Department of Mathematics

Indian Institute of Technology Guwahati

Real Analysis

Lecture Notes

MA224

Metric Spaces, Normed Spaces, Functions of Several Variables, and Lebesgue Integration

---

Rajesh K. Srivastava

Department of Mathematics, IIT Guwahati

Undergraduate Course

Updated: August 28, 2026

Preface

These lecture notes were prepared for the course *MA224: Real Analysis* as a unified undergraduate text for classroom use, guided reading, tutorial discussion, and examination preparation. Their central aim is to preserve the mathematical substance of the course materials while strengthening their logical architecture, sharpening the statements of results, regularizing the notation, and presenting the proofs so that the role of each hypothesis is transparent. A further aim is to produce a manuscript that can serve not only as a semester-long companion to the course but also as a dependable bridge from computational calculus to proof-based analysis.

The notes are written for a first serious encounter with real analysis. Their guiding conviction is that the subject is organized by a comparatively small collection of foundational ideas—order, approximation, convergence, continuity, compactness, differentiability, measure, and integration—whose full force becomes visible only when they are formulated with precision. Accordingly, the exposition gives pride of place to exact definitions, explicit quantifiers, careful proof structure, and examples that illuminate the scope and limitations of the theory. Counterexamples are included whenever they clarify why a natural assertion fails once a key assumption is weakened or removed.

The present edition has also been aligned carefully with the standard MA224 syllabus. It begins with the real line, countability, and the structure of sequences, then develops the language of metric spaces, open and closed sets, convergence, continuity, completeness, compactness, and fixed-point methods. It proceeds next to normed linear spaces and the fundamental inequalities that govern much of analysis, followed by sequences and series of functions, pointwise and uniform convergence, and the standard interchange problems involving limits, continuity, differentiation, and integration. A substantial portion of the text is devoted to functions of several variables, including partial derivatives, directional derivatives, Fréchet differentiability, the chain rule, higher-order expansions, and the inverse and implicit function theorems. The final part develops Lebesgue measure and integration, measurable functions, the basic convergence theorems, and the L^p spaces that form one of the principal meeting points of analysis, geometry, and applications.

The text has been shaped by three pedagogical considerations. First, the chapter sequence has been arranged so that conceptual dependencies emerge in a natural order: each major theorem is preceded by the definitions, lemmas, and model examples needed both for its proof and for its interpretation. Second, the exposition has been written so that the global structure of the subject remains visible. The reader is meant to see not merely that a theorem is true, but also why it belongs where it does, how it interacts with surrounding ideas, and which later developments depend on it. Third, the problem sets at the end of the volume have been organized to support several modes of study: routine exercises for consolidation, proof-based questions for conceptual development, and more synthetic problems for review and deeper understanding.

These notes are not intended to replace lectures. Rather, they are meant to complement them by providing a coherent written account of the material in a form suitable for repeated study. Students are encouraged to read actively: to compare examples with definitions, to test the sharpness of hypotheses, to reconstruct arguments independently, and to regard the exercises as part of the theory rather than as an optional appendix. If these notes help the reader see real analysis not as a collection of isolated techniques but as a connected and rigorous language for understanding limiting processes, then they will have achieved their purpose.

Contents

Preface i

Introduction vii

Conventions and Notation ix

I Foundations of Analysis 1

1 The Real Line, Sequences, and Countability 2
1.1 Chapter roadmap . . . 2
1.2 Countable and uncountable sets . . . 3
1.3 Completeness of $\mathbb{R}$ . . . 4
1.4 Sequences in $\mathbb{R}$ . . . 5

2 Metric Spaces and Topology 7
2.1 Chapter roadmap . . . 7
2.2 Metric spaces . . . 7
2.2.1 Metrics and basic examples . . . 7
2.2.2 Open balls and open sets . . . 11
2.2.3 Sequences in metric spaces . . . 14
2.2.4 Closed sets in metric spaces . . . 15
2.2.5 Interior points and interior of a set . . . 17
2.2.6 Closure and limit points . . . 18
2.2.7 Dense subsets . . . 19
2.3 Continuous maps between metric spaces . . . 20
2.3.1 Uniform continuity . . . 22

3 Completeness, Compactness, and Fixed Points 29
3.1 Chapter roadmap . . . 29
3.2 Complete metric spaces . . . 29
3.2.1 Completion of a metric space . . . 33
3.3 Compactness in metric spaces . . . 35
3.3.1 Total boundedness and metric characterizations . . . 37
3.3.2 Heine–Borel theorem in $\mathbb{R}^n$ . . . 37
3.4 The contraction mapping principle . . . 39

II Norms and Function Spaces **43**

4 Normed Linear Spaces and Fundamental Inequalities **44**
4.1 Chapter roadmap . . . 44
4.2 Normed linear spaces and fundamental inequalities . . . 44
4.2.1 Normed linear spaces . . . 44
4.2.2 Geometry of spheres in $(\mathbb{R}^n, \|\cdot\|_p)$ . . . 47
4.2.3 Young's inequality . . . 48
4.2.4 Hölder's inequality . . . 49
4.2.5 Minkowski's inequality . . . 50
4.2.6 Completeness and Banach spaces . . . 52
4.2.7 Bounded linear maps and the operator norm . . . 54
4.2.8 Equivalence of norms in finite dimensions (optional) . . . 54

5 Sequences of Functions and Uniform Convergence **55**
5.1 Chapter roadmap . . . 55
5.2 Uniform convergence . . . 55
5.2.1 Uniform convergence of sequences of functions . . . 55
5.3 Equicontinuity and the Arzelà–Ascoli theorem (brief) . . . 61
5.4 Power series and uniform convergence . . . 62

III Analysis in Several Variables **63**

6 Functions of Several Variables **64**
6.1 Chapter roadmap . . . 64
6.2 Limits and continuity . . . 65
6.2.1 Notation and basic definitions in Euclidean space . . . 65
6.2.2 Limits in Euclidean space . . . 66
6.2.3 Continuity in Euclidean space . . . 67
6.3 Differentiation in $\mathbb{R}^n$ . . . 69
6.3.1 Partial derivatives . . . 70
6.3.2 Directional derivatives . . . 70
6.3.3 Differentiability . . . 72
6.3.4 Chain rule . . . 75
6.3.5 Taylor's theorem . . . 83
6.3.6 Multi-index notation and the general Taylor formula . . . 83
6.3.7 Coordinate changes: a polar-coordinate template . . . 87
6.4 Inverse and implicit function theorems . . . 89
6.4.1 Inverse function theorem . . . 90
6.4.2 Implicit function theorem . . . 95

IV Measure and Integration **100**

7 Lebesgue Measure and Integral **101**

7.1 Chapter roadmap . . . 101
7.2 From Riemann to Lebesgue . . . 102
7.2.1 Limitations of the Riemann integral . . . 102
7.3 Lebesgue outer measure on $\mathbb{R}$ . . . 103
7.4 Lebesgue measurable sets and Lebesgue measure . . . 111
7.4.1 Lebesgue measurable sets . . . 112
7.4.2 Lebesgue measure and structural results . . . 122
7.5 Examples and pathologies . . . 125
7.5.1 The Cantor set . . . 125
7.5.2 Ternary representation of the Cantor set . . . 127
7.5.3 The Cantor set is uncountable . . . 128
7.5.4 Nonmeasurable sets . . . 129
7.6 Measurable functions and simple approximations . . . 132
7.6.1 Measurable functions . . . 133
7.6.2 Simple functions . . . 138
7.7 Lebesgue integration . . . 141
7.7.1 The Lebesgue integral . . . 141
7.7.2 Complex-valued measurable functions and integrals (optional) . . . 148
7.8 Convergence theorems and L^1 consequences . . . 149
7.8.1 Monotone convergence theorem . . . 149
7.8.2 Fatou's lemma . . . 152
7.8.3 Signed integrals and the space L^1 . . . 154
7.8.4 Chebyshev's inequality . . . 155
7.8.5 Dominated convergence theorem . . . 155
7.8.6 Bounded convergence theorem . . . 157
7.9 Relation with the Riemann integral . . . 157
7.9.1 Riemann and Lebesgue integration on compact intervals . . . 157
7.9.2 Improper Riemann integration . . . 161
7.9.3 Absolute convergence and Lebesgue integrability on unbounded intervals . . . 162
7.9.4 L^p spaces . . . 165
7.10 Further optional topics . . . 168
7.10.1 Egoroff's theorem and Lusin's theorem (optional) . . . 168
7.10.2 Product measures and Fubini–Tonelli (optional) . . . 171
7.10.3 Radon–Nikodym theorem (optional) . . . 172

V Problems and Revision **173**

Problem Set 1: The Real Line, Sequences, and Countability **174**

Problem Set 2: Metric Spaces and Topology **176**

Problem Set 3: Completeness, Compactness, and Fixed Points **178**

Problem Set 4: Normed Linear Spaces and Fundamental Inequalities **180**

Problem Set 5: Sequences of Functions and Uniform Convergence **182**

Problem Set 6: Functions of Several Variables **184**

Problem Set 7: Measure and Integration **191**

Selected References **196**

Introduction

Real analysis furnishes the rigorous foundation of calculus and, more broadly, of modern analysis. Its central concern is the disciplined study of limiting processes: convergence of sequences and functions, continuity and differentiability expressed through limits, and integration built on measurable structure rather than on partitions alone. The present text has been written to keep that logical architecture visible at every stage. Definitions are stated with explicit quantifiers, proofs are organized so that decisive hypotheses are easy to locate, and examples—including counterexamples—are chosen not merely for illustration but for diagnosis: they show what each hypothesis does and what fails when it is absent.

The material is organized around three closely related modules that together form a standard undergraduate introduction to modern analysis:

(1) Metrics and norms. Metric spaces, open and closed sets, convergence and continuity in the metric setting, Cauchy sequences and completeness, compactness in both sequential and open-cover forms, and the basic geometry of normed vector spaces.

(2) Functions of several variables. Limits and continuity in $\mathbb{R}^n$, partial and directional derivatives, Fréchet differentiability, the chain rule, Taylor's theorem with remainder, and the inverse and implicit function theorems.

(3) Lebesgue measure and integral. σ-algebras, measures and measurable spaces, Lebesgue measurable sets and functions, the Lebesgue integral, the fundamental convergence theorems, and the L^p spaces.

The notes are intended to support several styles of use. They may be read linearly alongside lectures; they may also be consulted selectively as a reference when a particular theorem or method is needed later in the course. Each chapter begins with a roadmap, and optional material is labeled as such. The problem sets at the end of the volume are designed to reinforce the text in layers, beginning with exercises that consolidate definitions and standard arguments and progressing toward problems that require synthesis, comparison of concepts, or construction of counterexamples.

A reader will profit most from the text if they already have a firm command of single-variable calculus, elementary linear algebra, and the basic language of sets and functions. What is newly demanded here is not technical machinery alone, but proof-writing: the ability to unpack definitions, control quantifiers, and distinguish local arguments from global ones. For that reason, many proofs are written at a level of detail meant to model good mathematical practice rather than to minimize length.

This edition is intended to serve as a single consolidated reference for the course. In preparing it, earlier handouts, assignments, and quizzes were checked against the present draft; whenever

they contained syllabus-relevant material not already integrated into the main exposition, that material was absorbed into the discussion, examples, and end-of-book problem sets.

The following schematic summarizes the logical movement of the course. It should be read from left to right: the early order and sequence material supplies the language of limiting processes; metric and normed-space arguments provide the abstract framework; the multivariable material develops differentiability as a linear approximation principle; and the measure-theoretic part replaces partition-based integration by a theory stable under limits.

Order and sequences	$\Longrightarrow$	Metric and normed spaces	$\Longrightarrow$	Several variables
Completeness and compactness	$\Longrightarrow$	Measure and measurability	$\Longrightarrow$	Integration and L^p spaces

Figure. A dependency map for the main ideas in the course.

The problem sets are correspondingly arranged in layers. Foundational questions test direct use of definitions; structural questions ask for proofs of implications and equivalences; construction questions develop examples and counterexamples; and synthesis questions require ideas from more than one chapter. This organization is meant to make the text suitable both for weekly study and for final revision.

Conventions and Notation

Throughout the notes, $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$, and $\mathbb{C}$ denote the sets of natural numbers, integers, rational numbers, real numbers, and complex numbers, respectively. Unless stated otherwise, Euclidean spaces are understood with their usual vector-space structure and with one of the equivalent norms introduced later in the text. The symbols $B_r(x)$ and $\overline{B}_r(x)$ denote, respectively, the open and closed balls of radius r centered at x. If $a < b$, the notation $\mathcal{R}[a, b]$ denotes the vector space of Riemann integrable real-valued functions on $[a, b]$. In the measure-theoretic chapters we write $\mathcal{M}$ for the Lebesgue σ-algebra on $\mathbb{R}$ and m for Lebesgue measure.

Proofs are written so that the function of each hypothesis is visible at the moment it is used. Statements involving compactness, completeness, differentiability, or measurability should therefore be read with close attention to the ambient space and to the regularity assumptions imposed on the relevant functions. Whenever a result is valid in a setting more general than the one immediately under discussion, the argument is usually arranged so that the underlying mechanism remains apparent.

For sequences, the index variable is typically denoted by n, and subsequences by (x_{n_k}). For functions defined on subsets of Euclidean space, the notation $Df(a)$ denotes the Fréchet derivative at a when it exists, whereas partial derivatives are written in the usual coordinate form. In the measure-theoretic chapters, equality almost everywhere is always understood with respect to the ambient measure unless a different measure is explicitly specified. All intervals are intervals in $\mathbb{R}$ unless another ordered set is explicitly named, and all vector spaces are over $\mathbb{R}$ unless the scalar field is stated otherwise.

Part I

Foundations of Analysis

Chapter 1

The Real Line, Sequences, and Countability

This chapter gathers several foundational facts that recur throughout the course. We begin with the set-theoretic notions of countability and uncountability, then recall the completeness of $\mathbb{R}$ through the least-upper-bound property. These ideas are applied immediately to the study of sequences in $\mathbb{R}$, including Cauchy sequences, monotone convergence, and the Bolzano–Weierstrass theorem. A central aim is to make explicit where *completeness enters standard existence and convergence arguments.*

Learning objectives.

- distinguish finite, countable, and uncountable sets, and prove that $\mathbb{Q}$ is countable while $\mathbb{R}$ is uncountable;
- use the least-upper-bound property and the Archimedean property as foundational completeness principles on $\mathbb{R}$;
- analyze convergence, boundedness, and monotonicity of real sequences with full control of quantifiers;
- prove and apply the monotone convergence theorem, Bolzano–Weierstrass theorem, and the Cauchy criterion in $\mathbb{R}$.

1.1 Chapter roadmap

This chapter is one of the first places where classical analysis opens into modern analysis. The aim is not merely to collect definitions, but to understand why those definitions are genuinely necessary. The roadmap below is intended to keep that broader perspective visible as the technical details accumulate.

Dependencies. This chapter supplies the sequential and order-theoretic tools used repeatedly in compactness, continuity, and measure theory.

- Countable and uncountable sets; countability of $\mathbb{Q}$ and uncountability of $\mathbb{R}$.
- Sequences in $\mathbb{R}$: limits, boundedness, monotone convergence, and Bolzano–Weierstrass.
- Cauchy sequences and the completeness of $\mathbb{R}$.

1.2 Countable and uncountable sets

Questions of size in analysis are not always numerical; before one measures length, area, or volume, one often first asks whether a set can even be listed. Countability marks the boundary between discrete and genuinely continuous behaviour, and it already shows why the real line is vastly richer than the rationals.

Definition 1.1 (Countability)**.** A set A is *countable* if it is finite or if there exists a bijection $A \to \mathbb{N}$. A set is *uncountable* if it is not countable.

Proposition 1.2. *If $\{A_n\}_{n\geq 1}$ is a sequence of countable sets, then $\bigcup_{n\geq 1} A_n$ is countable.*

Proof. Discard the empty members of the family, since they do not affect the union. For each remaining index n, choose a surjection $\phi_n : \mathbb{N} \to A_n$. Define

$$\Phi : \mathbb{N} \times \mathbb{N} \to \bigcup_{n\geq 1} A_n, \qquad \Phi(n,k) := \phi_n(k),$$

on the nonempty rows. Since $\mathbb{N} \times \mathbb{N}$ is countable (for instance, enumerate pairs along diagonals), the image of Φ is countable. Hence the union is countable. □

Theorem 1.3. *The set $\mathbb{Q}$ is countable, whereas $\mathbb{R}$ is uncountable.*

Proof. Countability of $\mathbb{Q}$. Every rational number can be written as p/q with $p \in \mathbb{Z}$ and $q \in \mathbb{N}$. Since $\mathbb{Z}$ and $\mathbb{N}$ are countable and $\mathbb{Z} \times \mathbb{N}$ is countable, it follows that $\mathbb{Q}$ is countable.

Uncountability of $\mathbb{R}$. It suffices to show that $(0,1)$ is uncountable. Suppose, for contradiction, that

$$(0,1) = \{x_1, x_2, \dots\}.$$

Write each x_n in a decimal expansion

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

choosing the expansion that does not terminate in repeating 9's. Define

$$y = 0.b_1b_2b_3\dots$$

by requiring $b_k \in \{1,2\}$ and $b_k \neq a_{kk}$ for every k. Then $y \in (0,1)$, but $y \neq x_n$ for every n, because the nth decimal digit of y differs from the nth decimal digit of x_n. This contradicts the assumed enumeration. Hence $(0,1)$, and therefore $\mathbb{R}$, is uncountable. □

1.3 Completeness of $\mathbb{R}$

The least-upper-bound property is the structural feature that separates $\mathbb{R}$ from $\mathbb{Q}$. Many existence theorems in analysis are, at bottom, consequences of completeness, so it is useful to isolate this principle before invoking it repeatedly in later arguments.

Theorem 1.4 (Least upper bound property)**.** *Every nonempty subset $E \subset \mathbb{R}$ that is bounded above has a supremum* $\sup E \in \mathbb{R}$.

Proposition 1.5 (Archimedean property)**.** *For every $x \in \mathbb{R}$ there exists $n \in \mathbb{N}$ with $n > x$. Equivalently, for every $\varepsilon > 0$ there exists $n \in \mathbb{N}$ with $1/n < \varepsilon$.*

Proof. If $x \leq 0$, take $n = 1$. If $x > 0$ and $n \leq x$ for all n, then $\mathbb{N}$ is bounded above by x, and therefore has $s = \sup \mathbb{N}$. But then $s - 1$ is not an upper bound, so there exists $m \in \mathbb{N}$ with $m > s - 1$, hence $m + 1 > s$, contradicting the fact that s is an upper bound. □

Theorem 1.6 (Nested interval theorem)**.** *Let $I_n = [a_n, b_n]$ be a nested sequence of nonempty closed intervals, i.e. $I_{n+1} \subset I_n$ for all n. Then $\bigcap_{n\geq 1} I_n \neq \varnothing$. If additionally $b_n - a_n \to 0$, then the intersection consists of a single point.*

Proof. The sequence (a_n) is increasing and bounded above by b_1, hence

$$a := \sup\{a_n : n \geq 1\}$$

exists. Similarly, (b_n) is decreasing and bounded below by a_1, so

$$b := \inf\{b_n : n \geq 1\}$$

exists. Since $a_n \leq b_n$ for every n, it follows that $a \leq b$.

We claim that

$$\bigcap_{n\geq 1} I_n = [a, b].$$

Indeed, if $x \in [a, b]$, then $a_n \leq a \leq x \leq b \leq b_n$ for every n, hence $x \in I_n$ for all n. Conversely, if $x \in \bigcap_n I_n$, then $a_n \leq x \leq b_n$ for every n; taking the supremum on the left and the infimum on the right gives $a \leq x \leq b$. Thus the intersection is nonempty. If, in addition, $b_n - a_n \to 0$, then $0 \leq b - a \leq b_n - a_n$ for every n, whence $a = b$, and the intersection consists of the single point $a = b$. □

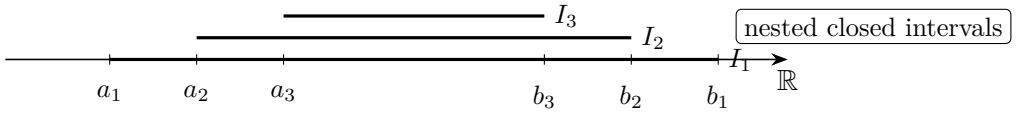


Figure 1.1: A schematic picture for the nested interval theorem.

1.4 Sequences in $\mathbb{R}$

Sequences provide the simplest laboratory in which limiting processes can be studied with full precision. In practice, many arguments in analysis reduce global questions to the behaviour of carefully chosen sequences, so the results collected here will be used throughout the course.

Definition 1.7 (Limit and Cauchy sequence)**.** A sequence (x_n) in $\mathbb{R}$ *converges* to $x \in \mathbb{R}$ if for every $\varepsilon > 0$ there exists N such that $|x_n - x| < \varepsilon$ for all $n \geq N$. It is *Cauchy* if for every $\varepsilon > 0$ there exists N such that $|x_n - x_m| < \varepsilon$ for all $n, m \geq N$.

Theorem 1.8 (Completeness of $\mathbb{R}$)**.** *Every Cauchy sequence in* $\mathbb{R}$ *converges.*

Proof. Let (x_n) be Cauchy. For each $m \in \mathbb{N}$ choose $N_m \in \mathbb{N}$ so large that

$$|x_n - x_k| < 2^{-m} \qquad \text{for all } n, k \geq N_m,$$

and arrange the indices so that $N_1 < N_2 < \cdots$. Define

$$I_m := \left[x_{N_m} - 2^{-m},\, x_{N_m} + 2^{-m}\right].$$

Then every tail element x_n with $n \geq N_m$ lies in I_m, since $|x_n - x_{N_m}| < 2^{-m}$. In particular, for $j \geq m$ we have $x_{N_j} \in I_m$, so the finite intersections

$$J_m := \bigcap_{r=1}^{m} I_r$$

are nonempty closed intervals. Moreover, (J_m) is nested and $J_m \subset I_m$, hence the length of J_m is at most 2^{1-m}. By the nested interval theorem, there exists a unique point

$$x \in \bigcap_{m \geq 1} J_m \subset \bigcap_{m \geq 1} I_m.$$

Now fix $\varepsilon > 0$ and choose m so large that $2^{2-m} < \varepsilon$. If $n \geq N_m$, then both x_n and x belong to I_m, so

$$|x_n - x| \leq 2^{-m} + 2^{-m} = 2^{1-m} < \varepsilon.$$

Thus $x_n \to x$, and the Cauchy sequence (x_n) converges in $\mathbb{R}$. □

Remark 1.9. This proof isolates a reusable template: if a Cauchy sequence has one convergent subsequence, then the entire sequence converges to the same limit. In later chapters the same idea reappears whenever compactness furnishes a subsequence and completeness upgrades it to convergence of the full sequence.

Theorem 1.10 (Monotone convergence)**.** *Every bounded monotone sequence in* $\mathbb{R}$ *converges.*

Proof. We treat the increasing case first. Assume that (x_n) is increasing and bounded above, and set

$$s := \sup\{x_n : n \in \mathbb{N}\}.$$

Given $\varepsilon > 0$, the number $s - \varepsilon$ is not an upper bound for the set $\{x_n : n \in \mathbb{N}\}$, so there exists $N \in \mathbb{N}$ such that

$$x_N > s - \varepsilon.$$

If $n \geq N$, then monotonicity gives $x_N \leq x_n \leq s$, and hence

$$s - \varepsilon < x_n \leq s,$$

which implies $|x_n - s| < \varepsilon$. Therefore $x_n \to s$.

If (x_n) is decreasing and bounded below, then $(-x_n)$ is increasing and bounded above. By the first part, $(-x_n)$ converges in $\mathbb{R}$, and consequently (x_n) also converges. Thus every bounded monotone sequence converges. □

Example 1.11. The sequence $x_n = 1 - \frac{1}{n}$ is increasing and bounded above by 1, so the theorem shows immediately that it converges to $1 = \sup\{x_n : n \in \mathbb{N}\}$.

Theorem 1.12 (Bolzano–Weierstrass)**.** *Every bounded sequence in $\mathbb{R}$ has a convergent subsequence.*

Proof. Let (x_n) be bounded, so that $x_n \in [a_1, b_1]$ for some closed interval. Bisect $[a_1, b_1]$ into two closed intervals and choose $[a_2, b_2]$ to be one that contains infinitely many terms of the sequence. Iterating this construction, we obtain nested closed intervals $[a_k, b_k]$ with $b_k - a_k = 2^{-(k-1)}(b_1 - a_1)$, each containing infinitely many x_n. By Theorem 1.6 there exists $x \in \bigcap_k [a_k, b_k]$. Choose n_k strictly increasing so that $x_{n_k} \in [a_k, b_k]$. Then $x_{n_k} \to x$ because $|x_{n_k} - x| \leq b_k - a_k \to 0$. □

Chapter 2

Metric Spaces and Topology

This chapter introduces metric spaces as the common framework for convergence, continuity, and compactness. After developing the topology induced by a metric—open and closed sets, closure, limit points, and density—we study continuous maps between metric spaces and the basic invariance principles they satisfy.

Learning objectives.

- work fluently with metrics, balls, open and closed sets, closures, interiors, and limit points;
- translate between metric and topological formulations of convergence and continuity;
- compare equivalent metrics and identify when a metric change preserves topology and completeness;
- use continuity to transport compactness, connectedness, and density-type properties between metric spaces.

2.1 Chapter roadmap

Dependencies. The main transition in this chapter is from pointwise distance estimates to genuinely topological statements about openness, closure, and continuity.

- Metrics and canonical examples; the metric topology via open balls.
- Convergence and Cauchy sequences in metric spaces.
- Open/closed sets, interior, closure, limit points, and density.
- Continuous and uniformly continuous maps between metric spaces.

2.2 Metric spaces

A metric abstracts the idea of distance while retaining exactly the structure needed to discuss convergence, continuity, and approximation. The point of this generality is not abstraction for its own sake, but the recognition that many arguments from calculus depend only on distance estimates and not on coordinates.

2.2.1 Metrics and basic examples

Definition 2.1 (Metric space)**.** Let X be a nonempty set. A function $d\colon X \times X \to [0, \infty)$ is called a *metric* on X if, for all $x, y, z \in X$,

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

(ii) $d(x,y) = d(y,x)$ (symmetry).

(iii) $d(x,z) \le d(x,y) + d(y,z)$ (triangle inequality).

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

Example 2.2 (Standard metrics on $\mathbb{R}^n$)**.** For $x = (x_1,\dots,x_n)$ and $y = (y_1,\dots,y_n)$ in $\mathbb{R}^n$, define

$$d_1(x,y) := \sum_{i=1}^{n} |x_i - y_i|,$$
$$d_2(x,y) := \Big(\sum_{i=1}^{n} |x_i - y_i|^2\Big)^{1/2},$$
$$d_\infty(x,y) := \max_{1\le i\le n} |x_i - y_i|.$$

Each of these functions defines a metric on $\mathbb{R}^n$.

Exercise 2.3. Let (X,d) be a metric space. Prove that $d'(x,y) := \min\{1, d(x,y)\}$ defines a metric on X.

Example 2.4 (Supremum metric on $C[0,1]$)**.** Let $X = C[0,1]$, the space of continuous real-valued functions on $[0,1]$. For $f, g \in X$, set

$$d_\infty(f,g) := \sup_{0\le t\le 1} |f(t) - g(t)|.$$

Since every continuous function on $[0,1]$ is bounded, the supremum is finite. The triangle inequality follows from

$$|f(t) - h(t)| \le |f(t) - g(t)| + |g(t) - h(t)|, \qquad t \in [0,1].$$

Thus d_∞ is a metric on $C[0,1]$.

Example 2.5 (Discrete metric)**.** Let $X \neq \emptyset$. Define, for $x, y \in X$,

$$d_0(x,y) := \begin{cases} 1, & x \neq y, \\ 0, & x = y. \end{cases}$$

Then d_0 is a metric on X. The metric space (X, d_0) is called the *discrete metric space*. In particular, every nonempty set admits at least one metric.

To verify the triangle inequality for the discrete metric, it suffices to consider the cases in which two of the points coincide and the case in which all three are distinct; in each case the inequality is immediate.

Example 2.6. Let (X,d) be a metric space. Then

$$\widetilde{d}(x,y) := \frac{d(x,y)}{1 + d(x,y)}$$

defines another metric on X.

Indeed, consider the function $f(t) = \dfrac{t}{1+t}$ on $[0,\infty)$. Since

$$f'(t) = \frac{1}{(1+t)^2} > 0,$$

the function f is strictly increasing and satisfies $f(0) = 0$. Moreover,

$$\frac{t+s}{1+t+s} < \frac{t}{1+t} + \frac{s}{1+s} \qquad (s,t \geq 0),$$

so f is subadditive. Writing $t = d(x,y)$ and $s = d(y,z)$, the triangle inequality for d gives

$$d(x,z) \leq t + s.$$

Since f is increasing, we obtain

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

Hence $\widetilde{d} = f \circ d$ is a metric.

Proposition 2.7 (Changing a metric by a subadditive function)**.** *Let (X,d) be a metric space and let $f\colon [0,\infty) \to [0,\infty)$ be an increasing function such that*

$$f(s+t) \leq f(s) + f(t) \qquad \text{for all } s,t \geq 0, \tag{2.1}$$

and $f(t) = 0$ if and only if $t = 0$. Then $d_f := f \circ d$ defines a metric on X.

Proof. Non-negativity and symmetry are immediate. If $d_f(x,y) = 0$, then $f(d(x,y)) = 0$, hence $d(x,y) = 0$ and $x = y$. Finally, by the triangle inequality for d and the monotonicity and subadditivity (2.1) of f,

$$\begin{aligned} d_f(x,z) &= f(d(x,z)) \\ &\leq f(d(x,y) + d(y,z)) \\ &\leq f(d(x,y)) + f(d(y,z)) \\ &= d_f(x,y) + d_f(y,z). \end{aligned}$$

□

Proposition 2.8 (Bounded metrics with the same open sets)**.** *Let (X,d) be a metric space and define*

$$d_1(x,y) := \min\{1, d(x,y)\}, \qquad d_2(x,y) := \frac{d(x,y)}{1+d(x,y)}.$$

Then the following hold.

(i) Both d_1 and d_2 are metrics on X.

(ii) A subset $U \subset X$ is open with respect to d if and only if it is open with respect to d_1, and this is also equivalent to being open with respect to d_2.

(iii) A sequence (x_n) converges to x with respect to d if and only if it converges to x with respect to d_1, and this is also equivalent to convergence with respect to d_2.

(iv) A sequence (x_n) is Cauchy with respect to d if and only if it is Cauchy with respect to d_1, and this is also equivalent to being Cauchy with respect to d_2. In particular, (X, d) is complete if and only if (X, d_1) is complete, and also if and only if (X, d_2) is complete.

Proof. Part (i) for d_2 is the previous example, and part (i) for d_1 follows from Theorem 2.7 applied to the increasing subadditive function $t \mapsto \min\{1, t\}$.

For part (ii), fix $x \in X$. If $0 < r < 1$, then

$$\{y : d_1(x, y) < r\} = \{y : d(x, y) < r\},$$

because $\min\{1, d(x, y)\} < r < 1$ holds exactly when $d(x, y) < r$. Similarly, if $0 < r < 1$, then

$$\{y : d_2(x, y) < r\} = \left\{y : d(x, y) < \frac{r}{1 - r}\right\},$$

since

$$\frac{d(x, y)}{1 + d(x, y)} < r \iff d(x, y) < \frac{r}{1 - r}.$$

Hence every d-open set is both d_1-open and d_2-open. Conversely, if U is d_1-open (or d_2-open) and $x \in U$, choose a radius $r > 0$ such that the corresponding d_1-ball (or d_2-ball) is contained in U. Replacing r by $\min\{r, \frac{1}{2}\}$ if necessary reduces to the previous formulas and produces an ordinary d-ball contained in U. Thus all three metrics define the same open sets.

Part (iii) follows immediately from the description of small balls: for every $\varepsilon \in (0, 1)$,

$$d_1(x_n, x) < \varepsilon \iff d(x_n, x) < \varepsilon,$$

and

$$d_2(x_n, x) < \varepsilon \iff d(x_n, x) < \frac{\varepsilon}{1 - \varepsilon}.$$

Hence $d(x_n, x) \to 0$, $d_1(x_n, x) \to 0$, and $d_2(x_n, x) \to 0$ are equivalent statements.

For part (iv), suppose first that (x_n) is d-Cauchy. Since $d_1 \leq d$ and $d_2 \leq d$, the sequence is automatically Cauchy for both transformed metrics. Conversely, assume (x_n) is d_1-Cauchy. Given $\varepsilon > 0$, choose N such that $d_1(x_m, x_n) < \min\{1, \varepsilon\}$ whenever $m, n \geq N$. Because the bound is < 1, this means $d(x_m, x_n) < \varepsilon$. Thus (x_n) is d-Cauchy. The same argument works for d_2: given $\varepsilon > 0$, choose $\delta \in (0, 1)$ so that $\delta/(1 - \delta) < \varepsilon$, for instance $\delta = \varepsilon/(1 + \varepsilon)$. If $d_2(x_m, x_n) < \delta$ for $m, n \geq N$, then

$$d(x_m, x_n) < \frac{\delta}{1 - \delta} = \varepsilon.$$

Hence the three Cauchy notions coincide. The completeness statement is now immediate from part (iii): a Cauchy sequence for any one of the three metrics is Cauchy for all of them, and therefore converges for one metric exactly when it converges for the others. □

Remark 2.9. This bounded-metric reduction is useful in practice: it allows one to replace an unbounded metric by a bounded one without changing the underlying topology or the notions of convergence and completeness. Several exercises in the problem sets exploit precisely this idea.

Example 2.10. Let H^∞ (the *Hilbert cube*) denote the set of all sequences

$$x = (x_n) = (x_1, x_2, \ldots) \qquad \text{with} \qquad |x_n| \leq 1 \ \text{ for all } n.$$

Then

$$d(x, y) = \sum_{n=1}^{\infty} \frac{|x_n - y_n|}{2^n}$$

defines a metric on H^∞.

(i) The series is finite for every pair $x, y \in H^\infty$, since $|x_n - y_n| \leq 2$ for every n and hence

$$d(x, y) \leq \sum_{n=1}^{\infty} \frac{2}{2^n} < \infty.$$

Non-negativity and symmetry are immediate. If $d(x, y) = 0$, then each nonnegative summand is zero, so $x_n = y_n$ for every n and therefore $x = y$.

(ii) For each $k \in \mathbb{N}$, the pointwise triangle inequality gives

$$\sum_{n=1}^{k} \frac{|x_n - z_n|}{2^n} \leq \sum_{n=1}^{k} \frac{|x_n - y_n|}{2^n} + \sum_{n=1}^{k} \frac{|y_n - z_n|}{2^n} \leq d(x, y) + d(y, z).$$

Letting $k \to \infty$ in the increasing sequence of partial sums yields

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

Exercise 2.11. Prove that $d(x, y) = \left|\frac{1}{x} - \frac{1}{y}\right|$ defines a metric on $(0, \infty)$.

2.2.2 Open balls and open sets

Definition 2.12 (Open and closed balls)**.** Let (X, d) be a metric space, let $x_0 \in X$, and let $r > 0$.

(i) The set

$$B_r(x_0) = \{y \in X : d(x_0, y) < r\}$$

is called the *open ball* of radius r centred at x_0.

(ii) The set

$$\overline{B_r(x_0)} = \{y \in X : d(x_0, y) \leq r\}$$

is called the *closed ball* of radius r centred at x_0.

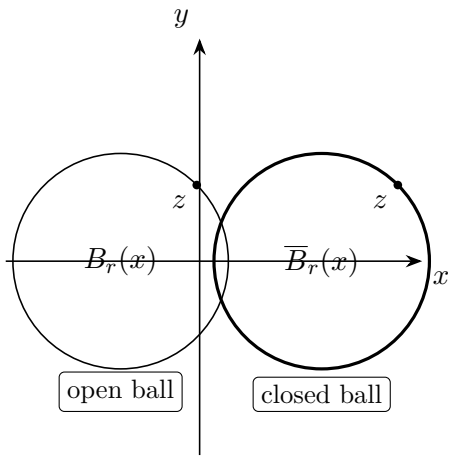


Figure 2.1: Open and closed metric balls (schematic in $\mathbb{R}^2$).

Definition 2.13 (Open sets)**.** A subset $O \subset X$ is called *open* if for every $x \in O$ there exists $r > 0$ such that

$$B_r(x) \subset O.$$

Proposition 2.14. *If $\{O_i : i \in I\}$ is any indexed family of open sets, then*
(i) $\bigcup_{i\in I} O_i$ is open (arbitrary union of open sets is open).
(ii) $\bigcap_{i=1}^{n} O_i$ is open (finite intersection of open sets is open).
Remark. Arbitrary intersection of open sets need not *be open.*

Example 2.15. Let $X = \mathbb{R}$ with the usual metric $u(x, y) = |x - y|$. Then

$$\bigcap_{n=1}^{\infty} \left(-\frac{1}{n}, \frac{1}{n}\right) = \{0\},$$

which is not open. This shows that an arbitrary intersection of open sets need not be open.

Example 2.16. Let $f : \mathbb{R} \to \mathbb{R}$ be continuous. Then the set

$$A := \{x \in \mathbb{R} : f(x) > 0\}$$

is open.

Proof. Fix $x \in A$. Then $f(x) > 0$. By continuity of f at x, with $\varepsilon = f(x)$, there exists $\delta > 0$ such that

$$|f(y) - f(x)| < f(x) \qquad \text{whenever } |y - x| < \delta.$$

Hence $f(y) > 0$ for every $y \in (x - \delta, x + \delta)$, so $(x - \delta, x + \delta) \subset A$. Therefore A is open. □

Open sets in $\mathbb{R}$. A countable union of open intervals is an open set. On the other hand, any open set in $\mathbb{R}$ can be written as a countable union of disjoint open intervals.

Theorem 2.17 (Decomposition of open sets in $\mathbb{R}$)**.** *Let $O \subset \mathbb{R}$ be open. Then there exists a countable family of pairwise disjoint open intervals $\{I_n\}_{n\geq 1}$ such that*

$$O = \bigcupdot_{n=1}^{\infty} I_n.$$

Moreover, this representation is unique up to a permutation of the intervals.

Proof. Fix $x \in O$. Since O is open, there exists an open interval (a, b) with $x \in (a, b) \subset O$. Define the *maximal* interval contained in O and containing x by

$$a_x := \inf\{a \in \mathbb{R} : (a, x] \subset O\}, \qquad b_x := \sup\{b \in \mathbb{R} : [x, b) \subset O\},$$

and set $I_x := (a_x, b_x)$. *Step 1: $I_x \subset O$.* Let $z \in (a_x, b_x)$. Choose $\eta > 0$ so small that $a_x + \eta < z < b_x - \eta$. By definition of the infimum, there exists $a < a_x + \eta$ with $(a, x] \subset O$, hence $(a_x + \eta, x] \subset O$. Similarly, by definition of the supremum there exists $b > b_x - \eta$ with $[x, b) \subset O$, hence $[x, b_x - \eta) \subset O$. Therefore $(a_x + \eta, b_x - \eta) \subset O$, and in particular $z \in O$. *Step 2: maximality and disjointness.* By construction, I_x is an open interval with $x \in I_x \subset O$, and it is maximal with respect to these properties. If $x, y \in O$ and $I_x \cap I_y \neq \varnothing$, then $I_x \cup I_y$ is an open interval contained in O that contains x. Maximality forces $I_y \subset I_x$, and by symmetry $I_x \subset I_y$, hence $I_x = I_y$. Consequently, the family $\{I_x : x \in O\}$ consists of pairwise disjoint open intervals, and

$$O = \bigcup_{x \in O} I_x.$$

Step 3: countability. Each nonempty open interval contains a rational number. Choose $q_x \in I_x \cap \mathbb{Q}$. If $I_x \neq I_y$, then $I_x \cap I_y = \varnothing$, so $q_x \neq q_y$. Thus the map $I_x \mapsto q_x$ is injective into $\mathbb{Q}$, and therefore the collection $\{I_x : x \in O\}$ is countable. Renaming these intervals as $\{I_n\}_{n\geq 1}$ gives the desired representation. *Step 4: uniqueness.* Suppose

$$O = \bigcupdot_{n=1}^{\infty} I_n = \bigcupdot_{m=1}^{\infty} J_m,$$

where $\{I_n\}$ and $\{J_m\}$ are pairwise disjoint families of open intervals. Fix n. Since $I_n \subset O$, we have

$$I_n = I_n \cap O = \bigcupdot_{m=1}^{\infty} (I_n \cap J_m).$$

The sets $I_n \cap J_m$ are pairwise disjoint, and each is (possibly empty) open in $\mathbb{R}$. Since I_n is a connected open set (an interval), only one of these intersections can be nonempty; hence

$I_n \subset J_{m_0}$ for some m_0. By the same argument with the roles reversed, $J_{m_0} \subset I_n$, and therefore $I_n = J_{m_0}$. This proves uniqueness up to a permutation. □

2.2.3 Sequences in metric spaces

Definition 2.18 (Convergent sequence)**.** A sequence (x_n) in a metric space (X, d) is said to *converge* to $x_0 \in X$ if for every $\varepsilon > 0$ there exists $N \in \mathbb{N}$ such that

$$n \geq N \quad \Longrightarrow \quad d(x_n, x_0) < \varepsilon.$$

Equivalently, $x_n \in B_\varepsilon(x_0)$ for all $n \geq N$.

