. Show that if $f_n \to f$ uniformly on D, then (f_n) is uniformly Cauchy on D. Prove the converse using Theorem 9.7.

Exercise 9.11. Let $f_n(x) = \frac{nx}{1+nx}$ on $[0,1]$. Find the pointwise limit and decide whether convergence is uniform.

9.2 Uniform Convergence and Continuity

9.2.1 Uniform Limit of Continuous Functions

Theorem 9.12 (Uniform limit theorem for continuity)**.** *Let $D \subseteq \mathbb{R}$ and let $f_n : D \to \mathbb{R}$ be continuous on D for all n. If $f_n \to f$ uniformly on D, then f is continuous on D.*

Proof. Fix $a \in D$ and $\varepsilon > 0$. By uniform convergence, choose N such that

$$\sup_{x\in D}|f_N(x) - f(x)| < \frac{\varepsilon}{3}.$$

Since f_N is continuous at a, choose $\delta > 0$ such that for all $x \in D$ with $|x - a| < \delta$,

$$|f_N(x) - f_N(a)| < \frac{\varepsilon}{3}.$$

Then for $|x - a| < \delta$ (with $x \in D$),

$$\begin{aligned} |f(x) - f(a)| &\leq |f(x) - f_N(x)| + |f_N(x) - f_N(a)| + |f_N(a) - f(a)| \\ &< \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon. \end{aligned}$$

This is exactly continuity of f at a. Since a was arbitrary, f is continuous on D. □

Remark 9.13. Example Theorem 9.8 explains *why* uniform convergence is needed: each x^n is continuous on $[0, 1]$, but the pointwise limit is discontinuous, so the convergence cannot be uniform.

9.2.2 Uniform Convergence on Compact Sets and Uniform Continuity

Corollary 9.14 (Uniform continuity on compact domains)**.** *Let $K \subseteq \mathbb{R}$ be compact. If $f_n : K \to \mathbb{R}$ are continuous and $f_n \to f$ uniformly on K, then f is uniformly continuous on K.*

Proof. By Theorem 9.12, f is continuous on K. A continuous function on a compact set is uniformly continuous (Heine–Cantor). □

9.2.3 Dini's Theorem (A Powerful Compactness Upgrade)

Theorem 9.15 (Dini's theorem)**.** *Let $K \subseteq \mathbb{R}$ be compact. Let (f_n) be a sequence of continuous functions $f_n : K \to \mathbb{R}$ such that:*

(a) $f_n(x) \to f(x)$ *pointwise on* K*,*

(b) f *is continuous on* K*,*

(c) the convergence is monotone in n*: for each* $x \in K$*, either* $f_1(x) \le f_2(x) \le \cdots \le f(x)$ *for all* n *(increasing case), or* $f_1(x) \ge f_2(x) \ge \cdots \ge f(x)$ *for all* n *(decreasing case).*

Then $f_n \to f$ *uniformly on* K*.*

Proof. We prove the increasing case; the decreasing case follows by applying the result to $-f_n$.

Fix $\varepsilon > 0$ and define

$$E_n := \{x \in K : \; f(x) - f_n(x) < \varepsilon\}.$$

Because $f - f_n$ is continuous (difference of continuous functions), E_n is open in K (relative topology). Also, since $f_n(x) \uparrow f(x)$ pointwise, for every $x \in K$ there exists n with $f(x) - f_n(x) < \varepsilon$, hence

$$K = \bigcup_{n=1}^{\infty} E_n.$$

Thus $\{E_n\}_{n\in\mathbb{N}}$ is an open cover of the compact set K, so there exists a finite subcover $K \subseteq E_{n_1} \cup \cdots \cup E_{n_m}$. Let $N = \max\{n_1, \ldots, n_m\}$. Since f_n is increasing in n, we have $f_N \ge f_{n_j}$ pointwise, hence

$$f(x) - f_N(x) \le f(x) - f_{n_j}(x) < \varepsilon \quad \text{whenever } x \in E_{n_j}.$$

Because the E_{n_j} cover K, it follows that $f(x) - f_N(x) < \varepsilon$ for all $x \in K$, i.e.

$$\sup_{x\in K} |f(x) - f_N(x)| = \sup_{x\in K} \big(f(x) - f_N(x)\big) < \varepsilon.$$

Therefore $f_n \to f$ uniformly on K. $\square$

Exercise 9.16. Let $K = [0, 1]$ and $f_n(x) = x^{1/n}$. Show that $f_n \to 1$ pointwise on $(0, 1]$ and $f_n(0) = 0$ for all n. Decide whether the convergence is uniform on $[0, 1]$ and explain which hypothesis of Dini's theorem fails.

9.3 Uniform Convergence and Integration

9.3.1 Interchanging Limit and Integral

Theorem 9.17 (Uniform convergence and the Riemann integral)**.** *Let (f_n) be a sequence of Riemann integrable functions on $[a,b]$. Assume $f_n \to f$ uniformly on $[a,b]$. Then:*

(a) f is Riemann integrable on $[a,b]$;

(b) $\displaystyle\int_a^b f(x)\,\mathrm{d}x = \lim_{n\to\infty}\int_a^b f_n(x)\,\mathrm{d}x$.

Proof. Let $\varepsilon > 0$. By uniform convergence, choose N such that for all $n \geq N$,

$$\|f_n - f\|_{\infty,[a,b]} < \frac{\varepsilon}{4(b-a)}.$$

Fix such an n. Since f_n is Riemann integrable, by the Darboux criterion there exists a partition P with

$$U(f_n, P) - L(f_n, P) < \frac{\varepsilon}{2}.$$

For each subinterval $[x_{i-1}, x_i]$ of P, the sup and inf of f differ from those of f_n by at most $\|f - f_n\|_\infty$:

$$M_i(f,P) \leq M_i(f_n,P) + \|f - f_n\|_\infty, \qquad m_i(f,P) \geq m_i(f_n,P) - \|f - f_n\|_\infty.$$

Hence

$$\begin{aligned}
U(f,P) - L(f,P) &= \sum_{i=1}^{n}\big(M_i(f,P) - m_i(f,P)\big)\Delta x_i \\
&\leq \sum_{i=1}^{n}\Big(\big(M_i(f_n,P) - m_i(f_n,P)\big) + 2\|f - f_n\|_\infty\Big)\Delta x_i \\
&= \big(U(f_n,P) - L(f_n,P)\big) + 2\|f - f_n\|_\infty(b-a) < \frac{\varepsilon}{2} + 2\cdot\frac{\varepsilon}{4(b-a)}(b-a) = \varepsilon.
\end{aligned}$$

So $U(f,P) - L(f,P) < \varepsilon$, and by the Darboux criterion f is integrable. This proves (a).

For (b), use the estimate

$$\left|\int_a^b f(x)\,\mathrm{d}x - \int_a^b f_n(x)\,\mathrm{d}x\right| = \left|\int_a^b (f(x) - f_n(x))\,\mathrm{d}x\right| \leq \int_a^b |f(x)-f_n(x)|\,\mathrm{d}x \leq (b-a)\|f-f_n\|_\infty,$$

where the last inequality is a direct integral bound. As $n \to \infty$, $\|f - f_n\|_\infty \to 0$, so the difference of integrals tends to 0, proving

$$\int_a^b f = \lim_{n\to\infty} \int_a^b f_n.$$

□

9.3.2 Uniform Convergence of Series and Term-by-Term Integration

Theorem 9.18 (Term-by-term integration for uniformly convergent series)**.** *Let $f_n : [a,b] \to \mathbb{R}$ be Riemann integrable for all n. Assume the series $\sum_{n=1}^{\infty} f_n$ converges uniformly on $[a,b]$ to a function f (i.e. its partial sums converge uniformly to f). Then f is Riemann integrable and*

$$\int_a^b f(x)\,\mathrm{d}x = \sum_{n=1}^{\infty} \int_a^b f_n(x)\,\mathrm{d}x,$$

where the series on the right converges in $\mathbb{R}$.

Proof. Let $s_N = \sum_{n=1}^{N} f_n$. Uniform convergence means $s_N \to f$ uniformly. Each s_N is Riemann integrable as a finite sum of integrable functions. Apply Theorem 9.17 to (s_N):

$$\int_a^b f = \lim_{N\to\infty} \int_a^b s_N = \lim_{N\to\infty} \sum_{n=1}^{N} \int_a^b f_n = \sum_{n=1}^{\infty} \int_a^b f_n,$$

where the last equality is the definition of convergence of a real series via partial sums. □

Exercise 9.19. Give an example of a sequence (f_n) that converges pointwise on $[0,1]$ to 0 but for which $\int_0^1 f_n$ does not converge to 0. Explain precisely why this does not contradict Theorem 9.17.

9.4 Uniform Convergence and Differentiation

9.4.1 Why Differentiation is Subtle

Remark 9.20. Unlike integration, differentiation is not continuous with respect to uniform convergence: it is possible that $f_n \to f$ uniformly while f_n' fails to converge, or converges to something different from f'. To justify exchanging limit and derivative, we need *additional structure*, typically uniform control of derivatives.