Definition 2.19 (Cauchy sequence)**.** A sequence (x_n) in (X, d) is called a *Cauchy sequence* if for every $\varepsilon > 0$ there exists $N \in \mathbb{N}$ such that

$$m, n \geq N \quad \Longrightarrow \quad d(x_n, x_m) < \varepsilon.$$

Example 2.20. Let $X = (0, 1)$ with the usual metric $d(x, y) = |x - y|$. The sequence $\left(\frac{1}{n}\right)$ is Cauchy because

$$\left|\frac{1}{n} - \frac{1}{m}\right| \to 0 \qquad \text{as } m, n \to \infty.$$

However, its limit in $\mathbb{R}$ is 0, and $0 \notin X$. Hence this sequence does not converge in the metric space $(0, 1)$.

Every convergent sequence in a metric space is Cauchy.

Definition 2.21 (Bounded set)**.** A subset $A \subset X$ is called *bounded* if there exist $x_0 \in X$ and $M > 0$ such that

$$d(a, x_0) \leq M \qquad \text{for all } a \in A.$$

Equivalently, A is bounded if and only if it is contained in some closed ball.

Example 2.22. The set $\{(x, y) : y = \sin(1/x),\ x \neq 0\} \cup (\{0\} \times [-1, 1])$ is unbounded, since it contains points $(n, \sin(1/n))$ whose Euclidean norm tends to infinity as $n \to \infty$.

Proposition 2.23. *Every Cauchy sequence is bounded.*

Proof. Let (x_n) be Cauchy. Taking $\varepsilon = 1$, we obtain $N \in \mathbb{N}$ such that

$$d(x_m, x_n) < 1 \qquad \text{for all } m, n \geq N.$$

In particular,

$$d(x_n, x_N) < 1 \qquad \text{for all } n \geq N.$$

Set

$$M := \max\{1, d(x_1, x_N), d(x_2, x_N), \ldots, d(x_{N-1}, x_N)\}.$$

Then $d(x_n, x_N) \leq M$ for every $n \geq 1$, so the sequence is contained in the closed ball $\overline{B}_M(x_N)$. Hence (x_n) is bounded. □

The converse need not hold. For example, in $(\mathbb{R}, u)$ with the usual metric, the sequence

$$x_n = (-1)^n$$

is bounded, but it is not Cauchy.

Proposition 2.24. *Let (x_n) be a Cauchy sequence in (X, d). If a subsequence (x_{n_k}) converges to x, then the full sequence converges to x.*

Proof. Let $\varepsilon > 0$. Since (x_n) is Cauchy, there exists $N_1 \in \mathbb{N}$ such that

$$d(x_n, x_m) < \frac{\varepsilon}{2} \qquad \text{for all } n, m \geq N_1.$$

Since $x_{n_k} \to x$, there exists $K \in \mathbb{N}$ such that

$$d(x_{n_k}, x) < \frac{\varepsilon}{2} \qquad \text{for all } k \geq K.$$

Choose k so large that $n_k \geq N_1$. Then, for every $m \geq N_1$,

$$d(x_m, x) \leq d(x_m, x_{n_k}) + d(x_{n_k}, x) < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

Hence $x_m \to x$, and therefore $x_n \to x$. □

Remark 2.25. If $X = (0, 1)$ is equipped with the usual metric $d(x, y) = |x - y|$, then the sequence $x_n = 1/n$ is Cauchy in X but does not converge in X. In particular, it has no convergent subsequence with limit in X.

2.2.4 Closed sets in metric spaces

Definition 2.26. A subset $F \subset X$ is called *closed* if its complement $F^c = X \setminus F$ is open. Equivalently, if

$$B_\varepsilon(x) \cap F \neq \varnothing \qquad \text{for every } \varepsilon > 0,$$

then $x \in F$.

Example 2.27. Consider the set

$$A := \{(x, y) \in \mathbb{R}^2 : y = \sin(1/x),\ x \neq 0\}.$$

Then A is neither open nor closed in $\mathbb{R}^2$. It is not closed because the sequence

$$\left(\frac{1}{n\pi}, 0\right) \in A, \qquad n \geq 1,$$

converges to $(0,0)$, while $(0,0) \notin A$. It is not open because no open ball centered at a point of A can lie entirely inside A; geometrically, A is only a curve in the plane.

Theorem 2.28 (Closed sets via convergent sequences)**.** *Let (X,d) be a metric space and let $F \subset X$. Then the following are equivalent:*

(i) F is closed.

(ii) If $B_\varepsilon(x) \cap F \neq \varnothing$ for every $\varepsilon > 0$, then $x \in F$.

(iii) Whenever (x_n) is a sequence in F and $x_n \to x$, one has $x \in F$.

Proof. (i) $\Rightarrow$ (ii). Assume that F is closed, and suppose that $B_\varepsilon(x) \cap F \neq \varnothing$ for every $\varepsilon > 0$. If $x \notin F$, then $x \in F^c$. Since F^c is open, there exists $\varepsilon_0 > 0$ such that $B_{\varepsilon_0}(x) \subset F^c$, which contradicts the assumption that every ball about x meets F. Hence $x \in F$.

(ii) $\Rightarrow$ (iii). Let $(x_n) \subset F$ and suppose that $x_n \to x$. Then every ball $B_\varepsilon(x)$ contains all sufficiently large terms of the sequence, so in particular it meets F. By (ii), it follows that $x \in F$.

(iii) $\Rightarrow$ (i). Assume that every convergent sequence in F has its limit in F. To prove that F is closed, it suffices to show that F^c is open. Suppose instead that F^c is not open. Then there exists $x \in F^c$ such that $B_{1/n}(x) \cap F \neq \varnothing$ for every $n \in \mathbb{N}$. Choose $x_n \in B_{1/n}(x) \cap F$. Then $x_n \to x$, and (iii) implies $x \in F$, a contradiction. Therefore F^c is open, and hence F is closed. □

Proposition 2.29 (A diameter estimate)**.** *Let (X,d) be a metric space and let $A \subset X$ be nonempty. Define*

$$\operatorname{diam}(A) := \sup\{d(a,b) : a,b \in A\} \in [0,\infty].$$

If $\operatorname{diam}(A) \leq r$ and $A \cap B_r(x) \neq \varnothing$ for some $x \in X$ and $r > 0$, then

$$A \subset B_{2r}(x).$$

Proof. Choose $a_0 \in A \cap B_r(x)$. Then $d(a_0,x) < r$. For any $a \in A$ we have

$$d(a,x) \leq d(a,a_0) + d(a_0,x) \leq \operatorname{diam}(A) + r \leq 2r,$$

so $a \in B_{2r}(x)$. Hence $A \subset B_{2r}(x)$. □

Remark 2.30. The older quiz formulation of Theorem 2.29 assumed in addition that A is closed. The proof shows that closedness is unnecessary: only the diameter bound and the nonempty intersection with $B_r(x)$ matter.

Example 2.31. Let $f\colon \mathbb{R} \to \mathbb{R}$ be continuous, and set

$$A := \{x \in \mathbb{R} : f(x) = 0\}.$$

Then A is closed. Indeed, if $(x_n) \subset A$ and $x_n \to x$, then $f(x_n) = 0$ for all n. By continuity of f, we obtain

$$f(x) = \lim_{n\to\infty} f(x_n) = 0,$$

so $x \in A$.

Example 2.32. Let c_{00} = space of all sequences having finitely many non-zero terms.

$$c_{00} = \{x = (x_1, x_2, \ldots, x_n, 0, 0, \ldots) : x_i \in \mathbb{R}\}$$

$$\|x\|_\infty := \max_{1 \le i \le n} |x_i| < \infty.$$

$$\implies c_{00} \subsetneq \ell^\infty \quad \text{(proper subspace)}.$$

Let

$$X^n = \left(1, \frac{1}{2}, \ldots, \frac{1}{n}, 0, 0, \ldots\right) \in c_{00}.$$

Let

$$X = \left(1, \frac{1}{2}, \ldots, \frac{1}{n}, \frac{1}{n+1}, \ldots\right) \in \ell^\infty.$$

then

$$\|X - X^n\|_\infty = \sup_{k \ge n} \frac{1}{k+1} = \frac{1}{n+1} \to 0.$$

but $X \notin c_{00}$. Hence c_{00} is not a closed subspace of ℓ^∞. In addition c_{00} is not open in ℓ^∞. For this, let $\varepsilon > 0$ be arbitrary and consider the sequence $y = (\frac{\varepsilon}{2}, \frac{\varepsilon}{4}, \frac{\varepsilon}{8}, \ldots) \in \ell^\infty$. Then $\|y\|_\infty = \frac{\varepsilon}{2} < \varepsilon$, so $y \in B_\varepsilon(0)$, but $y \notin c_{00}$. Therefore, $B_\varepsilon(0) \not\subseteq c_{00}$ for any $\varepsilon > 0$.
For $1 \le p < \infty$, $c_{00} \subsetneq \ell^p$ and c_{00} is neither closed nor open in ℓ^p. For this, let

$$x_n = \left(\frac{\varepsilon^p}{2^{n+1}}\right)^{1/p}, \ 1 \le p < \infty,$$

and consider $x = (x_1, x_2, \ldots) \in \ell^p$. Then $x \notin c_{00}$ and

$$\|x\|_p^p = \sum_{n=1}^{\infty} |x_n|^p = \sum_{n=1}^{\infty} \frac{\varepsilon^p}{2^{n+1}} = \frac{\varepsilon^p}{2},$$

so $\|x\|_p = \frac{\varepsilon}{2^{1/p}} < \varepsilon$. Hence $x \in B_\varepsilon(0)$, and therefore $B_\varepsilon(0) \not\subseteq c_{00}$ for any $\varepsilon > 0$. Consequently, c_{00} is not open in ℓ^p.
To see that c_{00} is not closed in ℓ^p, let $X^n = (x_1, x_2, \ldots, x_n, 0, 0, \ldots) \in c_{00}$. Then $X^n \to x$ in ℓ^p, since

$$\|X^n - x\|_p^p = \sum_{k=n+1}^{\infty} \frac{\varepsilon^p}{2^{k+1}} \to 0 \quad \text{as} \quad n \to \infty.$$

But $x \notin c_{00}$.

2.2.5 Interior points and interior of a set

Let $A \subset X$. The *interior* of A, denoted by $\mathrm{Int}(A)$ or A°, is the largest open set contained in A. Equivalently,

$$A^\circ = \bigcup \{O \subset X : O \text{ is open and } O \subseteq A\},$$

that is, A° is the union of all open balls contained in A.

2.2.6 Closure and limit points

Let $A \subset X$. The *closure* of A, denoted by $\overline{A}$ or $\mathrm{clos}(A)$, is the smallest closed set containing A. Equivalently,

$$\overline{A} = \bigcap \{F \subset X : F \text{ is closed and } A \subseteq F\}.$$

Definition 2.33 (Limit point / accumulation point)**.** Let $A \subset X$. A point $x \in X$ is a *limit point* of A if for every $r > 0$, $(B(x,r) \setminus \{x\}) \cap A \neq \emptyset$. The set of all limit points of A is denoted by A'.

Remark 2.34. Note that if x is a limit point of A, then for $n \in \mathbb{N}$, and $r = \frac{1}{n}$, there exists $x_n \in A$ such that $x_n \in (B(x, \frac{1}{n}) \setminus \{x\}) \cap A$. Consequently, $x_n \to x$. In particular, every ball $(B(x,r))$ contains infinitely many terms of the sequence (x_n).

Exercise 2.35. Finite set in a metric space has no limit points.

Exercise 2.36. Let

$$A = \left\{ \left(n, \frac{1}{n} \right) : n \in \mathbb{N} \right\} \subset \mathbb{R}^2.$$

Show that the closure of A in $(\mathbb{R}^2, u)$ is $\overline{A} = A$ and that $A^\circ = \emptyset$.

Example 2.37. 1. $A = \{(x,y) : |x| < 1, |y| < 1\}$. Then
$\overline{A} = \{(x,y) : |x| \leq 1, |y| \leq 1\}$.

2. $A = \{(x,y) : y = \sin\left(\frac{1}{x}\right), x \neq 0\}$. Then
$\overline{A} = \{(x,y) : y = \sin\left(\frac{1}{x}\right), x \neq 0\} \cup (\{0\} \times [-1,1])$.

Proposition 2.38 (Closure by neighbourhoods)**.** *Let $A \subset X$. Then $x \in \overline{A}$ if and only if $B_\varepsilon(x) \cap A \neq \varnothing$ for every $\varepsilon > 0$.*

Proof. Assume first that $x \in \overline{A}$. If there were $\varepsilon_0 > 0$ such that $B_{\varepsilon_0}(x) \cap A = \varnothing$, then $A \subset (B_{\varepsilon_0}(x))^c$. Since $(B_{\varepsilon_0}(x))^c$ is closed and contains A, the minimality of $\overline{A}$ would imply

$$\overline{A} \subset (B_{\varepsilon_0}(x))^c,$$

which is impossible because $x \in \overline{A}$ but $x \notin (B_{\varepsilon_0}(x))^c$.

Conversely, assume that $B_\varepsilon(x) \cap A \neq \varnothing$ for every $\varepsilon > 0$. Since $\overline{A}$ is closed and contains A, Theorem 2.28 implies that $x \in \overline{A}$. □

Proposition 2.39. *For every $A \subset X$,*

$$\overline{A} = A \cup A'.$$

Proof. If $x \in A'$, then by Theorem 2.28 we have $x \in \overline{A}$. Hence $A \cup A' \subseteq \overline{A}$. Conversely, let $x \in \overline{A}$. If $x \in A$, then $x \in A \cup A'$. If $x \notin A$, then every ball centered at x meets A, and because $x \notin A$ this means every ball centered at x meets $A \setminus \{x\}$. Thus $x \in A'$. Therefore $\overline{A} \subseteq A \cup A'$. □

Proposition 2.40. *A point $x \in X$ belongs to $\overline{A}$ if and only if there exists a sequence (x_n) in A such that $x_n \to x$.*

Proof. Assume that $x \in \overline{A}$. For each $n \in \mathbb{N}$, choose $x_n \in B_{1/n}(x) \cap A$. Then

$$d(x_n, x) < \frac{1}{n} \qquad \text{for all } n \in \mathbb{N},$$

and hence $x_n \to x$.

Conversely, suppose that $x_n \in A$ and $x_n \to x$. Let $\varepsilon > 0$. For all sufficiently large n, we have $d(x_n, x) < \varepsilon$, so $x_n \in B_\varepsilon(x) \cap A$. Therefore $B_\varepsilon(x) \cap A \neq \varnothing$ for every $\varepsilon > 0$, and Theorem 2.38 yields $x \in \overline{A}$. □

2.2.7 Dense subsets

A subset $A \subset X$ is called *dense* in X if $\overline{A} = X$. Equivalently, every point of X is the limit of a sequence from A, or, equivalently again, every nonempty open ball in X intersects A.

Example 2.41. The set $\mathbb{Q}$ is dense in $\mathbb{R}$ with respect to the usual metric $u(x, y) = |x - y|$.

Indeed, let $x \in \mathbb{R}$ and write $x = [x] + \alpha$, where $[x]$ denotes the integer part of x and $0 \leq \alpha < 1$. Choose a decimal expansion

$$\alpha = 0.x_1x_2x_3\cdots, \qquad x_i \in \{0, 1, \ldots, 9\}.$$

For each n, define

$$x_n = [x] + \frac{x_1}{10} + \frac{x_2}{10^2} + \cdots + \frac{x_n}{10^n}.$$

Then $x_n \in \mathbb{Q}$ for every n, and $x_n \to x$. Hence $\overline{\mathbb{Q}} = \mathbb{R}$.

Example 2.42. If $1 \leq p < \infty$, then $\overline{c_{00}} = \ell^p$.

Indeed, let $x = (x_1, x_2, \ldots) \in \ell^p$, and define the truncations

$$X^n := (x_1, x_2, \ldots, x_n, 0, 0, \ldots), \qquad n \geq 1.$$

Then $X^n \in c_{00}$ for every n, and

$$\|x - X^n\|_p = \left(\sum_{k=n+1}^{\infty} |x_k|^p \right)^{1/p} \longrightarrow 0 \qquad (n \to \infty).$$

Hence $X^n \to x$ in ℓ^p, which proves that c_{00} is dense in ℓ^p.

Example 2.43. We also have $\overline{c_{00}} = c_0$ with respect to the supremum norm. Let $x = (x_1, x_2, \ldots) \in c_0$. Since $x_n \to 0$, given $\varepsilon > 0$ there exists $N \in \mathbb{N}$ such that

$$|x_n| < \frac{\varepsilon}{2} \qquad \text{for all } n \geq N. \tag{2.2}$$

Define

$$X^n := (x_1, x_2, \dots, x_n, 0, 0, \dots), \qquad n \geq N.$$

Then $X^n \in c_{00}$ and

$$\|x - X^n\|_\infty = \sup_{k \geq n+1} |x_k| \leq \frac{\varepsilon}{2} \qquad \text{for all } n \geq N,$$

by (2.2). Hence $X^n \to x$ in c_0.

Remark: $\overline{c_{00}} = c_0 \subsetneq \ell^\infty$. That is, c_{00} is not dense in ℓ^∞.

2.3 Continuous maps between metric spaces

Continuity formalizes the principle that small changes in the input should produce small changes in the output. In metric spaces this idea admits several equivalent formulations, and one of the aims of this section is to move comfortably among pointwise estimates, sequential criteria, and topological descriptions.

A function $f\colon (X,d) \to (\mathbb{R},u)$ is called *continuous at* $x_0 \in X$ if for every $\varepsilon > 0$ there exists $\delta > 0$ such that

$$d(x_0, y) < \delta \quad \Longrightarrow \quad |f(y) - f(x_0)| < \varepsilon.$$

Equivalently,

$$f\big(B_\delta(x_0)\big) \subset \big(f(x_0) - \varepsilon,\ f(x_0) + \varepsilon\big).$$

Example 2.44. Fix $x_0 \in X$ and define $f\colon X \to \mathbb{R}$ by $f(y) := d(x_0, y)$. By the reverse triangle inequality,

$$|f(y) - f(z)| = |d(x_0, y) - d(x_0, z)| \leq d(y, z) \qquad \text{for all } y, z \in X.$$

Therefore f is Lipschitz, hence continuous.

Theorem 2.45 (Continuity of real-valued functions)**.** *Let $f : (X,d) \to (\mathbb{R},u)$. The following are equivalent:*

(i) f is continuous on X in the ε–δ sense;

(ii) whenever $x_n \to x$ in X, one has $f(x_n) \to f(x)$ in $\mathbb{R}$;

(iii) $f^{-1}(O)$ is open in X for every open set $O \subseteq \mathbb{R}$;

(iv) $f^{-1}(F)$ is closed in X for every closed set $F \subseteq \mathbb{R}$.

Proof. Recall that f is *continuous at* $x \in X$ if for every $\varepsilon > 0$ there exists $\delta > 0$ such that

$$d(x, y) < \delta \implies |f(y) - f(x)| < \varepsilon.$$

The map f is *continuous on* X if it is continuous at every $x \in X$. **(i) $\Rightarrow$ (ii).** Assume that f is continuous on X and let $(x_n) \subset X$ satisfy $x_n \to x$. Fix $\varepsilon > 0$. By continuity of f at x, there exists $\delta > 0$ such that $d(x, y) < \delta$ implies $|f(y) - f(x)| < \varepsilon$. Since $x_n \to x$, there exists $N \in \mathbb{N}$ such that $n \geq N$ implies $d(x_n, x) < \delta$. Hence for all $n \geq N$,

$$|f(x_n) - f(x)| < \varepsilon,$$

so $f(x_n) \to f(x)$. **(ii) $\Rightarrow$ (i).** Assume (ii). Fix $x \in X$ and suppose, for contradiction, that f is not continuous at x. Then there exists $\varepsilon_0 > 0$ such that for every $\delta > 0$ there exists $y \in X$ with

$$d(x, y) < \delta \quad \text{and} \quad |f(y) - f(x)| \geq \varepsilon_0.$$

For each $n \in \mathbb{N}$, apply this with $\delta = \frac{1}{n}$ to choose $y_n \in X$ satisfying

$$d(x, y_n) < \frac{1}{n}, \qquad |f(y_n) - f(x)| \geq \varepsilon_0.$$

Then $y_n \to x$ in X, but $f(y_n) \not\to f(x)$ in $\mathbb{R}$, contradicting (ii). Therefore f is continuous at x. Since x was arbitrary, f is continuous on X. **(i) $\Rightarrow$ (iii).** Assume f is continuous on X and let $O \subset \mathbb{R}$ be open. Take $x \in f^{-1}(O)$, so $f(x) \in O$. Since O is open, there exists $\varepsilon > 0$ such that

$$(f(x) - \varepsilon,\ f(x) + \varepsilon) \subset O.$$

By continuity of f at x, there exists $\delta > 0$ such that $d(x, y) < \delta$ implies $|f(y) - f(x)| < \varepsilon$, i.e. $f(y) \in (f(x) - \varepsilon, f(x) + \varepsilon) \subset O$. Thus $y \in f^{-1}(O)$ whenever $d(x, y) < \delta$, so

$$B_\delta(x) \subset f^{-1}(O).$$

Hence $f^{-1}(O)$ is open in X. **(iii) $\Rightarrow$ (i).** Assume (iii). Fix $x \in X$ and $\varepsilon > 0$. Let

$$O := (f(x) - \varepsilon,\ f(x) + \varepsilon),$$

which is open in $\mathbb{R}$. By (iii), $f^{-1}(O)$ is open in X, and plainly $x \in f^{-1}(O)$. Therefore, there exists $\delta > 0$ such that $B_\delta(x) \subset f^{-1}(O)$. If $d(x, y) < \delta$, then $y \in f^{-1}(O)$, so $f(y) \in O$, i.e. $|f(y) - f(x)| < \varepsilon$. This proves continuity of f at x. Since x was arbitrary, f is continuous on X. **(iii) $\Rightarrow$ (iv).** Assume (iii) and let $F \subset \mathbb{R}$ be closed. Then $\mathbb{R} \setminus F$ is open, so $f^{-1}(\mathbb{R} \setminus F)$ is open in X by (iii). But

$$f^{-1}(\mathbb{R} \setminus F) = X \setminus f^{-1}(F).$$

Hence $X \setminus f^{-1}(F)$ is open, which implies $f^{-1}(F)$ is closed. **(iv) $\Rightarrow$ (iii)** is similar to the above. □

Remark 2.46. More generally, if (X, d) and (Y, ρ) are metric spaces, a map $f\colon X \to Y$ is continuous at $x \in X$ if for every $\varepsilon > 0$ there exists $\delta > 0$ such that

$$d(x, y) < \delta \quad \Longrightarrow \quad \rho\big(f(x), f(y)\big) < \varepsilon.$$

2.3.1 Uniform continuity

Definition 2.47. A function $f\colon A \subset X \to Y$ is called *uniformly continuous* on A if for every $\varepsilon > 0$ there exists $\delta > 0$ such that

$$d(x,y) < \delta \quad \Longrightarrow \quad \rho(f(x), f(y)) < \varepsilon \qquad \text{for all } x, y \in A.$$

The important point is that δ depends only on ε, not on the particular choice of x and y.

Example 2.48. For $x_0 \in X$, let $f(x) = d(x, x_0)$. Then f is uniformly continuous on X. (Hint: $d(x, x_0) \leq d(x, y) + d(y, x_0) \implies f(x) - f(y) < d(x, y)$.) Similarly, by replacing x with y, it follows.

Example 2.49. For $x \in X$ and $A \subset X$, define

$$d(x, A) := \inf\{d(x, a) : a \in A\}.$$

This is called the *distance from x to the set A*. The map $x \mapsto d(x, A)$ is uniformly continuous.

Indeed, for every $a \in A$ the triangle inequality gives

$$d(x, a) \leq d(x, y) + d(y, a).$$

Taking the infimum over $a \in A$ yields

$$d(x, A) \leq d(x, y) + d(y, A).$$

Interchanging the roles of x and y gives

$$|d(x, A) - d(y, A)| \leq d(x, y).$$

Hence $x \mapsto d(x, A)$ is Lipschitz, and therefore uniformly continuous.

Example 2.50. The function $f : (0, 1) \to \mathbb{R}$ given by $f(x) = 1/x$ is continuous on $(0, 1)$ but not uniformly continuous.

Continuity at each fixed point is immediate from the usual algebraic rules for continuous functions. To see directly why uniform continuity fails, take

$$x_n = \frac{1}{n}, \qquad y_n = \frac{1}{n+1}.$$

Then $|x_n - y_n| = \frac{1}{n(n+1)} \to 0$, whereas

$$|f(x_n) - f(y_n)| = |n - (n+1)| = 1 \qquad \text{for all } n.$$

Therefore f cannot be uniformly continuous on $(0, 1)$. From the above argument, we obtain the following useful sequential characterization of uniform continuity.

Theorem 2.51 (Sequential criterion for uniform continuity)**.** *Let $f\colon A \subset X \to \mathbb{R}$. Then f is uniformly continuous on A if and only if for every pair of sequences $x_n, y_n \in A$ satisfying $d(x_n, y_n) \to 0$, one has $|f(x_n) - f(y_n)| \to 0$.*

Proof. Suppose first that f is uniformly continuous on A. Then for every $\varepsilon > 0$ there exists $\delta > 0$ such that

$$d(x, y) < \delta \quad \Longrightarrow \quad |f(x) - f(y)| < \varepsilon \qquad (x, y \in A). \tag{2.3}$$

Now let $x_n, y_n \in A$ and assume that $d(x_n, y_n) \to 0$. For the above δ, there exists $n_0 \in \mathbb{N}$ such that $d(x_n, y_n) < \delta$ whenever $n \geq n_0$. Hence (2.3) gives

$$|f(x_n) - f(y_n)| < \varepsilon \qquad \text{for all } n \geq n_0.$$

Therefore $|f(x_n) - f(y_n)| \to 0$.

Conversely, assume that the sequential condition holds and suppose, toward a contradiction, that f is not uniformly continuous on A. Then there exists $\varepsilon_0 > 0$ such that for every $\delta > 0$ one can find $x, y \in A$ with

$$d(x, y) < \delta \qquad \text{but} \qquad |f(x) - f(y)| \geq \varepsilon_0.$$

Choosing $\delta = \frac{1}{n}$ for each $n \in \mathbb{N}$, we obtain sequences $x_n, y_n \in A$ such that

$$d(x_n, y_n) < \frac{1}{n} \qquad \text{and} \qquad |f(x_n) - f(y_n)| \geq \varepsilon_0 \quad \text{for all } n.$$

Then $d(x_n, y_n) \to 0$, but $|f(x_n) - f(y_n)| \not\to 0$, contradicting the assumed sequential criterion. Hence f is uniformly continuous. □

Example 2.52. Let $f : (X, d) \to \mathbb{R}$ be uniformly continuous. Then f maps Cauchy sequences in X to Cauchy sequences in $\mathbb{R}$.

Indeed, let (x_n) be a Cauchy sequence in (X, d), and let $\varepsilon > 0$. By uniform continuity, there exists $\delta > 0$ such that

$$d(x, y) < \delta \quad \Longrightarrow \quad |f(x) - f(y)| < \varepsilon \qquad (x, y \in X).$$

Since (x_n) is Cauchy, there exists $N \in \mathbb{N}$ such that $d(x_n, x_m) < \delta$ whenever $n, m \geq N$. Hence

$$|f(x_n) - f(x_m)| < \varepsilon \qquad \text{for all } n, m \geq N,$$

so $(f(x_n))$ is a Cauchy sequence in $\mathbb{R}$.

Theorem 2.53 (Heine–Cantor theorem)**.** *Let $f\colon [a, b] \to \mathbb{R}$ be continuous. Then f is uniformly continuous on $[a, b]$.*

Proof. Suppose, on the contrary, that f is not uniformly continuous on $[a,b]$. Then there exists $\varepsilon_0 > 0$ such that for every $\delta > 0$ one can find $x, y \in [a,b]$ with

$$|x-y| < \delta \qquad \text{but} \qquad |f(x)-f(y)| \geq \varepsilon_0.$$

Taking $\delta = \frac{1}{n}$, we obtain sequences $x_n, y_n \in [a,b]$ such that

$$|x_n - y_n| < \frac{1}{n} \qquad \text{and} \qquad |f(x_n)-f(y_n)| \geq \varepsilon_0.$$

By the Bolzano–Weierstrass theorem, the bounded sequence (x_n) has a convergent subsequence, say $x_{n_k} \to x \in [a,b]$. Since

$$|x_{n_k} - y_{n_k}| < \frac{1}{n_k} \to 0,$$

it follows that $y_{n_k} \to x$ as well. By continuity of f,

$$f(x_{n_k}) \to f(x) \qquad \text{and} \qquad f(y_{n_k}) \to f(x),$$

and therefore

$$|f(x_{n_k}) - f(y_{n_k})| \to 0,$$

contradicting the lower bound $|f(x_{n_k}) - f(y_{n_k})| \geq \varepsilon_0$. Hence f is uniformly continuous. □

Theorem 2.54 (Uniform continuity on compact sets)**.** *Let K be compact and let $f : K \to Y$ be continuous between metric spaces. Then f is uniformly continuous on K.*

Proof. Assume f is not uniformly continuous. Then there exists $\varepsilon_0 > 0$ and sequences $x_n, y_n \in K$ such that

$$d(x_n, y_n) \to 0 \quad \text{and} \quad \rho(f(x_n), f(y_n)) \geq \varepsilon_0 \text{ for all } n.$$

By compactness, there exists a subsequence $x_{n_k} \to x \in K$. Since $d(x_{n_k}, y_{n_k}) \to 0$, it follows that $y_{n_k} \to x$. By continuity, $f(x_{n_k}) \to f(x)$ and $f(y_{n_k}) \to f(x)$, contradicting $\rho(f(x_{n_k}), f(y_{n_k})) \geq \varepsilon_0$. Hence f is uniformly continuous. □

Example 2.55. Let $a < b < c$, and let $f : (a,c) \to \mathbb{R}$ be a function whose restrictions to $(a,b]$ and (b,c) are uniformly continuous. Then f is uniformly continuous on (a,c).

Proof. Let $\varepsilon > 0$. By uniform continuity on $(a,b]$, there exists $\delta_1 > 0$ such that

$$x, y \in (a,b],\ |x-y| < \delta_1 \implies |f(x)-f(y)| < \frac{\varepsilon}{2}.$$

By uniform continuity on (b,c), there exists $\delta_2 > 0$ such that

$$x, y \in (b,c),\ |x-y| < \delta_2 \implies |f(x)-f(y)| < \frac{\varepsilon}{2}.$$

Set $\delta := \min\{\delta_1, \delta_2\}$. Let $x, y \in (a, c)$ with $|x - y| < \delta$.

If both x and y lie in $(a, b]$, or both lie in (b, c), the corresponding estimate applies immediately. Otherwise, after relabeling we may assume that $x \leq b < y$. Then

$$|x - b| \leq |x - y| < \delta \leq \delta_1, \qquad |y - b| \leq |x - y| < \delta \leq \delta_2.$$

Hence

$$|f(x) - f(y)| \leq |f(x) - f(b)| + |f(b) - f(y)| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

Therefore f is uniformly continuous on (a, c). □

Example 2.56. Let $f : \mathbb{R} \to \mathbb{R}$ be continuous and suppose that

$$\lim_{|x| \to \infty} f(x) = 0.$$

Then f is uniformly continuous on $\mathbb{R}$.

Proof. Let $\varepsilon > 0$. Choose $R > 0$ such that

$$|x| \geq R \implies |f(x)| < \frac{\varepsilon}{2}.$$

Since f is continuous on the compact interval $[-R - 1, R + 1]$, it is uniformly continuous there. Hence there exists $\delta_0 > 0$ such that

$$x, y \in [-R - 1, R + 1],\ |x - y| < \delta_0 \implies |f(x) - f(y)| < \varepsilon.$$

Set $\delta := \min\{1, \delta_0\}$, and let $x, y \in \mathbb{R}$ with $|x - y| < \delta$.

If both x and y lie in $[-R - 1, R + 1]$, the compact-set estimate applies. Otherwise one of them, say x, satisfies $|x| > R + 1$. Since $|x - y| < 1$, we also have $|y| > R$, and therefore

$$|f(x) - f(y)| \leq |f(x)| + |f(y)| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

Hence f is uniformly continuous on $\mathbb{R}$. □

In particular, every function in $C_0(\mathbb{R})$ is uniformly continuous. By contrast, a continuous bounded function on $\mathbb{R}$ need not be uniformly continuous.

Example 2.57. The function $f(x) = \sin(x^2)$ is continuous and bounded on $\mathbb{R}$, but it is not uniformly continuous. (Hint: take $x_n = \sqrt{n\pi}$ and $y_n = \sqrt{n\pi + \frac{\pi}{2}}$.)

Example 2.58. Let $f : \mathbb{R} \to \mathbb{R}$ be differentiable, and suppose that $|f'(x)| \leq M$ for all $x \in \mathbb{R}$. Then f is uniformly continuous on $\mathbb{R}$.

Proof. For any $x, y \in \mathbb{R}$, the Mean Value Theorem gives a point t between x and y such that

$$|f(x) - f(y)| = |f'(t)|\,|x - y| \leq M|x - y|.$$

Thus f is Lipschitz, hence uniformly continuous.

The converse is false: the function $f(x) = \sqrt{x}$ on $(0, \infty)$ is uniformly continuous, but

$$f'(x) = \frac{1}{2\sqrt{x}}$$

is unbounded near 0. □

Example 2.59. Let $f : \mathbb{R} \to \mathbb{R}$ be bounded, continuous, and monotone. Then f is uniformly continuous on $\mathbb{R}$.

Proof. It suffices to consider the case where f is monotone increasing. Let

$$L := \inf_{x \in \mathbb{R}} f(x), \qquad M := \sup_{x \in \mathbb{R}} f(x).$$

Given $\varepsilon > 0$, choose $x_0 < y_0$ such that

$$f(x_0) < L + \frac{\varepsilon}{2}, \qquad f(y_0) > M - \frac{\varepsilon}{2}.$$

Then for $x, y \leq x_0$ we have

$$|f(x) - f(y)| \leq f(x_0) - L < \frac{\varepsilon}{2},$$

and for $x, y \geq y_0$ we have

$$|f(x) - f(y)| \leq M - f(y_0) < \frac{\varepsilon}{2}.$$

Since f is continuous on the compact interval $[x_0 - 1, y_0 + 1]$, it is uniformly continuous there. Hence there exists $\delta_0 > 0$ such that

$$x, y \in [x_0 - 1, y_0 + 1],\ |x - y| < \delta_0 \implies |f(x) - f(y)| < \varepsilon.$$

Set $\delta := \min\{1, \delta_0\}$. Let $x, y \in \mathbb{R}$ with $|x - y| < \delta$.

If both points lie in $[x_0 - 1, y_0 + 1]$, the compact-set estimate applies. If $x < x_0 - 1$, then $y < x_0$ because $|x - y| < 1$; hence both points lie to the left of x_0 and the left-tail estimate applies. Similarly, if $x > y_0 + 1$, then $y > y_0$, and the right-tail estimate applies. Therefore $|f(x) - f(y)| < \varepsilon$ in all cases, so f is uniformly continuous on $\mathbb{R}$. □

We observe that a uniformly continuous function can be extended uniformly to the closure of the set.

Theorem 2.60 (Uniformly continuous extension to the closure)**.** *Let (X, d) be a metric space, let $A \subset X$, and let (Y, ρ) be a* complete *metric space. If $f \colon A \to Y$ is uniformly continuous, then there exists a unique uniformly continuous map $\widetilde{f} \colon \overline{A} \to Y$ such that $\widetilde{f}|_A = f$.*

Proof. Fix $x \in \overline{A}$. Choose a sequence $(x_n) \subset A$ with $x_n \to x$. Then (x_n) is Cauchy in X, hence $(f(x_n))$ is Cauchy in Y because f is uniformly continuous. Since Y is complete, the limit $\lim_{n\to\infty} f(x_n)$ exists. Define

$$\widetilde{f}(x) := \lim_{n\to\infty} f(x_n).$$

Well-definedness. If $(y_n) \subset A$ is another sequence with $y_n \to x$, then $d(x_n, y_n) \to 0$. By uniform continuity, $\rho(f(x_n), f(y_n)) \to 0$. Since both $(f(x_n))$ and $(f(y_n))$ converge, their limits must coincide, so $\widetilde{f}$ is well defined. *Extension property.* If $x \in A$, choose the constant sequence $x_n = x$. Then $\widetilde{f}(x) = f(x)$. *Uniform continuity of* $\widetilde{f}$. Let $\varepsilon > 0$. By uniform continuity of f, there exists $\delta > 0$ such that

$$d(u, v) < \delta \implies \rho(f(u), f(v)) < \varepsilon/3 \qquad (u, v \in A).$$

Set $\delta' := \delta/3$. Take $x, y \in \overline{A}$ with $d(x, y) < \delta'$. Choose sequences $(x_n), (y_n) \subset A$ such that $x_n \to x$ and $y_n \to y$. Pick N so large that

$$d(x_N, x) < \delta', \qquad d(y_N, y) < \delta', \qquad \rho(\widetilde{f}(x), f(x_N)) < \varepsilon/3, \qquad \rho(\widetilde{f}(y), f(y_N)) < \varepsilon/3.$$

Then

$$d(x_N, y_N) \le d(x_N, x) + d(x, y) + d(y, y_N) < \delta' + \delta' + \delta' = \delta,$$

hence $\rho(f(x_N), f(y_N)) < \varepsilon/3$. Therefore,

$$\rho(\widetilde{f}(x), \widetilde{f}(y)) \le \rho(\widetilde{f}(x), f(x_N)) + \rho(f(x_N), f(y_N)) + \rho(f(y_N), \widetilde{f}(y)) < \varepsilon.$$

This proves that $\widetilde{f}$ is uniformly continuous on $\overline{A}$. *Uniqueness.* If $g\colon \overline{A} \to Y$ is another uniformly continuous extension of f, then g is continuous. For $x \in \overline{A}$ and any sequence $(x_n) \subset A$ with $x_n \to x$, continuity gives

$$g(x) = \lim_{n\to\infty} g(x_n) = \lim_{n\to\infty} f(x_n) = \widetilde{f}(x),$$

so $g = \widetilde{f}$. □

Exercise 2.61 (Why completeness is needed)**.** Show that completeness of the target space cannot be omitted in Theorem 2.60. More precisely, consider $A = \mathbb{Q} \subset \mathbb{R}$ with the usual metric and the function $f\colon \mathbb{Q} \to \mathbb{Q}$ given by $f(q) = q$. Prove that f is uniformly continuous on $\mathbb{Q}$, but there is no continuous (hence no uniformly continuous) extension $\widetilde{f}\colon \mathbb{R} \to \mathbb{Q}$.

Next, we show that a uniformly continuous function grows at most linearly.

Theorem 2.62 (Linear growth of uniformly continuous functions)**.** *Let $f : \mathbb{R} \to \mathbb{R}$ be uniformly continuous. Then there exist constants $A, B \ge 0$ such that*

$$|f(x)| \le A|x| + B \qquad \textit{for all } x \in \mathbb{R}.$$

Proof. Apply uniform continuity with $\varepsilon = 1$. Then there exists $\delta > 0$ such that

$$|u - v| < \delta \implies |f(u) - f(v)| < 1.$$

Let

$$B_0 := \max_{|t| \le 1} |f(t)|.$$

Fix $x \ge 1$, and choose $n \in \mathbb{N}$ so that

$$n\delta \le x < (n+1)\delta.$$

Then

$$|f(x) - f(0)| \le |f(x) - f(n\delta)| + \sum_{j=0}^{n-1} |f((j+1)\delta) - f(j\delta)| < 1 + n.$$

Since $n \le x/\delta$, it follows that

$$|f(x)| \le |f(0)| + 1 + \frac{x}{\delta} \le \frac{x}{\delta} + B_0 + 1.$$

The same argument applied to $-x$ shows that for every $x \le -1$,

$$|f(x)| \le \frac{|x|}{\delta} + B_0 + 1.$$

Therefore, for all $|x| \ge 1$,

$$|f(x)| \le \frac{1}{\delta}|x| + B_0 + 1.$$

On the interval $[-1, 1]$ we also have $|f(x)| \le B_0$. Hence the conclusion holds with

$$A := \frac{1}{\delta}, \qquad B := B_0 + 1.$$

□

Example 2.63. The function $f(x) = x^2$ is not uniformly continuous on $\mathbb{R}$. This is consistent with the theorem above, since x^2 does not satisfy any global linear growth estimate. Equivalently, if $x_n = n$ and $y_n = n + \frac{1}{n}$, then $|x_n - y_n| \to 0$ while

$$|f(x_n) - f(y_n)| = 2 + \frac{1}{n^2} \not\to 0.$$

Chapter 3

Completeness, Compactness, and Fixed Points

Completeness formalizes the principle that a space already contains the limits suggested by its Cauchy sequences, and it is the decisive hypothesis behind many existence theorems. We study complete metric spaces, compactness through both sequential and open-cover formulations, and conclude with the contraction mapping principle as a canonical application of completeness.

Learning objectives.

- recognize completeness through Cauchy sequences and understand its role in existence arguments;
- prove the equivalence of the principal compactness criteria available in metric spaces;
- use total boundedness, sequential compactness, and open-cover compactness in the appropriate settings;
- apply the contraction mapping principle to derive existence, uniqueness, and iterative approximation results.

3.1 Chapter roadmap

Dependencies. The core proof pattern in this chapter is extraction: one builds subsequences or iterates that reveal hidden compactness or convergence.

- Completeness and basic permanence properties (closed subsets, products).
- Completion of a metric space (existence and universal property).
- Compactness: open covers, sequential compactness, and metric characterizations.
- Contraction mapping principle and existence/uniqueness of fixed points.

3.2 Complete metric spaces

A space is complete when the limiting behaviour suggested by its metric is already realized inside the space itself. This hypothesis is subtle but indispensable: many existence theorems in analysis amount to showing that a carefully constructed sequence is Cauchy and then invoking completeness to conclude that a genuine solution exists.

We have already encountered the basic obstruction: a Cauchy sequence need not converge within the ambient space. For example, in $((0,1),u)$ with the usual metric, the sequence $(\frac{1}{n})$ is Cauchy, but its limit is $0 \notin (0,1)$.

This suggests enlarging a metric space so as to include the limits of all its Cauchy sequences; that construction leads to the *completion* of a metric space, which we discuss later. For the moment, we focus on spaces for which no enlargement is necessary.

Definition 3.1. A metric space (X,d) is said to be *complete* if every Cauchy sequence in X converges to a point of X.

Example 3.2. The real line $(\mathbb{R},u)$ is complete.

Indeed, let (x_n) be a Cauchy sequence in $\mathbb{R}$. Then (x_n) is bounded, so by the Bolzano–Weierstrass theorem it has a subsequence $x_{n_k} \to x \in \mathbb{R}$. By Proposition 2.24, the entire sequence converges to the same limit x.

Example 3.3. The metric space $(\mathbb{R},d)$ with

$$d(x,y) = |\tan^{-1}(x) - \tan^{-1}(y)|$$

is not complete.

Remark 3.4 (Hint). Consider the sequence

$$x_n = \tan\left(\frac{\pi}{2} \cdot \frac{n}{n+1}\right).$$

It is Cauchy in the metric d, but it does not converge to any point of $\mathbb{R}$.

Example 3.5. Every discrete metric space is complete.

Indeed, let $X \neq \varnothing$ and equip it with the discrete metric

$$d_0(x,y) = \begin{cases} 1, & x \neq y, \\ 0, & x = y. \end{cases}$$

Suppose $(x_n) \subset X$ is Cauchy. Taking $\varepsilon = \frac{1}{2}$, there exists $N \in \mathbb{N}$ such that

$$d_0(x_n, x_m) < \frac{1}{2} \qquad \text{for all } n,m \geq N.$$

Since the only possible values of d_0 are 0 and 1, this forces

$$d_0(x_n, x_m) = 0 \qquad \text{for all } n,m \geq N.$$

Hence $x_n = x_m$ for all $n,m \geq N$, so the sequence is eventually constant. Therefore (x_n) converges in X, and X is complete.

Example 3.6. $(\mathbb{R}^n, \|\cdot\|_p)$ is complete for every $1 \leq p \leq \infty$.

First assume $1 \le p < \infty$, and let $x^k = (x_1^k, \dots, x_n^k)$ be a Cauchy sequence in $(\mathbb{R}^n, \|\cdot\|_p)$. For each fixed $j \in \{1, \dots, n\}$ we have

$$|x_j^k - x_j^\ell| \le \|x^k - x^\ell\|_p,$$

so $(x_j^k)_k$ is a Cauchy sequence in $\mathbb{R}$. Since $\mathbb{R}$ is complete, there exists $x_j \in \mathbb{R}$ such that $x_j^k \to x_j$ as $k \to \infty$. Set $x := (x_1, \dots, x_n) \in \mathbb{R}^n$.

Let $\varepsilon > 0$ be given. Since (x^k) is Cauchy, there exists $N \in \mathbb{N}$ such that

$$\|x^k - x^\ell\|_p < \frac{\varepsilon}{2} \qquad \text{for all } k, \ell \ge N.$$

Fix $k \ge N$ and let $\ell \to \infty$. For each coordinate, $x_j^\ell \to x_j$, and because the sum has only finitely many terms, we may pass to the limit to obtain

$$\|x^k - x\|_p^p = \sum_{j=1}^n |x_j^k - x_j|^p = \lim_{\ell \to \infty} \sum_{j=1}^n |x_j^k - x_j^\ell|^p \le \left(\frac{\varepsilon}{2}\right)^p.$$

Hence $\|x^k - x\|_p \le \varepsilon/2 < \varepsilon$ for every $k \ge N$, so $x^k \to x$ in $\mathbb{R}^n$.

For $p = \infty$, the argument is even simpler: from

$$\|x^k - x^\ell\|_\infty = \max_{1 \le j \le n} |x_j^k - x_j^\ell| < \varepsilon$$

we again deduce that each coordinate sequence is Cauchy, hence convergent, say $x_j^k \to x_j$. Writing $x = (x_1, \dots, x_n)$ and letting $\ell \to \infty$ in the inequality above shows that

$$\|x^k - x\|_\infty \le \varepsilon \qquad \text{for all sufficiently large } k,$$

so $x^k \to x$. Therefore $(\mathbb{R}^n, \|\cdot\|_p)$ is complete for all $1 \le p \le \infty$.

Example 3.7. Let $1 \le p \le \infty$. Then $(\ell^p, \|\cdot\|_p)$ is complete.

For $1 \le p < \infty$, let $x^k = (x_1^k, x_2^k, \dots)$ be a Cauchy sequence in $(\ell^p, \|\cdot\|_p)$. Given $\varepsilon > 0$, choose $n_0 \in \mathbb{N}$ such that

$$\|x^k - x^l\|_p < \varepsilon \qquad \text{for all } k, l \ge n_0.$$

Then for every $n \in \mathbb{N}$,

$$\sum_{j=1}^n |x_j^k - x_j^l|^p < \varepsilon^p. \tag{3.1}$$

For each fixed n, the vectors consisting of the first n coordinates lie in $(\mathbb{R}^n, \|\cdot\|_p)$, which is complete. Hence for each j, the scalar sequence $(x_j^k)_k$ converges to some x_j. Passing to the limit as $k \to \infty$ in (3.1), we obtain

$$\sum_{j=1}^n |x_j^l - x_j|^p < \varepsilon^p \qquad \text{for all } l \ge n_0.$$

Now let $n \to \infty$. Since the left-hand side is increasing in n, it follows that

$$\sum_{j=1}^{\infty} |x_j^l - x_j|^p \le \varepsilon^p \qquad \text{for all } l \ge n_0.$$

Thus $x = (x_1, x_2, \dots)$ belongs to ℓ^p and $x^l \to x$ in ℓ^p. The case $p = \infty$ is similar.

Proposition 3.8. *Every closed subset of a complete metric space is complete.*

Proof. Let F be a closed subset of a complete metric space (X, d), and let $(x_n) \subset F$ be Cauchy. Then (x_n) is also a Cauchy sequence in X, so completeness of X gives a point $x \in X$ such that $x_n \to x$. Since F is closed and each x_n belongs to F, we conclude that $x \in F$. Hence F is complete. □

Remark 3.9. If (X, d) is complete, then a subset $F \subset X$ is closed if and only if it is complete (with the induced metric).

Exercise 3.10. Show that $(c_0, \|\cdot\|_\infty)$ is a proper closed subspace of $(\ell^\infty, \|\cdot\|_\infty)$.

We already know that $c_0 \subsetneq \ell^\infty$. To prove that c_0 is closed, let $x^k = (x_1^k, x_2^k, \dots) \in c_0$ and assume that $x^k \to x = (x_1, x_2, \dots)$ in ℓ^∞. Then for every $\varepsilon > 0$ there exists $k_0 \in \mathbb{N}$ such that

$$\|x^k - x\|_\infty < \varepsilon \qquad \text{for all } k \ge k_0,$$

so in particular

$$|x_j^k - x_j| < \varepsilon \qquad \text{for every } j \ge 1 \text{ and every } k \ge k_0. \tag{3.2}$$

Since $x^{k_0} \in c_0$, there exists $j_0 \in \mathbb{N}$ such that

$$|x_j^{k_0}| < \varepsilon \qquad \text{for all } j \ge j_0. \tag{3.3}$$

Combining (3.2) and (3.3), we get

$$|x_j| \le |x_j - x_j^{k_0}| + |x_j^{k_0}| < 2\varepsilon \qquad \text{for all } j \ge j_0.$$

Hence $x_j \to 0$, so $x \in c_0$. Therefore c_0 is closed in ℓ^∞, and consequently it is complete with respect to $\|\cdot\|_\infty$.

Example 3.11. The space $(C[a, b], \|\cdot\|_\infty)$ is complete.

Let (f_n) be a Cauchy sequence in $(C[a, b], \|\cdot\|_\infty)$. Given $\varepsilon > 0$, choose $n_0 \in \mathbb{N}$ such that

$$\|f_n - f_m\|_\infty < \varepsilon \qquad \text{for all } n, m \ge n_0.$$

Then for every $t \in [a, b]$,

$$|f_n(t) - f_{n_0}(t)| < \varepsilon \qquad \text{for all } n \ge n_0. \tag{3.4}$$

Thus for each fixed t, the sequence $(f_n(t))$ is Cauchy in $\mathbb{R}$, and hence converges to a limit which we denote by $f(t)$. Letting $n \to \infty$ in (3.4), we obtain

$$|f(t) - f_{n_0}(t)| \leq \varepsilon \qquad \text{for all } t \in [a,b].$$

The index n_0 is independent of t, so the convergence is uniform. Since f_{n_0} is continuous, for each fixed $t \in [a,b]$ and each $\varepsilon > 0$ there exists $\delta > 0$ such that $|s-t| < \delta$ implies $|f_{n_0}(s) - f_{n_0}(t)| < \varepsilon$. Therefore,

$$\begin{aligned} |f(s) - f(t)| &\leq |f(s) - f_{n_0}(s)| + |f_{n_0}(s) - f_{n_0}(t)| + |f_{n_0}(t) - f(t)| \\ &< 3\varepsilon, \end{aligned}$$

which shows that f is continuous. Hence $f \in C[a,b]$ and $f_n \to f$ uniformly.

On the other hand, $(C[0,1], \|\cdot\|_1)$ is not complete. Consider

$$f_n(t) = \begin{cases} nt, & 0 \leq t < 1/n, \\ 1, & 1/n \leq t \leq 1. \end{cases}$$

For $1/m < 1/n$, one checks that

$$\|f_n - f_m\|_1 = \left(\int_0^{1/m} + \int_{1/m}^{1/n} + \int_{1/n}^{1} \right) |f_n(t) - f_m(t)|\, dt = \frac{1}{2}\left(\frac{1}{m} - \frac{1}{n} \right) \to 0$$

as $n < m \to \infty$. Thus (f_n) is a Cauchy sequence in $(C[0,1], \|\cdot\|_1)$. Its pointwise limit is

$$f(t) = \lim_{n\to\infty} f_n(t) = \begin{cases} 1, & 0 < t \leq 1, \\ 0, & t = 0, \end{cases}$$

which is not continuous. Therefore $(C[0,1], \|\cdot\|_1)$ is not complete.

3.2.1 Completion of a metric space

Every metric space admits an isometric embedding into a complete metric space whose image is dense. One standard construction realizes the completion as equivalence classes of Cauchy sequences.

Theorem 3.12 (Completion of a metric space)**.** *Let (X,d) be a metric space. Then there exist a complete metric space $(\widehat{X}, \widehat{d})$ and an isometric embedding $i : X \to \widehat{X}$ such that $i(X)$ is dense in $\widehat{X}$. In particular, every Cauchy sequence in X converges in $\widehat{X}$.*

Proof. Let $\mathcal{C}(X)$ denote the set of all Cauchy sequences in X. Define a relation $\sim$ on $\mathcal{C}(X)$ by

$$(x_n) \sim (y_n) \quad \Longleftrightarrow \quad d(x_n, y_n) \xrightarrow[n\to\infty]{} 0.$$

We first verify that $\sim$ is an equivalence relation. Reflexivity is immediate, since $d(x_n, x_n) = 0$ for every n, and symmetry follows from the symmetry of d. For transitivity, suppose that $(x_n) \sim (y_n)$ and $(y_n) \sim (z_n)$. Then, for every n,

$$d(x_n, z_n) \leq d(x_n, y_n) + d(y_n, z_n).$$

Passing to the limit as $n \to \infty$ gives $d(x_n, z_n) \to 0$, and hence $(x_n) \sim (z_n)$. Set

$$\widehat{X} := \mathcal{C}(X)/\sim,$$

and write $[x_n]$ for the equivalence class of a Cauchy sequence (x_n). **Step 1: Definition of the metric.** If (x_n) and (y_n) are Cauchy sequences in X, then the real sequence $\big(d(x_n, y_n)\big)$ is Cauchy in $\mathbb{R}$ (by the reverse triangle inequality) and hence convergent. For $\alpha = [x_n]$ and $\beta = [y_n]$ in $\widehat{X}$, define

$$\widehat{d}(\alpha, \beta) := \lim_{n\to\infty} d(x_n, y_n).$$

Well-definedness. Assume $(x_n) \sim (x'_n)$ and $(y_n) \sim (y'_n)$. Then

$$\big|d(x_n, y_n) - d(x'_n, y'_n)\big| \leq d(x_n, x'_n) + d(y_n, y'_n) \xrightarrow[n\to\infty]{} 0,$$

so $\lim_n d(x_n, y_n) = \lim_n d(x'_n, y'_n)$. Hence $\widehat{d}$ is independent of representatives. **Metric axioms.** Nonnegativity and symmetry are immediate. If $\widehat{d}([x_n], [y_n]) = 0$, then $\lim_n d(x_n, y_n) = 0$, so $(x_n) \sim (y_n)$ and hence $[x_n] = [y_n]$. The triangle inequality follows by passing to limits in

$$d(x_n, z_n) \leq d(x_n, y_n) + d(y_n, z_n).$$

Therefore, $(\widehat{X}, \widehat{d})$ is a metric space. **Step 2: Isometric embedding and density.** Define $i : X \to \widehat{X}$ by

$$i(x) := [(x, x, x, \dots)].$$

Then for $x, y \in X$,

$$\widehat{d}(i(x), i(y)) = \lim_{n\to\infty} d(x, y) = d(x, y),$$

so i is an isometric embedding. Let $\alpha = [x_n] \in \widehat{X}$. We claim that $i(x_n) \to \alpha$ in $(\widehat{X}, \widehat{d})$. Indeed,

$$\begin{aligned}\widehat{d}(i(x_n), \alpha) &= \widehat{d}([(x_n, x_n, \dots)], [x_1, x_2, \dots]) \\ &= \lim_{m\to\infty} d(x_n, x_m).\end{aligned}$$

Since (x_m) is Cauchy, given $\varepsilon > 0$ there exists N such that $d(x_n, x_m) < \varepsilon$ for all $m \geq n \geq N$. Thus for $n \geq N$ we have $\lim_{m\to\infty} d(x_n, x_m) \leq \varepsilon$, and hence $\widehat{d}(i(x_n), \alpha) \to 0$. Therefore every

$\alpha \in \widehat{X}$ is a limit of points in $i(X)$, so $i(X)$ is dense in $\widehat{X}$. **Step 3: Completeness of $(\widehat{X}, \widehat{d})$.** Let (α_k) be a Cauchy sequence in $\widehat{X}$. By density of $i(X)$, for each k choose $x_k \in X$ such that

$$\widehat{d}(\alpha_k, i(x_k)) < 2^{-k}. \tag{3.5}$$

Then (x_k) is a Cauchy sequence in (X, d). Indeed, for m, n,

$$\begin{aligned} d(x_m, x_n) &= \widehat{d}(i(x_m), i(x_n)) \\ &\leq \widehat{d}(i(x_m), \alpha_m) + \widehat{d}(\alpha_m, \alpha_n) + \widehat{d}(\alpha_n, i(x_n)), \end{aligned}$$

and by (3.5),

$$d(x_m, x_n) \leq 2^{-m} + \widehat{d}(\alpha_m, \alpha_n) + 2^{-n}.$$

Since (α_k) is Cauchy, the middle term tends to 0 as $m, n \to \infty$, and hence $d(x_m, x_n) \to 0$. Define $\alpha := [x_k] \in \widehat{X}$. Then by definition of $\widehat{d}$,

$$\begin{aligned} \widehat{d}(i(x_k), \alpha) &= \widehat{d}([(x_k, x_k, \dots)], [x_1, x_2, \dots]) \\ &= \lim_{m \to \infty} d(x_k, x_m) \xrightarrow[k \to \infty]{} 0, \end{aligned}$$

so $i(x_k) \to \alpha$. Finally,

$$\widehat{d}(\alpha_k, \alpha) \leq \widehat{d}(\alpha_k, i(x_k)) + \widehat{d}(i(x_k), \alpha) \xrightarrow[k \to \infty]{} 0$$

by (3.5). Hence $\alpha_k \to \alpha$, and $(\widehat{X}, \widehat{d})$ is complete. □

3.3 Compactness in metric spaces

Compactness is one of the organizing principles of analysis: it turns local control into global control and allows infinite information to be reduced to finite subcollections or convergent subsequences. Although the open-cover definition may appear indirect at first, the equivalent metric formulations make the concept extremely usable.

Definition 3.13 (Open cover)**.** Let $K \subset X$. A family of open sets $\{U_\alpha\}_{\alpha \in I}$ is called an *open cover* of K if

$$K \subset \bigcup_{\alpha \in I} U_\alpha.$$

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

$$K \subset U_{\alpha_1} \cup \cdots \cup U_{\alpha_m}.$$

Definition 3.14 (Compact set)**.** A set $K \subset X$ is called *compact* if every open cover of K admits a finite subcover.

Example 3.15. Every finite subset of a metric space is compact.

Example 3.16. The interval $(0,1) \subset \mathbb{R}$ is not compact. Indeed, the family

$$U_n := \Big(\frac{1}{n}, 1\Big), \qquad n \geq 2,$$

is an open cover of $(0,1)$ with no finite subcover.

Definition 3.17 (Sequential compactness)**.** A set $K \subset X$ is *sequentially compact* if every sequence (x_n) in K has a subsequence (x_{n_k}) converging to a limit $x \in K$.

Theorem 3.18. *Let (X,d) be a metric space and $K \subset X$. Then the following are equivalent:*

(i) K is compact;

(ii) K is sequentially compact.

Proof. **(i) $\Rightarrow$ (ii).** Let $(x_n) \subset K$ be a sequence. Assume, towards a contradiction, that (x_n) has no convergent subsequence. Then (x_n) has no limit point in K. Equivalently, for each $x \in K$ there exists $r_x > 0$ such that $B(x, r_x)$ contains only finitely many terms of (x_n). The family $\{B(x, r_x)\}_{x \in K}$ is an open cover of K. By compactness, there exist $x_1, \ldots, x_m \in K$ such that

$$K \subset \bigcup_{j=1}^{m} B(x_j, r_{x_j}).$$

Each ball $B(x_j, r_{x_j})$ contains only finitely many terms of the sequence; hence their union contains only finitely many terms. This contradicts the fact that (x_n) is an infinite sequence in K. Therefore (x_n) admits a convergent subsequence with limit in K. **(ii) $\Rightarrow$ (i).** Assume K is not compact. Then there exists an open cover $\{U_\alpha\}_{\alpha \in A}$ of K with no finite subcover. We construct inductively a sequence $(x_n) \subset K$ as follows. Choose $x_1 \in K$ and pick α_1 such that $x_1 \in U_{\alpha_1}$. Since $\{U_{\alpha_1}\}$ does not cover K, choose $x_2 \in K \setminus U_{\alpha_1}$ and pick α_2 with $x_2 \in U_{\alpha_2}$. Continuing in this way, we obtain a sequence such that

$$x_{n+1} \notin U_{\alpha_1} \cup \cdots \cup U_{\alpha_n} \qquad (n \in \mathbb{N}).$$

By sequential compactness, there exists a subsequence $x_{n_k} \to x \in K$. Since $\{U_\alpha\}$ covers K, there exists j such that $x \in U_{\alpha_j}$. As U_{α_j} is open and $x_{n_k} \to x$, we have $x_{n_k} \in U_{\alpha_j}$ for all sufficiently large k. However, for all k such that $n_k > j$, the construction implies $x_{n_k} \notin U_{\alpha_j}$, a contradiction. Hence K is compact. □

Remark 3.19. In $\mathbb{R}^n$ equipped with the Euclidean metric, the Heine–Borel theorem asserts that a set is compact if and only if it is closed and bounded.

3.3.1 Total boundedness and metric characterizations

Definition 3.20 (Totally bounded)**.** A subset $E \subset X$ is called *totally bounded* if for every $\varepsilon > 0$ there exist points $x_1, \dots, x_N \in X$ such that

$$E \subset \bigcup_{j=1}^{N} B(x_j, \varepsilon).$$

Equivalently, for every $\varepsilon > 0$ the set E can be covered by finitely many ε-balls.

Proposition 3.21. *Every compact subset of a metric space is totally bounded.*

Proof. Let $K \subset X$ be compact and fix $\varepsilon > 0$. The family of open balls $\{B(x, \varepsilon) : x \in K\}$ is an open cover of K, hence admits a finite subcover $K \subset \bigcup_{j=1}^{N} B(x_j, \varepsilon)$. □

Theorem 3.22 (Compactness ⇔ completeness + total boundedness)**.** *Let (X, d) be a metric space and $K \subset X$. Then K is compact if and only if K is complete (as a metric space with the restricted metric) and totally bounded.*

Proof. Assume first that K is compact. Then K is totally bounded by Proposition 3.21. Moreover, compactness implies sequential compactness, and in a sequentially compact metric space every Cauchy sequence converges. Indeed, if (x_n) is Cauchy in K, then it admits a convergent subsequence $x_{n_k} \to x \in K$; the Cauchy property forces the full sequence to converge to the same limit x. Thus K is complete.

Conversely, assume that K is complete and totally bounded. Let (x_n) be a sequence in K. Total boundedness yields, for $\varepsilon = 1$, a finite cover of K by balls of radius 1; one of these balls contains infinitely many terms of the sequence. Iterating with $\varepsilon = 1/2, 1/3, \dots$ and using a diagonal selection produces a Cauchy subsequence (x_{n_k}). Completeness of K implies that this subsequence converges in K. Hence K is sequentially compact, and therefore compact. □

Remark 3.23. Neither completeness nor total boundedness alone is enough. The interval $(0, 1)$ is totally bounded but not complete, while an infinite discrete metric space is complete but not totally bounded. Compactness is exactly the conjunction of the two.

3.3.2 Heine–Borel theorem in $\mathbb{R}^n$

Theorem 3.24 (Heine–Borel)**.** *A set $K \subset \mathbb{R}^n$ (with the Euclidean metric) is compact if and only if it is closed and bounded.*

Proof. If K is compact then it is bounded: otherwise, for each $m \in \mathbb{N}$ pick $x_m \in K$ with $\|x_m\|_2 > m$; the sequence (x_m) has no convergent subsequence, contradicting sequential compactness. Also, K is closed because limits of convergent sequences in K must lie in K (again by sequential compactness).

Conversely, suppose K is closed and bounded. Any sequence in K is bounded in $\mathbb{R}^n$, so by Bolzano–Weierstrass it has a convergent subsequence in $\mathbb{R}^n$. Since K is closed, the subsequential limit lies in K. Therefore K is sequentially compact, hence compact. □

Remark 3.25. The Heine–Borel criterion is special to finite-dimensional Euclidean space. In general metric spaces, closed and bounded sets need not be compact; the right replacement is total boundedness together with completeness.

Proposition 3.26. *Let $K \subset X$ be compact. If $F \subset K$ is closed in X, then F is compact.*

Proof. Let $\{U_\alpha\}_{\alpha\in A}$ be an open cover of F. Then $\{U_\alpha\}_{\alpha\in A} \cup \{X \setminus F\}$ is an open cover of K. By compactness of K, there is a finite subcover; removing $X \setminus F$ yields a finite subcover of F. □

Proposition 3.27. *If $K \subset X$ is compact, then K is closed.*

Proof. Let $(x_n) \subset K$ and suppose $x_n \to x$ in X. By compactness, (x_n) has a convergent subsequence $x_{n_k} \to y \in K$. Since $x_n \to x$, every subsequence converges to x, hence $y = x$, and therefore $x \in K$. □

Proposition 3.28. *If $K \subset X$ is compact, then K is bounded.*

Proof. Fix $x_0 \in K$. The family $\{B(x_0, n)\}_{n\in\mathbb{N}}$ is an open cover of K. By compactness, $K \subset B(x_0, N)$ for some $N \in \mathbb{N}$, hence K is bounded. □

Proposition 3.29. *If (K, d) is compact, then (K, d) is complete.*

Proof. Let (x_n) be a Cauchy sequence in K. By compactness, (x_n) has a convergent subsequence $x_{n_k} \to x \in K$. By Proposition 2.24, it follows that $x_n \to x$. Thus (K, d) is complete. □

Remark 3.30. In metric spaces, compactness implies closedness, boundedness, and completeness. The converses fail in general metric spaces, but hold in $\mathbb{R}^n$ (Heine–Borel).

Exercise 3.31. If K is a compact subset of $(\mathbb{R}, u)$, show that $\inf K$ and $\sup K$ both belong to K.

Hint: choose a sequence $(x_n) \subset K$ with $x_n \to \inf K$. Compactness yields a convergent subsequence $x_{n_k} \to x \in K$, and the uniqueness of limits forces $x = \inf K$. Treat $\sup K$ similarly.

Theorem 3.32. *Let $f : K \to Y$ be continuous between metric spaces. If K is compact, then $f(K)$ is compact.*

Proof. Let $\{V_\beta\}_{\beta\in B}$ be an open cover of $f(K)$. Then $\{f^{-1}(V_\beta)\}_{\beta\in B}$ is an open cover of K by continuity. Extract a finite subcover of K and apply f to obtain a finite subcover of $f(K)$. □

Theorem 3.33 (Extreme Value Theorem)**.** *Let K be compact and let $f : K \to \mathbb{R}$ be continuous. Then f attains its maximum and minimum on K: there exist $x_o, y_o \in K$ such that*

$$f(x_o) = \max_{x\in K} f(x), \qquad f(y_o) = \min_{x\in K} f(x).$$

Proof. Let $M = \sup f(K)$. Choose $x_n \in K$ such that $f(x_n) \to M$. By compactness, there exists a subsequence $x_{n_k} \to x \in K$. By continuity, $f(x_{n_k}) \to f(x)$, hence $f(x) = M$. The minimum is proved similarly. □

Exercise 3.34. Let $K \subset X$ be compact. Prove that every infinite subset of K has a limit point in K.

Exercise 3.35. Give an example of a metric space (X, d) and a subset $E \subset X$ which is closed and bounded but not compact.

Exercise 3.36. Let K be compact and let $f : K \to K$ be continuous. Prove that the sequence defined by $x_{n+1} = f(x_n)$ has a convergent subsequence in K.

3.4 The contraction mapping principle

The contraction mapping principle is one of the clearest examples of an abstract theorem with concrete computational force. It turns an iterative procedure into an existence-and-uniqueness theorem: once a map shortens distances uniformly, repeated application of the map forces the orbit to converge to the unique point left unchanged by the map.

A natural strategy for solving an equation of the form $\varphi(x) = x$ is to study the iterates of φ. Starting from $x_0 \in X$, we form the sequence

$$x_0,\ \varphi(x_0),\ \varphi^2(x_0),\ \ldots,$$

where φ^n denotes the n-fold iterate of φ. If this sequence converges to some $x \in X$ and if φ is continuous, then

$$\varphi(x) = \varphi\Big(\lim_{n\to\infty} \varphi^n(x_0)\Big) = \lim_{n\to\infty} \varphi^{n+1}(x_0) = x,$$

so x is a fixed point of φ.

The contraction mapping principle identifies a setting in which this argument always succeeds: the contraction condition yields a Cauchy sequence of iterates, and completeness supplies the limit.

Definition 3.37. A map $\varphi : (X, d) \to (X, d)$ is called a *contraction* if there exists a constant α with $0 < \alpha < 1$ such that

$$d(\varphi(x), \varphi(y)) \le \alpha\, d(x, y) \qquad \text{for all } x, y \in X.$$

Theorem 3.38 (Contraction mapping principle)**.** *Let (X, d) be a complete metric space. If $\varphi : (X, d) \to (X, d)$ is a contraction, then φ has a unique fixed point.*

Proof. Choose $\alpha \in (0, 1)$ such that

$$d(\varphi(x), \varphi(y)) \le \alpha\, d(x, y) \qquad \text{for all } x, y \in X.$$

Fix $x_0 \in X$ and define $x_n := \varphi^n(x_0)$ for $n \geq 0$.

By repeated application of the contraction estimate,

$$d(x_{n+1}, x_n) = d(\varphi(x_n), \varphi(x_{n-1})) \leq \alpha\, d(x_n, x_{n-1}) \leq \alpha^n d(x_1, x_0).$$

Hence, for $m > n$,

$$\begin{aligned} d(x_n, x_m) &\leq d(x_n, x_{n+1}) + \cdots + d(x_{m-1}, x_m) \\ &\leq (\alpha^n + \alpha^{n+1} + \cdots + \alpha^{m-1}) d(x_1, x_0) \\ &\leq \frac{\alpha^n}{1-\alpha}\, d(x_1, x_0). \end{aligned}$$

Since the right-hand side tends to 0 as $n \to \infty$, the sequence (x_n) is Cauchy. Completeness of X therefore yields a point $x \in X$ such that $x_n \to x$.

Because every contraction is continuous, passing to the limit in $x_{n+1} = \varphi(x_n)$ gives

$$x = \lim_{n\to\infty} x_{n+1} = \lim_{n\to\infty} \varphi(x_n) = \varphi(x).$$

Thus x is a fixed point of φ.

To prove uniqueness, suppose $y \in X$ is another fixed point. Then

$$d(x, y) = d(\varphi(x), \varphi(y)) \leq \alpha d(x, y).$$

Since $0 < \alpha < 1$, this is possible only if $d(x, y) = 0$, and hence $x = y$. □

Remark 3.39. If $\Omega \subset \mathbb{R}^n$ is open, then any contraction $f : \Omega \to \Omega$ can have at most one fixed point. Completeness is needed for existence, not for uniqueness.

For example,

$$\varphi : (0, \infty) \to (0, \infty)$$
$$\varphi(x) = \frac{1}{2}(x + \frac{a}{x}), \qquad a > 0$$

satisfies $\varphi(\sqrt{a}) = \sqrt{a}$.
Observe that φ above is not a contraction mapping, since

$$|\varphi(x) - \varphi(y)| = \frac{1}{2}\,|1 - \frac{a}{xy}|\,|x - y|$$

because the function $|1 - \frac{a}{xy}|$ is not bounded near zero.

Example 3.40. $\varphi : (0, 2\pi) \to (0, 2\pi),\ \varphi(x) = \sin\frac{x}{2}$.

$$|\varphi(x) - \varphi(y)| \leq \frac{1}{2}|x - y| \quad \text{(By Mean Value Theorem)}.$$

Thus φ is a contraction mapping, but it has no fixed point in $(0, 2\pi)$. This shows that completeness of the underlying metric space cannot be omitted in the contraction mapping principle.

Exercise 3.41. If (X, d) is a complete metric space and $f : X \to X$ is such that f^k is a contraction, then show that f has a unique fixed point.

Remark 3.42 (Hint). First consider the case $k = 2$. Use the fact that f^k cannot have two distinct fixed points. If $f^2(x_0) = x_0$ and $y_0 := f(x_0)$, then $f(y_0) = x_0$ and hence $f^2(y_0) = y_0$; uniqueness of the fixed point of f^2 forces $y_0 = x_0$.

Exercise 3.43. Let $T : C[0, 1] \to C[0, 1]$ be defined by

$$T(f)(x) = \int_0^x f(t)dt.$$

Show that T^2 is a contraction but T is not a contraction.

This phenomenon is also natural from the viewpoint of iteration: when studying the convergence of $\varphi^n(x_0)$, changing finitely many initial terms does not affect the limit behaviour.
We next consider the existence and uniqueness of the following initial value problem:

$$\begin{cases} y' = f(x, y) \\ y(0) = y_0 \end{cases} \tag{3.6}$$

with the help of the fixed point theorem.
Suppose f is continuous on a rectangle containing the point $(0, y_0)$ in its interior, and that f is Lipschitz in the second variable, i.e.,

$$|f(x, y_1) - f(x, y_2)| \leq K|y_1 - y_2|,$$

where K is a fixed constant. Then (3.6) has a unique solution in some neighbourhood of $x = 0$. Solving (3.6) is equivalent to solving

$$\int_0^x y'(t)dt = \int_0^x f(t, y(t))dt,$$

i.e.,

$$y(x) = y_0 + \int_0^x f(t, y(t))dt. \tag{3.7}$$

Thus the problem is to find a fixed point of the operator encoded by (3.7). In other words, we seek a fixed point of the map $\varphi \mapsto F(\varphi)$, where

$$F(\varphi)(x) = y_0 + \int_0^x f(t, \varphi(t))dt,$$

with $\varphi \in C[-\delta, \delta]$ for some $\delta > 0$, as we establish shortly. Now,

$$|F(\varphi)(x) - F(\psi)(x)| \leq \int_0^x |f(t, \varphi(t)) - f(t, \psi(t))|dt,$$

$$\leq K \int_0^x |\varphi(t) - \psi(t)| dt$$
$$\leq K \cdot 2\delta \cdot \|\varphi - \psi\|_\infty.$$

Thus $F : C[-\delta, \delta] \to C[-\delta, \delta]$ is a contraction whenever $2K\delta < 1$, for instance whenever $\delta < \frac{1}{2K}$. Hence F has a unique fixed point in $C[-\delta, \delta]$. In other words, (3.6) has a unique solution on a sufficiently small interval about 0.

Example 3.44. Consider $y' = 2x(1 + y), \qquad y(0) = 0$. Then

$$\varphi(x) = \int_0^x 2t(1 + \varphi(t))dt.$$

With the initial guess $\varphi^0 \equiv 0$, we get

$$\varphi^1(x) = \int_0^x 2t(1 + 0)\, dt = x^2,$$

$$\varphi^2(x) = \int_0^x 2t(1 + t^2)\, dt = x^2 + \frac{x^4}{2},$$

$$\varphi^3(x) = x^2 + \frac{x^4}{2} + \frac{x^6}{6}.$$

Thus by induction,

$$\varphi^n(x) = \sum_{k=1}^{n} \frac{x^{2k}}{k!} \longrightarrow e^{x^2} - 1, \tag{3.8}$$

and $\varphi(x) = e^{x^2} - 1$ is a solution, in agreement with the solution obtained by separation of variables. Note that the series in (3.8) converges uniformly on every interval $[-a, a]$, and hence on every compact interval. On the other hand, the differential equation $\varphi'(x) = 2x(1 + \varphi(x))$ has a unique local solution near any point x_0; for instance, one may work on an interval $[x_0 - \delta, x_0 + \delta]$ with $\delta < \frac{1}{4}$.

Remark 3.45 (Hint). Lipschitz constant = 2.

Part II

Norms and Function Spaces

Chapter 4

Normed Linear Spaces and Fundamental Inequalities

A norm equips a vector space with quantitative geometry and, in particular, induces a metric. We study basic examples of normed spaces and prove the inequalities of Young, Hölder, and Minkowski that underlie norm estimates in L^p theory and many convergence arguments.

Learning objectives.

- understand how a norm induces geometry and topology on a vector space;
- compare the standard norms on finite-dimensional spaces and prove their basic inequalities;
- derive Young, Hölder, and Minkowski inequalities in forms suitable for later L^p applications;
- use norm estimates systematically in convergence, continuity, and operator arguments.

4.1 Chapter roadmap

Dependencies. The inequalities proved here become the algebraic backbone for later work with function spaces and L^p norms.

- Normed spaces and the metric induced by a norm.
- Geometry in $\mathbb{R}^n$ under ℓ^p norms.
- Young, Hölder, and Minkowski inequalities (with applications to L^p estimates).
- Completeness in normed spaces (Banach spaces) and bounded linear operators.

4.2 Normed linear spaces and fundamental inequalities

Once a vector space is equipped with a norm, algebra and geometry begin to interact in a way that is central to analysis. The inequalities proved in this section are not merely technical estimates; they become the structural tools governing function spaces, operator bounds, and the geometry of L^p theory.

4.2.1 Normed linear spaces

Definition 4.1 (Norm)**.** Let X be a vector space over $\mathbb{R}$ or $\mathbb{C}$. A function $\|\cdot\|: X \to [0,\infty)$ is a *norm* if, for all $x, y \in X$ and all scalars α,

(i) $\|x\| = 0$ if and only if $x = 0$;

(ii) $\|\alpha x\| = |\alpha|\,\|x\|$;

(iii) $\|x + y\| \leq \|x\| + \|y\|$ *(triangle inequality).*

Every norm induces a metric via

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

Conversely, not every metric on a vector space arises from a norm.

Example 4.2. Let X be a vector space. The discrete metric on X is not induced by any norm.