9.4.2 A Standard Interchange Theorem

Theorem 9.21 (Uniform convergence of derivatives)**.** *Let (f_n) be a sequence of functions on $[a, b]$ such that:*

(a) each f_n is differentiable on $[a, b]$ and f_n' is continuous on $[a, b]$;

(b) (f_n') converges uniformly on $[a, b]$ to a function g;

(c) there exists $x_0 \in [a, b]$ such that $(f_n(x_0))$ converges in $\mathbb{R}$.

Then:

(i) there exists a function $f : [a, b] \to \mathbb{R}$ such that $f_n \to f$ uniformly on $[a, b]$;

(ii) f is differentiable on $[a, b]$ (on (a, b) in the classical sense) and $f'(x) = g(x)$ for all $x \in [a, b]$;

(iii) in particular, g is continuous (as a uniform limit of continuous functions).

Proof. Fix $x \in [a, b]$. By the Fundamental Theorem of Calculus applied to f_n' (continuous),

$$f_n(x) - f_n(x_0) = \int_{x_0}^{x} f_n'(t)\ \mathrm{d}t.$$

Let $n, m \in \mathbb{N}$. Subtract the corresponding identities:

$$f_n(x) - f_m(x) = \big(f_n(x_0) - f_m(x_0)\big) + \int_{x_0}^{x} \big(f_n'(t) - f_m'(t)\big)\ \mathrm{d}t.$$

Take absolute values and use the integral bound:

$$\begin{aligned} |f_n(x) - f_m(x)| &\le |f_n(x_0) - f_m(x_0)| + \int_{x_0}^{x} |f_n'(t) - f_m'(t)|\ \mathrm{d}t \\ &\le |f_n(x_0) - f_m(x_0)| + (b - a)\,\|f_n' - f_m'\|_{\infty,[a,b]}. \end{aligned}$$

Now take the supremum over $x \in [a, b]$:

$$\|f_n - f_m\|_{\infty,[a,b]} \le |f_n(x_0) - f_m(x_0)| + (b - a) \|f'_n - f'_m\|_{\infty,[a,b]}.$$

By hypothesis, $(f_n(x_0))$ converges, hence is Cauchy, so the first term can be made small for large m, n. Also, uniform convergence of f'_n implies (f'_n) is uniformly Cauchy, so the second term can be made small. Therefore (f_n) is uniformly Cauchy, and by Theorem 9.7 there exists f such that $f_n \to f$ uniformly. This proves (i).

Next, we show $f' = g$. Fix $x \in [a, b]$ and write, again using FTC,

$$f_n(x) - f_n(x_0) = \int_{x_0}^{x} f'_n(t) \, \mathrm{d}t.$$

Let $n \to \infty$. Uniform convergence $f_n \to f$ implies the left-hand side tends to $f(x) - f(x_0)$. Uniform convergence $f'_n \to g$ and Theorem 9.17 imply

$$\int_{x_0}^{x} f'_n(t) \, \mathrm{d}t \to \int_{x_0}^{x} g(t) \, \mathrm{d}t.$$

Hence for all $x \in [a, b]$,

$$f(x) - f(x_0) = \int_{x_0}^{x} g(t) \, \mathrm{d}t.$$

Define $G(x) = \int_{x_0}^{x} g(t) \, \mathrm{d}t$. Since g is continuous (uniform limit of continuous functions), FTC I yields $G'(x) = g(x)$ for $x \in (a, b)$. Because $f(x) = f(x_0) + G(x)$, we obtain $f'(x) = g(x)$ on (a, b). This proves (ii), and (iii) has already been noted. $\square$

9.4.3 A Cautionary Example

Example 9.22 (Uniform convergence does not justify termwise differentiation)**.** Define $f_n : [0, 1] \to \mathbb{R}$ by

$$f_n(x) = \sqrt{x^2 + \frac{1}{n}}.$$

Then $f_n \to f$ uniformly on $[0, 1]$ where $f(x) = x$. Indeed,

$$0 \le f_n(x) - x = \sqrt{x^2 + \frac{1}{n}} - x = \frac{\frac{1}{n}}{\sqrt{x^2 + \frac{1}{n}} + x} \le \frac{\frac{1}{n}}{\sqrt{\frac{1}{n}}} = \frac{1}{\sqrt{n}},$$

so $\|f_n - f\|_\infty \le 1/\sqrt{n} \to 0$.

However,

$$f_n'(x) = \frac{x}{\sqrt{x^2 + \frac{1}{n}}},$$

and while $f_n'(x) \to 1$ for every $x > 0$, we have $f_n'(0) = 0$ for all n. Thus f_n' does not converge pointwise to $f'(x) = 1$ on all of $[0, 1]$, and certainly not uniformly. This illustrates why Theorem 9.21 needs uniform convergence of derivatives (or comparable hypotheses).

Exercise 9.23. Let $f_n(x) = \frac{\sin(nx)}{n}$ on $[0, 2\pi]$. Show that $f_n \to 0$ uniformly, but $f_n'(x) = \cos(nx)$ does not converge pointwise.

9.5 Series of Functions and the Weierstrass M-Test

9.5.1 Uniform Convergence of Series

Definition 9.24 (Series of functions)**.** Let $f_n : D \to \mathbb{R}$. Define partial sums

$$s_N(x) := \sum_{n=1}^{N} f_n(x).$$

We say $\sum_{n=1}^{\infty} f_n$ *converges pointwise* on D if $s_N(x)$ converges for each fixed x. It *converges uniformly* on D if $s_N \to s$ uniformly on D for some function s.

9.5.2 Weierstrass M-Test

Theorem 9.25 (Weierstrass M-test)**.** *Let $f_n : D \to \mathbb{R}$ and suppose there exist numbers $M_n \geq 0$ such that for all $x \in D$,*

$$|f_n(x)| \leq M_n.$$

If the numerical series $\sum_{n=1}^{\infty} M_n$ converges, then $\sum_{n=1}^{\infty} f_n$ converges uniformly (and absolutely) on D.

Proof. Let $s_N = \sum_{n=1}^{N} f_n$. For $m > n$ and any $x \in D$,

$$|s_m(x) - s_n(x)| = |\sum_{k=n+1}^{m} f_k(x)| \leq \sum_{k=n+1}^{m} |f_k(x)| \leq \sum_{k=n+1}^{m} M_k.$$

Taking the supremum over $x \in D$ yields

$$\|s_m - s_n\|_{\infty,D} \leq \sum_{k=n+1}^{m} M_k.$$

Since $\sum M_k$ converges, its tails go to 0, hence (s_N) is uniformly Cauchy. By Theorem 9.7, s_N converges uniformly on D. $\square$

9.5.3 Consequences: Continuity and Integration Term-by-Term

Corollary 9.26 (Uniformly convergent series of continuous functions)**.** *If each f_n is continuous on D and $\sum f_n$ converges uniformly on D, then the sum function s is continuous on D.*

Proof. The partial sums s_N are continuous as finite sums of continuous functions. Uniform convergence $s_N \to s$ and Theorem 9.12 imply s is continuous. $\square$

Corollary 9.27 (Term-by-term integration on $[a, b]$)**.** *If each f_n is Riemann integrable on $[a, b]$ and $\sum f_n$ converges uniformly...*

$$\int_a^b s(x)\ \mathrm{d}x = \sum_{n=1}^{\infty} \int_a^b f_n(x)\ \mathrm{d}x.$$

Proof. This is exactly Theorem 9.18. $\square$

Exercise 9.28. Use the M-test to show that $\sum_{n=1}^{\infty} \frac{\sin(nx)}{n^2}$ converges uniformly on $\mathbb{R}$. Deduce that its sum is continuous on $\mathbb{R}$.

9.6 Power Series as a Primary Source of Uniform Convergence

9.6.1 Uniform Convergence on Compact Subintervals

Theorem 9.29 (Uniform convergence inside the radius). *Let $\sum_{n=0}^{\infty} c_n(x-x_0)^n$ be a power series with radius of convergence $R>0$. Fix r with $0<r<R$ and set $K=[x_0-r,x_0+r]$. Then the power series converges uniformly on K.*

Proof. Because $r<R$, the series $\sum_{n=0}^{\infty}|c_n|r^n$ converges (absolute convergence at any point with $|x-x_0|=r$ is not guaranteed, but convergence of $\sum|c_n|r^n$ follows from the definition of radius; equivalently, apply the root test with $|x-x_0|=r<R$).

For $x\in K$ we have $|x-x_0|\le r$, hence

$$|c_n(x-x_0)^n|\le |c_n|r^n.$$

By the Weierstrass M-test (Theorem 9.25) with $M_n=|c_n|r^n$, the power series converges uniformly on K. $\square$

9.6.2 Term-by-Term Differentiation and Integration of Power Series

Theorem 9.30 (Differentiating and integrating power series). *Let $\sum_{n=0}^{\infty} c_n(x-x_0)^n$ have radius of convergence $R>0$. Define*

$$S(x) \coloneqq \sum_{n=0}^{\infty} c_n(x-x_0)^n \qquad (|x-x_0|<R).$$

Then:

(a) S is differentiable on (x_0-R,x_0+R) and

$$S'(x)=\sum_{n=1}^{\infty} nc_n(x-x_0)^{n-1}, \qquad |x-x_0|<R.$$

(b) S is Riemann integrable on every $[x_0-r,x_0+r]$ with $0<r<R$ and

$$\int_{x_0}^{x} S(t)\ \mathrm{d}t=\sum_{n=0}^{\infty} c_n\int_{x_0}^{x}(t-x_0)^n\ \mathrm{d}t=\sum_{n=0}^{\infty} c_n\frac{(x-x_0)^{n+1}}{n+1}, \qquad |x-x_0|<R.$$

Moreover, the differentiated and integrated series have the same radius of convergence R.

Proof. Fix r with $0<r<R$ and work on $K=[x_0-r,x_0+r]$.

(a) Differentiation. Consider the partial sums

$$S_N(x) = \sum_{n=0}^{N} c_n (x - x_0)^n, \qquad S'_N(x) = \sum_{n=1}^{N} n c_n (x - x_0)^{n-1}.$$

Each S_N is C^1 and S'_N is continuous. We will show $S'_N \to G$ uniformly on K for

$$G(x) := \sum_{n=1}^{\infty} n c_n (x - x_0)^{n-1}.$$

For $x \in K$,

$$|n c_n (x - x_0)^{n-1}| \le n |c_n| \, r^{n-1}.$$

The numerical series $\sum_{n=1}^{\infty} n|c_n| \, r^{n-1}$ converges because the original power series converges absolutely for any ρ with $r < \rho < R$ and one can compare using standard ratio/root test arguments; concretely, since $\sum |c_n|\rho^n$ converges, we have $|c_n|\rho^n \to 0$, and the ratio test implies $\sum n|c_n|r^{n-1}$ converges for $r < \rho < R$. Thus by the M-test, $S'_N \to G$ uniformly on K.

Also, $S_N(x_0) = c_0$ for all N, so $(S_N(x_0))$ converges. Therefore, by Theorem 9.21, $S_N \to S$ uniformly on K and S is differentiable on K with $S' = G$ on the interior. Since $r < R$ was arbitrary, this holds for all $|x - x_0| < R$.

(b) Integration. By Theorem 9.29, $S_N \to S$ uniformly on K, and each S_N is continuous hence Riemann integrable. By Theorem 9.17,

$$\int_{x_0}^{x} S(t) \, \mathrm{d}t = \lim_{N\to\infty} \int_{x_0}^{x} S_N(t) \, \mathrm{d}t = \lim_{N\to\infty} \sum_{n=0}^{N} c_n \int_{x_0}^{x} (t - x_0)^n \, \mathrm{d}t,$$

and the integral of $(t - x_0)^n$ is $\frac{(x-x_0)^{n+1}}{n+1}$. Thus the integrated series representation follows.

The statement about radius of convergence is standard: differentiating or integrating a power series does not change the radius, because multiplying coefficients by n or dividing by $n + 1$ does not affect the $\limsup$ in the root test. □

Exercise 9.31. Let $S(x) = \sum_{n=0}^{\infty} x^n$ for $|x| < 1$. Use Theorem 9.30 to justify differentiating term-by-term and show

$$\sum_{n=1}^{\infty} n x^{n-1} = \frac{1}{(1-x)^2}, \qquad |x| < 1.$$

Then integrate to obtain $\sum_{n=1}^{\infty} \frac{x^n}{n} = -\log(1 - x)$ for $|x| < 1$.

Exercise 9.32. Let $f_n(x) = \frac{x^n}{n}$ on $[0, 1]$. Show that $\sum_{n=1}^{\infty} f_n(x)$ converges for each $x \in [0, 1)$ but not uniformly on $[0, 1)$. Decide whether it converges uniformly on $[0, r]$ for each $r < 1$.

10 Metric Spaces

Key Idea. A metric space is the minimal framework needed to discuss *distance*, *limits*, *continuity*, *Cauchy sequences*, *completeness*, and *compactness*. Many theorems from $\mathbb{R}$ extend verbatim once we identify which properties of $\mathbb{R}$ the proof actually uses (e.g. the triangle inequality, completeness, compactness).

10.1 Metrics and Basic Geometry

10.1.1 Definition and Examples

Definition 10.1 (Metric space)**.** A *metric space* is a pair (X, d) where X is a nonempty set and $d : X \times X \to [0, \infty)$ satisfies, for all $x, y, z \in X$:

(a) **Positivity:** $d(x, y) \geq 0$, and $d(x, y) = 0$ if and only if $x = y$.

(b) **Symmetry:** $d(x, y) = d(y, x)$.

(c) **Triangle inequality:** $d(x, z) \leq d(x, y) + d(y, z)$.

Proposition 10.2 (Reverse triangle inequality)**.** *For all $x, y, z \in X$,*

$$|d(x, z) - d(y, z)| \leq d(x, y).$$

Proof. By the triangle inequality,

$$d(x, z) \leq d(x, y) + d(y, z) \quad \Rightarrow \quad d(x, z) - d(y, z) \leq d(x, y).$$

Swap x and y to get $d(y, z) - d(x, z) \leq d(x, y)$. Combine both inequalities. □

Example 10.3 ($\mathbb{R}$ with the usual metric)**.** On $\mathbb{R}$, $d(x, y) = |x - y|$ is a metric.

Example 10.4 ($\mathbb{R}^n$ with common metrics)**.** For $x, y \in \mathbb{R}^n$ define

$$d_2(x, y) = \left(\sum_{k=1}^{n} (x_k - y_k)^2 \right)^{1/2}, \qquad d_1(x, y) = \sum_{k=1}^{n} |x_k - y_k|, \qquad d_\infty(x, y) = \max_{1 \leq k \leq n} |x_k - y_k|.$$

Each is a metric on $\mathbb{R}^n$.

Example 10.5 (Discrete metric)**.** On any set X, the function

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

is a metric. This space is called *discrete*.

Example 10.6 (Subspace metric)**.** If (X, d) is a metric space and $Y \subseteq X$, then

$$d_Y(y_1, y_2) \coloneqq d(y_1, y_2) \qquad (y_1, y_2 \in Y)$$

is a metric on Y. We call it the *subspace metric*.

10.1.2 Equivalent Metrics

Definition 10.7 (Equivalent metrics)**.** Two metrics d and ρ on the same set X are *(topologically) equivalent* if they generate the same open sets (equivalently, the same convergent sequences).

Proposition 10.8 (Bi-Lipschitz comparison implies equivalence)**.** *If there exist constants $c, C > 0$ such that for all $x, y \in X$,*

$$c\, d(x, y) \leq \rho(x, y) \leq C\, d(x, y),$$

then d and ρ are equivalent.

Proof. Fix $x \in X$ and $r > 0$. If $\rho(x, y) < r$, then $c\, d(x, y) \leq \rho(x, y) < r$, hence $d(x, y) < r/c$, so

$$B_\rho(x, r) \subseteq B_d\Big(x, \frac{r}{c}\Big).$$

Similarly, if $d(x, y) < r$, then $\rho(x, y) \leq C\, d(x, y) < Cr$, so

$$B_d(x, r) \subseteq B_\rho(x, Cr).$$

Thus each ρ-ball contains a d-ball around the same center and vice versa, which implies the induced open sets coincide. $\square$

Remark 10.9. On $\mathbb{R}^n$, the metrics d_1, d_2, d_∞ are equivalent. For instance, for all $v \in \mathbb{R}^n$,

$$\|v\|_\infty \leq \|v\|_2 \leq \sqrt{n}\, \|v\|_\infty, \qquad \|v\|_2 \leq \|v\|_1 \leq \sqrt{n}\, \|v\|_2,$$

which yields inequalities of the form in Theorem 10.8.

10.2 Open Balls, Open Sets, and Closed Sets

10.2.1 Balls and the Induced Topology

Definition 10.10 (Open and closed balls)**.** Let (X,d) be a metric space, $x \in X$, and $r > 0$.

$$B(x,r) \coloneqq \{y \in X : \ d(x,y) < r\}, \qquad \overline{B}(x,r) \coloneqq \{y \in X : \ d(x,y) \le r\}.$$

Proposition 10.11 (Open balls are open)**.** *For every $x \in X$ and $r > 0$, the set $B(x,r)$ is open.*

Proof. Let $y \in B(x,r)$, so $d(x,y) < r$. Define $\varepsilon \coloneqq r - d(x,y) > 0$. If $z \in B(y,\varepsilon)$ then by the triangle inequality,

$$d(x,z) \le d(x,y) + d(y,z) < d(x,y) + \varepsilon = r,$$

so $z \in B(x,r)$. Hence $B(y,\varepsilon) \subseteq B(x,r)$, proving $B(x,r)$ is open. ☐

Definition 10.12 (Open set)**.** A set $U \subseteq X$ is *open* if for every $x \in U$ there exists $r > 0$ such that $B(x,r) \subseteq U$.