Proof. Suppose, to the contrary, that there were a norm $\|\cdot\|$ on X such that

$$d_0(x, y) = \|x - y\| \qquad (x, y \in X),$$

where d_0 denotes the discrete metric. Fix $x \neq 0$. Then $\|x\| = d_0(x, 0) = 1$. For every nonzero scalar α we also have

$$\|\alpha x\| = d_0(\alpha x, 0) = 1.$$

By homogeneity of the norm,

$$1 = \|\alpha x\| = |\alpha|\,\|x\| = |\alpha|,$$

which is impossible for arbitrary α. Hence the discrete metric cannot arise from a norm. □

Proposition 4.3 (Characterization of norm-induced metrics)**.** *Let X be a vector space, and let d be a metric on X such that*

$$d(x, y) = d(x - y, 0) \qquad \textit{and} \qquad d(\alpha x, \alpha y) = |\alpha|\, d(x, y)$$

for all $x, y \in X$ and all scalars α. Then

$$\|x\| := d(x, 0)$$

defines a norm on X.

Proof. The positivity axiom follows from the fact that d is a metric: $\|x\| = 0$ if and only if $d(x, 0) = 0$, which holds if and only if $x = 0$. The homogeneity axiom is immediate:

$$\|\alpha x\| = d(\alpha x, 0) = d(\alpha x, \alpha 0) = |\alpha|\, d(x, 0) = |\alpha|\,\|x\|.$$

For the triangle inequality, use the metric inequality together with translation invariance:

$$\|x + y\| = d(x + y, 0) = d(x, -y) \leq d(x, 0) + d(-y, 0) = \|x\| + \|y\|.$$

Thus $\|\cdot\|$ is a norm on X. □

Example 4.4. Let ℓ^1 denote the space of all real (or complex) sequences $x = (x_n)$ such that

$$\sum_{n=1}^{\infty} |x_n| < \infty.$$

Then

$$\|x\|_1 = \sum_{n=1}^{\infty} |x_n|$$

defines a norm on ℓ^1. The pair $(\ell^1, \|\cdot\|_1)$ is a normed linear space. For simplicity, we write ℓ^1 for $(\ell^1, \|\cdot\|_1)$.

Remark 4.5 (Hint). For each $k \in \mathbb{N}$,

$$\sum_{n=1}^{k} |x_n + y_n| \leq \sum_{n=1}^{k} |x_n| + \sum_{n=1}^{k} |y_n| \leq \|x\|_1 + \|y\|_1.$$

Example 4.6. ℓ^2 denotes the space of all sequences $x = (x_n)$ over $\mathbb{R}$ (or $\mathbb{C}$) such that

$$\sum_{n=1}^{\infty} |x_n|^2 < \infty.$$

The formula

$$\|x\|_2 := \left(\sum_{n=1}^{\infty} |x_n|^2 \right)^{1/2}$$

defines a norm on ℓ^2.

Remark 4.7 (Hint).

$$\sum_{n=1}^{k} |x_n + y_n|^2 \leq \Big(\Big(\sum_{n=1}^{k} |x_n|^2 \Big)^{1/2} + \Big(\sum_{n=1}^{k} |y_n|^2 \Big)^{1/2} \Big)^2.$$

Example 4.8. ℓ^∞ denotes the space of all sequences on $\mathbb{R}$ (or $\mathbb{C}$) such that

$$\sup_{n \in \mathbb{N}} |x_n| < \infty.$$

The function

$$\|x\|_\infty = \sup_{n \in \mathbb{N}} |x_n|$$

defines a norm on ℓ^∞.

Example 4.9. c_0 denotes the space of all sequences on $\mathbb{R}$ (or $\mathbb{C}$) such that

$$\lim_{n \to \infty} x_n = 0.$$

Then (x_n) must be bounded. Hence

$$\|x\|_\infty = \sup_{n\in\mathbb{N}} |x_n| < \infty.$$

Therefore $(c_0, \|\cdot\|_\infty)$ is a normed linear space.

Exercise 4.10. If $x = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$ (or $\mathbb{C}^n$), then

$$\|x\|_\infty \le \|x\|_1 \le \sqrt{n}\|x\|_2 \le n\|x\|_\infty.$$

4.2.2 Geometry of spheres in $(\mathbb{R}^n, \|\cdot\|_p)$

For $1 \le p \le \infty$ and $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$, define

$$\|x\|_p := \begin{cases} \left(\sum_{i=1}^n |x_i|^p\right)^{1/p}, & 1 \le p < \infty, \\ \max_{1\le i\le n} |x_i|, & p = \infty. \end{cases}$$

For $1 \le p \le \infty$, the mapping $\|\cdot\|_p$ is a norm and induces the metric $d_p(x, y) = \|x - y\|_p$. For $0 < p < 1$, the expression $\|x\|_p$ is no longer a norm (the triangle inequality fails), but $d_p(x, y) := \|x - y\|_p^p$ defines a metric.

In $\mathbb{R}^2$ the *unit balls*

$$B_p(0, 1) := \{x \in \mathbb{R}^2 : \ \|x\|_p \le 1\}$$

have characteristic shapes; see Figure 4.1.

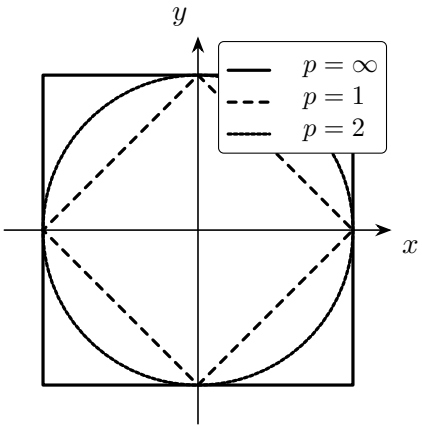


Figure 4.1: Unit balls in $(\mathbb{R}^2, \|\cdot\|_p)$ for selected values of p.

For $0 < p < 1$ the "unit ball" $\{x : \|x\|_p \le 1\}$ is star-shaped but not convex; this reflects the fact that $\|\cdot\|_p$ is not a norm in that regime.

Proposition 4.11 (Basic inclusions among sequence spaces)**.** *The following strict inclusions hold:*

$$\ell^1 \subsetneq \ell^2 \subsetneq c_0 \subsetneq \ell^\infty.$$

Moreover, if $x = (x_n) \in \ell^1$, then $x \in \ell^\infty$ and

$$\|x\|_2^2 \le \|x\|_\infty \, \|x\|_1.$$

Proof. If $x \in \ell^1$, then the series $\sum_{n=1}^\infty |x_n|$ converges, so in particular $|x_n| \to 0$; hence (x_n) is bounded and belongs to ℓ^∞. Also,

$$\sum_{n=1}^\infty |x_n|^2 \le \Big(\sup_{n\in\mathbb{N}} |x_n|\Big) \sum_{n=1}^\infty |x_n| = \|x\|_\infty \, \|x\|_1 < \infty,$$

so $x \in \ell^2$ and the stated estimate follows.

If $x \in \ell^2$, then $x_n \to 0$; indeed, if not, there would exist $\varepsilon > 0$ and infinitely many indices n such that $|x_n| \ge \varepsilon$, forcing $\sum |x_n|^2$ to diverge. Thus $\ell^2 \subset c_0$.

Finally, every null sequence is bounded, so $c_0 \subset \ell^\infty$.

The inclusions are strict. For example,

$$\left(\frac{1}{n}\right)_{n\ge1} \in \ell^2 \setminus \ell^1, \qquad \left(\frac{1}{\sqrt{n}}\right)_{n\ge1} \in c_0 \setminus \ell^2, \qquad (1)_{n\ge1} \in \ell^\infty \setminus c_0.$$

□

Exercise 4.12. If $1 < p < \infty$, then for $\sum_{n=1}^\infty |x_n|^p < \infty$, one can define a norm $\|\cdot\|_p$ on ℓ^p via

$$\|x\|_p = \left(\sum_{n=1}^\infty |x_n|^p\right)^{1/p}.$$

The proof of the triangle inequality is the essential point; it is supplied by Young's, H"older's, and Minkowski's inequalities below.

4.2.3 Young's inequality

Theorem 4.13 (Young's inequality)**.** *Let $1 < p < \infty$, and let q be the conjugate exponent, so that*

$$\frac{1}{p} + \frac{1}{q} = 1.$$

Then for all $a, b \ge 0$,

$$ab \le \frac{a^p}{p} + \frac{b^q}{q}.$$

Equality holds if and only if $a^p = b^q$.

Proof. Consider the convex function $\varphi(t) = t^p/p$ on $[0, \infty)$. Its derivative is $\varphi'(t) = t^{p-1}$. By the supporting-line inequality for convex functions,

$$\varphi(a) \ge \varphi\big(b^{q-1}\big) + \varphi'\big(b^{q-1}\big)\big(a - b^{q-1}\big).$$

Since $(q-1)p = q$ and $(q-1)(p-1) = 1$, this becomes

$$\frac{a^p}{p} \geq \frac{b^q}{p} + b\big(a - b^{q-1}\big) = ab - \frac{b^q}{q}.$$

Rearranging gives the inequality. Because φ is strictly convex, equality holds precisely when $a = b^{q-1}$, which is equivalent to $a^p = b^q$. □

4.2.4 Hölder's inequality

Theorem 4.14 (Hölder's inequality for sequences)**.** *Let $1 \leq p \leq \infty$, and let q be the conjugate exponent. If $x = (x_n) \in \ell^p$ and $y = (y_n) \in \ell^q$, then the termwise product $(x_n y_n)$ belongs to ℓ^1 and*

$$\sum_{n=1}^{\infty} |x_n y_n| \leq \|x\|_p \, \|y\|_q.$$

Proof. If $p = 1$ and $q = \infty$, then

$$\sum_{n=1}^{\infty} |x_n y_n| \leq \Big(\sup_{n\in\mathbb{N}} |y_n|\Big) \sum_{n=1}^{\infty} |x_n| = \|x\|_1 \, \|y\|_\infty.$$

The case $p = \infty$, $q = 1$ is symmetric.

Now assume $1 < p < \infty$. If $x = 0$ or $y = 0$, the estimate is immediate, so suppose both norms are nonzero. For each n define

$$a_n := \frac{|x_n|}{\|x\|_p}, \qquad b_n := \frac{|y_n|}{\|y\|_q}.$$

By Young's inequality,

$$a_n b_n \leq \frac{a_n^p}{p} + \frac{b_n^q}{q} \qquad (n \in \mathbb{N}).$$

Summing from 1 to N gives

$$\sum_{n=1}^{N} \frac{|x_n y_n|}{\|x\|_p \, \|y\|_q} \leq \frac{1}{p} \sum_{n=1}^{N} \frac{|x_n|^p}{\|x\|_p^p} + \frac{1}{q} \sum_{n=1}^{N} \frac{|y_n|^q}{\|y\|_q^q} \leq \frac{1}{p} + \frac{1}{q} = 1.$$

Hence, for every N,

$$\sum_{n=1}^{N} |x_n y_n| \leq \|x\|_p \, \|y\|_q.$$

The partial sums on the left are increasing and uniformly bounded above, so they converge, and passing to the limit as $N \to \infty$ yields

$$\sum_{n=1}^{\infty} |x_n y_n| \leq \|x\|_p \, \|y\|_q.$$

Thus $(x_n y_n) \in \ell^1$. □

4.2.5 Minkowski's inequality

Theorem 4.15 (Minkowski's inequality for sequences)**.** *Let $1 \leq p \leq \infty$. If $x, y \in \ell^p$, then $x + y \in \ell^p$ and*

$$\|x + y\|_p \leq \|x\|_p + \|y\|_p.$$

Consequently, $(\ell^p, \|\cdot\|_p)$ is a normed linear space for every $1 \leq p \leq \infty$.

Proof. For $p = 1$, the inequality follows directly from

$$\sum_{n=1}^{\infty} |x_n + y_n| \leq \sum_{n=1}^{\infty} |x_n| + \sum_{n=1}^{\infty} |y_n|.$$

For $p = \infty$, one has the pointwise estimate

$$|x_n + y_n| \leq |x_n| + |y_n| \leq \|x\|_\infty + \|y\|_\infty,$$

so taking the supremum over n gives

$$\|x + y\|_\infty \leq \|x\|_\infty + \|y\|_\infty.$$

Now let $1 < p < \infty$, and let $q = p/(p-1)$ be the conjugate exponent. The elementary bound

$$|x_n + y_n|^p \leq 2^{p-1}\big(|x_n|^p + |y_n|^p\big)$$

shows that $x + y \in \ell^p$. If $x + y = 0$, the desired estimate is immediate. Otherwise,

$$\begin{aligned}
\|x + y\|_p^p &= \sum_{n=1}^{\infty} |x_n + y_n|^p \\
&= \sum_{n=1}^{\infty} |x_n + y_n|^{p-1} |x_n + y_n| \\
&\leq \sum_{n=1}^{\infty} |x_n|\, |x_n + y_n|^{p-1} + \sum_{n=1}^{\infty} |y_n|\, |x_n + y_n|^{p-1}.
\end{aligned}$$

Apply Hölder's inequality to each sum. Since $(p-1)q = p$, we obtain

$$\sum_{n=1}^{\infty} |x_n|\, |x_n + y_n|^{p-1} \leq \|x\|_p \Big(\sum_{n=1}^{\infty} |x_n + y_n|^{(p-1)q}\Big)^{1/q} = \|x\|_p\, \|x + y\|_p^{p-1},$$

and similarly,

$$\sum_{n=1}^{\infty} |y_n|\, |x_n + y_n|^{p-1} \leq \|y\|_p\, \|x + y\|_p^{p-1}.$$

Therefore

$$\|x + y\|_p^p \leq \big(\|x\|_p + \|y\|_p\big) \|x + y\|_p^{p-1}.$$

Dividing by $\|x+y\|_p^{p-1}$ yields Minkowski's inequality.

The remaining norm axioms are straightforward: positivity and homogeneity are immediate from the definition of $\|\cdot\|_p$. Hence $(\ell^p, \|\cdot\|_p)$ is a normed linear space. □

Theorem 4.16 (Hölder and Minkowski inequalities for Riemann integrable functions)**.** *Let $f, g \in \mathcal{R}[a,b]$. For $1 \le p < \infty$ define*

$$\|f\|_p := \left(\int_a^b |f(t)|^p \, dt \right)^{1/p},$$

and for $p = \infty$ define

$$\|f\|_\infty := \sup_{t \in [a,b]} |f(t)|.$$

Then the following hold.

(i) If $1 \le p \le \infty$ and q is the conjugate exponent, then

$$\|fg\|_1 \le \|f\|_p \, \|g\|_q.$$

(ii) If $1 \le p < \infty$, then

$$\|f+g\|_p \le \|f\|_p + \|g\|_p.$$

(iii) $(\mathcal{R}[a,b], \|\cdot\|_\infty)$ is a normed linear space.

Proof. For $p = 1$ and $q = \infty$ we have

$$\int_a^b |fg| \, dt \le \|g\|_\infty \int_a^b |f| \, dt = \|f\|_1 \, \|g\|_\infty,$$

and the case $p = \infty$, $q = 1$ is symmetric. Now assume $1 < p < \infty$, and first suppose that $\|f\|_p \neq 0$ and $\|g\|_q \neq 0$. Young's inequality applied pointwise to

$$a(t) := \frac{|f(t)|}{\|f\|_p}, \qquad b(t) := \frac{|g(t)|}{\|g\|_q}$$

gives

$$\frac{|f(t)g(t)|}{\|f\|_p \, \|g\|_q} \le \frac{|f(t)|^p}{p \, \|f\|_p^p} + \frac{|g(t)|^q}{q \, \|g\|_q^q}.$$

Integrating over $[a,b]$ yields

$$\frac{\|fg\|_1}{\|f\|_p \, \|g\|_q} \le \frac{1}{p} \frac{\int_a^b |f|^p}{\|f\|_p^p} + \frac{1}{q} \frac{\int_a^b |g|^q}{\|g\|_q^q} = \frac{1}{p} + \frac{1}{q} = 1,$$

which proves Hölder's inequality. If one of the norms is zero, the inequality is immediate.

For Minkowski's inequality, the case $p = 1$ is just the triangle inequality for the integral:

$$\|f + g\|_1 = \int_a^b |f + g|\, dt \le \int_a^b |f|\, dt + \int_a^b |g|\, dt = \|f\|_1 + \|g\|_1.$$

Now let $1 < p < \infty$ and let $q = \frac{p}{p-1}$. If $\|f + g\|_p = 0$, there is nothing to prove. Otherwise,

$$\begin{aligned}\|f + g\|_p^p &= \int_a^b |f + g|\, |f + g|^{p-1}\, dt \\ &\le \int_a^b |f|\, |f + g|^{p-1}\, dt + \int_a^b |g|\, |f + g|^{p-1}\, dt.\end{aligned}$$

Applying Hölder's inequality to each term gives

$$\int_a^b |f|\, |f + g|^{p-1}\, dt \le \|f\|_p \Big(\int_a^b |f + g|^{(p-1)q}\, dt \Big)^{1/q} = \|f\|_p\, \|f + g\|_p^{p-1},$$

and similarly for g. Hence

$$\|f + g\|_p^p \le (\|f\|_p + \|g\|_p)\, \|f + g\|_p^{p-1}.$$

Dividing by $\|f + g\|_p^{p-1}$ proves

$$\|f + g\|_p \le \|f\|_p + \|g\|_p.$$

Finally, for $p = \infty$ the norm axioms are immediate from the pointwise inequalities $|cf| = |c|\, |f|$ and $|f + g| \le |f| + |g|$. Thus $(\mathcal{R}[a, b], \|\cdot\|_\infty)$ is a normed linear space. □

4.2.6 Completeness and Banach spaces

Definition 4.17 (Banach space)**.** A normed linear space $(X, \|\cdot\|)$ is called a *Banach space* if it is complete with respect to the metric $d(x, y) := \|x - y\|$.

Example 4.18. $\mathbb{R}^n$ (or $\mathbb{C}^n$) equipped with any norm is a Banach space. In particular, $(\mathbb{R}^n, \|\cdot\|_p)$ is Banach for every $1 \le p \le \infty$.

Example 4.19. The space $C([a, b])$ of continuous functions on a compact interval, equipped with the supremum norm $\|f\|_\infty := \sup_{x \in [a,b]} |f(x)|$, is a Banach space.

Theorem 4.20 (Finite-dimensional completeness)**.** *Every finite-dimensional normed linear space is complete.*

Proof. Let X be a finite-dimensional normed space, and choose a basis $e_1, \dots, e_n$ of X. Every $x \in X$ can be written uniquely as

$$x = \sum_{j=1}^n a_j e_j.$$

We first claim that there exists a constant $C > 0$ such that

$$\max_{1\le j\le n} |a_j| \le C\Big\|\sum_{j=1}^{n} a_j e_j\Big\|.$$

Suppose not. Then for each $m \in \mathbb{N}$ there exist scalars $a_1^{(m)}, \dots, a_n^{(m)}$ with

$$\max_{1\le j\le n} |a_j^{(m)}| = 1 \qquad \text{and} \qquad \Big\|\sum_{j=1}^{n} a_j^{(m)} e_j\Big\| < \frac{1}{m}.$$

Each scalar sequence $(a_j^{(m)})_{m\ge 1}$ is bounded, so by repeated use of Bolzano–Weierstrass in $\mathbb{R}$ we may pass to a subsequence (not relabelled) such that $a_j^{(m)} \to a_j$ for every j. Then $\max_j |a_j| = 1$, so the vector

$$x := \sum_{j=1}^{n} a_j e_j$$

is nonzero. On the other hand,

$$\Big\|x - \sum_{j=1}^{n} a_j^{(m)} e_j\Big\| \le \sum_{j=1}^{n} |a_j - a_j^{(m)}|\,\|e_j\| \longrightarrow 0,$$

while $\|\sum_{j=1}^{n} a_j^{(m)} e_j\| \to 0$. Hence $\|x\| = 0$, a contradiction. This proves the claim.

Now let (x_k) be a Cauchy sequence in X, and write

$$x_k = \sum_{j=1}^{n} a_j^{(k)} e_j.$$

Then for all k, ℓ and each j,

$$|a_j^{(k)} - a_j^{(\ell)}| \le \max_{1\le i\le n} |a_i^{(k)} - a_i^{(\ell)}| \le C\|x_k - x_\ell\|.$$

Thus each scalar sequence $(a_j^{(k)})_{k\ge 1}$ is Cauchy in $\mathbb{R}$, hence convergent. Let $a_j := \lim_{k\to\infty} a_j^{(k)}$ and set

$$x := \sum_{j=1}^{n} a_j e_j.$$

Then

$$\|x_k - x\| \le \sum_{j=1}^{n} |a_j^{(k)} - a_j|\,\|e_j\| \longrightarrow 0.$$

Therefore $x_k \to x$ in X, so X is complete. □

4.2.7 Bounded linear maps and the operator norm

Definition 4.21 (Bounded linear operator)**.** Let X, Y be normed spaces and $T\colon X \to Y$ be linear. We say that T is *bounded* if there exists $C \geq 0$ such that

$$\|Tx\|_Y \leq C\|x\|_X \qquad \text{for all } x \in X.$$

The least such constant is the *operator norm* and is denoted

$$\|T\|_{\mathrm{op}} := \sup_{\|x\|_X \leq 1} \|Tx\|_Y.$$

Proposition 4.22. *For a linear map $T\colon X \to Y$ between normed spaces, the following are equivalent:*

(i) T is continuous at 0;

(ii) T is continuous on X;

(iii) T is bounded.

Proof. $(iii) \Rightarrow (ii)$ is immediate from $\|T(x) - T(y)\| = \|T(x-y)\| \leq \|T\|_{\mathrm{op}}\|x-y\|$. $(ii) \Rightarrow (i)$ is immediate. For $(i) \Rightarrow (iii)$, continuity at 0 gives $\delta > 0$ such that $\|x\|_X < \delta \Rightarrow \|Tx\|_Y < 1$. For $x \neq 0$, apply this to $\frac{\delta}{2\|x\|_X} x$ to obtain $\|Tx\|_Y \leq \frac{2}{\delta}\|x\|_X$. □

4.2.8 Equivalence of norms in finite dimensions (optional)

Theorem 4.23 (Equivalence of norms on $\mathbb{R}^n$)**.** *Let $\|\cdot\|$ be any norm on $\mathbb{R}^n$. Then there exist constants $c, C > 0$ such that*

$$c\|x\|_2 \leq \|x\| \leq C\|x\|_2 \qquad \textit{for all } x \in \mathbb{R}^n.$$

In particular, all norms on a finite-dimensional vector space define the same notions of convergence, continuity, and compactness.

Proof. Consider the Euclidean unit sphere $S = \{x \in \mathbb{R}^n : \|x\|_2 = 1\}$, which is compact. The function $x \mapsto \|x\|$ is continuous, hence attains a minimum $m > 0$ and a maximum M on S. For $x \neq 0$, write $x = \|x\|_2 \frac{x}{\|x\|_2}$ with $\frac{x}{\|x\|_2} \in S$ to obtain $m\|x\|_2 \leq \|x\| \leq M\|x\|_2$. □

Chapter 5

Sequences of Functions and Uniform Convergence

Uniform convergence is a mode of convergence for functions that is strong enough to justify passing limits through several continuous operations. We develop the basic theory and record standard applications, including continuity of uniform limits, interchange of limit and integral in simple settings, and uniform convergence of power series on compact sets.

Learning objectives.

- distinguish pointwise from uniform convergence and test each notion on concrete examples;
- determine when uniform convergence preserves continuity, integrability, and differentiability;
- use equicontinuity and compactness ideas to extract uniformly convergent subsequences;
- apply local uniform convergence to power series and justify termwise operations rigorously.

5.1 Chapter roadmap

Dependencies. The central question is which analytic operations commute with passage to the limit, and under which mode of convergence.

- Pointwise versus uniform convergence; uniform Cauchy criterion.
- Continuity and uniform limits; exchanging limits with continuous operations.
- Series of functions and the Weierstrass M-test.
- Termwise integration and differentiation under uniform hypotheses (optional addenda).

5.2 Uniform convergence

Pointwise convergence records what happens at each individual point, but analysis often requires control that is simultaneous over an entire domain. Uniform convergence is precisely the mode of convergence strong enough to preserve continuity and to justify several important exchanges of limits with other operations.

5.2.1 Uniform convergence of sequences of functions

In the previous exercises, we established that $(C([0,1]), \|\cdot\|_\infty)$ is complete. That is, if $\|f_n - f_m\|_\infty \to 0$, then there exists $f \in C([0,1])$ such that $\|f_n - f\|_\infty \to 0$. But then,

$$|f_n(t) - f(t)| < \|f_n - f\|_\infty \to 0, \quad \forall t \in [0,1],$$

i.e., $f_n(t) \to f(t)$ for each $t \in [0,1]$. We say that $f_n \to f$ uniformly if

$$\sup_t |f_n(t) - f(t)| \to 0.$$

However, there are sequences of functions that converge pointwise but not uniformly.

Example 5.1. Let $f_n(t) = t^n, \ \ t \in [0,1]$. Then

$$f(t) = \lim_{n\to\infty} f_n(t) = \begin{cases} 0 & 0 \le t < 1 \\ 1 & t = 1 \end{cases}$$

So,

$$\sup_t |f_n(t) - f(t)| = 1 \not\to 0.$$

Example 5.2. Let $f_n : \mathbb{R} \to \mathbb{R}$ be given by

$$f_n(t) = e^{-nt^2}, \quad n \in \mathbb{N}.$$

Then

$$f(t) = \lim_{n\to\infty} f_n(t) = \begin{cases} 1 & t = 0 \\ 0 & |t| > 0 \end{cases}$$

At $t = 0$ we have $|f_n(0) - f(0)| = 0$ for every n. If $|t_0| > 0$, then $t_0^2 > 0$, and to ensure that $|f_n(t_0) - 0| < \varepsilon$ it suffices to require

$$e^{-nt_0^2} < \varepsilon \implies n > \frac{\log \frac{1}{\varepsilon}}{t_0^2}.$$

Let $n_0 = \left\lceil \frac{\log(1/\varepsilon)}{t_0^2} \right\rceil + 1$. Then

$$|f_n(t_0) - f(t_0)| < \varepsilon \qquad \text{for all } n \ge n_0.$$

Observe that $n_0 = n_0(\varepsilon, t_0)$ depends on the point t_0, and becomes large as $|t_0|$ approaches 0. Hence the choice of n_0 cannot be made independently of t_0. Therefore, $f_n \to f$ pointwise but not uniformly. Also,

$$\|f_n - f\|_\infty = \sup_{t\in\mathbb{R}} e^{-nt^2} = 1 \not\to 0.$$

If $f_n(t) = e^{-nt}$ for $t \in [1, \infty)$, then

$$\sup_t |f_n(t) - 0| = e^{-n} \to 0 \implies e^{-nt} \xrightarrow[[1,\infty)]{\text{unif.}} 0.$$

Uniform convergence is particularly useful because it interacts well with basic analytic operations. In particular, uniform limits preserve boundedness and continuity (and, under additional

hypotheses, allow the interchange of limits with integration or differentiation). The results below formalize these stability properties.

Exercise 5.3. Let $f_n, f : A(\subseteq \mathbb{R}) \to \mathbb{R}$ be such that $f_n \to f$ uniformly on A. If each f_n is bounded, then f is bounded.

Remark 5.4 (Hint). Choose n_0 so that $\sup_{t\in A} |f_{n_0}(t) - f(t)| < \varepsilon$. Then for every $t \in A$,

$$|f(t)| \leq |f(t) - f_{n_0}(t)| + |f_{n_0}(t)| < \varepsilon + M_{n_0} < \infty.$$

Example 5.5. Uniform convergence preserves continuity and, with the same proof, also preserves uniform continuity.

Theorem 5.6 (Uniform limit theorem for continuity)**.** *Let $A \subset \mathbb{R}$ and let $f, f_n : A \to \mathbb{R}$ be such that $f_n \to f$ uniformly. If each f_n is continuous, then f is continuous.*

Proof. Fix $t \in A$ and let $\varepsilon > 0$. By uniform convergence, there exists $n_0 \in \mathbb{N}$ such that

$$\sup_{x\in A} |f_{n_0}(x) - f(x)| < \varepsilon.$$

Since f_{n_0} is continuous at t, there exists $\delta > 0$ such that

$$|s - t| < \delta \quad \Longrightarrow \quad |f_{n_0}(s) - f_{n_0}(t)| < \varepsilon.$$

Therefore, whenever $|s - t| < \delta$,

$$\begin{aligned} |f(s) - f(t)| &\leq |f(s) - f_{n_0}(s)| + |f_{n_0}(s) - f_{n_0}(t)| + |f_{n_0}(t) - f(t)| \\ &< 3\varepsilon. \end{aligned}$$

Hence f is continuous at t, and since t was arbitrary, f is continuous on A. □

Theorem 5.7 (Uniform convergence and Riemann integration)**.** *Let $f_n, f \in \mathcal{R}[a, b]$ and suppose that $f_n \to f$ uniformly on $[a, b]$. Then*

$$\int_a^b f_n \to \int_a^b f$$

that is,

$$\lim_{n\to\infty} \int_a^b f_n = \int_a^b \lim_{n\to\infty} f_n$$

Proof.

$$\begin{aligned} \left| \int_a^b (f_n - f) \right| &\leq \int_a^b |f_n - f| \\ &\leq \|f_n - f\|_\infty (b - a) \to 0. \end{aligned}$$

□

Corollary 5.8. *If $f_n \in \mathcal{R}[a,b]$ and the partial sums*

$$S_n := f_1 + f_2 + \cdots + f_n$$

converge uniformly on $[a,b]$ to S, then

$$\int_a^b \sum_{n=1}^{\infty} f_n = \int_a^b S = \sum_{n=1}^{\infty} \int_a^b f_n.$$

Proof. Since $S_n \to S$ uniformly and each S_n is Riemann integrable, the preceding theorem gives

$$\int_a^b S_n \, dt \longrightarrow \int_a^b S \, dt.$$

But

$$\int_a^b S_n \, dt = \sum_{k=1}^{n} \int_a^b f_k \, dt.$$

Passing to the limit in n yields the desired identity. □

Theorem 5.9 (Uniform convergence of derivatives)**.** *Let $f_n \in C^1[a,b]$ be such that $f_n' \to g$ uniformly. If there exists $x_0 \in [a,b]$ such that $f_n(x_0)$ converges, then there exists $f \in C^1[a,b]$ such that $f_n \to f$ uniformly and $f' = g$.*

Proof. Since each f_n' is continuous and $f_n' \to g$ uniformly, the preceding theorem shows that g is continuous on $[a,b]$. Define

$$f(x) := \lim_{n\to\infty} f_n(x_0) + \int_{x_0}^{x} g(t) \, dt, \qquad x \in [a,b].$$

Then $f \in C^1[a,b]$ and $f' = g$.

For $x \in [a,b]$, the fundamental theorem of calculus gives

$$f_n(x) - f_m(x) = (f_n(x_0) - f_m(x_0)) + \int_{x_0}^{x} (f_n'(t) - f_m'(t)) \, dt.$$

Hence

$$|f_n(x) - f_m(x)| \le |f_n(x_0) - f_m(x_0)| + (b-a)\|f_n' - f_m'\|_\infty.$$

Taking the supremum over $x \in [a,b]$, we obtain

$$\|f_n - f_m\|_\infty \le |f_n(x_0) - f_m(x_0)| + (b-a)\|f_n' - f_m'\|_\infty.$$

The right-hand side tends to 0 because $(f_n(x_0))$ converges and (f_n') converges uniformly. Thus (f_n) is Cauchy in $(C[a,b], \|\cdot\|_\infty)$, so by completeness there exists $h \in C[a,b]$ such that $f_n \to h$ uniformly.

Finally, for each $x \in [a, b]$,

$$f_n(x) - f_n(x_0) = \int_{x_0}^{x} f_n'(t)\, dt.$$

Passing to the limit, using uniform convergence of f_n' and convergence of $f_n(x_0)$, we get

$$h(x) - f(x_0) = \int_{x_0}^{x} g(t)\, dt = f(x) - f(x_0).$$

Therefore $h = f$, and consequently $f_n \to f$ uniformly. □

Remark 5.10. Convergence of $(f_n(x_0))$ is necessary in the above result. Consider

$$f_n(t) = \sqrt{t+n}, \qquad t \in [0, 1].$$

Then f_n does not converge at any point of $[0, 1]$, but

$$f_n'(t) = \frac{1}{2\sqrt{t+n}} \xrightarrow{\text{unif.}} 0$$

Since

$$\sup_{t\in[0,1]} |f_n'(t) - 0| = \sup_{t\in[0,1]} \frac{1}{2\sqrt{t+n}} = \frac{1}{2\sqrt{n}} \to 0.$$

Exercise 5.11. Let $f_n : \mathbb{R} \to \mathbb{R}$. Check for uniform convergence of f_n to some f:

1. $f_n(t) = \frac{\sin(nt)}{\sqrt{n}}$.
2. $f_n(t) = n^2 t(1 - t^2)^n$.
3. $f_n(t) = te^{-nt}$.

Also, verify for term-by-term integration and differentiation for each of the above.

Theorem 5.12 (Uniform convergence and passage of limits)**.** *Let $E \subseteq \mathbb{R}$, and suppose that $f_n \to f$ uniformly on E. Let x be a limit point of E, and assume that*

$$\lim_{t\to x} f_n(t) = A_n \qquad \text{for each } n. \tag{5.1}$$

Then (A_n) converges and

$$\lim_{t\to x} f(t) = \lim_{n\to\infty} A_n.$$

Equivalently, whenever the stated limits exist,

$$\lim_{t\to x} \lim_{n\to\infty} f_n(t) = \lim_{n\to\infty} \lim_{t\to x} f_n(t).$$

Proof. Since $f_n \to f$ uniformly on E, the sequence (f_n) is uniformly Cauchy: for every $\varepsilon > 0$ there exists $N \in \mathbb{N}$ such that

$$|f_n(t) - f_m(t)| < \varepsilon \qquad (n, m \geq N, t \in E). \tag{5.2}$$

Letting $t \to x$ in (5.2) and using (5.1), we obtain

$$|A_n - A_m| \leq \varepsilon \qquad (n, m \geq N).$$

Thus (A_n) is Cauchy and therefore converges to some $A \in \mathbb{R}$.

It remains to identify the limit of f at x. Fix $\varepsilon > 0$. Choose $n \in \mathbb{N}$ so large that

$$\sup_{t \in E} |f(t) - f_n(t)| < \varepsilon \qquad \text{and} \qquad |A_n - A| < \varepsilon.$$

Since $\lim_{t \to x} f_n(t) = A_n$, there exists $\delta > 0$ such that

$$t \in E,\ 0 < |t - x| < \delta \quad \Longrightarrow \quad |f_n(t) - A_n| < \varepsilon.$$

For such t,

$$|f(t) - A| \leq |f(t) - f_n(t)| + |f_n(t) - A_n| + |A_n - A| < 3\varepsilon.$$

Hence $\lim_{t \to x} f(t) = A = \lim_{n \to \infty} A_n$. □

Theorem 5.13 (Term-by-term differentiation for series)**.** *Let $f_k \in C^1[a, b]$ for $k \in \mathbb{N}$. Suppose that the numerical series $\sum_{k=1}^{\infty} f_k(x_0)$ converges for some $x_0 \in [a, b]$ and that the series of derivatives*

$$\sum_{k=1}^{\infty} f_k'$$

converges uniformly on $[a, b]$. Then the series $\sum_{k=1}^{\infty} f_k$ converges uniformly on $[a, b]$ to a function $S \in C^1[a, b]$, and

$$S'(x) = \sum_{k=1}^{\infty} f_k'(x) \qquad (x \in [a, b]).$$

Proof. Let $S_n = \sum_{k=1}^{n} f_k$. Then $S_n \in C^1[a, b]$, the sequence $(S_n(x_0))$ converges by hypothesis, and

$$S_n' = \sum_{k=1}^{n} f_k'$$

converges uniformly on $[a, b]$. Applying the preceding differentiation theorem for sequences to (S_n), we obtain a function $S \in C^1[a, b]$ such that $S_n \to S$ uniformly and $S' = \lim_{n \to \infty} S_n'$. This is precisely the asserted term-by-term differentiation formula. □

Remark 5.14 (Newton–Leibniz formula). The correct fundamental identity for a differentiable function is

$$f(x) - f(a) = \int_a^x f'(t)\,dt. \tag{5.3}$$

If $f' \in C[a,b]$, then (5.3) follows immediately from the Fundamental Theorem of Calculus. More generally, the same identity remains valid whenever f' is Riemann integrable on $[a,b]$. Thus continuity of f' is sufficient but not necessary; the real obstruction is not mere discontinuity of f', but failure of the derivative to be integrable in the relevant sense.

5.3 Equicontinuity and the Arzelà–Ascoli theorem (brief)

The preceding discussion concerns the convergence of a specified sequence of functions. Arzelà–Ascoli addresses a subtler question: under what hypotheses does an entire family of functions contain a uniformly convergent subsequence? The additional hypothesis of equicontinuity is what turns boundedness into compactness in this setting.

Definition 5.15 (Equicontinuity)**.** A family $\mathcal{F} \subset C(K)$, where K is a metric space, is *equicontinuous* if for every $\varepsilon > 0$ there exists $\delta > 0$ such that $d(x,y) < \delta$ implies $|f(x) - f(y)| < \varepsilon$ for all $f \in \mathcal{F}$.

Theorem 5.16 (Arzelà–Ascoli)**.** *Let K be compact and $\mathcal{F} \subset C(K)$. If $\mathcal{F}$ is uniformly bounded and equicontinuous, then every sequence in $\mathcal{F}$ has a uniformly convergent subsequence.*

Proof. Let (f_n) be a sequence in $\mathcal{F}$. Since K is compact metric, it is totally bounded and separable. Choose a countable dense set $D = \{x_1, x_2, \dots\} \subset K$. Because the family is uniformly bounded, the scalar sequence $(f_n(x_1))$ is bounded, so it has a convergent subsequence. Passing inductively to subsequences, we obtain a diagonal subsequence (g_n) such that for every $j \in \mathbb{N}$ the sequence $(g_n(x_j))$ converges.

We claim that (g_n) is uniformly Cauchy on K. Fix $\varepsilon > 0$. By equicontinuity there exists $\delta > 0$ such that

$$d(x,y) < \delta \Longrightarrow |h(x) - h(y)| < \varepsilon/3 \qquad \text{for all } h \in \mathcal{F}.$$

By total boundedness, there exist points $y_1, \dots, y_N \in D$ such that

$$K \subset \bigcup_{i=1}^{N} B(y_i, \delta).$$

For each i, the sequence $(g_n(y_i))$ converges, hence is Cauchy. Therefore there exists N_0 such that for all $n, m \geq N_0$,

$$|g_n(y_i) - g_m(y_i)| < \varepsilon/3, \qquad i = 1, \dots, N.$$

Now let $x \in K$. Choose i with $d(x, y_i) < \delta$. Then for $n, m \geq N_0$,

$$\begin{aligned} |g_n(x) - g_m(x)| &\leq |g_n(x) - g_n(y_i)| + |g_n(y_i) - g_m(y_i)| + |g_m(y_i) - g_m(x)| \\ &< \varepsilon. \end{aligned}$$

Thus (g_n) is uniformly Cauchy. Since $C(K)$ is complete under the supremum norm, there exists $g \in C(K)$ such that $g_n \to g$ uniformly on K. Hence (f_n) has a uniformly convergent subsequence, as claimed. □

Remark 5.17. The proof isolates the two key mechanisms behind compactness in $C(K)$: diagonal extraction on a dense set and equicontinuity, which upgrades pointwise Cauchy control to uniform Cauchy control.

5.4 Power series and uniform convergence

Power series lie at the boundary between algebra and analysis: formally they resemble infinite polynomials, but analytically they behave far better than arbitrary series of functions. Their local uniform convergence explains why differentiation and integration may often be carried out term by term inside the radius of convergence.