Proposition 10.13 (Open set axioms)**.** *(a) $\varnothing$ and X are open.*

(b) Arbitrary unions of open sets are open.

(c) Finite intersections of open sets are open.

Proof. (a) is immediate. (b) If $x \in \bigcup_\alpha U_\alpha$, then $x \in U_{\alpha_0}$ for some α_0; take a ball inside U_{α_0}. (c) If $x \in U \cap V$ with U, V open, choose radii r_U, r_V with $B(x, r_U) \subseteq U$ and $B(x, r_V) \subseteq V$. Then $B(x, \min\{r_U, r_V\}) \subseteq U \cap V$. ☐

10.2.2 Closed Sets and a Sequential Characterization

Definition 10.14 (Closed set)**.** A set $F \subseteq X$ is *closed* if $X \setminus F$ is open.

Proposition 10.15 (Closed $\Leftrightarrow$ sequentially closed)**.** *A set $F \subseteq X$ is closed if and only if it contains the limits of all convergent sequences from F: whenever $x_n \in F$ and $x_n \to x$ in X, we have $x \in F$.*

Proof. ($\Rightarrow$) Suppose F is closed and $x_n \in F$ with $x_n \to x$. If $x \notin F$, then $x \in X \setminus F$, which is open, so there exists $r > 0$ with $B(x,r) \subseteq X \setminus F$. But $x_n \to x$ implies $x_n \in B(x,r)$ for all large n, contradicting $x_n \in F$. Hence $x \in F$.

($\Leftarrow$) Suppose the sequential property holds and $X \setminus F$ is not open. Then there exists $x \in X \setminus F$ such that for every n there exists $x_n \in F$ with $d(x_n, x) < 1/n$. Thus $x_n \to x$ with $x_n \in F$, forcing $x \in F$, a contradiction. Hence $X \setminus F$ is open and F is closed. ☐

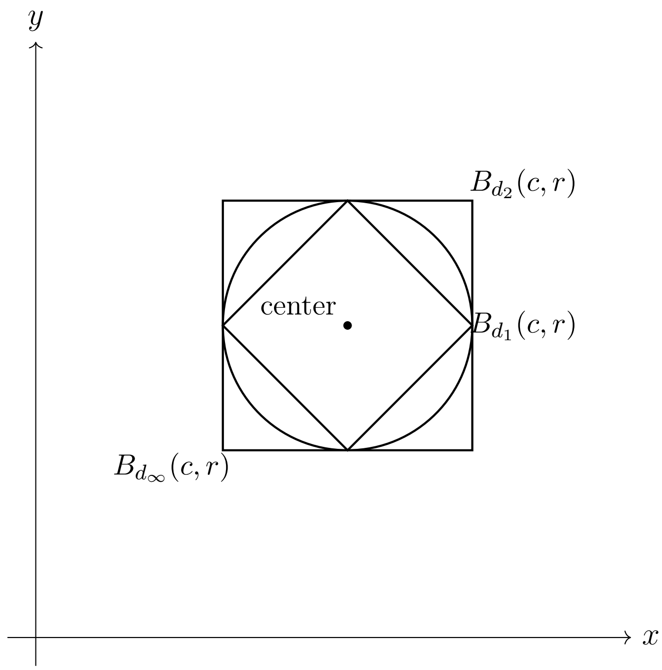


Figure 21: In $\mathbb{R}^2$, the shape of a ball depends on the metric: circle (d_2), diamond (d_1), square (d_∞). These metrics are equivalent, so they generate the same open sets.

10.3 Convergence, Continuity, and Completeness

10.3.1 Convergent and Cauchy Sequences

Definition 10.16 (Convergence)**.** A sequence (x_n) in (X, d) converges to $x \in X$ (written $x_n \to x$) if

$$\forall \varepsilon > 0\ \exists N\ \forall n \geq N, \quad d(x_n, x) < \varepsilon.$$

Proposition 10.17 (Limits are unique)**.** *If $x_n \to x$ and $x_n \to y$, then $x = y$.*

Proof. Fix $\varepsilon > 0$. Choose N_1 with $d(x_n, x) < \varepsilon/2$ for $n \geq N_1$ and N_2 with $d(x_n, y) < \varepsilon/2$ for $n \geq N_2$. Then for $n \geq \max\{N_1, N_2\}$,

$$d(x, y) \leq d(x, x_n) + d(x_n, y) < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

Since ε is arbitrary, $d(x, y) = 0$ and $x = y$. $\square$

Definition 10.18 (Cauchy sequence)**.** A sequence (x_n) is *Cauchy* if

$$\forall \varepsilon > 0\ \exists N\ \forall m, n \geq N, \quad d(x_n, x_m) < \varepsilon.$$

Proposition 10.19 (Convergent $\Rightarrow$ Cauchy). *Every convergent sequence in a metric space is Cauchy.*

Proof. If $x_n \to x$, choose N with $d(x_n, x) < \varepsilon/2$ for $n \geq N$. Then for $m, n \geq N$,

$$d(x_n, x_m) \leq d(x_n, x) + d(x, x_m) < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

□

10.3.2 Continuity

Definition 10.20 (Continuity between metric spaces). Let (X, d_X) and (Y, d_Y) be metric spaces and $f : X \to Y$. We say f is *continuous at* $x \in X$ if

$$\forall \varepsilon > 0 \; \exists \delta > 0 \; \forall x' \in X, \quad d_X(x', x) < \delta \Rightarrow d_Y\big(f(x'), f(x)\big) < \varepsilon.$$

Theorem 10.21 (Sequential characterization of continuity). *f is continuous at x if and only if for every sequence $x_n \to x$ in X, we have $f(x_n) \to f(x)$ in Y.*

Proof. ($\Rightarrow$) Let $x_n \to x$ and fix $\varepsilon > 0$. Choose δ from continuity at x. Then for large n, $d_X(x_n, x) < \delta$, hence $d_Y(f(x_n), f(x)) < \varepsilon$.

($\Leftarrow$) If f is not continuous at x, there exists $\varepsilon_0 > 0$ such that for every $k \in \mathbb{N}$ there is x_k with $d_X(x_k, x) < 1/k$ but $d_Y(f(x_k), f(x)) \geq \varepsilon_0$. Then $x_k \to x$ but $f(x_k) \not\to f(x)$, contradicting the sequential property. □

Theorem 10.22 (Open-set characterization). *A function $f : X \to Y$ is continuous on X if and only if for every open set $U \subseteq Y$, the preimage $f^{-1}(U)$ is open in X.*

Proof. ($\Rightarrow$) Let U be open in Y and $x \in f^{-1}(U)$. Then $f(x) \in U$, so there exists $\varepsilon > 0$ with $B_Y(f(x), \varepsilon) \subseteq U$. Continuity gives $\delta > 0$ such that $d_X(x', x) < \delta \Rightarrow d_Y(f(x'), f(x)) < \varepsilon$, hence $f(x') \in U$. Thus $B_X(x, \delta) \subseteq f^{-1}(U)$, so $f^{-1}(U)$ is open.

($\Leftarrow$) Fix $x \in X$ and $\varepsilon > 0$. The ball $B_Y(f(x), \varepsilon)$ is open, so $f^{-1}(B_Y(f(x), \varepsilon))$ is open and contains x. Thus it contains some $B_X(x, \delta)$, which is exactly continuity at x. □

10.3.3 Completeness

Definition 10.23 (Complete metric space). A metric space (X, d) is *complete* if every Cauchy sequence in X converges to a limit in X.

Example 10.24 ($\mathbb{R}$ is complete; $\mathbb{Q}$ is not)**.** $(\mathbb{R}, |\cdot|)$ is complete. $(\mathbb{Q}, |\cdot|)$ is not complete: there are Cauchy sequences of rationals converging in $\mathbb{R}$ to irrational limits (e.g. to $\sqrt{2}$).

Proposition 10.25 (Closed subspaces of complete spaces are complete)**.** *Let (X, d) be complete and let $F \subseteq X$ with the subspace metric. Then F is complete if and only if F is closed in X.*

Proof. ($\Leftarrow$) If F is closed and (x_n) is Cauchy in F, then it is Cauchy in X and hence converges to some $x \in X$. Since F is closed and $x_n \in F$, Theorem 10.15 implies $x \in F$. Thus F is complete.

($\Rightarrow$) If F is complete and $x_n \in F$ with $x_n \to x$ in X, then (x_n) is Cauchy in F. Completeness gives a limit $y \in F$ with $x_n \to y$ in F, hence also in X. By uniqueness of limits (Theorem 10.17), $x = y \in F$, so F is closed. $\square$

10.4 Compactness in Metric Spaces

10.4.1 Definitions

Definition 10.26 (Compactness (open covers))**.** A subset $K \subseteq X$ is *compact* if every open cover of K has a finite subcover: if $\{U_\alpha\}_{\alpha \in A}$ are open in X and

$$K \subseteq \bigcup_{\alpha \in A} U_\alpha,$$

then there exist $\alpha_1, \dots, \alpha_m$ such that

$$K \subseteq U_{\alpha_1} \cup \cdots \cup U_{\alpha_m}.$$

Definition 10.27 (Sequential compactness)**.** A set $K \subseteq X$ is *sequentially compact* if every sequence in K has a convergent subsequence whose limit lies in K.

10.4.2 Compactness and Sequential Compactness

Lemma 10.28 (Infinite sets in compact metric spaces have accumulation points)**.** *If K is compact and $A \subseteq K$ is infinite, then A has an accumulation point in K: there exists $x \in K$ such that every ball $B(x, r)$ contains infinitely many points of A.*

Proof. Assume for contradiction that A has no accumulation point in K. Then for each $x \in K$ there exists $r_x > 0$ such that $B(x, r_x)$ contains only finitely many points of A. The family $\{B(x, r_x)\}_{x \in K}$ is an open cover of K, hence has a finite subcover $K \subseteq \bigcup_{j=1}^{m} B(x_j, r_{x_j})$. But then

$$A \subseteq \bigcup_{j=1}^{m} \big(A \cap B(x_j, r_{x_j})\big),$$

a finite union of finite sets, hence finite, contradicting that A is infinite. $\square$

Theorem 10.29 (Compact $\iff$ sequentially compact (metric spaces))**.** *Let* (X, d) *be a metric space and* $K \subseteq X$*. Then*

$$K \text{ is compact} \quad \iff \quad K \text{ is sequentially compact.}$$

Proof. ($\Rightarrow$) Assume K is compact and let (x_n) be a sequence in K. If $\{x_n : n \in \mathbb{N}\}$ is finite, some value repeats infinitely often, giving a constant subsequence. If it is infinite, apply Theorem 10.28 to the infinite set $A = \{x_n : n \in \mathbb{N}\}$ to obtain an accumulation point $x \in K$. For each $k \in \mathbb{N}$, the ball $B(x, 1/k)$ contains infinitely many points of A, so we can choose indices $n_1 < n_2 < \cdots$ such that $x_{n_k} \in B(x, 1/k)$. Then $d(x_{n_k}, x) < 1/k \to 0$, so $x_{n_k} \to x$ with $x \in K$. Hence K is sequentially compact.

($\Leftarrow$) Assume K is sequentially compact and let $\{U_\alpha\}_{\alpha \in A}$ be an open cover of K. Suppose no finite subcover exists. We inductively construct a sequence (x_n) in K as follows: choose $x_1 \in K$ and pick U_{α_1} with $x_1 \in U_{\alpha_1}$. Since $\{U_\alpha\}$ has no finite subcover, $K \setminus U_{\alpha_1} \neq \varnothing$, so choose $x_2 \in K \setminus U_{\alpha_1}$ and pick U_{α_2} with $x_2 \in U_{\alpha_2}$. Continue: having chosen x_n and U_{α_n}, since $U_{\alpha_1} \cup \cdots \cup U_{\alpha_n}$ does not cover K, choose

$$x_{n+1} \in K \setminus \big(U_{\alpha_1} \cup \cdots \cup U_{\alpha_n}\big),$$

and then select $U_{\alpha_{n+1}}$ containing x_{n+1}.

By construction, for each fixed j, all terms x_n with $n > j$ lie outside U_{α_j}. If (x_{n_k}) were a convergent subsequence with limit $x \in K$, then $x \in U_{\alpha_j}$ for some j because the sets cover K. Since U_{α_j} is open, eventually $x_{n_k} \in U_{\alpha_j}$, contradicting that all x_n for $n > j$ lie outside U_{α_j}. Thus the sequence has no convergent subsequence, contradicting sequential compactness. Therefore a finite subcover must exist and K is compact. $\square$

10.4.3 Total Boundedness and a Useful Characterization

Definition 10.30 (Totally bounded)**.** A subset $K \subseteq X$ is *totally bounded* if for every $\varepsilon > 0$ there exist finitely many points $x_1, \ldots, x_N \in X$ such that

$$K \subseteq \bigcup_{j=1}^{N} B(x_j, \varepsilon).$$

Proposition 10.31 (Compact $\Rightarrow$ totally bounded)**.** *If K is compact, then K is totally bounded.*

Proof. Fix $\varepsilon > 0$. The family $\{B(x, \varepsilon)\}_{x \in K}$ is an open cover of K. Compactness gives a finite subcover, which is exactly total boundedness. $\square$

Proposition 10.32 (Compact $\Rightarrow$ complete)**.** *If K is compact (with the subspace metric), then K is complete.*

Proof. By Theorem 10.29, K is sequentially compact. Let (x_n) be Cauchy in K. Sequential compactness gives a convergent subsequence $x_{n_k} \to x \in K$. Fix $\varepsilon > 0$ and choose N such that $d(x_n, x_m) < \varepsilon/2$ for all $n, m \geq N$. Choose k with $n_k \geq N$ and $d(x_{n_k}, x) < \varepsilon/2$. Then for $n \geq N$,

$$d(x_n, x) \leq d(x_n, x_{n_k}) + d(x_{n_k}, x) < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon,$$

so $x_n \to x$. Thus every Cauchy sequence converges in K. $\square$

Theorem 10.33 (Complete + totally bounded $\iff$ compact)**.** *A subset $K \subseteq X$ is compact if and only if it is complete and totally bounded (with the subspace metric).*

Proof. ($\Rightarrow$) If K is compact, it is totally bounded by Theorem 10.31 and complete by Theorem 10.32.

($\Leftarrow$) Assume K is complete and totally bounded. Let (x_n) be any sequence in K. We will build a Cauchy subsequence.

By total boundedness with $\varepsilon = 1$, cover K by finitely many balls of radius 1. One such ball contains infinitely many terms of (x_n); pass to a subsequence inside it. Next cover K by finitely many balls of radius $1/2$ and pass to a further subsequence contained in one such ball. Continue inductively to obtain nested subsequences; choose a diagonal subsequence (y_k) such that for each m, all sufficiently large k satisfy $y_k \in B(z_m, 2^{-m})$ for some center z_m.

Then for k, ℓ sufficiently large (depending on m), both y_k and y_ℓ lie in the same ball of radius 2^{-m}, hence

$$d(y_k, y_\ell) \leq 2 \cdot 2^{-m} = 2^{-(m-1)}.$$

Given $\varepsilon > 0$, choose m such that $2^{-(m-1)} < \varepsilon$; then $d(y_k, y_\ell) < \varepsilon$ for all large k, ℓ. So (y_k) is Cauchy. Completeness of K implies (y_k) converges to some $y \in K$. Thus every sequence in K has a convergent subsequence with limit in K, meaning K is sequentially compact, hence compact by Theorem 10.29. $\square$

10.5 Continuous Images of Compact Sets

Theorem 10.34 (Continuous image of compact is compact)**.** *Let $f : (X, d_X) \to (Y, d_Y)$ be continuous and let $K \subseteq X$ be compact. Then $f(K)$ is compact in Y.*

Proof. Let $\{V_\alpha\}$ be an open cover of $f(K)$ in Y. Then $\{f^{-1}(V_\alpha)\}$ is an open cover of K in X by Theorem 10.22. By compactness of K, select a finite subcover

$$K \subseteq f^{-1}(V_{\alpha_1}) \cup \cdots \cup f^{-1}(V_{\alpha_m}).$$

Applying f yields

$$f(K) \subseteq V_{\alpha_1} \cup \cdots \cup V_{\alpha_m},$$

so $f(K)$ is compact. □

Corollary 10.35 (Extreme value theorem (metric-space form))**.** *If K is compact and $f : K \to \mathbb{R}$ is continuous, then f attains a maximum and a minimum on K.*

Proof. By Theorem 10.34, $f(K)$ is compact in $\mathbb{R}$. Compact subsets of $\mathbb{R}$ are closed and bounded, hence have min and max, which must be achieved by values of f on K. □

Corollary 10.36 (Heine–Cantor)**.** *If K is compact and $f : K \to (Y, d_Y)$ is continuous, then f is uniformly continuous on K.*

Proof. Assume not. Then there exists $\varepsilon_0 > 0$ and sequences $(x_n), (y_n)$ in K with

$$d_X(x_n, y_n) \to 0, \qquad d_Y\big(f(x_n), f(y_n)\big) \geq \varepsilon_0 \text{ for all } n.$$

By compactness, extract a subsequence $x_{n_k} \to x \in K$. Then

$$d_X(y_{n_k}, x) \leq d_X(y_{n_k}, x_{n_k}) + d_X(x_{n_k}, x) \to 0,$$

so $y_{n_k} \to x$ as well. Continuity gives $f(x_{n_k}) \to f(x)$ and $f(y_{n_k}) \to f(x)$, hence $d_Y(f(x_{n_k}), f(y_{n_k})) \to 0$, contradicting the uniform lower bound ε_0. □

10.6 Function Spaces and Uniform Convergence

10.6.1 The Sup Metric

Example 10.37 (Bounded functions with the sup metric)**.** Let D be any set and let $\mathcal{B}(D)$ be the set of bounded functions $f : D \to \mathbb{R}$. Define

$$d_\infty(f, g) := \sup_{x \in D} |f(x) - g(x)|.$$

Then d_∞ is a metric on $\mathcal{B}(D)$, and convergence in d_∞ is exactly uniform convergence on D.