Theorem 5.18 (Uniform convergence on compact subintervals)**.** *Let $\sum_{n\geq 0} a_n(x - x_0)^n$ be a power series with radius of convergence $R > 0$. Then for every $0 < r < R$ the series converges uniformly on $[x_0 - r, x_0 + r]$.*

Proof. Fix $0 < r < R$. By definition of the radius of convergence, the numerical series $\sum_{n\geq 0} |a_n| r^n$ converges. For $|x - x_0| \leq r$ we have $|a_n(x - x_0)^n| \leq |a_n| r^n$, so uniform convergence follows from the Weierstrass M-test. □

Part III

Analysis in Several Variables

Chapter 6

Functions of Several Variables

This chapter develops the basic analytic framework for functions on $\mathbb{R}^n$. We begin with limits and continuity in the Euclidean metric, then introduce partial and directional derivatives and, more importantly, the correct notion of differentiability as linear approximation in the Fréchet sense. The chain rule is stated and proved in forms suited to composition and coordinate change, after which we discuss higher derivatives and mixed partials. We then establish Taylor's theorem as a systematic higher-order approximation scheme. The chapter culminates in the inverse and implicit function theorems, which provide precise criteria for local invertibility and for representing level sets as graphs.

Learning objectives.

- formulate limits and continuity in $\mathbb{R}^n$ with clear control over paths, norms, and neighborhoods;
- distinguish partial, directional, and Fréchet differentiability, and understand their logical relations;
- use Jacobians, the chain rule, and Taylor expansions as the central computational tools of multivariable analysis;
- apply the inverse and implicit function theorems in a precise local form.

6.1 Chapter roadmap

Dependencies. The decisive distinction throughout this chapter is between partial information and full differentiability; most counterexamples are designed to expose that gap.

This chapter corresponds to the **Functions of Several Variables** unit of MA224. The guiding viewpoint is that differentiability means *linear approximation with a remainder that is small relative to* $\|h\|$. The chapter develops this idea in a sequence that mirrors the syllabus:

(1) **Limits and continuity in $\mathbb{R}^n$:** the ε–δ definition and its sequential characterization, together with basic permanence properties.

(2) **Differentiability and the Fréchet derivative:** partial and directional derivatives, differentiability as the existence of a best linear approximation, and the chain rule.

(3) **Taylor's theorem and quantitative estimates:** higher-order expansions and remainder bounds that control approximation errors.

(4) **Local structure of smooth maps:** the inverse function theorem and implicit function theorem, with applications to local parametrizations and constraint manifolds.

6.2 Limits and continuity

Before one can study differentiability in several variables, one must first understand what it means for points to approach one another in $\mathbb{R}^n$. The main novelty is that there are many possible directions of approach, so examples and counterexamples become especially instructive.

6.2.1 Notation and basic definitions in Euclidean space

Throughout this chapter we work in $\mathbb{R}^n$ with the standard inner product

$$\langle \mathbf{x}, \mathbf{y} \rangle := \sum_{j=1}^{n} x_j y_j, \qquad \mathbf{x} = (x_1, \dots, x_n),\ \mathbf{y} = (y_1, \dots, y_n),$$

and the associated Euclidean norm

$$\|\mathbf{x}\| := \langle \mathbf{x}, \mathbf{x} \rangle^{1/2} = \left(x_1^2 + \cdots + x_n^2\right)^{1/2}. \tag{6.1}$$

We will use (6.1) as the default norm on $\mathbb{R}^n$ (and on $\mathbb{R}^m$ when needed). The Euclidean distance is

$$d(\mathbf{x}, \mathbf{y}) := \|\mathbf{x} - \mathbf{y}\|.$$

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

$$B_r(\mathbf{x}) := \{\mathbf{y} \in \mathbb{R}^n : \|\mathbf{y} - \mathbf{x}\| < r\}.$$

Conventions. Vectors are written in boldface, and $\mathbf{x}_k \to \mathbf{x}$ always means $\|\mathbf{x}_k - \mathbf{x}\| \to 0$. When the meaning is clear from context we may suppress boldface.

We record two inequalities that will be used repeatedly.

Theorem 6.1 (Cauchy–Schwarz inequality)**.** *For all $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$,*

$$|\langle \mathbf{x}, \mathbf{y} \rangle| \le \|\mathbf{x}\|\, \|\mathbf{y}\|.$$

Moreover, equality holds if and only if $\mathbf{x}$ and $\mathbf{y}$ are linearly dependent.

Proof. If $\mathbf{y} = \mathbf{0}$, the conclusion is immediate. Otherwise, consider the quadratic polynomial

$$\begin{aligned} p(t) :&= \|\mathbf{x} - t\mathbf{y}\|^2 \\ &= \langle \mathbf{x} - t\mathbf{y}, \mathbf{x} - t\mathbf{y} \rangle \\ &= \|\mathbf{x}\|^2 - 2t\langle \mathbf{x}, \mathbf{y} \rangle + t^2 \|\mathbf{y}\|^2. \end{aligned}$$

Since $p(t) \ge 0$ for all $t \in \mathbb{R}$, its discriminant is non-positive:

$$4\langle \mathbf{x}, \mathbf{y} \rangle^2 - 4\|\mathbf{x}\|^2 \|\mathbf{y}\|^2 \le 0.$$

This is exactly $|\langle \mathbf{x}, \mathbf{y} \rangle| \le \|\mathbf{x}\| \, \|\mathbf{y}\|$. Equality holds if and only if $p(t)$ has a real root, i.e. $\mathbf{x} - t\mathbf{y} = \mathbf{0}$ for some t. □

Proposition 6.2 (Triangle inequality)**.** *For all* $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$*,*

$$\|\mathbf{x} + \mathbf{y}\| \le \|\mathbf{x}\| + \|\mathbf{y}\|.$$

Proof. By expanding the square and applying Cauchy–Schwarz,

$$\begin{aligned} \|\mathbf{x} + \mathbf{y}\|^2 &= \|\mathbf{x}\|^2 + \|\mathbf{y}\|^2 + 2\langle \mathbf{x}, \mathbf{y} \rangle \\ &\le (\|\mathbf{x}\| + \|\mathbf{y}\|)^2. \end{aligned}$$

Taking square roots gives the result. □

Theorem 6.3 (Bolzano–Weierstrass in $\mathbb{R}^n$)**.** *Every bounded sequence in* $\mathbb{R}^n$ *has a convergent subsequence.*

Proof. Let $\mathbf{x}_k = (x_1^k, \dots, x_n^k)$ be bounded. Then each coordinate sequence $(x_j^k)_{k \ge 1}$ is bounded in $\mathbb{R}$. Apply the one-dimensional Bolzano–Weierstrass theorem to (x_1^k) to obtain a subsequence along which the first coordinate converges. Restrict to that subsequence and repeat for the second coordinate, and so on. After n steps we obtain a subsequence $(\mathbf{x}_{k_\ell})$ such that each coordinate converges; hence $\mathbf{x}_{k_\ell} \to \mathbf{x} \in \mathbb{R}^n$. □

6.2.2 Limits in Euclidean space

The definition of a limit in $\mathbb{R}^n$ uses the Euclidean distance.

Definition. Let $D \subseteq \mathbb{R}^n$ and let $f : D \to \mathbb{R}^m$. Given $x_0 \in \mathbb{R}^n$ that is a limit point of D, we write

$$\lim_{x \to x_0} f(x) = L$$

if for every $\varepsilon > 0$ there exists $\delta > 0$ such that

$$0 < \|x - x_0\| < \delta \ \text{ and } \ x \in D \quad \Longrightarrow \quad \|f(x) - L\| < \varepsilon.$$

Sequential criterion. The limit equals L if and only if for every sequence $(x_k) \subset D$ with $x_k \to x_0$ and $x_k \neq x_0$, one has $f(x_k) \to L$. We will freely use either formulation, depending on convenience.

Paths and polar coordinates in $\mathbb{R}^2$. When $n = 2$ and $x_0 = (0, 0)$, a common way to *disprove* the existence of a limit is to find two paths approaching $(0, 0)$ along which f approaches different values. A common way to *compute* limits is to use polar coordinates $x = r\cos\theta$, $y = r\sin\theta$; then $(x, y) \to (0, 0)$ is equivalent to $r \to 0^+$. In many examples, showing that $\lim_{r \to 0^+} f(r\cos\theta, r\sin\theta)$ exists and is independent of θ suffices.

A caution about the path test. Agreement along a few special curves never *proves* the existence of a multivariable limit; it only rules out some obvious obstructions. To establish existence, one needs an estimate that works for *all* nearby points—for example, a bound in terms of $\sqrt{x^2+y^2}$, or a representation in polar coordinates whose absolute value tends to 0 uniformly in θ.

6.2.3 Continuity in Euclidean space

Let $D \subseteq \mathbb{R}^n$ and $f : D \to \mathbb{R}^m$.

Definition 6.4. The function f is *continuous at* $x_0 \in D$ if for every $\varepsilon > 0$ there exists $\delta > 0$ such that

$$\|x - x_0\| < \delta \ \text{ and } \ x \in D \quad \Longrightarrow \quad \|f(x) - f(x_0)\| < \varepsilon.$$

Equivalently, f is continuous at x_0 if and only if $\lim_{x \to x_0} f(x) = f(x_0)$.

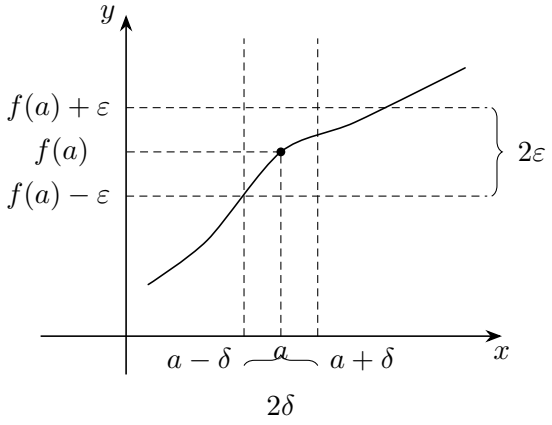


Figure 6.1: Geometric interpretation of the ε–δ definition of continuity at a.

Sequential criterion. The function f is continuous at x_0 if and only if for every sequence $(x_k) \subset D$ with $x_k \to x_0$ one has $f(x_k) \to f(x_0)$.

Componentwise continuity. If $f = (f_1, \dots, f_m) : D \to \mathbb{R}^m$, then f is continuous at x_0 if and only if each component f_i is continuous at x_0. Indeed, $\|f(x) - f(x_0)\| < \varepsilon$ implies $|f_i(x) - f_i(x_0)| \le \|f(x) - f(x_0)\| < \varepsilon$, while conversely one can estimate

$$\|f(x) - f(x_0)\|^2 = \sum_{i=1}^{m} |f_i(x) - f_i(x_0)|^2.$$

Negation of continuity. The function f fails to be continuous at x_0 if and only if there exists $\varepsilon_0 > 0$ such that for every $\delta > 0$ there exists $x \in D$ with $\|x - x_0\| < \delta$ but $\|f(x) - f(x_0)\| \ge \varepsilon_0$.

Example 6.5. Define

$$f(x,y) = \begin{cases} 1, & xy \neq 0, \\ 0, & xy = 0. \end{cases}$$

Then $\lim_{(x,y)\to(0,0)} f(x,y)$ does not exist, because along the path $y = 0$ the function is identically 0, whereas along the path $y = x$ it is identically 1 for $x \neq 0$. By contrast, if one defines

$$g(x,y) = \begin{cases} 1, & xy = 1, \\ 0, & xy \neq 1, \end{cases}$$

then $\lim_{(x,y)\to(0,0)} g(x,y) = 0$, since the condition $xy = 1$ cannot hold near $(0,0)$.

Exercise 6.6. Let $f : \mathbb{R}^2 \to \mathbb{R}$, check the continuity of f at $(0,0)$.

1. $f(x,y) = \begin{cases} \dfrac{xy}{\sqrt{x^2+y^2}} & \text{if } x^2+y^2 \neq 0 \\ 0 & \text{otherwise} \end{cases}$
2. $f(x,y) = \dfrac{\sin^2(x-y)}{\sqrt{x^2+y^2}}, \qquad f(0,0) = 0.$
3. $f(x,y) = \begin{cases} \dfrac{x^2 y}{x^2+y} & \text{if } x^2+y \neq 0 \\ 0 & \text{otherwise} \end{cases}$
4. $f(x,y) = \begin{cases} \dfrac{x^2 y}{x^4+y^2} & \text{if } x^4+y^2 \neq 0 \\ 0 & \text{otherwise} \end{cases}$
5. $f(x,y) = \begin{cases} \dfrac{\sin xy}{xy} & \text{if } xy \neq 0 \\ 0 & \text{otherwise} \end{cases}$

Remark 6.7 (A typical ε–δ negation). Consider the function

$$f(x,y) := \begin{cases} \dfrac{xy}{x^2+y^2}, & (x,y) \neq (0,0), \\ 0, & (x,y) = (0,0). \end{cases}$$

Along the diagonal $y = x$ we have $f(x,x) = \frac{1}{2}$ for all $x \neq 0$. Hence for $\varepsilon_0 := \frac{1}{4}$ and for every

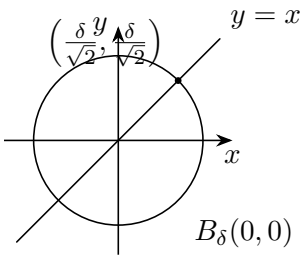


Figure 6.2: The diagonal path $y = x$ intersects every ball $B_\delta(0,0)$, which forces the negation estimate in Remark 6.7.

$\delta > 0$, the point

$$(x_\delta, y_\delta) := \Big(\frac{\delta}{\sqrt{2}}, \frac{\delta}{\sqrt{2}}\Big)$$

satisfies $\sqrt{x_\delta^2 + y_\delta^2} = \delta$ but

$$|f(x_\delta, y_\delta) - f(0,0)| = \frac{1}{2} \geq \varepsilon_0.$$

Therefore f is not continuous at $(0,0)$, and consequently $\lim_{(x,y)\to(0,0)} f(x,y)$ does not exist.

Proposition 6.8 (Composition preserves continuity)**.** *Let $D \subseteq \mathbb{R}^n$ and let $f\colon D \to \mathbb{R}^m$ be continuous at $x_0 \in D$. Let $I \subseteq \mathbb{R}^m$ and let $g\colon I \to \mathbb{R}^k$ be continuous at $f(x_0)$, and assume that $f(D) \subseteq I$. Then the composition $g \circ f\colon D \to \mathbb{R}^k$ is continuous at x_0.*

Proof. Fix $\varepsilon > 0$. Since g is continuous at $f(x_0)$, there exists $\eta > 0$ such that

$$\|u - f(x_0)\| < \eta \implies \|g(u) - g(f(x_0))\| < \varepsilon, \qquad u \in I. \tag{6.2}$$

Since f is continuous at x_0, there exists $\delta > 0$ such that

$$\|x - x_0\| < \delta \implies \|f(x) - f(x_0)\| < \eta, \qquad x \in D. \tag{6.3}$$

Combining (6.3) with (6.2) yields

$$\|x - x_0\| < \delta \implies \|g(f(x)) - g(f(x_0))\| < \varepsilon.$$

Hence $g \circ f$ is continuous at x_0. □

Remark 6.9 (Sequential proof). Alternatively, if $x_n \to x_0$ in D, then $f(x_n) \to f(x_0)$ by continuity of f, and hence $g(f(x_n)) \to g(f(x_0))$ by continuity of g.

Example 6.10.

$$f(x,y) = \begin{cases} \dfrac{\sin xy}{xy} & \text{if } xy \neq 0 \\ 1 & \text{otherwise} \end{cases}$$

$$f(x,y) = p \circ g(x,y), \quad \text{where } p(t) = \begin{cases} \dfrac{\sin t}{t} & t \neq 0 \\ 1 & t = 0 \end{cases}$$

6.3 Differentiation in $\mathbb{R}^n$

Partial derivatives record one-coordinate behaviour, but the genuine analytic notion is stronger: differentiability means that the function is well approximated by a single linear map in every direction at once. Much of multivariable calculus consists of learning to distinguish these ideas and to understand when partial information is sufficient.

6.3.1 Partial derivatives

Let $D \subseteq \mathbb{R}^2$ be open and $f : D \to \mathbb{R}$. Fix $(x_0, y_0) \in D$.

$$\frac{\partial f}{\partial x}(x_0, y_0) := \lim_{h \to 0} \frac{f(x_0 + h, y_0) - f(x_0, y_0)}{h}, \tag{6.4}$$

provided the limit exists. Equation (6.4) is the defining limit for the partial derivative in the x-direction. Similarly,

$$\frac{\partial f}{\partial y}(x_0, y_0) := \lim_{k \to 0} \frac{f(x_0, y_0 + k) - f(x_0, y_0)}{k},$$

when the limit exists. We also use the notation $f_x(x_0, y_0)$ and $f_y(x_0, y_0)$.

First-order expansion along coordinate lines. If $f_x(x_0, y_0)$ exists, then

$$f(x_0 + h, y_0) = f(x_0, y_0) + h\, f_x(x_0, y_0) + o(h) \qquad (h \to 0),$$

and analogously for f_y. The important point is that the existence of $\partial f/\partial x$ depends only on the values of f along the line segment $\{(x_0 + h, y_0) : |h| \text{ small}\}$, not on the behavior of f in two dimensions.

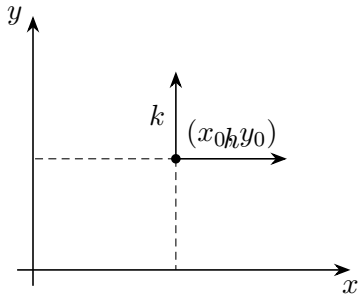


Figure 6.3: Partial derivatives measure change along coordinate directions: f_x varies x while holding $y = y_0$ fixed (horizontal arrow), and f_y varies y while holding $x = x_0$ fixed (vertical arrow).

Warning. The existence of partial derivatives at a point does *not* imply continuity or differentiability at that point.

Example 6.11. $f(x, y) = \dfrac{xy}{x^2 + y^2}$, $f(0, 0) = 0$. Then $f_x(0, 0) = 0 = f_y(0, 0)$ but f is not continuous at $(0, 0)$.

6.3.2 Directional derivatives

The directional derivative measures the instantaneous rate of change of f along a prescribed direction.

Let $D \subseteq \mathbb{R}^n$ be open and $f : D \to \mathbb{R}$. Fix $x_0 \in D$ and a unit vector $\mathbf{v} \in \mathbb{R}^n$ (i.e. $\|\mathbf{v}\| = 1$). The *directional derivative* of f at x_0 in the direction $\mathbf{v}$ is

$$D_{\mathbf{v}} f(x_0) := \lim_{t \to 0} \frac{f(x_0 + t\mathbf{v}) - f(x_0)}{t}, \tag{6.5}$$

provided the limit exists. As in (6.5), only the values of f along the line segment $\{x_0 + t\mathbf{v} : |t| \text{ small}\}$ matter.

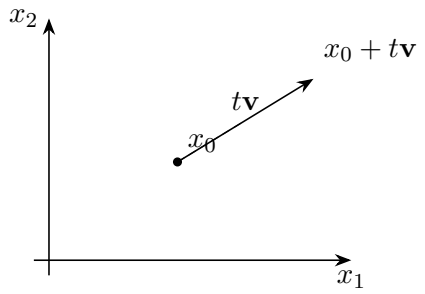


Figure 6.4: Directional derivatives probe f along the line $t \mapsto x_0 + t\mathbf{v}$. Only the behavior on this one-dimensional slice is relevant.

Example 6.12. Define

$$f(x,y) = \begin{cases} \dfrac{x^2 y}{x^4 + y^2} & \text{if } (x,y) \neq (0,0), \\ 0 & \text{if } (x,y) = (0,0). \end{cases}$$

Then every directional derivative $D_{\mathbf{v}} f(0,0)$ exists at $(0,0)$. Indeed, for $\mathbf{v} = (v_1, v_2)$ with $\|\mathbf{v}\| = 1$,

$$\frac{f(t\mathbf{v}) - f(0,0)}{t} = \frac{v_1^2 v_2}{t^2 v_1^4 + v_2^2} \xrightarrow[t \to 0]{} \begin{cases} 0, & v_2 = 0, \\ \dfrac{v_1^2}{v_2}, & v_2 \neq 0. \end{cases}$$

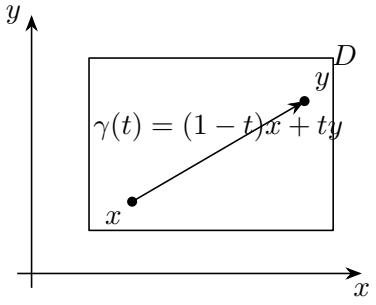


Figure 6.5: On a convex set D, the straight-line path from x to y stays inside D, so the one-variable mean value theorem applied to $t \mapsto f(\gamma(t))$ yields constancy when $\nabla f \equiv 0$.

6.3.3 Differentiability

Partial and directional derivatives capture *one-dimensional* behavior of f along lines. The correct multivariable notion is differentiability as *linear approximation*.

Let $D \subseteq \mathbb{R}^n$ be open and let $f : D \to \mathbb{R}^m$. Fix $x_0 \in D$ and write $H \in \mathbb{R}^n$ for an increment.

Definition 6.13 (Fréchet differentiability)**.** We say that f is *differentiable at* x_0 if there exists a linear map $A : \mathbb{R}^n \to \mathbb{R}^m$ such that

$$\frac{\|f(x_0 + H) - f(x_0) - AH\|}{\|H\|} \longrightarrow 0 \qquad \text{as } \|H\| \to 0. \tag{6.6}$$

The linear map A is unique; it is called the *derivative* of f at x_0 and is denoted by $Df(x_0)$ or $f'(x_0)$.

Uniqueness of the derivative. Suppose A and B both satisfy (6.6). Then

$$\frac{\|(A-B)H\|}{\|H\|} \le \frac{\|f(x_0+H) - f(x_0) - AH\|}{\|H\|} + \frac{\|f(x_0+H) - f(x_0) - BH\|}{\|H\|} \longrightarrow 0.$$

Taking $H = t\mathbf{v}$ with $\mathbf{v} \neq \mathbf{0}$ and letting $t \to 0$ shows $\|(A-B)\mathbf{v}\| = 0$ for every $\mathbf{v}$, hence $A = B$. □

Consequences. If f is differentiable at x_0, then f is continuous at x_0, and every directional derivative exists with

$$D_{\mathbf{v}} f(x_0) = f'(x_0)\mathbf{v} \qquad (\mathbf{v} \in \mathbb{R}^n,\ \|\mathbf{v}\| = 1).$$

In particular, for scalar-valued $f : D \to \mathbb{R}$ the derivative $f'(x_0)$ can be identified with the gradient row vector $\nabla f(x_0)$, and $f'(x_0)\mathbf{v} = \nabla f(x_0) \cdot \mathbf{v}$.

Example 6.14 (Locally bounded partial derivatives imply continuity)**.** Let $D \subseteq \mathbb{R}^2$ be open and let $f : D \to \mathbb{R}$. Suppose that the partial derivatives f_x and f_y exist and are bounded on a neighbourhood of each point of D. Then f is locally Lipschitz on D and therefore continuous on D.

Proof. Fix $(x_0, y_0) \in D$. Since D is open, there is $r > 0$ such that the closed square

$$Q = [x_0 - r, x_0 + r] \times [y_0 - r, y_0 + r]$$

is contained in D. By the local boundedness hypothesis, after decreasing r if necessary there is $M < \infty$ such that $|f_x| \le M$ and $|f_y| \le M$ on Q. If $|h| < r$ and $|k| < r$, the two coordinate

segments joining (x_0, y_0) to $(x_0 + h, y_0 + k)$ lie in Q. The one-variable mean value theorem applied along these segments gives

$$\begin{aligned} |f(x_0+h, y_0+k) - f(x_0, y_0)| &\le |f(x_0+h, y_0+k) - f(x_0, y_0+k)| \\ &\quad + |f(x_0, y_0+k) - f(x_0, y_0)| \\ &\le M|h| + M|k| \\ &\le \sqrt{2}\, M\sqrt{h^2+k^2}. \end{aligned}$$

This proves a Lipschitz estimate in a neighbourhood of (x_0, y_0) and hence continuity there. Since the point was arbitrary, f is continuous on D. □

Exercise 6.15. Even if $f_x(x_0)$ and $f_y(x_0)$ exist, f need not be differentiable at x_0. Find an explicit example (or verify one of the examples below).

Example 6.16 (Directional derivatives may all exist without differentiability)**.** Define

$$f(x,y) = \begin{cases} \dfrac{y}{|y|}\sqrt{x^2+y^2}, & y \neq 0, \\ 0, & y = 0. \end{cases}$$

Then f is continuous at $(0,0)$, every two-sided directional derivative $D_{\mathbf{v}} f(0,0)$ exists, but f is not differentiable at $(0,0)$ since the map $\mathbf{v} \mapsto D_{\mathbf{v}} f(0,0)$ is not linear.

Proof. Continuity at the origin follows from $|f(x,y)| \le \sqrt{x^2+y^2}$. For $\mathbf{v} = (v_1, v_2)$ with $\|\mathbf{v}\| = 1$,

$$\frac{f(t\mathbf{v}) - f(0,0)}{t} = \begin{cases} 1, & v_2 > 0, \\ -1, & v_2 < 0, \\ 0, & v_2 = 0, \end{cases} \qquad (t \to 0).$$

Thus $D_{\mathbf{v}} f(0,0)$ exists for every $\mathbf{v}$. If f were differentiable at $(0,0)$, then $D_{\mathbf{v}} f(0,0) = Df(0,0)\mathbf{v}$ would depend linearly on $\mathbf{v}$, which is impossible. □

Theorem 6.17 (A sufficient condition for differentiability)**.** *Let $D \subseteq \mathbb{R}^2$ be an open set, and let $(x_0, y_0) \in D$. Suppose that the partial derivatives f_x and f_y exist in a neighborhood of (x_0, y_0) and are continuous at (x_0, y_0). Then f is differentiable at (x_0, y_0). Moreover,*

$$f'(x_0, y_0) = \big(f_x(x_0, y_0),\, f_y(x_0, y_0)\big).$$

Proof. Write $X_0 = (x_0, y_0)$ and let $H = (h, k)$. Since D is open, there exists $r > 0$ such that the closed ball $\overline{B(X_0, r)}$ is contained in D. We may therefore assume that $\|H\| < r$, so that all points used below lie in D.

Decompose the increment as

$$f(x_0+h,y_0+k)-f(x_0,y_0) = \big[f(x_0+h,y_0+k)-f(x_0,y_0+k)\big] + \big[f(x_0,y_0+k)-f(x_0,y_0)\big].$$

By the one-variable mean value theorem, there exist $\theta_1,\theta_2 \in (0,1)$ such that

$$f(x_0+h,y_0+k)-f(x_0,y_0) = f_x(x_0+\theta_1 h, y_0+k)\,h + f_y(x_0,y_0+\theta_2 k)\,k.$$

Subtracting $f_x(x_0,y_0)h + f_y(x_0,y_0)k$, we get

$$\begin{aligned} &f(x_0+h,y_0+k)-f(x_0,y_0)-f_x(x_0,y_0)h-f_y(x_0,y_0)k \\ &= \big[f_x(x_0+\theta_1 h,y_0+k)-f_x(x_0,y_0)\big]h \\ &\quad + \big[f_y(x_0,y_0+\theta_2 k)-f_y(x_0,y_0)\big]k. \end{aligned}$$

Now let $\epsilon>0$. Since f_x and f_y are continuous at (x_0,y_0), there exists $\delta>0$ such that

$$\|(u,v)-(x_0,y_0)\|<\delta \quad\Longrightarrow\quad |f_x(u,v)-f_x(x_0,y_0)|<\epsilon, \quad |f_y(u,v)-f_y(x_0,y_0)|<\epsilon.$$

If $\|H\|<\delta$, then both points $(x_0+\theta_1 h, y_0+k)$ and $(x_0,y_0+\theta_2 k)$ lie within distance at most $\|H\|$ of (x_0,y_0). Therefore

$$\begin{aligned} &\big|f(x_0+h,y_0+k)-f(x_0,y_0)-f_x(x_0,y_0)h-f_y(x_0,y_0)k\big| \\ &\le \epsilon|h|+\epsilon|k| \le \sqrt{2}\,\epsilon\,\|(h,k)\|. \end{aligned}$$

Thus

$$\frac{\big|f(X_0+H)-f(X_0)-\nabla f(X_0)\cdot H\big|}{\|H\|} \le \sqrt{2}\,\epsilon$$

whenever $0<\|H\|<\delta$. □

Exercise 6.18. Prove that

$$f(x,y)=\begin{cases}(x^2+y^2)\,\sin\!\left(\dfrac{1}{x^2+y^2}\right), & x^2+y^2\neq 0,\\[2mm] 0, & (x,y)=(0,0),\end{cases}$$

is differentiable at $(0,0)$ and $f'(0,0)=\mathbf{0}$, but neither f_x nor f_y is continuous at $(0,0)$.

Geometric interpretation (scalar case). If $f:\mathbb{R}^n\to\mathbb{R}$ is differentiable at X_0, then the graph of f has a tangent hyperplane at $(X_0,f(X_0))$ given by

$$z = f(X_0)+f'(X_0)\,(X-X_0) = f(X_0)+\nabla f(X_0)\cdot(X-X_0).$$

For $n = 1$ this is the familiar tangent line; for $n = 2$ it is a tangent plane. In general, $f'(X_0)$ is the best linear approximation to f near X_0.

Proposition 6.19 (Tangent line, tangent plane, and tangent hyperplane)**.** *If $f : \mathbb{R}^n \to \mathbb{R}$ is differentiable at X_0, then*

$$f(X_0 + H) = f(X_0) + f'(X_0)\, H + o(\|H\|),$$

so the graph of f near $(X_0, f(X_0))$ is approximated by the affine hyperplane

$$z = f(X_0) + f'(X_0)\,(X - X_0).$$

For $n = 1$ this reduces to the tangent line

$$z = f(x_0) + f'(x_0)(x - x_0),$$

and for $n = 2$ it becomes the tangent plane

$$z = f(x_0, y_0) + f_x(x_0, y_0)(x - x_0) + f_y(x_0, y_0)(y - y_0).$$

Proof. The differentiability relation

$$f(X_0 + H) - f(X_0) - f'(X_0)\, H = o(\|H\|)$$

says precisely that the affine function

$$L(X) := f(X_0) + f'(X_0)\,(X - X_0)$$

approximates $f(X)$ with an error negligible compared to $\|X - X_0\|$. Thus L is the tangent hyperplane to the graph at $(X_0, f(X_0))$. The formulas for $n = 1$ and $n = 2$ are the corresponding coordinate expressions. □

6.3.4 Chain rule

One-variable chain rule. Let $I, J \subseteq \mathbb{R}$ be intervals, $g : I \to J$, and $f : J \to \mathbb{R}$. If g is differentiable at $x \in I$ and f is differentiable at $g(x) \in J$, then $f \circ g$ is differentiable at x and

$$(f \circ g)'(x) = f'(g(x))\, g'(x).$$

Proof. Set $y = g(x)$ and $k = g(x+h) - g(x)$. Since f is differentiable at y, there exists a function η with $\eta(k) \to 0$ as $k \to 0$ such that

$$f(y + k) - f(y) - f'(y)k = k\,\eta(k).$$

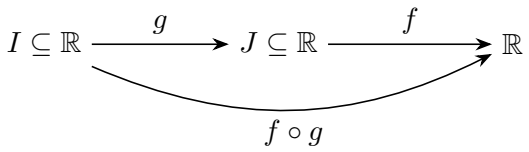


Figure 6.6: The chain rule differentiates the composition $f \circ g$ by combining the derivatives of g and f at the appropriate points.

Since g is differentiable at x, we may write $k = hg'(x) + h\mu(h)$ with $\mu(h) \to 0$ as $h \to 0$. Consider the error term

$$\varepsilon(h) := \frac{f(g(x+h)) - f(g(x)) - f'(g(x))\, g'(x)\, h}{h}. \tag{6.7}$$

We will show that $\varepsilon(h) \to 0$ as $h \to 0$; this is exactly the statement that the difference quotient in (6.7) tends to $f'(g(x))g'(x)$. Substituting the expansions above and simplifying gives

$$\varepsilon(h) = \eta(k)\big(g'(x) + \mu(h)\big) + f'(y)\mu(h).$$

As $h \to 0$ we have $k \to 0$, hence $\eta(k) \to 0$; also $\mu(h) \to 0$. Therefore $\varepsilon(h) \to 0$, which means that the difference quotient tends to $f'(g(x))g'(x)$. □

Notation 6.20 (Matrix spaces and the operator norm). Let $L_n(\mathbb{R})$ denote the space of linear maps $\mathbb{R}^n \to \mathbb{R}^n$ (equivalently, real $n \times n$ matrices), and let $GL_n(\mathbb{R}) \subset L_n(\mathbb{R})$ be the group of invertible matrices.

For a linear map $A : \mathbb{R}^n \to \mathbb{R}^m$ we use the (subordinate) *operator norm*

$$\|A\| := \sup_{\|x\|=1} \|Ax\| = \sup_{x \neq 0} \frac{\|Ax\|}{\|x\|}.$$

It satisfies the basic estimates

$$\|Ax\| \le \|A\|\,\|x\|, \qquad \|AB\| \le \|A\|\,\|B\|.$$

Examples.

(1) If $A : \mathbb{R}^2 \to \mathbb{R}$ is given by $A(x, y) = 4x + 3y$, then by Cauchy–Schwarz, $|A(x, y)| \le \sqrt{4^2 + 3^2}\,\sqrt{x^2 + y^2} = 5\|(x, y)\|$, and equality is attained at $(x, y) = (4, 3)/5$. Hence $\|A\| = 5$.

(2) If $A : \mathbb{R}^2 \to \mathbb{R}^2$ is given by $A(x, y) = (3x, 4y)$, then $\|A(x, y)\|^2 = 9x^2 + 16y^2 \le 16(x^2 + y^2)$, with equality when $x = 0$. Hence $\|A\| = 4$.

Since $L_n(\mathbb{R})$ is finite-dimensional, all norms on $L_n(\mathbb{R})$ are equivalent, and the choice affects only constants, not topological statements.

Proposition 6.21 (Existence of the operator norm)**.** *Let $A : \mathbb{R}^n \to \mathbb{R}^m$ be linear. Then the quantity*

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

is finite, and in fact the supremum is attained:

$$\|A\| = \max_{\|x\|=1} \|Ax\|.$$

Equivalently,

$$\|A\| = \sup_{x \neq 0} \frac{\|Ax\|}{\|x\|},$$

and therefore

$$\|Ax\| \leq \|A\| \, \|x\| \qquad \text{for all } x \in \mathbb{R}^n.$$

Proof. Write $A = (R_1, \ldots, R_m)^t$, where $R_i \in \mathbb{R}^n$ is the ith row of the matrix of A. For any $x \in \mathbb{R}^n$, Cauchy–Schwarz gives

$$|(Ax)_i| = |R_i \cdot x| \leq \|R_i\| \, \|x\|, \qquad i = 1, \ldots, m.$$

Hence

$$\|Ax\|^2 = \sum_{i=1}^{m} |(Ax)_i|^2 \leq \left(\sum_{i=1}^{m} \|R_i\|^2 \right) \|x\|^2,$$

so $\|Ax\|$ is bounded on the unit sphere $S^{n-1} = \{x \in \mathbb{R}^n : \|x\| = 1\}$. The map $x \mapsto \|Ax\|$ is continuous, and S^{n-1} is compact; therefore a maximum is attained on S^{n-1}.

If $x \neq 0$, write $x = \|x\|u$ with $\|u\| = 1$; then

$$\frac{\|Ax\|}{\|x\|} = \|Au\| \leq \max_{\|v\|=1} \|Av\| = \|A\|.$$

Taking the supremum over $x \neq 0$ gives

$$\sup_{x \neq 0} \frac{\|Ax\|}{\|x\|} \leq \|A\|.$$

The reverse inequality follows by restricting the supremum to vectors with $\|x\| = 1$, so the two formulas agree. The estimate $\|Ax\| \leq \|A\| \, \|x\|$ is immediate. □

Chain rule in $\mathbb{R}^n$. Let $U \subseteq \mathbb{R}^n$, $V \subseteq \mathbb{R}^m$ be open. If $g : U \to V$ is differentiable at $x \in U$ and $f : V \to \mathbb{R}^k$ is differentiable at $g(x)$, then $f \circ g$ is differentiable at x and

$$(f \circ g)'(x) = f'(g(x)) \, g'(x),$$

where the right-hand side is ordinary matrix multiplication of the Jacobians.

Proof. Set $y = g(x)$. Since f is differentiable at y, there exists a function $\eta : \mathbb{R}^m \to \mathbb{R}^k$ such that

$$f(y+k) - f(y) - f'(y)k = \|k\|\,\eta(k), \qquad \eta(k) \to 0 \quad (k \to 0).$$

Since g is differentiable at x, there exists a function $\varepsilon : \mathbb{R}^n \to \mathbb{R}^m$ such that

$$g(x+h) - g(x) = g'(x)h + \|h\|\,\varepsilon(h), \qquad \varepsilon(h) \to 0 \quad (h \to 0).$$

Now set

$$k := g(x+h) - g(x).$$

Because differentiability implies continuity, g is continuous at x, hence $\|h\| \to 0$ implies $\|k\| \to 0$. Moreover,

$$\|k\| = \|g'(x)h + \|h\|\,\varepsilon(h)\| \le \|g'(x)\|\,\|h\| + \|h\|\,\|\varepsilon(h)\|.$$

Consider

$$\mu(h) := \frac{(f \circ g)(x+h) - (f \circ g)(x) - f'(g(x))\,g'(x)\,h}{\|h\|}.$$

Using $y = g(x)$ and $k - g'(x)h = \|h\|\,\varepsilon(h)$, we obtain

$$\begin{aligned}
\mu(h) &= \frac{f(y+k) - f(y) - f'(y)\,g'(x)h}{\|h\|} \\
&= \frac{f(y+k) - f(y) - f'(y)k + f'(y)\big(k - g'(x)h\big)}{\|h\|} \\
&= \frac{\|k\|\,\eta(k) + \|h\|\,f'(y)\varepsilon(h)}{\|h\|}.
\end{aligned}$$

Hence

$$\|\mu(h)\| \le \frac{\|k\|}{\|h\|}\,\|\eta(k)\| + \|f'(y)\|\,\|\varepsilon(h)\| \le \|\eta(k)\|\big(\|g'(x)\| + \|\varepsilon(h)\|\big) + \|f'(y)\|\,\|\varepsilon(h)\|.$$

As $h \to 0$, we have $k \to 0$, so $\eta(k) \to 0$, and also $\varepsilon(h) \to 0$. Therefore $\mu(h) \to 0$. This proves that $f \circ g$ is differentiable at x and

$$(f \circ g)'(x) = f'(g(x))\,g'(x).$$

□

Proposition 6.22 (Vanishing gradient implies constancy on convex sets)**.** *Let $D \subseteq \mathbb{R}^n$ be open and convex, and let $f \in C^1(D)$ satisfy $\nabla f \equiv \mathbf{0}$ on D. Then f is constant on D.*

Proof. Fix $x, y \in D$. Since D is convex, the line segment $\gamma(t) = (1-t)x + ty$ lies in D for $t \in [0,1]$. Define $\varphi(t) = f(\gamma(t))$. By the chain rule,

$$\varphi'(t) = \nabla f(\gamma(t)) \cdot (y - x) = 0 \quad \text{for all } t \in [0,1].$$

Hence φ is constant, so $f(y) = \varphi(1) = \varphi(0) = f(x)$. □

Worked examples and a useful identity.

Example 6.23 (Radial compositions)**.** Let $\phi : \mathbb{R} \to \mathbb{R}$ be differentiable and define $F : \mathbb{R}^n \to \mathbb{R}$ by $F(x) = \phi(\|x\|^2)$. Then F is differentiable and

$$DF(x) = \phi'(\|x\|^2)\, D(\|x\|^2) = 2\phi'(\|x\|^2)\, x^t, \qquad \text{so} \qquad \nabla F(x) = 2\phi'(\|x\|^2)\, x.$$

More generally, for an integer $k \geq 1$ and $G(x) = \phi(\|x\|^{2k})$ one has (for $x \neq 0$)

$$\nabla G(x) = 2k\, \|x\|^{2k-2}\, \phi'(\|x\|^{2k})\, x.$$

Theorem 6.24 (Euler's homogeneous function theorem)**.** *Let $U \subseteq \mathbb{R}^n \setminus \{0\}$ be open and let $f : U \to \mathbb{R}$ be differentiable. Assume that f is* homogeneous of degree $\alpha \in \mathbb{R}$*, meaning that*

$$f(tx) = t^\alpha f(x) \qquad \textit{for all } x \in U \textit{ and all } t > 0 \textit{ with } tx \in U.$$

Then for every $x \in U$,

$$x \cdot \nabla f(x) = \alpha f(x).$$

Proof. Fix $x \in U$ and consider the one-variable function $\psi(t) = f(tx)$ for t near 1. By the chain rule, $\psi'(t) = f'(tx)\, x = \nabla f(tx) \cdot x$. The homogeneity hypothesis gives $\psi(t) = t^\alpha f(x)$, so $\psi'(t) = \alpha t^{\alpha-1} f(x)$. Evaluating at $t = 1$ yields $\nabla f(x) \cdot x = \alpha f(x)$. □

Corollary 6.25 (Homogeneous maps near the origin)**.** *Suppose $f : \mathbb{R}^n \setminus \{0\} \to \mathbb{R}$ is homogeneous of degree $\alpha > 1$ and extends continuously to 0 by setting $f(0) := 0$. Then f is differentiable at 0 and $f'(0) = 0$.*

Proof. Write $h = \|h\| v$ with $\|v\| = 1$. By homogeneity, $|f(h)| = \|h\|^\alpha |f(v)|$. Continuity on the compact unit sphere implies $M := \sup_{\|v\|=1} |f(v)| < \infty$, hence

$$\frac{|f(h) - f(0)|}{\|h\|} = \frac{|f(h)|}{\|h\|} \leq M\, \|h\|^{\alpha-1} \xrightarrow[h \to 0]{} 0.$$

This is exactly differentiability at 0 with derivative 0. □

Mean value theorem on convex domains. Let $\mathcal{D} \subseteq \mathbb{R}^n$ be open and convex, and let $f : \mathcal{D} \to \mathbb{R}$ be differentiable. Then for any $x, y \in \mathcal{D}$ there exists a point c on the open line segment (x, y) such that

$$f(y) - f(x) = f'(c)\, (y - x) = \nabla f(c) \cdot (y - x).$$

Proof. Define $\varphi(t) = f\big((1-t)x + ty\big)$ for $t \in [0, 1]$. By the chain rule, φ is differentiable on $(0, 1)$ and

$$\varphi'(t) = f'\big((1-t)x + ty\big)\,(y-x).$$

Applying the one-variable mean value theorem to φ yields $\lambda \in (0,1)$ such that

$$\begin{aligned} f(y) - f(x) &= \varphi(1) - \varphi(0) \\ &= \varphi'(\lambda) \\ &= f'\big((1-\lambda)x + \lambda y\big)\,(y-x). \end{aligned}$$

Set $c = (1-\lambda)x + \lambda y$. □

Exercise 6.26. Let U be an open and connected subset of $\mathbb{R}^n$, and let $f : U \to \mathbb{R}$ be differentiable with $f'(x) = 0$ for every $x \in U$. Show that f is constant on U.

Jacobian matrix for vector-valued maps. Let $U \subseteq \mathbb{R}^n$ be open and $f = (f_1, \ldots, f_m) : U \to \mathbb{R}^m$ be differentiable at $x_0 \in U$. Then each partial derivative $\partial f_i/\partial x_j(x_0)$ exists and

$$f'(x_0) = \left(\frac{\partial f_i}{\partial x_j}(x_0)\right)_{1\le i\le m,\ 1\le j\le n}.$$

Proof. Let e_j be the jth standard basis vector of $\mathbb{R}^n$. Differentiability gives

$$\frac{\|f(x_0 + he_j) - f(x_0) - f'(x_0)(he_j)\|}{|h|} \longrightarrow 0 \qquad (h \to 0).$$

Since $f'(x_0)(he_j) = h\,f'(x_0)e_j$, dividing by h shows that

$$\lim_{h\to 0} \frac{f(x_0 + he_j) - f(x_0)}{h} = f'(x_0)e_j \in \mathbb{R}^m,$$

so every component limit exists; equivalently, each $\partial f_i/\partial x_j(x_0)$ exists. Moreover, the vector $f'(x_0)e_j$ is exactly the jth column of the Jacobian matrix, hence $f'(x_0)$ is the matrix of partial derivatives. □

Example 6.27 (Jacobian at a point does not guarantee differentiability)**.** Define $F : \mathbb{R}^2 \to \mathbb{R}^2$ by

$$F(x,y) := \begin{cases} \left(\dfrac{x^2 y}{x^2+y^2},\ \dfrac{xy^2}{x^2+y^2}\right), & (x,y) \neq (0,0), \\ (0,0), & (x,y) = (0,0). \end{cases}$$

Then all first partial derivatives of the component functions exist at $(0,0)$ and the Jacobian matrix $F'(0,0)$ is the zero matrix, but F is *not* differentiable at $(0,0)$.

Proof. Along the coordinate axes we have $F(h,0) = (0,0)$ and $F(0,k) = (0,0)$, so each partial derivative at $(0,0)$ exists and equals 0, hence $F'(0,0) = 0$. If F were differentiable at $(0,0)$ with

derivative 0, we would have $\|F(h,k)\|/\sqrt{h^2+k^2} \to 0$ as $(h,k) \to (0,0)$. However, along the diagonal $y = x$,

$$F(t,t) = \left(\frac{t}{2}, \frac{t}{2}\right), \qquad t \neq 0.$$

Thus

$$\frac{\|F(t,t)\|}{\|(t,t)\|} = \frac{\sqrt{(t/2)^2 + (t/2)^2}}{\sqrt{t^2+t^2}} = \frac{1}{2}$$

for all $t \neq 0$, which does not tend to 0 as $t \to 0$. Therefore F is not differentiable at $(0,0)$. □

Higher derivatives and the Hessian (scalar case). If $f : U \subseteq \mathbb{R}^n \to \mathbb{R}$ is differentiable on U, then each component of ∇f may itself be differentiated. The matrix of second partial derivatives, when it exists, is called the *Hessian*:

$$f''(x) = \left(\frac{\partial^2 f}{\partial x_j \, \partial x_k}(x)\right)_{1 \le j,k \le n}.$$

The next theorem shows that, under mild regularity assumptions, mixed second partials agree.

Example 6.28 (Mixed partials can disagree without continuity)**.** Define $f : \mathbb{R}^2 \to \mathbb{R}$ by

$$f(x,y) := \begin{cases} \dfrac{xy(x^2-y^2)}{x^2+y^2}, & (x,y) \neq (0,0), \\ 0, & (x,y) = (0,0). \end{cases}$$

Then the mixed partials at $(0,0)$ exist but are different: $f_{yx}(0,0) = 1$ whereas $f_{xy}(0,0) = -1$. Indeed, for $h \neq 0$ one computes

$$f_y(h,0) = \lim_{k \to 0} \frac{f(h,k) - f(h,0)}{k} = \lim_{k \to 0} h \, \frac{h^2 - k^2}{h^2 + k^2} = h,$$

so $f_{yx}(0,0) = \lim_{h \to 0} \frac{f_y(h,0) - f_y(0,0)}{h} = 1$. Similarly, for $k \neq 0$,

$$f_x(0,k) = \lim_{h \to 0} \frac{f(h,k) - f(0,k)}{h} = \lim_{h \to 0} k \, \frac{h^2 - k^2}{h^2 + k^2} = -k,$$

so $f_{xy}(0,0) = \lim_{k \to 0} \frac{f_x(0,k) - f_x(0,0)}{k} = -1$.

Notation 6.29 (Smoothness classes). Let $\mathcal{D} \subset \mathbb{R}^n$ be open. For $k \in \mathbb{N}$, we write $C^k(\mathcal{D})$ for the space of functions $f : \mathcal{D} \to \mathbb{R}$ whose partial derivatives of order $\le k$ exist and are continuous on $\mathcal{D}$. In particular, $C^1(\mathcal{D})$ consists of continuously differentiable functions and $C^2(\mathcal{D})$ consists of twice continuously differentiable functions.

Theorem 6.30 (Clairaut–Schwarz)**.** *Let $\mathcal{D} \subset \mathbb{R}^2$ be open and let $f \in C^2(\mathcal{D})$. Then for every $(x_0, y_0) \in \mathcal{D}$,*

$$\frac{\partial^2 f}{\partial x \, \partial y}(x_0, y_0) = \frac{\partial^2 f}{\partial y \, \partial x}(x_0, y_0).$$

More generally, if $\mathcal{D} \subset \mathbb{R}^n$ is open and $f \in C^2(\mathcal{D})$, then

$$\frac{\partial^2 f}{\partial x_j \, \partial x_k} = \frac{\partial^2 f}{\partial x_k \, \partial x_j}, \qquad 1 \le j, k \le n.$$

Proof. Fix $(x_0, y_0) \in \mathcal{D}$. Since $\mathcal{D}$ is open, there exists $r > 0$ such that the rectangle $[x_0 - r, x_0 + r] \times [y_0 - r, y_0 + r]$ is contained in $\mathcal{D}$. For (x, y) in this rectangle define

$$F(x, y) := f(x, y) - f(x_0, y) - f(x, y_0) + f(x_0, y_0).$$

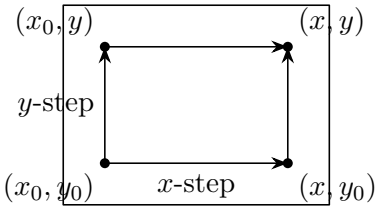


Figure 6.7: The rectangle used in the proof of Clairaut–Schwarz: compare increments by moving in the x-direction then y-direction, versus y then x.

Note that $F(x_0, y) = 0$ and $F(x, y_0) = 0$. Fix $x \neq x_0$ and apply the one-variable mean value theorem to the map $y \mapsto f(x, y) - f(x_0, y)$ to obtain a point η between y and y_0 such that

$$F(x, y) = (y - y_0)\big(f_y(x, \eta) - f_y(x_0, \eta)\big).$$

Now fix $y \neq y_0$ and apply the mean value theorem to the map $x \mapsto f_y(x, \eta)$ to obtain a point ξ between x and x_0 such that

$$f_y(x, \eta) - f_y(x_0, \eta) = (x - x_0) f_{xy}(\xi, \eta).$$

Consequently,

$$\frac{F(x, y)}{(x - x_0)(y - y_0)} = f_{xy}(\xi, \eta).$$

Letting $(x, y) \to (x_0, y_0)$ forces $(\xi, \eta) \to (x_0, y_0)$, hence by continuity of f_{xy} we obtain

$$\lim_{(x,y) \to (x_0, y_0)} \frac{F(x, y)}{(x - x_0)(y - y_0)} = f_{xy}(x_0, y_0).$$

Repeating the argument with the roles of x and y interchanged yields the same limit equal to $f_{yx}(x_0, y_0)$. Therefore $f_{xy}(x_0, y_0) = f_{yx}(x_0, y_0)$. □

6.3.5 Taylor's theorem

Theorem 6.31 (Second-order Taylor formula in $\mathbb{R}^n$)**.** *Let $\mathcal{D} \subseteq \mathbb{R}^n$ be open and let $f \in C^2(\mathcal{D})$. Fix $X \in \mathcal{D}$. Then there exists $\delta > 0$ such that for every $H \in \mathbb{R}^n$ with $\|H\| < \delta$ there exists $\lambda \in (0,1)$ for which*

$$f(X+H) = f(X) + f'(X)\,H + \frac{1}{2}\,H^t\,f''(C)\,H, \qquad C = X + \lambda H,$$

where $f'(X)$ is the derivative (gradient row vector) and $f''(C)$ is the Hessian matrix at C.

How to read the formula. For scalar-valued f, the derivative $f'(X)$ can be identified with the gradient row vector $\nabla f(X)$, so the first-order term is $f'(X)\,H = \nabla f(X) \cdot H$. The matrix $f''(C)$ is the Hessian at C, and $H^t f''(C) H$ is the associated quadratic form.

Proof. Fix $X \in \mathcal{D}$ and choose $\delta > 0$ so that $X + tH \in \mathcal{D}$ for all $t \in [0,1]$ whenever $\|H\| < \delta$. For such H, define $g : [0,1] \to \mathbb{R}$ by $g(t) = f(X + tH)$. Then $g \in C^2([0,1])$, and by the chain rule,

$$g'(t) = f'(X + tH)\,H, \qquad g''(t) = H^t\,f''(X + tH)\,H.$$

Apply the one-variable Taylor theorem (with Lagrange remainder) at $t = 0$ to obtain $\lambda \in (0,1)$ such that

$$g(1) = g(0) + g'(0) + \frac{1}{2}\,g''(\lambda).$$

Substituting the expressions for g, g', g'' and writing $C = X + \lambda H$ gives the stated formula. □

6.3.6 Multi-index notation and the general Taylor formula

For higher-order expansions in $\mathbb{R}^n$, it is convenient to use *multi-indices*. A multi-index is a vector $\alpha = (\alpha_1, \ldots, \alpha_n) \in \mathbb{N}_0^n$. We write $|\alpha| := \alpha_1 + \cdots + \alpha_n$, $\alpha! := \alpha_1! \cdots \alpha_n!$, and for $h = (h_1, \ldots, h_n) \in \mathbb{R}^n$,

$$h^\alpha := h_1^{\alpha_1} \cdots h_n^{\alpha_n}.$$

For a sufficiently differentiable function f, define

$$D^\alpha f := \frac{\partial^{|\alpha|} f}{\partial x_1^{\alpha_1} \cdots \partial x_n^{\alpha_n}}.$$

Theorem 6.32 (Taylor's theorem in $\mathbb{R}^n$ (multi-index form))**.** *Let $U \subset \mathbb{R}^n$ be open, let $k \in \mathbb{N}$, and let $f \in C^{k+1}(U)$. Fix $a \in U$. Then there exists a neighborhood V of a such that $a + h \in U$ for all $h \in V$, and for each such h,*

$$f(a+h) = \sum_{|\alpha| \le k} \frac{D^\alpha f(a)}{\alpha!}\,h^\alpha + R_k(a,h),$$

where the remainder satisfies

$$\lim_{h\to 0} \frac{R_k(a,h)}{\|h\|_2^k} = 0.$$

In fact, under the stated C^{k+1} hypothesis one has the sharper estimate

$$R_k(a,h) = O(\|h\|_2^{k+1}) \qquad (h \to 0).$$

In particular, the polynomial part gives the best k-th order approximation of f near a in the sense of little-o.

Proof. Because U is open, there exists $r > 0$ such that $a + th \in U$ whenever $\|h\|_2 < r$ and $0 \le t \le 1$. Fix such an h and define the one-variable function

$$\phi_h(t) := f(a + th), \qquad 0 \le t \le 1.$$

Since $f \in C^{k+1}(U)$, we have $\phi_h \in C^{k+1}[0,1]$. By repeated use of the chain rule,

$$\phi_h^{(j)}(t) = \sum_{|\alpha|=j} \frac{j!}{\alpha!}\, D^\alpha f(a + th)\, h^\alpha, \qquad j = 0, 1, \ldots, k+1.$$

Applying the one-variable Taylor theorem at $t = 0$ gives

$$\phi_h(1) = \sum_{j=0}^{k} \frac{\phi_h^{(j)}(0)}{j!} + \frac{\phi_h^{(k+1)}(\theta_h)}{(k+1)!}$$

for some $\theta_h \in (0,1)$. Substituting the expression for $\phi_h^{(j)}(0)$ yields

$$\phi_h(1) = \sum_{j=0}^{k} \sum_{|\alpha|=j} \frac{D^\alpha f(a)}{\alpha!}\, h^\alpha + \frac{1}{(k+1)!}\phi_h^{(k+1)}(\theta_h),$$

which is exactly

$$f(a+h) = \sum_{|\alpha|\le k} \frac{D^\alpha f(a)}{\alpha!}\, h^\alpha + R_k(a,h)$$

with

$$R_k(a,h) = \frac{1}{(k+1)!}\phi_h^{(k+1)}(\theta_h).$$

Now

$$\phi_h^{(k+1)}(\theta_h) = \sum_{|\alpha|=k+1} \frac{(k+1)!}{\alpha!}\, D^\alpha f(a + \theta_h h)\, h^\alpha.$$

Since the derivatives $D^\alpha f$ are continuous, they are bounded on a small closed ball around a. Hence there exists $C > 0$ such that

$$|R_k(a,h)| \le C \sum_{|\alpha|=k+1} |h^\alpha| \le C' \|h\|_2^{k+1}$$

for h sufficiently small. Therefore $R_k(a,h) = O(\|h\|_2^{k+1})$, and in particular

$$\frac{R_k(a,h)}{\|h\|_2^k} \longrightarrow 0 \qquad (h \to 0).$$

□

Remark 6.33. There are several equivalent remainder formulations (integral form, mean-value form, and little-o form). For many applications in analysis, the multi-index expansion above is the most efficient way to organize the local information carried by the derivatives.

A mean value inequality for vector-valued functions.

Theorem 6.34. *Let $f : [a,b] \to \mathbb{R}^n$ be differentiable on (a,b) and continuous on $[a,b]$. Then there exists $\lambda \in (a,b)$ such that*

$$\|f(b) - f(a)\| \le \|f'(\lambda)\|\,(b-a).$$

Proof. Consider the scalar function $g(t) = (f(b) - f(a)) \cdot f(t)$ on $[a,b]$. Then g is continuous on $[a,b]$ and differentiable on (a,b), with

$$g'(t) = (f(b) - f(a)) \cdot f'(t).$$

By the one-variable mean value theorem, there exists $\lambda \in (a,b)$ such that

$$g(b) - g(a) = g'(\lambda)\,(b-a).$$

Since $g(b) - g(a) = (f(b) - f(a)) \cdot (f(b) - f(a)) = \|f(b) - f(a)\|^2$, we obtain

$$\begin{aligned} \|f(b) - f(a)\|^2 &= (b-a)\,(f(b) - f(a)) \cdot f'(\lambda) \\ &\le (b-a)\,\|f(b) - f(a)\|\,\|f'(\lambda)\|, \end{aligned}$$

and the conclusion follows (trivially if $f(b) = f(a)$). □

Theorem 6.35 (Mean value inequality along line segments)**.** *Let $\mathcal{D} \subseteq \mathbb{R}^n$ be open and let $f : \mathcal{D} \to \mathbb{R}^m$ be differentiable. Fix $X \in \mathcal{D}$. Then there exists $\varepsilon > 0$ such that whenever $\|H\| < \varepsilon$ we can find $\lambda \in (0,1)$ with*

$$\|f(X+H) - f(X)\| \le \|f'(X + \lambda H)\|\,\|H\|.$$

Proof. Choose $\varepsilon > 0$ so that $X + tH \in \mathcal{D}$ for all $t \in [0,1]$ whenever $\|H\| < \varepsilon$. Fix such H and define $g(t) = f(X + tH)$ for $t \in [0,1]$. Then $g : [0,1] \to \mathbb{R}^m$ is differentiable and $g'(t) = f'(X + tH)\, H$. Apply Theorem 6.34 to g on $[0,1]$ to obtain $\lambda \in (0,1)$ such that

$$\|f(X+H) - f(X)\| = \|g(1) - g(0)\| \leq \|g'(\lambda)\| \leq \|f'(X + \lambda H)\|\, \|H\|.$$

□

Remark 6.36 (Equality need not hold in the vector mean value inequality). The scalar mean value theorem produces an exact identity, but for vector-valued functions the correct statement is generally only an inequality. For example, define $g : [-1,1] \to \mathbb{R}^2$ by

$$g(t) = (t^3, 1 - t^2).$$

Then

$$g(1) - g(-1) = (2,0), \qquad \|g(1) - g(-1)\| = 2.$$

Also

$$g'(t) = (3t^2, -2t), \qquad \|g'(t)\| = \sqrt{9t^4 + 4t^2} = |t|\sqrt{9t^2 + 4}.$$

If one had an exact vector-valued mean value formula

$$g(1) - g(-1) = g'(c)\,(1 - (-1))$$

for some $c \in (-1,1)$, then comparison of components would give

$$2 = 6c^2, \qquad 0 = -4c.$$

The first equation forces $c^2 = 1/3$, whereas the second forces $c = 0$, a contradiction. Thus no single tangent vector need reproduce the secant vector exactly.

Geometric picture. The image of g is the planar curve

$$g(t) = (t^3, 1 - t^2), \qquad -1 \leq t \leq 1,$$

which joins $g(-1) = (-1,0)$ to $g(1) = (1,0)$. The secant vector $g(1) - g(-1)$ is horizontal, whereas at the point $c = \frac{1}{\sqrt{3}}$ forced by the first component of the impossible exact formula, the tangent direction $g'(c) = (1, -\frac{2}{\sqrt{3}})$ is not horizontal. The figure below makes this mismatch visible.

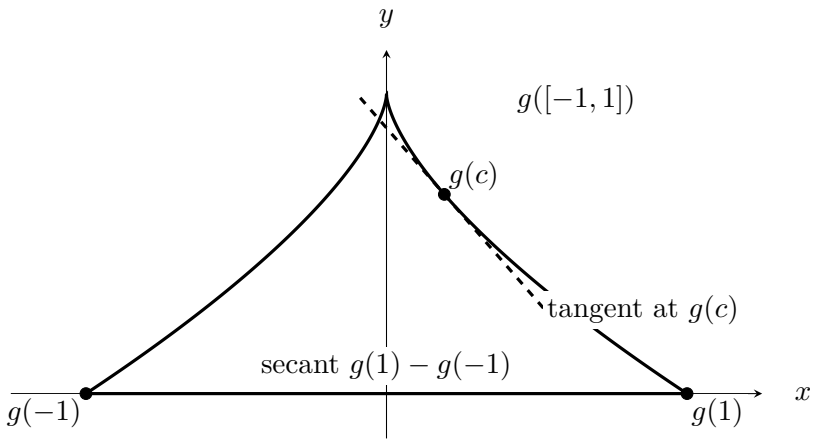


Figure for Remark 6.36: the secant direction determined by the endpoints is horizontal, but the tangent direction at the point forced by the first component is not.

Thus the one-dimensional equality has no direct vector-valued analogue in general; the norm inequality is the correct replacement.

6.3.7 Coordinate changes: a polar-coordinate template

The multivariable chain rule becomes especially transparent when one rewrites a function in new coordinates. A standard instance is the polar-coordinate change of variables

$$x = r\cos\theta, \qquad y = r\sin\theta, \qquad r > 0.$$

If $u = u(x, y)$ is defined on a punctured neighborhood of the origin, we write

$$U(r, \theta) := u(r\cos\theta, r\sin\theta).$$

Proposition 6.37 (Chain rule in polar coordinates)**.** *Assume that u is differentiable at $(x, y) \neq (0, 0)$ and let $x = r\cos\theta$, $y = r\sin\theta$ with $r > 0$. Then*

$$U_r(r, \theta) = u_x(x, y)\cos\theta + u_y(x, y)\sin\theta,$$

and

$$U_\theta(r, \theta) = -r\, u_x(x, y)\sin\theta + r\, u_y(x, y)\cos\theta.$$

If, moreover, $u \in C^2$ on a punctured neighborhood of the origin, then

$$u_{xx}(x, y) + u_{yy}(x, y) = U_{rr}(r, \theta) + \frac{1}{r}\, U_r(r, \theta) + \frac{1}{r^2}\, U_{\theta\theta}(r, \theta).$$

Proof. Since $U(r, \theta) = u(x(r, \theta), y(r, \theta))$, the chain rule gives

$$U_r = u_x\, x_r + u_y\, y_r \qquad \text{and} \qquad U_\theta = u_x\, x_\theta + u_y\, y_\theta.$$

Now

$$x_r = \cos\theta, \quad y_r = \sin\theta, \quad x_\theta = -r\sin\theta, \quad y_\theta = r\cos\theta,$$

which yields the first two formulas.

To derive the Laplacian identity, first solve the linear system above for u_x and u_y:

$$u_x = \cos\theta\, U_r - \frac{\sin\theta}{r}\, U_\theta, \qquad u_y = \sin\theta\, U_r + \frac{\cos\theta}{r}\, U_\theta.$$

Differentiate once more and simplify. The mixed terms cancel, and one obtains

$$u_{xx} + u_{yy} = U_{rr} + \frac{1}{r}\, U_r + \frac{1}{r^2}\, U_{\theta\theta}.$$

□

Example 6.38 (The Laplacian in polar form)**.** If $u(x, y) = x^2 + y^2$, then $U(r, \theta) = r^2$. Hence

$$U_r = 2r, \qquad U_{rr} = 2, \qquad U_{\theta\theta} = 0,$$

and Theorem 6.37 gives

$$u_{xx} + u_{yy} = 2 + \frac{1}{r}(2r) + 0 = 4,$$

which matches the direct computation $u_{xx} = 2$, $u_{yy} = 2$.

Example 6.39 (A classical local-solvability test)**.** Consider

$$F(x, y) = x^2 + ye^x - \cos(xy).$$

At a point (x_0, y_0) satisfying $F(x_0, y_0) = 0$, the equation $F(x, y) = 0$ can be solved locally for y as a C^1 function of x whenever

$$F_y(x_0, y_0) = e^{x_0} + x_0 \sin(x_0 y_0) \neq 0.$$

This is exactly the nondegeneracy condition singled out by the implicit function theorem.

Proposition 6.40 (Derivative of the inverse in one dimension)**.** *Let $I \subset \mathbb{R}$ be an open interval, let $f : I \to \mathbb{R}$ be continuously differentiable, and fix $x_0 \in I$ with $f'(x_0) \neq 0$. Then there are open intervals $U \subset I$ containing x_0 and V containing $y_0 = f(x_0)$ such that $f : U \to V$ is bijective. Its inverse is differentiable at y_0, and*

$$(f^{-1})'(y_0) = \frac{1}{f'(x_0)}.$$

Proof. By continuity of f', there is an open interval $U \subset I$ containing x_0 on which f' has the same sign as $f'(x_0)$ and satisfies

$$|f'(x)| \geq m > 0 \qquad (x \in U).$$

Thus f is strictly monotone on U and maps U bijectively onto the open interval $V = f(U)$.

Let $k \to 0$ with $y_0 + k \in V$, and put

$$h := f^{-1}(y_0 + k) - x_0.$$

Then $x_0 + h \in U$ and

$$k = f(x_0 + h) - f(x_0).$$

Since the local inverse maps points of V near y_0 to points of U near x_0, we have $h \to 0$ as $k \to 0$. Therefore

$$\frac{f^{-1}(y_0 + k) - f^{-1}(y_0)}{k} = \frac{h}{f(x_0 + h) - f(x_0)} = \left(\frac{f(x_0 + h) - f(x_0)}{h}\right)^{-1} \longrightarrow \frac{1}{f'(x_0)}.$$

This proves the formula for the derivative of the inverse. □

Example 6.41 (Invertible Jacobian does not imply global one-to-one behavior)**.** Define $F : \mathbb{R}^2 \to \mathbb{R}^2$ by

$$F(x, y) = \big(e^x \cos y,\ e^x \sin y\big).$$

Then

$$DF(x, y) = \begin{pmatrix} e^x \cos y & -e^x \sin y \\ e^x \sin y & e^x \cos y \end{pmatrix}, \qquad \det DF(x, y) = e^{2x} > 0$$

for every $(x, y) \in \mathbb{R}^2$. Thus the Jacobian matrix is invertible at every point. Nevertheless F is not injective on $\mathbb{R}^2$, because

$$F(x, y + 2\pi) = F(x, y) \qquad \text{for all } (x, y) \in \mathbb{R}^2.$$

This example is a useful warning before the inverse function theorem: invertibility of the derivative guarantees *local* invertibility, not a global inverse on the whole domain.

6.4 Inverse and implicit function theorems

Differentiability becomes especially powerful when the derivative is nondegenerate. The inverse and implicit function theorems formalize this principle: an invertible derivative forces a map to admit local coordinates, while an invertible block of the derivative allows an equation to be solved locally for selected variables.

We now turn to two fundamental results describing the local structure of C^1 maps: the inverse function theorem, which yields local invertibility, and the implicit function theorem, which identifies suitable level sets locally as graphs.

Proposition 6.42 (Stability of invertibility and continuity of inversion)**.** *Let $A \in GL_n(\mathbb{R})$ and $B \in L_n(\mathbb{R})$ satisfy*

$$\|B - A\| < \frac{1}{\|A^{-1}\|}.$$

Then $B \in GL_n(\mathbb{R})$. Moreover, the map $A \mapsto A^{-1}$ is continuous on $GL_n(\mathbb{R})$.

Proof. Set $\alpha = \frac{1}{\|A^{-1}\|}$ and $\beta = \|B - A\|$, so $\beta < \alpha$. For any $x \in \mathbb{R}^n$,

$$\begin{aligned}
\alpha\|x\| &= \alpha\|A^{-1}Ax\| \\
&\leq \alpha\|A^{-1}\|\,\|Ax\| \\
&= \|Ax\| \\
&\leq \|(A-B)x\| + \|Bx\| \\
&\leq \beta\|x\| + \|Bx\|.
\end{aligned}$$

Hence $(\alpha-\beta)\|x\| \leq \|Bx\|$ for all x. If $Bx = 0$, this forces $x = 0$, so B is injective and therefore invertible (finite dimensions). For continuity, note that for $y \neq 0$ and $x = B^{-1}y$,

$$\begin{aligned}
(\alpha-\beta)\|B^{-1}y\| \leq \|y\| & \\
\Rightarrow \quad \|B^{-1}\| & \\
\leq \frac{1}{\alpha-\beta}. &
\end{aligned}$$

Then

$$\begin{aligned}
\|B^{-1} - A^{-1}\| &= \|B^{-1}(A-B)A^{-1}\| \\
&\leq \|B^{-1}\|\,\|A-B\|\,\|A^{-1}\| \\
&\leq \frac{\|A-B\|}{\alpha(\alpha-\beta)} \xrightarrow[B\to A]{} 0.
\end{aligned}$$

Thus inversion is continuous on $GL_n(\mathbb{R})$. □

6.4.1 Inverse function theorem

Theorem 6.43 (Inverse function theorem)**.** *Let $\Omega \subset \mathbb{R}^n$ be open, let $f : \Omega \to \mathbb{R}^n$ be of class C^1, and let $x_0 \in \Omega$ satisfy*

$$\det f'(x_0) \neq 0.$$

Then there exist open neighborhoods U of x_0 and V of $f(x_0)$ such that

$$f : U \to V$$

is bijective. Moreover, the inverse map $f^{-1} : V \to U$ is of class C^1, and

$$(f^{-1})'(f(x_0)) = (f'(x_0))^{-1}.$$

Idea. Near x_0, the map f is well-approximated by its linearization $A = f'(x_0)$. The hypothesis $\det A \neq 0$ means this linear approximation is invertible. The proof turns the equation $f(x) = y$ into a fixed-point problem whose solution follows from the contraction mapping principle, uniformly for y in a small neighborhood of $f(x_0)$.

Proof. Let $A = f'(x_0)$. Since $\det A \neq 0$, the matrix A is invertible. **Step 0 (fixed-point reformulation).** For $y \in \mathbb{R}^n$, define

$$\varphi_y(x) := x + A^{-1}(y - f(x)), \qquad x \in \Omega. \tag{6.8}$$

Then $\varphi_y(x) = x$ if and only if $f(x) = y$. In other words, solving $f(x) = y$ is equivalent to finding a fixed point of φ_y. **Step 1 (a uniform contraction estimate).** Because Ω is open and f' is continuous at x_0, we may choose $\delta > 0$ such that $\overline{B}_\delta(x_0) \subset \Omega$ and

$$\|x - x_0\| < \delta \implies \|f'(x) - A\| < \frac{1}{2\|A^{-1}\|}.$$

Set $U := B_\delta(x_0)$ and $V := f(U)$. By the stability of invertibility, $f'(x)$ is invertible for every $x \in U$. Moreover, for $x \in U$ we have

$$\begin{aligned} \|\varphi_y'(x)\| &= \|I - A^{-1}f'(x)\| \\ &= \|A^{-1}(A - f'(x))\| \\ &\leq \|A^{-1}\|\,\|A - f'(x)\| < \frac{1}{2}. \end{aligned}$$

By the mean value estimate for C^1 maps on convex sets (applied to φ_y on U), it follows that

$$\|\varphi_y(x_1) - \varphi_y(x_2)\| \leq \frac{1}{2}\|x_1 - x_2\| \qquad \text{for all } x_1, x_2 \in U. \tag{6.9}$$

Thus φ_y is a contraction on U (with contraction constant $1/2$), uniformly in y. **(i) f is one-to-one on U.** If $f(x_1) = f(x_2) = y$ with $x_1, x_2 \in U$, then both x_1 and x_2 are fixed points of φ_y. By (6.9), a contraction has at most one fixed point, hence $x_1 = x_2$. **(ii) V is open (and $f : U \to V$ is bijective).** Fix $y^* \in V$. Choose $x^* \in U$ with $y^* = f(x^*)$ and pick $r > 0$ such that the closed ball $\overline{B}_r(x^*) \subset U$. We claim that

$$\|y - y^*\| < \frac{r}{2\|A^{-1}\|} \quad \implies \quad y \in V. \tag{6.10}$$

Indeed, for such y we have

$$\|\varphi_y(x^*) - x^*\| = \|A^{-1}(y - y^*)\| \le \|A^{-1}\|\,\|y - y^*\| < \frac{r}{2}.$$

If $x \in \overline{B}_r(x^*)$, then using (6.9) we obtain

$$\begin{aligned}\|\varphi_y(x) - x^*\| &\le \|\varphi_y(x) - \varphi_y(x^*)\| + \|\varphi_y(x^*) - x^*\| < \frac{1}{2}\|x - x^*\| + \frac{r}{2}\\ &\le r.\end{aligned}$$

Hence $\varphi_y(\overline{B}_r(x^*)) \subset \overline{B}_r(x^*)$. Since $\overline{B}_r(x^*)$ is complete and φ_y is a contraction there, the contraction mapping principle gives a unique fixed point $x \in \overline{B}_r(x^*)$. This fixed point satisfies $f(x) = y$, so $y \in f(U) = V$, proving (6.10). Thus every point of V is an interior point, so V is open. Since $V = f(U)$ and f is injective on U, the restriction $f : U \to V$ is bijective. **(iii) Differentiability of f^{-1} at $f(x_0)$.** Let $y_0 = f(x_0)$ and let k be small enough that $y_0 + k \in V$. Set

$$h := f^{-1}(y_0 + k) - x_0, \qquad \text{so that} \qquad k = f(x_0 + h) - f(x_0).$$

Using (6.8) with $y = y_0 + k$, the fixed-point property $\varphi_{y_0+k}(x_0 + h) = x_0 + h$ and $\varphi_{y_0+k}(x_0) = x_0 + A^{-1}k$ yield

$$h - A^{-1}k = \varphi_{y_0+k}(x_0 + h) - \varphi_{y_0+k}(x_0).$$

By the contraction estimate (6.9),

$$\|h - A^{-1}k\| \le \frac{1}{2}\,\|h\|.$$

Consequently,

$$\frac{1}{2}\,\|h\| \le \|A^{-1}k\| \le \|A^{-1}\|\,\|k\|, \qquad \text{so} \qquad \|h\| \le 2\|A^{-1}\|\,\|k\|. \tag{6.11}$$

Since f is differentiable at x_0, we can write

$$f(x_0 + h) - f(x_0) - Ah = \|h\|\,\rho(h), \qquad \rho(h) \to 0 \text{ as } \|h\| \to 0.$$

But $k = f(x_0 + h) - f(x_0)$, hence

$$\begin{aligned}k - Ah &= \|h\|\,\rho(h), \qquad \text{and therefore} \qquad h - A^{-1}k\\ &= -A^{-1}(k - Ah)\\ &= -\|h\|\,A^{-1}\rho(h).\end{aligned}$$

Divide by $\|k\|$ and use (6.11):

$$\frac{\|h - A^{-1}k\|}{\|k\|} \le \frac{\|h\|\,\|A^{-1}\|\,\|\rho(h)\|}{\|k\|}$$
$$\le 2\|A^{-1}\|^2\,\|\rho(h)\| \longrightarrow 0.$$

Thus f^{-1} is differentiable at y_0 with

$$(f^{-1})'(y_0) = A^{-1} = (f'(x_0))^{-1}.$$

(iv) f^{-1} is C^1 on V. The above argument applies at any $y \in V$ with $x = f^{-1}(y)$ and $A = f'(x)$ (which is invertible on U by the choice of U). Hence

$$(f^{-1})'(y) = \big(f'(f^{-1}(y))\big)^{-1}, \qquad y \in V.$$

The map $y \mapsto f^{-1}(y)$ is continuous, the derivative f' is continuous on U, and the inversion map $A \mapsto A^{-1}$ is continuous on $GL_n(\mathbb{R})$. Therefore $y \mapsto (f^{-1})'(y)$ is continuous on V, i.e., $f^{-1} \in C^1(V)$. □

Example 6.44. Let $f : \mathbb{R}^2 \to \mathbb{R}^2$ be defined by

$$f(x, y) = (x - e^{-y},\, y - e^{x}).$$

Then

$$f'(0,0) = \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}, \qquad \det f'(0,0) = 2 \neq 0.$$

Hence the inverse function theorem applies at $(0,0)$: there exist neighborhoods U of $(0,0)$ and V of $(-1,-1)$ such that $f : U \to V$ is a C^1-diffeomorphism. Moreover,

$$(f^{-1})'(-1,-1) = (f'(0,0))^{-1} = \frac{1}{2}\begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}.$$

Example 6.45 (Local vs. global invertibility)**.** Define $F : \mathbb{R}^2 \to \mathbb{R}^2$ by

$$F(x, y) = (e^x \cos y,\, e^x \sin y).$$

Then F is C^1 and

$$\det F'(x, y) = e^{2x} \neq 0 \qquad \text{for all } (x, y) \in \mathbb{R}^2.$$

Hence by the inverse function theorem, F is locally one-to-one at every point. Nevertheless,

$$F(x, y + 2\pi) = F(x, y) \qquad \text{for all } (x, y) \in \mathbb{R}^2,$$

so F is not globally injective.

Remark 6.46. The determinant condition detects local invertibility of the derivative; it does not detect global injectivity of the map itself. This distinction is essential whenever one passes from local coordinates to global geometry.

Example 6.47 (An everywhere locally invertible map)**.** Define

$$F(x,y) = \big(x + e^{y},\ y + e^{-x}\big), \qquad (x,y) \in \mathbb{R}^2.$$

Then

$$F'(x,y) = \begin{pmatrix} 1 & e^{y} \\ -e^{-x} & 1 \end{pmatrix}, \qquad \det F'(x,y) = 1 + e^{\,y-x} > 0$$

for every $(x,y) \in \mathbb{R}^2$. Hence the inverse function theorem applies at every point: F is locally invertible everywhere on $\mathbb{R}^2$. Moreover, for each $(x_0,y_0) \in \mathbb{R}^2$,

$$(F^{-1})'(F(x_0,y_0)) = \frac{1}{1+e^{\,y_0-x_0}} \begin{pmatrix} 1 & -e^{y_0} \\ e^{-x_0} & 1 \end{pmatrix}.$$

Proposition 6.48 (Jacobian tests for local invertibility: worked examples)**.** *Consider the following maps.*

(a)

$$F_1(x,y) = (x^2+y^2, xy), \qquad F_2(x,y) = (x^2+xy+y^2, xy).$$

Each of these maps is locally invertible at (x,y) if and only if $x^2 \neq y^2$.

(b)

$$F_3(x,y) = (\cos x + \cos y,\ \sin x + \sin y).$$

The map F_3 is locally invertible at (x,y) if and only if $y - x \notin \pi\mathbb{Z}$.

(c)

$$G(x,y,z) = (x+y,\ xy+z,\ y+z).$$

The map G is locally invertible at (x,y,z) if and only if $x - y - 1 \neq 0$.