Proof. Positivity and symmetry are immediate. For the triangle inequality, for each $x \in D$,

$$|f(x) - h(x)| \le |f(x) - g(x)| + |g(x) - h(x)|.$$

Taking $\sup_{x \in D}$ and using $\sup(a_x + b_x) \le \sup a_x + \sup b_x$ gives

$$d_\infty(f, h) \le d_\infty(f, g) + d_\infty(g, h).$$

Finally, $d_\infty(f_n, f) \to 0$ means $\sup_{x \in D} |f_n(x) - f(x)| \to 0$, which is uniform convergence by definition. □

Remark 10.38. This viewpoint explains why uniform convergence is so robust: it is simply convergence in a metric space.

10.7 Contraction Mappings and Fixed Points

10.7.1 Banach Fixed Point Theorem

Definition 10.39 (Contraction)**.** A map $T : (X, d) \to (X, d)$ is a *contraction* if there exists $L \in [0, 1)$ such that

$$d\big(T(x), T(y)\big) \le L\, d(x, y) \qquad \text{for all } x, y \in X.$$

Theorem 10.40 (Banach fixed point theorem)**.** *Let (X, d) be complete and let $T : X \to X$ be a contraction with constant $L \in [0, 1)$. Then:*

(a) T has a unique fixed point $x^* \in X$.

(b) For any $x_0 \in X$, the iteration $x_{n+1} = T(x_n)$ converges to x^*.

(c) Moreover, for all $n \ge 0$,

$$d(x_n, x^*) \le \frac{L^n}{1 - L}\, d(x_1, x_0).$$

Proof. Define $x_{n+1} = T(x_n)$. Then

$$d(x_{n+1}, x_n) = d\big(T(x_n), T(x_{n-1})\big) \le L\, d(x_n, x_{n-1}),$$

so by induction,

$$d(x_{n+1}, x_n) \le L^n\, d(x_1, x_0).$$

For $m > n$, the triangle inequality yields

$$d(x_m, x_n) \le \sum_{k=n}^{m-1} d(x_{k+1}, x_k) \le d(x_1, x_0) \sum_{k=n}^{m-1} L^k \le d(x_1, x_0) \sum_{k=n}^{\infty} L^k$$
$$= d(x_1, x_0)\, \frac{L^n}{1-L}.$$

Since $L^n/(1-L) \to 0$, (x_n) is Cauchy, hence converges to some $x^* \in X$ by completeness.

Because T is Lipschitz with constant L, it is continuous, so

$$T(x^*) = T\Big(\lim_{n\to\infty} x_n\Big) = \lim_{n\to\infty} T(x_n) = \lim_{n\to\infty} x_{n+1} = x^*,$$

so x^* is a fixed point.

If y^* is another fixed point, then

$$d(x^*, y^*) = d(T(x^*), T(y^*)) \le L\, d(x^*, y^*),$$

which forces $d(x^*, y^*) = 0$ and hence $x^* = y^*$.

Finally, letting $m \to \infty$ in the earlier estimate for $d(x_m, x_n)$ gives

$$d(x^*, x_n) \le \frac{L^n}{1-L}\, d(x_1, x_0).$$

□

Example 10.41 (A classical fixed point iteration)**.** Let $T : [0,1] \to [0,1]$ be $T(x) = \cos x$. Then $T([0,1]) \subseteq [0,1]$ and, for $x, y \in [0,1]$, the Mean Value Theorem gives

$$|T(x) - T(y)| = |-\sin(\xi)|\, |x-y| \le (\sin 1)\, |x-y|$$

for some ξ between x and y. Since $\sin 1 < 1$, T is a contraction on the complete space $([0,1], |\cdot|)$. By Theorem 10.40, there is a unique $x^* \in [0,1]$ solving $x^* = \cos x^*$, and the iteration $x_{n+1} = \cos(x_n)$ converges to x^* for every $x_0 \in [0,1]$.

Exercise 10.42. Prove that for any fixed $x_0 \in X$, the function $x \mapsto d(x, x_0)$ is continuous on X. (Hint: use Theorem 10.2.)

Exercise 10.43. In the discrete metric space from Theorem 10.5, characterize the compact subsets.

Exercise 10.44. Let $C([0,1])$ be the space of continuous functions on $[0,1]$ with the sup metric d_∞. Show that if (f_n) is Cauchy in d_∞, then it converges uniformly to a continuous function. (Interpret this as: $C([0,1])$ is complete under d_∞.)

A Common Counterexamples in Real Analysis

Key Idea. In real analysis, many theorems are "best possible." The easiest way to see why a hypothesis is necessary is to know a small library of counterexamples. This appendix collects standard examples, together with short, fully rigorous verifications of the claimed properties.

A.1 Pointwise Convergence Does Not Preserve Continuity

Example A.1 (Continuous f_n with a discontinuous pointwise limit)**.** Let $f_n : [0,1] \to \mathbb{R}$ be $f_n(x) = x^n$. Then each f_n is continuous on $[0,1]$ and

$$f_n(x) \to f(x) := \begin{cases} 0, & 0 \le x < 1, \\ 1, & x = 1, \end{cases}$$

pointwise. The limit function f is discontinuous at 1.

Proof. Fix $x \in [0,1)$. Since $0 \le x < 1$, we have $x^n \to 0$ as $n \to \infty$. Also $f_n(1) = 1$ for all n, so $f_n(1) \to 1$. The function f has $\lim_{x\uparrow 1} f(x) = 0 \neq f(1) = 1$, hence is discontinuous at 1. □

Remark A.2. This example also shows why *uniform* convergence is the correct notion for preserving continuity: a uniform limit of continuous functions is continuous, but a pointwise limit need not be.

A.2 Pointwise Convergence Does Not Commute with Integration

Example A.3 (Pointwise $f_n \to 0$ but integrals do not converge to 0)**.** Define $f_n : [0,1] \to \mathbb{R}$ by

$$f_n(x) = \begin{cases} n, & 0 < x < \frac{1}{n}, \\ 0, & \text{otherwise.} \end{cases}$$

Then f_n is Riemann integrable on $[0,1]$, $f_n(x) \to 0$ for every $x \in [0,1]$, but

$$\int_0^1 f_n(x)\, \mathrm{d}x = 1 \quad \text{for all } n,$$

so $\int_0^1 f_n \not\to 0$.

Proof. **(1) Pointwise convergence.** Fix $x \in (0,1]$. Choose $N > \frac{1}{x}$, so for $n \geq N$ we have $\frac{1}{n} < x$, hence $x \notin (0, 1/n)$ and $f_n(x) = 0$. Also $f_n(0) = 0$ for all n. Thus $f_n(x) \to 0$ for every $x \in [0,1]$.

(2) Riemann integrability for each n. Each f_n is a bounded step function with only finitely many discontinuities (at 0 and $1/n$), hence is Riemann integrable.

(3) Integral value. Since $f_n = n$ on $(0, 1/n)$ and 0 elsewhere,

$$\int_0^1 f_n(x)\, \mathrm{d}x = n \cdot \left(\frac{1}{n} - 0 \right) = 1.$$

□

Remark A.4. The failure occurs because the convergence is not uniform. In fact, $\|f_n\|_\infty = n \to \infty$. Uniform convergence (or other stronger hypotheses) restores interchange of limit and integral.

A.3 Uniform Convergence Does Not Commute with Differentiation

Example A.5 (Uniform $f_n \to f$ but (f_n') does not converge)**.** Let $f_n : [0, 2\pi] \to \mathbb{R}$ be $f_n(x) = \frac{\sin(nx)}{n}$. Then $f_n \to 0$ uniformly on $[0, 2\pi]$, but $f_n'(x) = \cos(nx)$ does not converge pointwise.

Proof. For all x, $|f_n(x)| \leq \frac{1}{n}$, so

$$\|f_n - 0\|_{\infty,[0,2\pi]} \leq \frac{1}{n} \to 0,$$

hence $f_n \to 0$ uniformly.

But $f_n'(x) = \cos(nx)$. At $x = 0$ we have $\cos(n \cdot 0) = 1$ for all n. At $x = \pi$ we have $\cos(n\pi) = (-1)^n$, which does not converge. So f_n' fails to converge pointwise on $[0, 2\pi]$. □

Remark A.6. To justify differentiating a limit $f = \lim f_n$, one typically needs uniform convergence of (f_n') (or a comparable hypothesis) plus a normalization at one point.