Proof. For F_1,

$$F_1'(x,y) = \begin{pmatrix} 2x & 2y \\ y & x \end{pmatrix}, \qquad \det F_1'(x,y) = 2(x^2-y^2).$$

Hence F_1 is locally invertible exactly where $x^2 - y^2 \neq 0$.

For F_2,

$$F_2'(x,y) = \begin{pmatrix} 2x+y & x+2y \\ y & x \end{pmatrix}, \qquad \det F_2'(x,y) = 2(x^2-y^2).$$

Thus the same criterion appears.

For F_3,

$$F_3'(x,y) = \begin{pmatrix} -\sin x & -\sin y \\ \cos x & \cos y \end{pmatrix}, \qquad \det F_3'(x,y) = \sin(y-x).$$

Therefore F_3 is locally invertible exactly when $\sin(y-x) \neq 0$, i.e. when $y-x$ is not an integer multiple of π.

Finally,

$$G'(x,y,z) = \begin{pmatrix} 1 & 1 & 0 \\ y & x & 1 \\ 0 & 1 & 1 \end{pmatrix}, \qquad \det G'(x,y,z) = x - y - 1.$$

So G is locally invertible exactly at the points where $x - y - 1 \neq 0$.

Each conclusion follows directly from the inverse function theorem. □

Remark 6.49. These examples are worth comparing side by side. The theorem itself is qualitative, but in concrete problems the workflow is almost always the same: compute the Jacobian, take its determinant, and identify the region on which that determinant does not vanish.

6.4.2 Implicit function theorem

A motivating example. Consider $f : \mathbb{R}^2 \to \mathbb{R}$ defined by

$$f(x,y) = x^2 + y^2 - 1.$$

The level set $\{(x,y) : f(x,y) = 0\}$ is the unit circle.

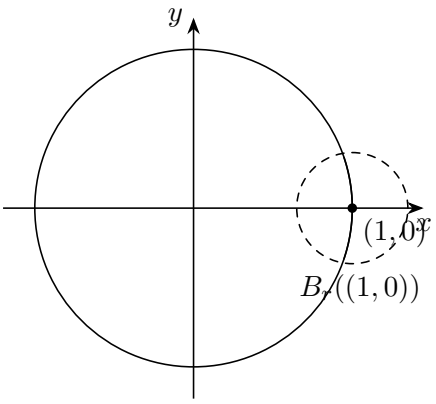


Figure 6.8: Near $(1,0)$, the circle $x^2 + y^2 = 1$ is the graph $x = \sqrt{1-y^2}$, but it cannot be written as a single-valued graph $y = \psi(x)$ near $x = 1$.

We have $f'(x,y) = (2x, 2y)$, so at $(1,0)$

$$\left.\frac{\partial f}{\partial x}\right|_{(1,0)} = 2 \neq 0, \qquad \left.\frac{\partial f}{\partial y}\right|_{(1,0)} = 0.$$

Because $\partial f/\partial x \neq 0$ at $(1,0)$, we can solve $f(x,y)=0$ locally for x as a C^1 function of y:

$$x = \varphi(y) = \sqrt{1-y^2}\,, \qquad |y| < r,$$

for some small $r > 0$. On the other hand, we *cannot* solve locally for y as a single-valued function of x near $x = 1$, since for x slightly less than 1 the equation $x^2 + y^2 = 1$ has two solutions $y = \pm\sqrt{1-x^2}$. This illustrates that which variable can be solved for depends on a non-degeneracy condition on the relevant partial derivative.

A linear-algebra model. Let

$$A : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^n$$

be a linear map. Writing $(h,k) = (h,0) + (0,k)$, we may decompose

$$A(h,k) = A(h,0) + A(0,k) = A_x h + A_y k,$$

where $A_x : \mathbb{R}^n \to \mathbb{R}^n$ and $A_y : \mathbb{R}^m \to \mathbb{R}^n$ are linear.

Lemma 6.50. *If A_x is invertible, then for each $k \in \mathbb{R}^m$ there exists a unique $h \in \mathbb{R}^n$ such that $A(h,k) = 0$, namely*

$$h = -A_x^{-1} A_y k.$$

Proof. The equation $A(h,k) = 0$ is equivalent to $A_x h + A_y k = 0$. If A_x is invertible, we can solve uniquely for h: $h = -A_x^{-1} A_y k$. □

Block form of the derivative. If $\Omega \subset \mathbb{R}^n \times \mathbb{R}^m$ is open and $f : \Omega \to \mathbb{R}^n$ is differentiable, then for $(x,y) \in \Omega$ the derivative $f'(x,y)$ is an $n \times (n+m)$ matrix. It is convenient to write it in block form

$$f'(x,y) = \begin{pmatrix} A_x & A_y \end{pmatrix},$$

where

$$A_x = \left(\frac{\partial f_i}{\partial x_j}(x,y) \right)_{1 \le i \le n,\ 1 \le j \le n}, \qquad A_y = \left(\frac{\partial f_i}{\partial y_k}(x,y) \right)_{1 \le i \le n,\ 1 \le k \le m}.$$

Theorem 6.51 (Implicit Function Theorem)**.** *Let Ω be an open subset of $\mathbb{R}^n \times \mathbb{R}^m$ and let $f : \Omega \to \mathbb{R}^n$ be a C^1 map. Assume that $f(x_0,y_0) = 0$ for some $(x_0,y_0) \in \Omega$ and that the $n \times n$ block*

$$A_x := \big[f'(x_0,y_0) \big]_x$$

is invertible (equivalently, $\det A_x \neq 0$*). Then:*

(i) There exist open sets $U \subset \mathbb{R}^n \times \mathbb{R}^m$ with $(x_0,y_0) \in U$ and $W \subset \mathbb{R}^m$ with $y_0 \in W$ such that for every $y \in W$ there exists a unique $x \in \mathbb{R}^n$ with $(x,y) \in U$ and $f(x,y) = 0$.

(ii) Writing this unique solution as $x = g(y)$, we obtain a C^1 map $g : W \to \mathbb{R}^n$ with $g(y_0) = x_0$ and $f(g(y), y) = 0$ for all $y \in W$. Moreover,

$$g'(y_0) = -A_x^{-1} A_y, \qquad \text{where} \qquad A_y := \left[f'(x_0, y_0)\right]_y .$$

In particular, near (x_0, y_0) the zero set $\{(x, y) \in \Omega : f(x, y) = 0\}$ is the graph of the function $x = g(y)$.

Geometric meaning. The equation $f(x, y) = 0$ describes a level set in $\mathbb{R}^{n+m}$. The invertibility of the x-block of the Jacobian says that, near (x_0, y_0), the equation can be solved uniquely for x as a C^1 function of y. In other words, the level set is locally the graph of a function $x = \varphi(y)$. Figure 6.9 illustrates the inverse-function-theorem construction used in the proof.

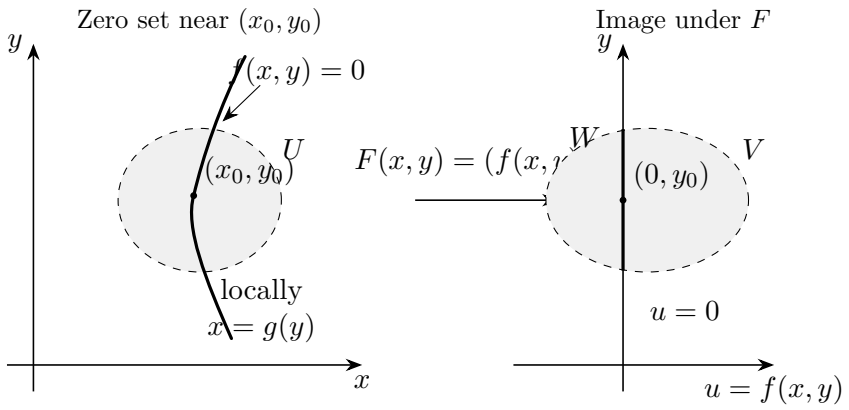


Figure 6.9: The zero set of $f(x, y) = 0$ is locally the graph $x = g(y)$ near (x_0, y_0). The proof applies the map $F(x, y) = (f(x, y), y)$, which straightens this local curve into the vertical slice $\{(0, y) : y \in W\} \subset V$.

Proof. (i) Define $F : \Omega \to \mathbb{R}^n \times \mathbb{R}^m$ by

$$F(x, y) := (f(x, y), y).$$

Then F is a C^1 map and its derivative at (x_0, y_0) has block form

$$F'(x_0, y_0) = \begin{bmatrix} A_x & A_y \\ 0 & I \end{bmatrix}.$$

Hence $\det F'(x_0, y_0) = \det(A_x) \neq 0$. By the inverse function theorem, there exist open sets $U \subset \mathbb{R}^n \times \mathbb{R}^m$ and $V \subset \mathbb{R}^n \times \mathbb{R}^m$ such that

$$F : U \to V$$

is a C^1 bijection and $F^{-1} : V \to U$ is C^1. Let

$$W := \{\, y \in \mathbb{R}^m : \ (0, y) \in V \,\}.$$

Since V is open, W is open in $\mathbb{R}^m$. For any $y \in W$, we have $(0, y) \in V$, and since F is onto V there exists $(x, y) \in U$ such that

$$(0, y) = F(x, y) = (f(x, y), y).$$

Thus $f(x, y) = 0$. If also $(x', y) \in U$ satisfies $f(x', y) = 0$, then $F(x', y) = (0, y) = F(x, y)$, and injectivity of F on U forces $x' = x$. This proves existence and uniqueness of x for each $y \in W$.

(ii) Define $g : W \to \mathbb{R}^n$ by declaring $g(y)$ to be the unique x with $(x, y) \in U$ and $f(x, y) = 0$. Then $(g(y), y) \in U$ and

$$F(g(y), y) = (0, y), \qquad y \in W.$$

Equivalently, $F^{-1}(0, y) = (g(y), y)$ for $y \in W$. Since F^{-1} is C^1, it follows that g is C^1. To compute $g'(y_0)$, differentiate the identity $f(g(y), y) = 0$ at $y = y_0$ and use the chain rule:

$$0 = \frac{d}{dy}\, f(g(y), y)\Big|_{y=y_0} = A_x\, g'(y_0) + A_y.$$

Since A_x is invertible, this gives $g'(y_0) = -A_x^{-1} A_y$.

□

Exercise 6.52. Prove that $x^2 + ye^x - \sin(xy) = 0$ can be solved for y in a neighborhood of $(0, 0)$, but cannot be solved for x in any neighborhood of $(0, 0)$.

$$f(x, y) = x^2 + ye^x - \sin(xy)$$

(i) $f(0, 0) = 0$, $\frac{\partial f}{\partial y}|_{(0,0)} = 1 \neq 0$. By the implicit function theorem, there exists a neighborhood of $(0, 0)$ and a function g such that

$$f(x, g(x)) = 0$$

for $|x| < r$, equivalently $y = g(x)$ near 0.

(ii) $\frac{\partial f}{\partial x}|_{(0,0)} = 0$. Hence the implicit function theorem *cannot* be applied.
On the contrary, suppose that $x = \phi(y)$ near $y = 0$. Then $\phi(0) = 0$, and

$$(\phi(y))^2 + ye^{\phi(y)} - \sin(\phi(y)y) = 0$$

for $|y| < r$ for some $r > 0$. Then

$$2\phi(0)\phi'(0) + 1 \cdot e^{\phi(0)} + 0 \cdot e^{\phi(0)}\phi'(0) - \cos(\phi(0)0)\,(\phi'(0)0 + \phi(0) \cdot 1) = 0$$

$$\implies 1 = 0 \quad \text{(contradiction)}$$

Example 6.53. Let $f : \mathbb{R} \times \mathbb{R}^2 \to \mathbb{R}^2$

$$f(x, y, z) = (xe^y + ye^z,\ xe^z + ze^y)$$

Then f is a C^1-map.

$$f'(x, y, z) = \begin{pmatrix} e^y & xe^y + e^z & ye^z \\ e^z & ze^y & xe^z + e^y \end{pmatrix}$$

$$f(-1, 1, 1) = (0, 0)$$

Let $f = (f_1, f_2)$. Then

$$\begin{pmatrix} \dfrac{\partial f_1}{\partial y} & \dfrac{\partial f_1}{\partial z} \\ \dfrac{\partial f_2}{\partial y} & \dfrac{\partial f_2}{\partial z} \end{pmatrix} (-1, 1, 1) = \begin{pmatrix} 0 & e \\ e & 0 \end{pmatrix}$$

By the implicit function theorem, there exists an open ball U in $\mathbb{R}^3$ and open ball V in $\mathbb{R}^2$, such that

$$(y, z) = (\phi(x), \psi(x)), \quad |x| < r \quad \text{for some } r > 0.$$

Exercise 6.54. Let $f : \mathbb{R}^2 \to \mathbb{R}$ be a C^1-map such that $f(0, 0) = 0, f_x(0, 0) = 1$. Let $F(x, y) = (f(x, y), y)$. Prove that F is injective in some neighborhood of $(0, 0)$. Does F remain injective in any neighborhood of $(0, 0)$?

Remark: Condition in implicit function theorem or inverse mapping theorem on derivatives are sufficient.

Example 6.55. $f : \mathbb{R}^2 \to \mathbb{R}$, $f(x, y) = x^2 - y^3$.

$$f(0, 0) = 0,$$
$$\frac{\partial f}{\partial y}(0, 0) = 0,$$

but $y = x^{2/3}$ is a solution of $f(x, y) = 0$ near $(0, 0)$.

Example 6.56. Let $f : \mathbb{R}^2 \to \mathbb{R}^2$, $f(x, y) = (x^3, y^3)$. Then $\det f'(0, 0) = 0$ but f is one-to-one, onto.

Part IV

Measure and Integration

Chapter 7

Lebesgue Measure and Integral

This chapter develops Lebesgue measure and integration on the real line from first principles. We begin by identifying the limitations of the Riemann integral and then construct Lebesgue outer measure by countable interval coverings. The measurable sets are introduced through regular approximation and are then characterized by Carathéodory's criterion; on this σ-algebra, outer measure becomes countably additive. The second half of the chapter develops measurable functions, approximation by simple functions, the Lebesgue integral, and the fundamental convergence theorems. The chapter concludes by explaining the relation with Riemann integration and by recording the basic structure of the spaces L^p.

Learning objectives.

- construct Lebesgue outer measure and prove its basic properties;
- characterize Lebesgue measurable sets by regular approximation and by Carathéodory's criterion;
- use simple functions to approximate measurable functions and to define the Lebesgue integral;
- apply the Monotone Convergence Theorem, Fatou's lemma, and the Dominated Convergence Theorem;
- relate Lebesgue integration to Riemann integration and understand the basic norm structure of L^p spaces.

7.1 Chapter roadmap

Dependencies. This chapter is arranged in a natural conceptual order: we begin with the limitations of Riemann integration, then construct Lebesgue outer measure on $\mathbb{R}$, identify the measurable sets and the resulting measure-theoretic structure, and finally move to measurable functions, the Lebesgue integral, and the main convergence theorems before returning to the relation with the Riemann theory.

This chapter corresponds to the **Lebesgue measure and integral** unit of MA224. Its guiding theme is that limiting processes become tractable once size is assigned to sets and functions are integrated according to the measure of their level sets. The material is organized as follows:

(1) **Why Riemann integration is not enough:** bounded nowhere-continuous examples, unbounded integrands, and unbounded domains.

(2) **Lebesgue outer measure on $\mathbb{R}$:** coverings by intervals, basic properties of m^*, intervals, countable sets, open sets, regular approximation, and translation invariance.

(3) **Lebesgue measurable sets and Lebesgue measure:** regular approximation by open and closed sets, the measurable σ-algebra, Borel sets, continuity properties of measure, local criteria, and Carathéodory's characterization.

(4) **Key examples and pathologies:** the Cantor set and the existence of nonmeasurable sets.

(5) **Measurable functions:** closure properties, continuity versus measurability, and approximation by simple functions.

(6) **Lebesgue integration:** nonnegative functions first, then signed and optional complex-valued integrals.

(7) **Convergence principles and L^1:** the monotone convergence theorem, Fatou's lemma, Chebyshev's inequality, the dominated convergence theorem, and the bounded convergence theorem.

(8) **Back to Riemann integration:** criteria for Riemann integrability and equality of the Riemann and Lebesgue integrals on compact intervals, and L^p spaces.

(9) **Further optional topics:** Egoroff's theorem, Lusin's theorem, product measures, Fubini–Tonelli, and the Radon–Nikodym theorem.

Standing notation. Throughout this chapter, unless explicitly stated otherwise, all sets are subsets of $\mathbb{R}$. We write m^* for Lebesgue outer measure, $\mathcal{M}$ for the σ-algebra of Lebesgue measurable sets, and

$$m := m^*|_{\mathcal{M}}$$

for Lebesgue measure. Equalities involving measures or integrals are understood in the extended sense unless finiteness is explicitly asserted.

7.2 From Riemann to Lebesgue

Lebesgue integration does not arise by replacing the Riemann integral arbitrarily; it arises because certain natural analytic questions fall outside the scope of the older theory. The examples below are meant to isolate the precise points at which the Riemann framework becomes too restrictive and to motivate the passage to measure.

7.2.1 Limitations of the Riemann integral

The Riemann integral is defined (in its basic form) only for bounded functions on a compact interval. Even in that setting it fails for certain highly oscillatory functions. A precise characterization is given by *Lebesgue's criterion*: if $f\colon [a,b] \to \mathbb{R}$ is bounded, then f is Riemann integrable if and only if its set of discontinuities has Lebesgue measure zero. The following examples motivate the measure-theoretic approach.

(I) **A bounded function that is nowhere continuous.** Define $f\colon [0,1] \to \mathbb{R}$ by

$$f(x) = \begin{cases} 1, & x \in (\mathbb{R} \setminus \mathbb{Q}) \cap [0,1], \\ 0, & x \in \mathbb{Q} \cap [0,1]. \end{cases}$$

Every subinterval of $[0,1]$ contains both rational and irrational points, so for every partition P we have $L(P,f) = 0$ and $U(P,f) = 1$. Hence f is not Riemann integrable on $[0,1]$.

(II) **An unbounded integrand.** The function $f(t) = t^{-1/2}$ is not bounded near $t = 0$, so it does not fit the basic definition of the Riemann integral on $[0,1]$. One therefore introduces the *improper* integral

$$\int_0^1 \frac{1}{\sqrt{t}}\,dt \;:=\; \lim_{\varepsilon \downarrow 0} \int_\varepsilon^1 \frac{1}{\sqrt{t}}\,dt.$$

In this case the limit exists and equals 2.

(III) **Integration over an unbounded domain.** Similarly, for $g(t) = \frac{1}{1+t^2}$ one defines

$$\int_0^\infty \frac{1}{1+t^2}\,dt \;:=\; \lim_{R \to \infty} \int_0^R \frac{1}{1+t^2}\,dt = \lim_{R \to \infty} \tan^{-1}(R) = \frac{\pi}{2}.$$

Lebesgue measure and the Lebesgue integral provide a unified framework that treats (I)–(III) systematically and extends naturally to higher dimensions and L^p spaces.

7.3 Lebesgue outer measure on $\mathbb{R}$

The first step toward Lebesgue integration is to measure sets rather than functions. The guiding idea is to approximate an arbitrary set from the outside by open intervals and to define its size as the smallest total length of such coverings.

For an open interval $I = (a,b) \subset \mathbb{R}$, define its length by $\ell(I) := b - a$. For an arbitrary set $A \subset \mathbb{R}$, define the *outer measure* of A by

$$m^*(A) = \inf \left\{ \sum_{n=1}^\infty \ell(I_n) \;:\; A \subseteq \bigcup_{n=1}^\infty I_n \text{ and each } I_n \text{ is an open interval} \right\}.$$

Intuitively, we cover A by countably many open intervals and take the smallest possible total length.

Proposition 7.1 (Basic properties of m^*)**.** *For all $A, B \subset \mathbb{R}$:*

(i) $m^*(\emptyset) = 0$*;*

(ii) if $A \subseteq B$*, then* $m^*(A) \le m^*(B)$*;*

*(iii) (*countable subadditivity*) for any sequence* $(A_n)_{n \ge 1}$*,*

$$m^*\left(\bigcup_{n=1}^\infty A_n \right) \le \sum_{n=1}^\infty m^*(A_n).$$

Proof. (i) Since $\emptyset$ is covered by the empty collection, we have $m^*(\emptyset) \leq 0$, and by definition $m^* \geq 0$, hence $m^*(\emptyset) = 0$. (ii) If $A \subseteq B$ and (I_n) is any open-interval cover of B, then it is also an open-interval cover of A. Taking infima over all covers of A and B gives $m^*(A) \leq m^*(B)$. (iii) Let $(A_n)_{n\geq 1} \subset \mathbb{R}$ and fix $\varepsilon > 0$. For each n, choose an open-interval cover $A_n \subseteq \bigcup_{j\geq 1} I_{n,j}$ such that

$$\sum_{j\geq 1} \ell(I_{n,j}) \leq m^*(A_n) + \frac{\varepsilon}{2^n}.$$

Then $\bigcup_{n\geq 1} A_n \subseteq \bigcup_{n,j} I_{n,j}$, so

$$\begin{aligned} m^*\left(\bigcup_{n\geq 1} A_n\right) &\leq \sum_{n\geq 1}\sum_{j\geq 1} \ell(I_{n,j}) \\ &\leq \sum_{n\geq 1}\left(m^*(A_n) + \frac{\varepsilon}{2^n}\right) \\ &= \sum_{n\geq 1} m^*(A_n) + \varepsilon. \end{aligned}$$

Since $\varepsilon > 0$ was arbitrary, the desired inequality follows. □

Example 7.2 (Countable sets have outer measure 0)**.** If $A \subset \mathbb{R}$ is countable, then $m^*(A) = 0$. Indeed, list $A = \{a_1, a_2, \ldots\}$ and fix $\varepsilon > 0$. Cover each point a_n by the open interval

$$I_n := \left(a_n - \frac{\varepsilon}{2^{n+1}},\ a_n + \frac{\varepsilon}{2^{n+1}}\right),$$

whose length is $\ell(I_n) = \varepsilon/2^n$. Then $A \subseteq \bigcup_{n\geq 1} I_n$ and

$$m^*(A) \leq \sum_{n\geq 1} \ell(I_n) = \sum_{n\geq 1} \frac{\varepsilon}{2^n} = \varepsilon.$$

Letting $\varepsilon \downarrow 0$ yields $m^*(A) = 0$. In particular, $\mathbb{Q}$ is countable, hence $m^*(\mathbb{Q}) = 0$.

Proposition 7.3 (Outer measure of intervals)**.** *If I is any bounded interval with endpoints $a < b$, then*

$$m^*(I) = b - a.$$

Proof. We first treat the closed interval $[a, b]$.

Upper bound. For every $n \geq 1$,

$$[a, b] \subset \left(a - \frac{1}{n}, b + \frac{1}{n}\right),$$

and hence

$$m^*([a, b]) \leq b - a + \frac{2}{n}.$$

Letting $n \to \infty$ gives $m^*([a, b]) \leq b - a$.

Lower bound. Let $\{J_n\}_{n\geq 1}$ be any open-interval cover of $[a,b]$. By compactness there is a finite subcover, which after relabeling we write as $J_1,\ldots,J_k$. We claim that

$$b-a \leq \sum_{j=1}^{k} \ell(J_j).$$

To justify this, merge overlapping intervals in the finite family $J_1,\ldots,J_k$. The union then becomes a finite disjoint union of open intervals, one of whose components contains the connected set $[a,b]$; that component must therefore contain all of $[a,b]$ and has length at least $b-a$. Since merging overlapping intervals never increases the sum of lengths, the claim follows. Hence

$$b-a \leq \sum_{j=1}^{k} \ell(J_j) \leq \sum_{n=1}^{\infty} \ell(J_n).$$

Taking the infimum over all open-interval covers gives $b-a \leq m^*([a,b])$.

Thus $m^*([a,b]) = b-a$. For the other bounded intervals with endpoints a and b, use

$$(a,b) \subset [a,b] \subset (a-\varepsilon, b+\varepsilon)$$

and the fact that singletons have outer measure zero. Hence

$$m^*((a,b)) = m^*([a,b)) = m^*((a,b]) = m^*([a,b]) = b-a.$$

□

Example 7.4. Let $A \subset \mathbb{R}$ and $x \in \mathbb{R}$. For the translate

$$A + x := \{a + x : a \in A\},$$

we have

$$m^*(A+x) = m^*(A).$$

Let $A \subset \bigcup_n I_n$. Then $A + x \subset \bigcup_n (I_n + x)$ that is, $\{I_n + x\}_{n=1}^{\infty}$ is a covering of $A + x$. Hence

$$m^*(A+x) \leq \sum \ell(I_n + x) = \sum \ell(I_n)$$

for all cover $\{I_n\}$ of A. Therefore, $m^*(A+x) \leq m^*(A)$. Replacing x by $-x$ gives

$$m^*(A-x) \leq m^*(A).$$

Now replace A by $A+x$ in this inequality to obtain $m^*(A) \leq m^*(A+x)$. Hence

$$m^*(A+x) = m^*(A),$$

so outer measure is translation invariant.

Proposition 7.5. *Let $A \subset \mathbb{R}$ and $\varepsilon > 0$. Then there exists an open set $O \supset A$ such that*

$$m^*(O) < m^*(A) + \varepsilon.$$

Equivalently,

$$m^*(A) = \inf\{m^*(O) : O \supset A,\ O \text{ open}\}.$$

Proof. If $m^*(A) = \infty$, the conclusion is immediate. Assume therefore that $m^*(A) < \infty$. By the definition of outer measure, there exists a countable family of open intervals $\{I_n\}_{n\geq 1}$ such that

$$A \subset \bigcup_{n=1}^{\infty} I_n \qquad \text{and} \qquad \sum_{n=1}^{\infty} \ell(I_n) < m^*(A) + \varepsilon.$$

Set

$$O := \bigcup_{n=1}^{\infty} I_n.$$

Then O is open and $A \subset O$. By countable subadditivity,

$$m^*(O) \leq \sum_{n=1}^{\infty} \ell(I_n) < m^*(A) + \varepsilon.$$

Since every open set containing A has outer measure at least $m^*(A)$ by monotonicity, the displayed approximation immediately yields

$$m^*(A) = \inf\{m^*(O) : O \supset A,\ O \text{ open}\}.$$

□

Theorem 7.6 (Outer regular G_δ envelope)**.** *For every set $A \subset \mathbb{R}$ there exists a G_δ set $G \supset A$ such that*

$$m^*(G) = m^*(A).$$

Proof. If $m^*(A) = \infty$, take $G = \mathbb{R}$. Assume now that $m^*(A) < \infty$. For each $n \geq 1$, choose an open set $O_n \supset A$ such that

$$m^*(O_n) < m^*(A) + \frac{1}{n}.$$

Set $G := \bigcap_{n=1}^{\infty} O_n$. Then G is a G_δ set and $A \subset G \subset O_n$ for every n. Hence, by monotonicity,

$$m^*(A) \leq m^*(G) \leq m^*(O_n) < m^*(A) + \frac{1}{n} \qquad (n \geq 1).$$

Letting $n \to \infty$ gives $m^*(G) = m^*(A)$. □

Proposition 7.7 (Countable unions of null sets)**.** *Let $E = \bigcup_{n=1}^{\infty} E_n$, where each $E_n \subset \mathbb{R}$. Then*

$$m^*(E) = 0 \quad \Longleftrightarrow \quad m^*(E_n) = 0 \text{ for every } n \in \mathbb{N}.$$

Proof. If $m^*(E_n) = 0$ for every n, then by countable subadditivity,

$$m^*(E) = m^*\left(\bigcup_{n=1}^{\infty} E_n\right) \leq \sum_{n=1}^{\infty} m^*(E_n) = 0,$$

so $m^*(E) = 0$.

Conversely, suppose that $m^*(E) = 0$ and that $m^*(E_{n_0}) > 0$ for some $n_0 \in \mathbb{N}$. Set

$$\varepsilon := \frac{1}{2}\, m^*(E_{n_0}) > 0.$$

Since $m^*(E) = 0$, by the definition of outer measure there exists a countable collection of open intervals $\{I_k\}_{k\geq 1}$ covering E such that

$$\sum_{k=1}^{\infty} \ell(I_k) < m^*(E) + \varepsilon = \frac{1}{2}\, m^*(E_{n_0}).$$

But $E_{n_0} \subset E \subset \bigcup_{k\geq 1} I_k$, so the same family also covers E_{n_0}. Hence, by the definition of outer measure,

$$m^*(E_{n_0}) \leq \sum_{k=1}^{\infty} \ell(I_k) < \frac{1}{2}\, m^*(E_{n_0}),$$

which is impossible. Therefore no such n_0 can exist, and thus $m^*(E_n) = 0$ for every $n \in \mathbb{N}$. □

Proposition 7.8 (Metric additivity of outer measure)**.** *If $A, B \subset \mathbb{R}$ satisfy $\operatorname{dist}(A, B) > 0$, then*

$$m^*(A \cup B) = m^*(A) + m^*(B).$$

Consequently, m^* is finitely additive on every finite family of pairwise separated sets.

Proof. Let $\delta := \operatorname{dist}(A, B) > 0$. The inequality

$$m^*(A \cup B) \leq m^*(A) + m^*(B)$$

is subadditivity. For the reverse inequality, let (I_n) be any countable open-interval cover of $A \cup B$. Put

$$I_n' := I_n \cap \{x : \operatorname{dist}(x, A) < \delta/3\}, \qquad I_n'' := I_n \cap \{x : \operatorname{dist}(x, B) < \delta/3\}.$$

Then (I'_n) covers A and (I''_n) covers B. Moreover, I'_n and I''_n are disjoint subsets of I_n, and their total interval length is at most $\ell(I_n)$ after decomposing them into intervals. Hence

$$m^*(A) + m^*(B) \leq \sum_{n\geq 1} \ell(I_n).$$

Taking the infimum over all covers of $A \cup B$ gives

$$m^*(A) + m^*(B) \leq m^*(A \cup B).$$

The equality follows. The finite statement follows by induction. □

Proposition 7.9 (Outer measure of an open set)**.** *Let $O \subset \mathbb{R}$ be open, and write its connected-component decomposition as*

$$O = \bigcup_{n=1}^{\infty} I_n,$$

where the I_n are pairwise disjoint open intervals. Then

$$m^*(O) = \sum_{n=1}^{\infty} \ell(I_n).$$

Proof. The inequality

$$m^*(O) \leq \sum_{n=1}^{\infty} \ell(I_n)$$

follows directly from the definition of m^*, since the intervals I_n cover O.

For the reverse inequality, first suppose that some component I_n is unbounded. Then O contains compact intervals of arbitrarily large length, so $m^*(O) = \infty$, and the formula is immediate. We may therefore assume that all I_n have finite length.

Fix $N \in \mathbb{N}$ and $\varepsilon > 0$. For each $1 \leq n \leq N$, choose a closed interval $J_n \subset I_n$ such that

$$\ell(J_n) > \ell(I_n) - \frac{\varepsilon}{N}.$$

The intervals $J_1, \ldots, J_N$ are pairwise separated compact intervals. Hence, by Theorem 7.8,

$$m^*\left(\bigcup_{n=1}^{N} J_n\right) = \sum_{n=1}^{N} \ell(J_n) > \sum_{n=1}^{N} \ell(I_n) - \varepsilon.$$

Since $\bigcup_{n=1}^{N} J_n \subset O$, monotonicity yields

$$\sum_{n=1}^{N} \ell(I_n) \leq m^*(O) + \varepsilon.$$

Letting $\varepsilon \downarrow 0$ and then $N \to \infty$ gives

$$\sum_{n=1}^{\infty} \ell(I_n) \le m^*(O).$$

The two inequalities prove the formula. □

Corollary 7.10 (Countable additivity for disjoint open sets)**.** *If $\{O_i\}_{i=1}^{\infty}$ is a family of pairwise disjoint open sets in $\mathbb{R}$, then*

$$m^*\left(\bigcup_{i=1}^{\infty}{}^{\cdot}\, O_i\right) = \sum_{i=1}^{\infty} m^*(O_i).$$

Proof. Write each open set O_i as a countable disjoint union of open intervals,

$$O_i = \bigcup_{n\ge 1}{}^{\cdot}\, I_{i,n}.$$

Because the sets O_i are pairwise disjoint, the entire family $\{I_{i,n} : i, n \ge 1\}$ is pairwise disjoint. By Theorem 7.9,

$$m^*(O_i) = \sum_{n\ge 1} \ell(I_{i,n}) \qquad \text{for each } i.$$

Applying Theorem 7.9 again to the open set $\bigcup_{i\ge 1} O_i$, we obtain

$$m^*\left(\bigcup_{i\ge 1} O_i\right) = \sum_{i\ge 1}\sum_{n\ge 1} \ell(I_{i,n}) = \sum_{i\ge 1} m^*(O_i).$$

□

Proposition 7.11 (Finite additivity for separated closed intervals)**.** *If $[a, b]$ and $[c, d]$ are disjoint closed intervals, then*

$$m^*([a, b] \cup [c, d]) = m^*([a, b]) + m^*([c, d]).$$

Proof. After relabeling the intervals if necessary, assume that $b < c$. Set

$$\delta := c - b > 0.$$

Let $(I_n)_{n\ge 1}$ be any countable family of open intervals covering $[a, b] \cup [c, d]$. Define

$$I_n' := I_n \cap (-\infty, b + \delta/2), \qquad I_n'' := I_n \cap (c - \delta/2, \infty).$$

Then (I_n') covers $[a, b]$, (I_n'') covers $[c, d]$, and I_n' and I_n'' are disjoint subintervals of I_n for each n. Hence

$$\ell(I_n') + \ell(I_n'') \le \ell(I_n).$$

Using the formula for the outer measure of intervals, we get

$$\begin{aligned} m^*([a,b]) + m^*([c,d]) &\leq \sum_{n\geq 1} \ell(I'_n) + \sum_{n\geq 1} \ell(I''_n) \\ &\leq \sum_{n\geq 1} \ell(I_n). \end{aligned}$$

Taking the infimum over all such covers gives

$$m^*([a,b]) + m^*([c,d]) \leq m^*([a,b] \cup [c,d]).$$

The reverse inequality follows from subadditivity of m^*. □

Question 7.12. What are *all* those sets for which m^* is *countably additive*, that is,

$$m^*\left(\bigcup_{n=1}^{\infty} E_n\right) = \sum_{n=1}^{\infty} m^*(E_n)?$$

Example 7.13 (Inner regularity for bounded open sets)**.** Let $G \subset \mathbb{R}$ be open and bounded. Then for every $\varepsilon > 0$ there exists a compact set $K \subset G$ such that

$$m^*(G \setminus K) < \varepsilon \qquad \text{and hence} \qquad m^*(K) > m^*(G) - \varepsilon.$$

Proof. Write

$$G = \bigcupdot_{n=1}^{\infty} I_n, \qquad I_n = (a_n, b_n),$$

as a disjoint union of open intervals. Since G is bounded, there exist $\alpha < \beta$ such that $G \subset [\alpha, \beta]$; hence $m^*(G) \leq \beta - \alpha < \infty$. By the preceding result on disjoint open sets,

$$m^*(G) = \sum_{n=1}^{\infty} \ell(I_n) < \infty.$$

Now fix $\varepsilon > 0$. Choose $N \in \mathbb{N}$ so that

$$\sum_{n=N+1}^{\infty} \ell(I_n) < \frac{\varepsilon}{2}.$$

For each $n = 1, \ldots, N$, define

$$\delta_n := \min\left\{\frac{\varepsilon}{4N}, \frac{b_n - a_n}{4}\right\} \quad \text{and} \quad K_n := [a_n + \delta_n,\, b_n - \delta_n].$$

Then each K_n is a compact interval and $K_n \subset I_n \subset G$. Put

$$K := \bigcup_{n=1}^{N} K_n.$$

Since K is a finite union of compact sets, it is compact, and, by construction, $K \subset G$.

Also,

$$G \setminus K \subset \Big(\bigcup_{n=N+1}^{\infty} I_n \Big) \cup \Big(\bigcup_{n=1}^{N} (I_n \setminus K_n) \Big).$$

Therefore, by subadditivity of m^*,

$$\begin{aligned} m^*(G \setminus K) &\leq \sum_{n=N+1}^{\infty} \ell(I_n) + \sum_{n=1}^{N} m^*(I_n \setminus K_n) \\ &= \sum_{n=N+1}^{\infty} \ell(I_n) + \sum_{n=1}^{N} 2\delta_n \\ &\leq \frac{\varepsilon}{2} + N \cdot \frac{\varepsilon}{2N} = \varepsilon. \end{aligned}$$

Hence

$$m^*(K) \geq m^*(G) - m^*(G \setminus K) > m^*(G) - \varepsilon.$$

This is exactly the desired inner approximation of G by a compact set. □

Remark 7.14 (Two-sided regular approximation for bounded open sets). If G is open and bounded, then for every $\varepsilon > 0$ there are a compact set K and an open set O such that

$$K \subset G \subset O \qquad \text{and} \qquad m^*(O \setminus K) < \varepsilon.$$

This is the prototype for the regular approximation used in the definition of Lebesgue measurability below.

7.4 Lebesgue measurable sets and Lebesgue measure

Outer measure is defined for every subset of $\mathbb{R}$, but it is too coarse by itself to support a satisfactory integration theory. The next step is to isolate those sets for which size behaves well with respect to complements, decompositions, and approximations; these are the measurable sets.

In these notes we *define* Lebesgue measurability through regular approximation by open and closed sets. An equivalent and extremely useful formulation is Carathéodory's criterion; see Theorem 7.31. Once measurability is established, the *Lebesgue measure* is the restriction of outer measure to the σ-algebra $\mathcal{M}$ of measurable sets.

7.4.1 Lebesgue measurable sets

Definition 7.15 (Lebesgue measurable set)**.** A set $E \subset \mathbb{R}$ is said to be *Lebesgue measurable* if it admits arbitrarily accurate two-sided regular approximation: for every $\varepsilon > 0$ there exist a closed set F and an open set O such that

$$F \subset E \subset O \qquad \text{and} \qquad m^*(O \setminus F) < \varepsilon.$$

The collection of all Lebesgue measurable subsets of $\mathbb{R}$ will be denoted by $\mathcal{M}$. This regularity-based definition is equivalent, as proved below, to Carathéodory's splitting criterion.

Remark 7.16 (Outer and inner approximation)*.* If $F \subset E \subset O$, then $O \setminus E \subset O \setminus F$ and $E \setminus F \subset O \setminus F$. Consequently,

$$m^*(O \setminus E) \leq m^*(O \setminus F) \qquad \text{and} \qquad m^*(E \setminus F) \leq m^*(O \setminus F).$$

Thus a measurable set may be approximated arbitrarily well, in outer measure, from above by open sets and from below by closed sets.

Proposition 7.17 (Elementary closure properties of measurable sets)**.** *Let $\mathcal{M}$ denote the class of Lebesgue measurable subsets of $\mathbb{R}$. Then:*

(i) *if $E \in \mathcal{M}$, then $E^c \in \mathcal{M}$;*

(ii) *if $E_1, \ldots, E_N \in \mathcal{M}$, then $\bigcup_{n=1}^{N} E_n \in \mathcal{M}$;*

(iii) *if $A \subset \mathbb{R}$ and $m^*(A) = 0$, then $A \in \mathcal{M}$; consequently, every subset of a measurable null set is measurable;*

(iv) *if $E_1, E_2 \in \mathcal{M}$, then $E_1 \cap E_2 \in \mathcal{M}$ and $E_1 \setminus E_2 \in \mathcal{M}$.*

In particular, $\mathcal{M}$ is an algebra of sets.

Proof. (i) Let $\varepsilon > 0$ be given. Since $E \in \mathcal{M}$, there exist a closed set F and an open set O such that

$$F \subset E \subset O \qquad \text{and} \qquad m^*(O \setminus F) < \varepsilon.$$

Passing to complements yields

$$O^c \subset E^c \subset F^c,$$

where O^c is closed and F^c is open. Moreover,

$$F^c \setminus O^c = O \setminus F,$$

and therefore

$$m^*(F^c \setminus O^c) = m^*(O \setminus F) < \varepsilon.$$

Hence $E^c \in \mathcal{M}$.