A.4 Discontinuity Sets and Riemann Integrability

Example A.7 (A bounded function that is not Riemann integrable)**.** Define $f : [0,1] \to \mathbb{R}$ by $f(x) = 1$ for $x \in \mathbb{Q}$ and $f(x) = 0$ for $x \notin \mathbb{Q}$. Then f is not Riemann integrable on $[0,1]$.

Proof. Every subinterval of $[0, 1]$ contains both rational and irrational numbers, hence on each subinterval the infimum of f is 0 and the supremum is 1. Therefore, for every partition P,

$$L(f, P) = 0, \qquad U(f, P) = 1,$$

so the lower and upper integrals differ. Hence f is not integrable. □

Example A.8 (Countably many discontinuities but still Riemann integrable)**.** Thomae's function $t : [0, 1] \to \mathbb{R}$ defined by

$$t(0) = 1, \qquad t(x) = \begin{cases} \frac{1}{q}, & x = \frac{p}{q} \in (0, 1] \text{ in lowest terms,} \\ 0, & x \notin \mathbb{Q}, \end{cases}$$

is Riemann integrable on $[0, 1]$ and satisfies $\int_0^1 t(x)\, \mathrm{d}x = 0$.

Remark A.9. This example is often paired with the Dirichlet function: both are defined using rationals/irrationals, but the size of the rational values matters dramatically.

A.5 Compactness Pitfalls on the Real Line

Example A.10 (Closed does not imply compact in $\mathbb{R}$)**.** The set $[0, \infty)$ is closed in $\mathbb{R}$ but not compact.

Proof. The sequence $x_n = n$ lies in $[0, \infty)$ but has no convergent subsequence in $\mathbb{R}$ (it diverges to $+\infty$). A compact subset of a metric space must be sequentially compact, so $[0, \infty)$ is not compact. □

Example A.11 (Bounded does not imply compact)**.** The open interval $(0, 1)$ is bounded but not compact.

Proof. The sequence $x_n = \frac{1}{n}$ lies in $(0, 1)$ but converges to $0 \notin (0, 1)$. Thus $(0, 1)$ is not sequentially compact, hence not compact. □

Example A.12 (Continuous functions can fail to be bounded without compactness)**.** The function $f(x) = \frac{1}{x}$ is continuous on $(0, 1)$ but unbounded on $(0, 1)$.

Proof. For $n \in \mathbb{N}$, $f(1/n) = n \to \infty$, so f is unbounded. □

B Standard Epsilon Proof Templates

Key Idea. Most proofs involving limits reduce to the same few patterns: split an error term, estimate each piece, and choose parameters (N, δ, etc.) to control them. This appendix provides reliable templates you can reuse with minimal modification.

B.1 Sequence Limits: ε–N

Template B.1 (To prove $a_n \to L$)**.** *Goal:*

$$\forall \varepsilon > 0 \; \exists N \; \forall n \geq N, \quad |a_n - L| < \varepsilon.$$

Common strategy:

(1) Start with an arbitrary $\varepsilon > 0$.

(2) Manipulate $|a_n - L|$ into something you can control (often by algebra or inequalities).

(3) Choose N large enough so that your bound is $< \varepsilon$ whenever $n \geq N$.

Example B.1 (Geometric decay)**.** If $|r| < 1$, then $r^n \to 0$.

Proof. Fix $\varepsilon > 0$. Since $|r| < 1$, the real sequence $|r|^n$ decreases to 0. Equivalently, using logarithms: choose N such that $N > \frac{\log \varepsilon}{\log |r|}$ (note $\log |r| < 0$), then $|r|^N < \varepsilon$, and for $n \geq N$ we have $|r|^n \leq |r|^N < \varepsilon$. Thus $r^n \to 0$. □

B.2 Function Limits and Continuity: ε–δ

Template B.2 (To prove $\lim_{x \to a} f(x) = L$)**.** *Goal:*

$$\forall \varepsilon > 0 \; \exists \delta > 0 \; \forall x, \quad 0 < |x - a| < \delta \Rightarrow |f(x) - L| < \varepsilon.$$

Typical steps:

(1) Fix $\varepsilon > 0$.

(2) Estimate $|f(x) - L|$ in terms of $|x - a|$.

(3) Choose δ so that the estimate becomes $< \varepsilon$ whenever $|x - a| < \delta$.

Template B.3 (To prove continuity at a)**.** *Show*

$$\forall \varepsilon > 0 \; \exists \delta > 0, \quad |x - a| < \delta \Rightarrow |f(x) - f(a)| < \varepsilon.$$

Often this is done by expressing $f(x) - f(a)$ in a factored form and bounding it.

Example B.2 (Continuity of polynomials)**.** Every polynomial $p(x)$ is continuous on $\mathbb{R}$.

Proof. It suffices to prove continuity for monomials x^k and then use algebra of continuous functions. Fix $a \in \mathbb{R}$ and consider $f(x) = x^k$. For $x \neq a$,

$$x^k - a^k = (x-a)\sum_{j=0}^{k-1} x^{k-1-j}a^j.$$

Let $\varepsilon > 0$ and choose $\delta \leq 1$ so that $|x-a| < \delta$ implies $|x| \leq |a| + 1$. Then for $|x-a| < \delta$,

$$|x^k - a^k| \leq |x-a|\sum_{j=0}^{k-1}|x|^{k-1-j}|a|^j \leq |x-a|\sum_{j=0}^{k-1}(|a|+1)^{k-1-j}|a|^j \leq C(a,k)\,|x-a|,$$

where $C(a,k)$ is a finite constant depending only on a and k. Now choose $\delta < \varepsilon / C(a,k)$. Then $|x-a| < \delta$ implies $|x^k - a^k| < \varepsilon$. So x^k is continuous at a, hence polynomials are continuous. □

B.3 Uniform Continuity: "One δ for All Points"

Template B.4 (To prove uniform continuity on E)**.** *Goal:*

$$\forall \varepsilon > 0\ \exists \delta > 0\ \forall x, y \in E, \quad |x-y| < \delta \Rightarrow |f(x) - f(y)| < \varepsilon.$$

Key difference from ordinary continuity: δ may depend on ε but not *on the base point.*

Example B.3 (A standard non-example)**.** $f(x) = 1/x$ is not uniformly continuous on $(0,1)$.

Proof. Take $\varepsilon = \frac{1}{2}$. For any $\delta > 0$, choose $n \in \mathbb{N}$ such that $\frac{1}{n(n+1)} < \delta$ (possible because $\frac{1}{n(n+1)} \to 0$). Let

$$x = \frac{1}{n+1}, \qquad y = \frac{1}{n}.$$

Then $|x-y| = \frac{1}{n(n+1)} < \delta$, but

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

Hence the uniform continuity condition fails. □

B.4 Uniform Convergence: Triangle Inequality Splitting

Template B.5 (To prove $f_n \to f$ uniformly)**.** *Goal:*

$$\forall \varepsilon > 0\ \exists N\ \forall n \geq N\ \forall x \in D, \quad |f_n(x) - f(x)| < \varepsilon.$$

Two practical methods:

- *Show* $\sup_{x \in D} |f_n(x) - f(x)| \to 0$.
- *Or use a uniform bound: find numbers* $M_n \to 0$ *with* $|f_n(x) - f(x)| \leq M_n$ *for all* $x \in D$.

Template B.6 (The $\varepsilon/3$ trick)**.** *To show a limit statement for* f *when you only know it for* f_N*:*

$$|f(x) - f(a)| \leq |f(x) - f_N(x)| + |f_N(x) - f_N(a)| + |f_N(a) - f(a)|.$$

Choose N *so the first and third terms are* $< \varepsilon/3$ *uniformly, then choose* δ *for continuity of* f_N *to make the middle term* $< \varepsilon/3$.

B.5 Cauchy Criteria

Template B.7 (Cauchy criterion in $\mathbb{R}$)**.** *A sequence* (a_n) *converges in* $\mathbb{R}$ *if and only if it is Cauchy:*

$$\forall \varepsilon > 0\ \exists N\ \forall m, n \geq N, \quad |a_n - a_m| < \varepsilon.$$

Template B.8 (Uniform Cauchy criterion for functions)**.** *A function sequence* (f_n) *converges uniformly on* D *if and only if it is uniformly Cauchy:*

$$\forall \varepsilon > 0\ \exists N\ \forall m, n \geq N\ \forall x \in D, \quad |f_n(x) - f_m(x)| < \varepsilon.$$

C Frequently Used Theorems

Key Idea. This appendix is a concise "toolbox" of theorems used repeatedly across real analysis. Each statement is given in a form that is directly usable in proofs.

C.1 Completeness and Compactness on the Real Line

Theorem C.1 (Completeness / Least Upper Bound Property). *Every nonempty subset $S \subseteq \mathbb{R}$ that is bounded above has a least upper bound $\sup S \in \mathbb{R}$. Equivalently, every Cauchy sequence in $\mathbb{R}$ converges in $\mathbb{R}$.*

Theorem C.2 (Archimedean property). *For every $x \in \mathbb{R}$ there exists $n \in \mathbb{N}$ such that $n > x$. Equivalently, for every $\varepsilon > 0$ there exists $n \in \mathbb{N}$ such that $\frac{1}{n} < \varepsilon$.*

Theorem C.3 (Monotone Convergence Theorem). *Every monotone (nondecreasing or nonincreasing) sequence that is bounded converges in $\mathbb{R}$.*

Theorem C.4 (Bolzano–Weierstrass Theorem). *Every bounded sequence in $\mathbb{R}$ has a convergent subsequence.*

Theorem C.5 (Heine–Borel Theorem). *A set $K \subseteq \mathbb{R}$ is compact if and only if it is closed and bounded.*

Theorem C.6 (Sequential compactness). *A set $K \subseteq \mathbb{R}$ is compact if and only if every sequence in K has a convergent subsequence with limit in K.*

C.2 Continuity

Theorem C.7 (Intermediate Value Theorem). *If $f : [a, b] \to \mathbb{R}$ is continuous and y lies between $f(a)$ and $f(b)$, then there exists $c \in [a, b]$ such that $f(c) = y$.*

Theorem C.8 (Extreme Value Theorem). *If $f : [a, b] \to \mathbb{R}$ is continuous, then f attains a maximum and a minimum on $[a, b]$.*

Theorem C.9 (Heine–Cantor Theorem). *If $K \subseteq \mathbb{R}$ is compact and $f : K \to \mathbb{R}$ is continuous, then f is uniformly continuous on K.*

C.3 Differentiation

Theorem C.10 (Rolle's Theorem). *If $f : [a, b] \to \mathbb{R}$ is continuous on $[a, b]$, differentiable on (a, b), and $f(a) = f(b)$, then there exists $c \in (a, b)$ such that $f'(c) = 0$.*

Theorem C.11 (Mean Value Theorem). *If $f : [a,b] \to \mathbb{R}$ is continuous on $[a,b]$ and differentiable on (a,b), then there exists $c \in (a,b)$ such that*

$$f'(c) = \frac{f(b) - f(a)}{b - a}.$$

Theorem C.12 (Taylor's theorem with remainder (Lagrange form)). *If $f^{(n+1)}$ exists on (a,b) and is continuous on $[a,b]$, then for each $x \in [a,b]$ there exists ξ between a and x such that*

$$f(x) = \sum_{k=0}^{n} \frac{f^{(k)}(a)}{k!}(x-a)^k + \frac{f^{(n+1)}(\xi)}{(n+1)!}(x-a)^{n+1}.$$

C.4 Riemann Integration

Theorem C.13 (Darboux criterion). *A bounded function $f : [a,b] \to \mathbb{R}$ is Riemann integrable if and only if for every $\varepsilon > 0$ there exists a partition P such that*

$$U(f,P) - L(f,P) < \varepsilon.$$

Theorem C.14 (Continuous and monotone functions are integrable). *If f is continuous on $[a,b]$, then f is Riemann integrable on $[a,b]$. If f is monotone on $[a,b]$, then f is Riemann integrable on $[a,b]$.*

Theorem C.15 (Fundamental Theorem of Calculus). *If f is continuous on $[a,b]$ and $F(x) = \int_a^x f(t)\,\mathrm{d}t$, then $F'(x) = f(x)$ for $x \in (a,b)$. Conversely, if F is differentiable on $[a,b]$ with continuous derivative F', then*

$$\int_a^b F'(x)\,\mathrm{d}x = F(b) - F(a).$$

C.5 Uniform Convergence and Function Series

Theorem C.16 (Uniform limit theorem for continuity). *If f_n are continuous on D and $f_n \to f$ uniformly on D, then f is continuous on D.*

Theorem C.17 (Uniform convergence and integration). *If f_n are Riemann integrable on $[a,b]$ and $f_n \to f$ uniformly on $[a,b]$, then f is Riemann integrable and*

$$\int_a^b f = \lim_{n\to\infty} \int_a^b f_n.$$

Theorem C.18 (Weierstrass M-test). *If $|f_n(x)| \le M_n$ for all $x \in D$ and $\sum_{n=1}^{\infty} M_n$ converges, then $\sum_{n=1}^{\infty} f_n$ converges uniformly on D.*

Theorem C.19 (Differentiation under uniform convergence of derivatives)**.** *If $f_n \in C^1([a,b])$, (f'_n) converges uniformly on $[a,b]$ to g, and $f_n(x_0)$ converges for some $x_0 \in [a,b]$, then f_n converges uniformly to a differentiable f with $f' = g$ on $[a,b]$ (in the interior sense).*

C.6 Metric Spaces

Theorem C.20 (Compactness in metric spaces)**.** *In a metric space, compactness is equivalent to sequential compactness. Also, a set is compact if and only if it is complete and totally bounded (with the subspace metric).*

Theorem C.21 (Banach fixed point theorem)**.** *A contraction on a complete metric space has a unique fixed point, and iteration converges to it.*

D Notation Index

Key Idea. This index lists common symbols and conventions used throughout the notes. When reading or writing proofs, always check whether a symbol denotes a *number*, a *set*, or a *function*—many errors come from mixing these roles.

D.1 Sets and Number Systems

$\mathbb{N}$ Positive integers $\{1, 2, 3, \dots\}$ (sometimes $\mathbb{N}_0 = \{0, 1, 2, \dots\}$).

$\mathbb{Z}$ Integers $\{\dots, -2, -1, 0, 1, 2, \dots\}$.

$\mathbb{Q}$ Rational numbers.

$\mathbb{R}$ Real numbers.

$\mathbb{C}$ Complex numbers (if used).

$A \subseteq B$ A is a subset of B.

$A \setminus B$ Set difference: $\{x \in A : x \notin B\}$.

A^c Complement (context-dependent ambient set).

$A \cup B$ Union; $A \cap B$ intersection.

$\varnothing$ Empty set.

$\times$ Cartesian product.

D.2 Intervals and Order

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

$[a, b]$ Closed interval $\{x \in \mathbb{R} : a \le x \le b\}$.

$(a, b]$ **or** $[a, b)$ Half-open intervals.

$|x|$ Absolute value.

$\max\{a, b\}$**,** $\min\{a, b\}$ Maximum / minimum of two numbers.

$\sup S$**,** $\inf S$ Least upper bound / greatest lower bound of a set $S \subseteq \mathbb{R}$.

D.3 Sequences, Limits, and Asymptotics

(a_n) A sequence of real numbers.

$a_n \to L$ Convergence of a sequence to L.

$\lim_{n\to\infty} a_n$ Limit of a sequence (if it exists).

$\limsup a_n$, $\liminf a_n$ Limit superior / inferior of a sequence.

$o(1)$ A quantity that tends to 0 in the relevant limit.

$o(h)$ **as** $h \to 0$ A function $r(h)$ with $\frac{r(h)}{h} \to 0$.

$O(1)$ Bounded (in the relevant limit).

D.4 Functions and Continuity

$f : D \to \mathbb{R}$ Function with domain D.

f^{-1} Inverse function (exists only if f is bijective onto its range).

$f \circ g$ Composition: $(f \circ g)(x) = f(g(x))$.

$\lim_{x\to a} f(x) = L$ Limit of f at a.

f **continuous at** a $\lim_{x\to a} f(x) = f(a)$.

f **uniformly continuous on** E One $\delta = \delta(\varepsilon)$ works for all $x, y \in E$.

D.5 Differentiation

$f'(a)$ Derivative of f at a:

$$f'(a) = \lim_{h\to 0} \frac{f(a+h) - f(a)}{h}.$$

$f^{(n)}$ nth derivative (when it exists).

T_n Taylor polynomial of order n.

D.6 Integration

$\int_a^b f(x)\,\mathrm{d}x$ Riemann integral (when it exists).

$P = \{x_0, \dots, x_n\}$ Partition of $[a, b]$ with $a = x_0 < \cdots < x_n = b$.

$\|P\|$ Mesh of a partition: $\max_i(x_i - x_{i-1})$.

$U(f, P)$, $L(f, P)$ Upper sum and lower sum (Darboux sums).

D.7 Uniform Convergence and Function Spaces

$f_n \to f$ **uniformly on** D $\sup_{x \in D} |f_n(x) - f(x)| \to 0$.

$\|f\|_{\infty,D}$ Sup norm: $\sup_{x \in D} |f(x)|$ (for bounded f).

$\sum_{n=1}^{\infty} f_n$ Series of functions; uniform convergence refers to uniform convergence of partial sums.

D.8 Metric Spaces

(X, d) Metric space.

$B(x, r)$ Open ball $\{y \in X : \ d(x, y) < r\}$.

Cauchy sequence (x_n) with $\forall \varepsilon > 0\, \exists N\, \forall m, n \geq N,\ d(x_n, x_m) < \varepsilon$.

Complete Every Cauchy sequence converges in the space.

Compact Every open cover has a finite subcover (equivalently: sequentially compact in metric spaces).