(ii) It suffices to treat the case of two sets, the general finite case following by induction. Let $E_1, E_2 \in \mathcal{M}$, and fix $\varepsilon > 0$. Choose closed sets F_i and open sets O_i such that

$$F_i \subset E_i \subset O_i \qquad \text{and} \qquad m^*(O_i \setminus F_i) < \frac{\varepsilon}{2} \qquad (i = 1, 2).$$

Set

$$F := F_1 \cup F_2, \qquad O := O_1 \cup O_2.$$

Then F is closed, O is open, and

$$F \subset E_1 \cup E_2 \subset O.$$

Furthermore,

$$O \setminus F \subset (O_1 \setminus F_1) \cup (O_2 \setminus F_2),$$

so by subadditivity,

$$m^*(O \setminus F) \leq m^*(O_1 \setminus F_1) + m^*(O_2 \setminus F_2) < \varepsilon.$$

Therefore $E_1 \cup E_2 \in \mathcal{M}$.

(iii) Let $\varepsilon > 0$. If $m^*(A) = 0$, choose an open set $O \supset A$ such that

$$m^*(O) < \varepsilon.$$

Taking $F := \emptyset$, which is closed, we have $F \subset A \subset O$ and

$$m^*(O \setminus F) = m^*(O) < \varepsilon.$$

Thus $A \in \mathcal{M}$. Now suppose that $N \in \mathcal{M}$ with $m(N) = 0$, and let $A \subset N$. By monotonicity,

$$m^*(A) \leq m^*(N) = 0.$$

Hence $A \in \mathcal{M}$ by the first part.

(iv) Since

$$E_1 \cap E_2 = (E_1^c \cup E_2^c)^c,$$

part (i) together with part (ii) shows that $E_1 \cap E_2 \in \mathcal{M}$. Moreover,

$$E_1 \setminus E_2 = E_1 \cap E_2^c,$$

so differences of measurable sets are measurable as well. □

Proposition 7.18 (Intervals are measurable)**.** *Let $a, b \in \mathbb{R}$ with $a < b$. Then every interval having endpoints a and b is Lebesgue measurable. In particular,*

$$(a, b), \quad [a, b), \quad (a, b], \quad [a, b] \in \mathcal{M}.$$

Proof. We first show that (a, b) is measurable. Let $\varepsilon > 0$ be given, and choose

$$0 < \delta < \min\{(b-a)/2, \varepsilon/2\}.$$

Then

$$[a+\delta, b-\delta] \subset (a,b) \subset (a,b),$$

where $[a+\delta, b-\delta]$ is closed and (a, b) is open. Moreover,

$$(a,b) \setminus [a+\delta, b-\delta] = (a, a+\delta) \cup (b-\delta, b),$$

so

$$m^*\big((a,b) \setminus [a+\delta, b-\delta]\big) \leq 2\delta < \varepsilon.$$

Hence $(a, b) \in \mathcal{M}$.

Since singletons have outer measure zero, they are measurable by Theorem 7.17(iii). Therefore

$$[a,b) = (a,b) \cup \{a\}, \qquad (a,b] = (a,b) \cup \{b\}, \qquad [a,b] = (a,b) \cup \{a,b\},$$

and each of these intervals is measurable by finite unions. □

Proposition 7.19 (Closure under countable unions)**.** *Let* $(E_n)_{n\geq 1} \subset \mathcal{M}$*. Then*

$$\bigcup_{n=1}^{\infty} E_n \in \mathcal{M}.$$

Proof. Set

$$E := \bigcup_{n=1}^{\infty} E_n.$$

By Theorem 7.17, finite intersections and finite differences of measurable sets are measurable. Replacing E_n by

$$A_1 := E_1, \qquad A_n := E_n \setminus \bigcup_{j=1}^{n-1} E_j \qquad (n \geq 2),$$

we may therefore assume, without loss of generality, that the family $(E_n)_{n\geq 1}$ is pairwise disjoint.

Assume first that E is bounded. Fix $\varepsilon > 0$. For each $n \geq 1$, choose a closed set F_n and an open set O_n such that

$$F_n \subset E_n \subset O_n \qquad \text{and} \qquad m^*(O_n \setminus F_n) < \frac{\varepsilon}{2^{n+1}}.$$

Because E is bounded, every F_n is compact. For each $k \in \mathbb{N}$, the compact sets $F_1, \dots, F_k$ are pairwise disjoint, hence separated by a positive distance. By Theorem 7.8, m^* is finitely additive on $\{F_1, \dots, F_k\}$, and therefore

$$\sum_{n=1}^{k} m^*(F_n) = m^*\Big(\bigcup_{n=1}^{k} F_n\Big) \le m^*(E).$$

Since

$$O_n \subset F_n \cup (O_n \setminus F_n),$$

we obtain

$$\sum_{n=1}^{k} m^*(O_n) \le \sum_{n=1}^{k} m^*(F_n) + \sum_{n=1}^{k} m^*(O_n \setminus F_n) \le m^*(E) + \frac{\varepsilon}{2} \qquad (k \in \mathbb{N}).$$

Hence the series $\sum_{n=1}^{\infty} m^*(O_n)$ converges. Choose $N \in \mathbb{N}$ so large that

$$\sum_{n=N+1}^{\infty} m^*(O_n) < \frac{\varepsilon}{2}.$$

Set

$$O := \bigcup_{n=1}^{\infty} O_n, \qquad F := \bigcup_{n=1}^{N} F_n.$$

Then O is open, F is closed, and

$$F \subset E \subset O.$$

Moreover,

$$O \setminus F \subset \Big(\bigcup_{n=1}^{N} (O_n \setminus F_n)\Big) \cup \Big(\bigcup_{n=N+1}^{\infty} O_n\Big),$$

so

$$m^*(O \setminus F) \le \sum_{n=1}^{N} m^*(O_n \setminus F_n) + \sum_{n=N+1}^{\infty} m^*(O_n) < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

Thus $E \in \mathcal{M}$ whenever E is bounded.

We now treat the general case. For each $k \in \mathbb{Z}$, set

$$E^{(k)} := E \cap (k, k+1].$$

Each set $E_n \cap (k, k+1]$ is measurable, and the bounded case therefore implies that $E^{(k)}$ is measurable. Fix $\varepsilon > 0$. For every $k \in \mathbb{Z}$, choose a closed set F_k and an open set O_k such that

$$F_k \subset E^{(k)} \subset O_k \qquad \text{and} \qquad m^*(O_k \setminus F_k) < \frac{\varepsilon}{2^{|k|+3}}.$$

Define

$$O := \bigcup_{k\in\mathbb{Z}} O_k, \qquad F := \bigcup_{k\in\mathbb{Z}} F_k.$$

Then O is open and $F \subset E \subset O$. It remains to show that F is closed. Let $(x_m) \subset F$ be a sequence converging to some $x \in \mathbb{R}$. Choose the unique integer k such that $x \in (k, k+1]$. For all sufficiently large m we have $|x_m - x| < \frac{1}{2}$, and hence $x_m \in (k-1, k+2]$. Since $F_j \subset (j, j+1]$, it follows that, for all sufficiently large m,

$$x_m \in F_{k-1} \cup F_k \cup F_{k+1}.$$

The latter set is closed, being a finite union of closed sets. Passing to the limit gives $x \in F_{k-1} \cup F_k \cup F_{k+1} \subset F$. Therefore F is closed.

Finally,

$$m^*(O \setminus F) \leq \sum_{k\in\mathbb{Z}} m^*(O_k \setminus F_k) < \sum_{k\in\mathbb{Z}} \frac{\varepsilon}{2^{|k|+3}} < \varepsilon.$$

Hence $E \in \mathcal{M}$.

Combining this result with closure under complements from Theorem 7.17, we conclude that $\mathcal{M}$ is a σ-algebra. □

Corollary 7.20 (Open, closed, and Borel sets are measurable)**.** *Every open subset of $\mathbb{R}$ and every closed subset of $\mathbb{R}$ is Lebesgue measurable. Consequently, if $\mathcal{B}(\mathbb{R})$ denotes the Borel σ-algebra on $\mathbb{R}$, then*

$$\mathcal{B}(\mathbb{R}) \subset \mathcal{M}.$$

Proof. Every open interval is measurable by the preceding proposition. Since every open subset of $\mathbb{R}$ is a countable union of open intervals, Theorem 7.19 implies that every open set is measurable. Closed sets are measurable because $\mathcal{M}$ is closed under complements. Finally, $\mathcal{B}(\mathbb{R})$ is the σ-algebra generated by the open sets, whereas $\mathcal{M}$ is a σ-algebra containing every open set. Therefore $\mathcal{B}(\mathbb{R}) \subset \mathcal{M}$. □

Example 7.21. Let $A, B \subset \mathbb{R}$ and assume that $m^*(A) = 0$. Then

$$m^*(A \cup B) = m^*(B).$$

Indeed, subadditivity gives

$$m^*(A \cup B) \leq m^*(A) + m^*(B) = m^*(B),$$

whereas monotonicity yields

$$m^*(B) \leq m^*(A \cup B).$$

Combining the two inequalities proves the claim.

Proposition 7.22 (Translation invariance of measurability)**.** *Let $x \in \mathbb{R}$ and $E \in \mathcal{M}$. Then $x + E \in \mathcal{M}$.*

Proof. Let $\varepsilon > 0$. Choose a closed set F and an open set O such that

$$F \subset E \subset O \qquad \text{and} \qquad m^*(O \setminus F) < \varepsilon.$$

Then $x + F$ is closed and $x + O$ is open, with

$$x + F \subset x + E \subset x + O.$$

Since outer measure is translation invariant,

$$m^*\big((x + O) \setminus (x + F)\big) = m^*(O \setminus F) < \varepsilon.$$

Therefore $x + E \in \mathcal{M}$. □

Exercise 7.23. Verify the following translation identities:

(i) $(F + x)^c = F^c + x$.

(ii) $(O + x) \cap (F + x)^c = (O \cap F^c) + x$.

Remark 7.24 (Hint). For (i), begin with

$$z \notin F + x \iff z - x \notin F \iff z - x \in F^c \iff z \in F^c + x.$$

Then use part (i) to deduce part (ii).

The following regularity theorem is one of the central structural facts about Lebesgue measurable sets: every such set differs from a Borel set by a null set.

Theorem 7.25 (Regularity characterizations of measurable sets)**.** *Let $E \subseteq \mathbb{R}$. Then the following are equivalent:*

(i) E is Lebesgue measurable.

(ii) For every $\varepsilon > 0$, there exists an open set $O \supseteq E$ such that

$$m^*(O \setminus E) < \varepsilon.$$

(iii) For every $\varepsilon > 0$, there exists a closed set $F \subseteq E$ such that

$$m^*(E \setminus F) < \varepsilon.$$

(iv) There exists a G_δ set $G \supseteq E$ such that

$$m^*(G \setminus E) = 0.$$

(v) There exists an F_σ set $F \subseteq E$ such that

$$m^*(E \setminus F) = 0.$$

Proof. We prove

$$(i) \Rightarrow (ii) \Rightarrow (iv) \Rightarrow (i) \qquad \text{and} \qquad (i) \Rightarrow (iii) \Rightarrow (v) \Rightarrow (i).$$

If (i) holds, then the defining approximation of measurability gives an open set $O \supset E$ with $m^*(O \setminus E) < \varepsilon$, because $O \setminus E \subset O \setminus F$. This proves $(i) \Rightarrow (ii)$. Applying the same argument to E^c and then taking complements gives a closed set $F \subset E$ with $m^*(E \setminus F) < \varepsilon$, proving $(i) \Rightarrow (iii)$.

Assume (ii). For each $n \geq 1$, choose an open set $O_n \supseteq E$ such that

$$m^*(O_n \setminus E) < \frac{1}{n}.$$

Set $G := \bigcap_{n=1}^{\infty} O_n$. Then G is a G_δ set, $E \subseteq G$, and for every n,

$$G \setminus E \subseteq O_n \setminus E.$$

Hence

$$m^*(G \setminus E) \leq m^*(O_n \setminus E) < \frac{1}{n},$$

so $m^*(G \setminus E) = 0$. This proves $(ii) \Rightarrow (iv)$.

If (iv) holds, then $G \setminus E$ is measurable because every null set is measurable. Since G is Borel, hence measurable, we obtain

$$E = G \setminus (G \setminus E),$$

so E is measurable. Thus $(iv) \Rightarrow (i)$.

Now assume (iii). For each $n \geq 1$, choose a closed set $F_n \subseteq E$ such that

$$m^*(E \setminus F_n) < \frac{1}{n}.$$

Set $F := \bigcup_{n=1}^{\infty} F_n$. Then F is an F_σ set, $F \subseteq E$, and for every n,

$$E \setminus F \subseteq E \setminus F_n.$$

Therefore

$$m^*(E \setminus F) \leq m^*(E \setminus F_n) < \frac{1}{n},$$

which implies $m^*(E \setminus F) = 0$. Hence $(iii) \Rightarrow (v)$.

Finally, if (v) holds, then $E \setminus F$ is measurable because it is null, and F is measurable because it is Borel. Therefore

$$E = F \cup (E \setminus F)$$

is measurable. This proves $(v) \Rightarrow (i)$. □

The outer measure $m^* : \mathcal{P}(\mathbb{R}) \to [0, \infty]$ restricts to a measure on the σ-algebra $\mathcal{M}$. For $E \in \mathcal{M}$ we write

$$m(E) := m^*(E).$$

Theorem 7.26 (Countable additivity on measurable sets)**.** *Let $(E_n)_{n\geq 1} \subset \mathcal{M}$ be pairwise disjoint, and set $E := \bigcup_{n\geq 1} E_n$. Then $E \in \mathcal{M}$, and*

$$m(E) = \sum_{n\geq 1} m(E_n).$$

Proof. By Theorem 7.19, the set

$$E = \bigcup_{n=1}^{\infty} E_n$$

is measurable. Thus it remains to show that

$$m(E) = \sum_{n=1}^{\infty} m(E_n).$$

The inequality

$$m(E) \leq \sum_{n=1}^{\infty} m(E_n)$$

follows immediately from countable subadditivity. We prove the reverse inequality in two steps.

Step 1: the case where E is bounded. Fix $\varepsilon > 0$. Since each E_n is measurable, we may choose a closed set F_n and an open set O_n such that

$$F_n \subset E_n \subset O_n \qquad \text{and} \qquad m(O_n \setminus F_n) < \frac{\varepsilon}{2^n} \qquad (n \geq 1).$$

Because E is bounded and $F_n \subset E_n \subset E$, each F_n is compact.

For every n,

$$E_n = (E_n \setminus F_n) \cup F_n \subset (O_n \setminus F_n) \cup F_n,$$

so by subadditivity,

$$m(E_n) \leq m(F_n) + m(O_n \setminus F_n).$$

Hence, for each $k \in \mathbb{N}$,

$$\begin{aligned}\sum_{n=1}^{k} m(E_n) &\le \sum_{n=1}^{k} m(F_n) + \sum_{n=1}^{k} m(O_n \setminus F_n) \\ &< \sum_{n=1}^{k} m(F_n) + \sum_{n=1}^{k} \frac{\varepsilon}{2^n} < \sum_{n=1}^{k} m(F_n) + \varepsilon.\end{aligned}$$

Now $F_1, \dots, F_k$ are pairwise disjoint compact sets. Hence they are separated by a positive distance, and therefore Theorem 7.8 gives

$$\sum_{n=1}^{k} m(F_n) = m\Big(\bigcup_{n=1}^{k} F_n\Big) \le m(E).$$

Consequently,

$$\sum_{n=1}^{k} m(E_n) < m(E) + \varepsilon \qquad (k \in \mathbb{N}).$$

Letting $k \to \infty$ and then $\varepsilon \downarrow 0$, we obtain

$$\sum_{n=1}^{\infty} m(E_n) \le m(E).$$

Together with the already proved opposite inequality, this gives

$$m(E) = \sum_{n=1}^{\infty} m(E_n)$$

whenever E is bounded.

Step 2: the general case. For each $k \in \mathbb{Z}$, define

$$A_k := E \cap (k, k+1], \qquad E_{n,k} := E_n \cap (k, k+1].$$

Then

$$E = \bigcup_{k \in \mathbb{Z}} A_k, \qquad E_n = \bigcup_{k \in \mathbb{Z}} E_{n,k}, \qquad A_k = \bigcup_{n=1}^{\infty} E_{n,k}.$$

For each fixed k, the set A_k is bounded, and the sets $E_{n,k}$ are pairwise disjoint and measurable. By Step 1,

$$m(A_k) = \sum_{n=1}^{\infty} m(E_{n,k}) \qquad (k \in \mathbb{Z}).$$

On the other hand, since $E_n = \bigcup_{k \in \mathbb{Z}} E_{n,k}$, countable subadditivity gives

$$m(E_n) \le \sum_{k \in \mathbb{Z}} m(E_{n,k}) \qquad (n \ge 1).$$

Summing over n and using nonnegativity to rearrange the series, we get

$$\sum_{n=1}^{\infty} m(E_n) \le \sum_{n=1}^{\infty} \sum_{k\in\mathbb{Z}} m(E_{n,k}) = \sum_{k\in\mathbb{Z}} \sum_{n=1}^{\infty} m(E_{n,k}) = \sum_{k\in\mathbb{Z}} m(A_k).$$

Now, for each $\ell \in \mathbb{N}$, the sets $A_{-\ell}, \dots, A_\ell$ are pairwise disjoint measurable subsets of E, so finite additivity yields

$$\sum_{k=-\ell}^{\ell} m(A_k) = m\Big(\bigcup_{k=-\ell}^{\ell} A_k \Big) \le m(E).$$

Therefore the partial sums of $\sum_{k\in\mathbb{Z}} m(A_k)$ are bounded above by $m(E)$; if $m(E) = \infty$ this is automatic, and if $m(E) < \infty$ we pass to the limit to obtain

$$\sum_{k\in\mathbb{Z}} m(A_k) \le m(E).$$

Combining the last two displays, we obtain

$$\sum_{n=1}^{\infty} m(E_n) \le m(E).$$

Together with countable subadditivity,

$$m(E) \le \sum_{n=1}^{\infty} m(E_n),$$

we conclude that

$$m(E) = \sum_{n=1}^{\infty} m(E_n).$$

□

Corollary 7.27 (Difference formula)**.** *Let $A, B \in \mathcal{M}$ with $A \subset B$. Then*

$$m(B) = m(A) + m(B \setminus A).$$

In particular, if $m(B) < \infty$, then

$$m(B \setminus A) = m(B) - m(A).$$

Proof. The sets A and $B \setminus A$ are measurable, disjoint, and satisfy

$$B = A \cup (B \setminus A).$$

By countable additivity on measurable sets,

$$m(B) = m(A) + m(B \setminus A).$$

If $m(B) < \infty$, the second identity follows immediately. □

7.4.2 Lebesgue measure and structural results

Theorem 7.28 (Continuity from below)**.** *Let $(E_n)_{n\geq 1} \subset \mathcal{M}$ be an increasing sequence and set $E := \bigcup_{n\geq 1} E_n$. Then*

$$m(E) = \lim_{n\to\infty} m(E_n).$$

Proof. Define $F_1 := E_1$ and $F_n := E_n \setminus E_{n-1}$ for $n \geq 2$. Then the sets $(F_n)_{n\geq 1}$ are pairwise disjoint, measurable, and $E = \bigcup_{n\geq 1} F_n$. By countable additivity,

$$m(E) = \sum_{n\geq 1} m(F_n) \quad \text{and} \quad m(E_k) = \sum_{n=1}^{k} m(F_n).$$

Taking $k \to \infty$ yields $m(E) = \lim_{k\to\infty} m(E_k)$ (possibly $+\infty$). □

Theorem 7.29 (Continuity from above)**.** *Let $(E_n)_{n\geq 1} \subset \mathcal{M}$ be a decreasing sequence such that $m(E_1) < \infty$, and set $E := \bigcap_{n\geq 1} E_n$. Then*

$$m(E) = \lim_{n\to\infty} m(E_n).$$

Proof. Set $A_n := E_1 \setminus E_n$. Then $(A_n)_{n\geq 1}$ is an increasing sequence in $\mathcal{M}$ and $\bigcup_{n\geq 1} A_n = E_1 \setminus E$. By continuity from below,

$$\begin{aligned} m(E_1 \setminus E) &= \lim_{n\to\infty} m(E_1 \setminus E_n) \\ &= \lim_{n\to\infty} \big(m(E_1) - m(E_n)\big), \end{aligned}$$

where the difference is justified by $E_n \subset E_1$ and finiteness of $m(E_1)$. Rearranging gives $m(E) = \lim_{n\to\infty} m(E_n)$. □

Theorem 7.30 (Local criterion for measurability)**.** *For a set $E \subset \mathbb{R}$, the following are equivalent:*

(i) $E \in \mathcal{M}$*;*

(ii) $E \cap (a, b) \in \mathcal{M}$ *for every open interval (a, b);*

(iii) $E \cap (k, k+1] \in \mathcal{M}$ *for every $k \in \mathbb{Z}$.*

Proof. The implication (i)⇒(ii) is immediate because intervals are measurable and $\mathcal{M}$ is a σ-algebra. To prove (ii)⇒(iii), observe that for each $k \in \mathbb{Z}$,

$$E \cap (k, k+1] = \big(E \cap (k, k+2)\big) \setminus \big(E \cap (k+1, k+2)\big),$$

and the two sets on the right are measurable by (ii). Finally,

$$E = \bigcup_{k\in\mathbb{Z}} (E \cap (k, k+1]),$$

so (iii)⇒(i) follows from closure of $\mathcal{M}$ under countable unions. □

Theorem 7.31 (Carathéodory's criterion)**.** *A set $E \subset \mathbb{R}$ is Lebesgue measurable (equivalently, $E \in \mathcal{M}$) if and only if for every $A \subset \mathbb{R}$,*

$$m^*(A) = m^*(A \cap E) + m^*(A \setminus E). \tag{7.1}$$

Proof. Assume first that $E \in \mathcal{M}$, and let $A \subset \mathbb{R}$. Since

$$A = (A \cap E) \cup (A \setminus E),$$

countable subadditivity gives

$$m^*(A) \leq m^*(A \cap E) + m^*(A \setminus E).$$

To prove the reverse inequality, it is enough to consider the case $m^*(A) < \infty$. Choose a G_δ set $G \supset A$ such that

$$m^*(G) = m^*(A).$$

Because every G_δ set is Borel, hence measurable, and $E \in \mathcal{M}$, we have

$$G = (G \cap E) \cup (G \setminus E),$$

so

$$m^*(G) = m(G) = m(G \cap E) + m(G \setminus E) = m^*(G \cap E) + m^*(G \setminus E).$$

By monotonicity,

$$m^*(A \cap E) + m^*(A \setminus E) \leq m^*(G \cap E) + m^*(G \setminus E) = m^*(G) = m^*(A).$$

Hence

$$m^*(A) = m^*(A \cap E) + m^*(A \setminus E)$$

for every $A \subset \mathbb{R}$.

Conversely, assume that (7.1) holds for every $A \subset \mathbb{R}$. We must show that $E \in \mathcal{M}$.

First suppose that $m^*(E) < \infty$. Choose a G_δ set $G \supset E$ such that

$$m^*(G) = m^*(E).$$

Applying (7.1) with $A = G$ and using $G \cap E = E$, we obtain

$$m^*(G) = m^*(G \cap E) + m^*(G \setminus E) = m^*(E) + m^*(G \setminus E).$$

Since $m^*(G) = m^*(E)$, it follows that

$$m^*(G \setminus E) = 0.$$

Therefore $G \setminus E \in \mathcal{M}$ by Theorem 7.17, and hence

$$E = G \setminus (G \setminus E)$$

is measurable.

Now suppose that $m^*(E) = \infty$. For each $k \in \mathbb{Z}$, set

$$E_k := E \cap (k, k+1].$$

We claim that each E_k also satisfies (7.1). Indeed, fix $A \subset \mathbb{R}$. Since $(k, k+1]$ is measurable, we have

$$m^*(A) = m^*\big(A \cap (k, k+1]\big) + m^*\big(A \setminus (k, k+1]\big).$$

Applying (7.1) to $A \cap (k, k+1]$ gives

$$m^*\big(A \cap (k, k+1]\big) = m^*(A \cap E_k) + m^*\big(A \cap (k, k+1] \setminus E\big).$$

Also, applying the measurability of $(k, k+1]$ to the set $A \setminus E_k$, we get

$$m^*(A \setminus E_k) = m^*\big(A \cap (k, k+1] \setminus E\big) + m^*\big(A \setminus (k, k+1]\big).$$

Combining the last three identities yields

$$m^*(A) = m^*(A \cap E_k) + m^*(A \setminus E_k).$$

Thus each E_k satisfies (7.1). Since $E_k \subset (k, k+1]$, we have $m^*(E_k) \leq 1$, so the finite-measure case already proved shows that $E_k \in \mathcal{M}$ for every $k \in \mathbb{Z}$. Finally,

$$E = \bigcup_{k \in \mathbb{Z}} E_k,$$

so $E \in \mathcal{M}$ by Theorem 7.30. □

Remark 7.32. The null-set result already proved in Theorem 7.17(iii) can also be recovered immediately from Carathéodory's criterion: if $m^*(A) = 0$, then for every $E \subset \mathbb{R}$ one has $m^*(E \cap A) = 0$, so

$$m^*(E) = m^*(E \cap A) + m^*(E \setminus A).$$

Interpretation. The identity (7.1) means that E does not "distort" the outer measure: every set A splits into the part inside E and the part outside E without loss or gain of measure. In particular, if $E \in \mathcal{M}$ and $A \subset \mathbb{R}$, then $m^*(A) \geq m^*(A \cap E)$ and $m^*(A) \geq m^*(A \setminus E)$, while (7.1) upgrades these inequalities to an exact decomposition.

7.5 Examples and pathologies

Examples in measure theory do more than illustrate definitions: they calibrate intuition. The Cantor set shows that topological size and measure-theoretic size can diverge dramatically, while the Vitali construction shows that measurability is a genuine structural condition rather than an automatic property of all subsets of $\mathbb{R}$.

7.5.1 The Cantor set

The *Cantor set* is obtained from the interval $[0, 1]$ by iteratively removing open middle thirds. It is a central example in real analysis and measure theory: it is compact, nowhere dense, uncountable, and nevertheless has Lebesgue measure zero. In particular, it illustrates vividly that topological largeness and measure-theoretic largeness are fundamentally different notions.

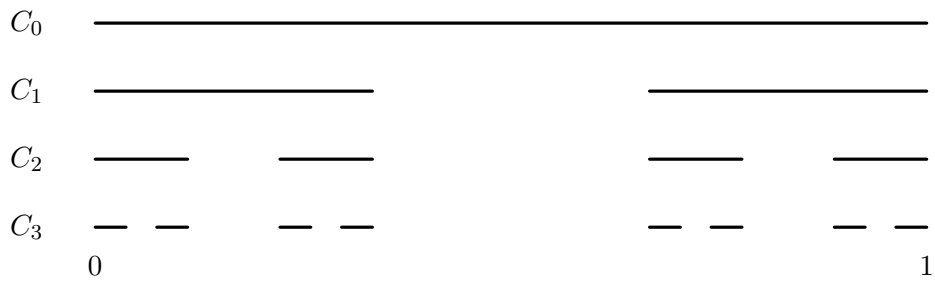


Figure 7.1: The first stages in the construction of the Cantor set. At each stage one removes the open middle third from every remaining interval.

Construction. Define

$$C_0 := [0, 1].$$

After removing the open middle third $(\frac{1}{3}, \frac{2}{3})$, we obtain

$$C_1 = \left[0, \frac{1}{3}\right] \cup \left[\frac{2}{3}, 1\right].$$

Removing the middle third from each connected component of C_1 yields

$$C_2 = \left[0, \frac{1}{9}\right] \cup \left[\frac{2}{9}, \frac{1}{3}\right] \cup \left[\frac{2}{3}, \frac{7}{9}\right] \cup \left[\frac{8}{9}, 1\right].$$

Proceeding inductively, once C_n has been defined, the set C_{n+1} is obtained from C_n by deleting the open middle third of each closed interval component of C_n. The *Cantor set* is then defined by

$$C := \bigcap_{n=0}^{\infty} C_n.$$

Proposition 7.33. *For each $n \geq 0$, the set C_n is the union of 2^n pairwise disjoint closed intervals, each of length 3^{-n}.*

Proof. We argue by induction on n. The assertion is immediate for $n = 0$, since $C_0 = [0, 1]$ is a single closed interval of length $1 = 3^0$.

Assume that the conclusion holds for some $n \geq 0$. Then C_n is a union of 2^n pairwise disjoint closed intervals, each of length 3^{-n}. Removing the open middle third from each such interval leaves precisely two closed intervals, each of length $3^{-(n+1)}$. Consequently, C_{n+1} is the union of $2 \cdot 2^n = 2^{n+1}$ pairwise disjoint closed intervals, each of length $3^{-(n+1)}$. □

Proposition 7.34 (Basic properties of the Cantor set)**.** *The Cantor set C enjoys the following properties.*

(i) The family $\{C_n\}_{n\geq 0}$ is a decreasing sequence of compact measurable sets. In particular, C is compact and Lebesgue measurable.

(ii) The set C contains every endpoint of every interval arising in the construction.

(iii) The complement of C in $[0, 1]$ is the union of all deleted middle-third intervals:

$$[0, 1] \setminus C = \left(\frac{1}{3}, \frac{2}{3}\right) \cup \left(\frac{1}{9}, \frac{2}{9}\right) \cup \left(\frac{7}{9}, \frac{8}{9}\right) \cup \cdots.$$

(iv) The Cantor set has Lebesgue measure zero.

(v) The Cantor set has empty interior; since it is closed, it is nowhere dense in $[0, 1]$.

Proof. (i) Each C_n is a finite union of closed intervals, and is therefore compact and measurable. Moreover, by construction,

$$C_0 \supset C_1 \supset C_2 \supset \cdots.$$

Hence

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

is an intersection of compact measurable sets. Since $C \subset [0, 1]$, it is closed and bounded, hence compact; in particular, it is Lebesgue measurable.

(ii) At every stage of the construction one removes only open middle thirds. Thus no endpoint of a remaining interval is ever deleted. Once an endpoint appears, it belongs to every subsequent C_n, and therefore lies in C.

(iii) A point $x \in [0,1]$ fails to belong to C if and only if it is removed at some finite stage. Accordingly, $[0,1] \setminus C$ is exactly the union of all open middle-third intervals deleted during the construction.

(iv) Since $C \subset C_n$ for every $n \geq 0$, we have

$$m(C) \leq m(C_n) = 2^n \cdot 3^{-n} = \left(\frac{2}{3}\right)^n.$$

Letting $n \to \infty$, we obtain $m(C) = 0$.

(v) If C had nonempty interior, then some open interval I would satisfy $I \subset C$. This is impossible, because

$$0 < m(I) \leq m(C) = 0.$$

Hence C has empty interior. Since C is closed, its closure is itself, so C is nowhere dense in $[0,1]$.

□

7.5.2 Ternary representation of the Cantor set

The ternary description of C provides a precise symbolic model for the construction. Recall, for example, that

$$\frac{1}{3} = \frac{0}{3} + \frac{2}{3^2} + \frac{2}{3^3} + \cdots = (0.0222\ldots)_3, \qquad \frac{2}{3} = (0.2)_3.$$

Thus certain points admit two distinct ternary expansions. More generally, every endpoint of a deleted interval has two ternary representations, one terminating and one nonterminating.

Let E denote the set of all endpoints of deleted intervals. Since the ambiguity in ternary expansion occurs precisely on E, it is convenient to exclude this set when discussing uniqueness.

Proposition 7.35. *Let*

$$F := \left\{ x \in [0,1] : x = \sum_{i=1}^{\infty} \frac{a_i}{3^i},\ a_i \in \{0,1,2\} \right\} \cap E^c.$$

Then each $x \in F$ has a unique ternary expansion. Moreover,

$$C = \left\{ x \in [0,1] : x = \sum_{i=1}^{\infty} \frac{a_i}{3^i},\ a_i \in \{0,2\} \right\}.$$

Proof. Let

$$x = \sum_{i=1}^{\infty} \frac{a_i}{3^i}, \qquad a_i \in \{0,1,2\}.$$

If $a_1 = 1$, then $x \in (\frac{1}{3}, \frac{2}{3})$, so $x \notin C_1$ and hence $x \notin C$. More generally, if the first $i-1$ ternary digits are fixed and $a_i = 1$, then x lies in one of the open middle-third intervals removed at stage i, which implies $x \notin C_i$ and therefore $x \notin C$. Thus, whenever some ternary digit of x is equal to 1, we necessarily have $x \notin C$.

Conversely, suppose $x \notin C$. Then $x \notin C_i$ for some i, and hence x belongs to an open middle-third interval deleted at stage i. This forces the i-th ternary digit of x to be 1. We conclude that $x \in C$ if and only if its ternary expansion involves only the digits 0 and 2.

The uniqueness assertion on F is the standard uniqueness of base-3 expansions away from the exceptional endpoints, which have been excluded by the condition $x \in E^c$. □

Proposition 7.36 (Uniqueness of ternary representation on C)**.** *If*

$$x = \sum_{i=1}^{\infty} \frac{a_i}{3^i} = \sum_{i=1}^{\infty} \frac{b_i}{3^i}, \qquad a_i, b_i \in \{0, 2\},$$

then $a_i = b_i$ for every i.

Proof. Assume, to obtain a contradiction, that the two sequences are different. Let i_0 be the smallest index such that $a_{i_0} \neq b_{i_0}$. Then $a_i = b_i$ for all $i < i_0$. Without loss of generality, $a_{i_0} = 0$ and $b_{i_0} = 2$. Subtracting the common initial segment and multiplying by 3^{i_0-1}, we may reduce to the case $i_0 = 1$. In that case,

$$x \leq \frac{0}{3} + \frac{2}{3^2} + \frac{2}{3^3} + \cdots = \frac{1}{3}, \qquad x \geq \frac{2}{3},$$

which is impossible. Therefore the representation is unique. □

7.5.3 The Cantor set is uncountable

We now establish the uncountability of C by constructing a surjection from C onto $[0, 1]$. For each $x \in C$, fix once and for all a ternary expansion

$$x = \sum_{k=1}^{\infty} \frac{a_k}{3^k}, \qquad a_k \in \{0, 2\},$$

taking the nonterminating expansion for endpoints. Define

$$F : C \to [0, 1], \qquad F(x) := \sum_{k=1}^{\infty} \frac{a_k/2}{2^k}.$$

Since $a_k/2 \in \{0, 1\}$ for every k, the expression on the right defines a point of $[0, 1]$.

Proposition 7.37. *The map $F : C \to [0, 1]$ is surjective. Consequently, the Cantor set is uncountable.*

Proof. Let $y \in [0,1]$. Choose a binary expansion

$$y = \sum_{k=1}^{\infty} \frac{b_k}{2^k}, \qquad b_k \in \{0,1\},$$

and define

$$x := \sum_{k=1}^{\infty} \frac{2b_k}{3^k}.$$

By the ternary characterization established above, $x \in C$. Moreover,

$$F(x) = \sum_{k=1}^{\infty} \frac{b_k}{2^k} = y.$$

Thus F is surjective. Since $[0,1]$ is uncountable, it follows immediately that C is uncountable. □

Exercise 7.38. Show that $F(x) = F(y)$ if and only if x and y are endpoints of one of the deleted middle-third intervals.

Under the convention above, the map F takes the explicit form

$$F(x) = \sum_{i=1}^{\infty} \frac{q_i/2}{2^i}, \qquad x = \sum_{i=1}^{\infty} \frac{q_i}{3^i}, \quad q_i \in \{0,2\}.$$

That is, F is obtained by replacing each ternary digit 0 or 2 by the binary digit 0 or 1. The function is monotone increasing and surjective onto $[0,1]$. Extending F so that it is constant on each deleted middle-third interval yields the classical *Cantor function*.

The ternary model also yields a useful comparison of cardinalities. The Borel σ-algebra $\mathcal{B}(\mathbb{R})$ has cardinality $\mathfrak{c}$, whereas the Cantor set has both cardinality $\mathfrak{c}$ and Lebesgue measure zero. Since every subset of a null set is Lebesgue measurable, it follows that

$$\#\mathcal{M}(\mathbb{R}) \geq \#\mathcal{P}(C) = 2^{\mathfrak{c}},$$

and hence there are strictly more Lebesgue measurable sets than Borel sets. In particular,

$$\mathcal{B}(\mathbb{R}) \subsetneq \mathcal{M}(\mathbb{R}).$$

7.5.4 Nonmeasurable sets

The existence of nonmeasurable sets shows that Lebesgue measurability is a genuine structural condition, not a property enjoyed by all subsets of $\mathbb{R}$. We present the classical Vitali construction.

Definition 7.39 (Rational equivalence)**.** Define a relation $\sim$ on $\mathbb{R}$ by

$$x \sim y \quad \Longleftrightarrow \quad x - y \in \mathbb{Q}.$$

This is an equivalence relation on $\mathbb{R}$. For $x \in \mathbb{R}$, the corresponding equivalence class is

$$x + \mathbb{Q} := \{x + q : \ q \in \mathbb{Q}\}.$$

Every equivalence class intersects $[0, 1]$. Indeed, if $x \in \mathbb{R}$, then there exists $n \in \mathbb{Z}$ such that $x - n \in [0, 1]$, and since $n \in \mathbb{Q}$ we have $x \sim x - n$.

Choose a set $E \subset [0, 1]$ that contains *exactly one* point from each equivalence class that meets $[0, 1]$. Such a set is called a *Vitali set*. Enumerate

$$\mathbb{Q} \cap [-1, 1] = \{r_1, r_2, \dots\},$$

and for each $i \in \mathbb{N}$ define the translate

$$E_i := E + r_i = \{x + r_i : \ x \in E\}.$$

Proposition 7.40. *The family $(E_i)_{i\geq 1}$ has the following properties.*

1. *The sets E_i are pairwise disjoint.*
2. *One has*
$$[0, 1] \subset \bigcup_{i=1}^{\infty} E_i \subset [-1, 2].$$

Proof. (1) Suppose that $z \in E_i \cap E_j$. Then

$$z = x + r_i = y + r_j$$

for some $x, y \in E$. Hence

$$x - y = r_j - r_i \in \mathbb{Q},$$

so $x \sim y$. Since E contains exactly one representative from each equivalence class meeting $[0, 1]$, it follows that $x = y$, and therefore $r_i = r_j$. Thus $i = j$, proving that the sets are pairwise disjoint.

(2) Let $t \in [0, 1]$. By the choice of E, there exists a unique $y \in E$ such that $t \sim y$. Hence $t - y \in \mathbb{Q}$. Because $t, y \in [0, 1]$, we also have $t - y \in [-1, 1]$, so $t - y = r_k$ for some k. Therefore

$$t = y + r_k \in E_k,$$

which proves that $[0, 1] \subset \bigcup_{i=1}^{\infty} E_i$.

On the other hand, if $z \in E_i$, then $z = x + r_i$ for some $x \in E \subset [0, 1]$ and some $r_i \in [-1, 1] \cap \mathbb{Q}$. Hence

$$-1 \leq z \leq 2,$$

so $\bigcup_{i=1}^{\infty} E_i \subset [-1, 2]$. □

Theorem 7.41 (Existence of a nonmeasurable set)**.** *The set E is not Lebesgue measurable.*

Proof. Assume, for contradiction, that $E \in \mathcal{M}$. Since Lebesgue measure is translation invariant, each E_i is measurable and

$$m(E_i) = m(E) \qquad (i \geq 1).$$

Because the sets E_i are pairwise disjoint, Theorem 7.26 gives

$$m\left(\bigcup_{i=1}^{\infty} E_i\right) = \sum_{i=1}^{\infty} m(E_i) = \sum_{i=1}^{\infty} m(E).$$

By Theorem 7.40,

$$1 = m([0,1]) \leq m\left(\bigcup_{i=1}^{\infty} E_i\right) \leq m([-1,2]) = 3.$$

Therefore

$$1 \leq \sum_{i=1}^{\infty} m(E) \leq 3.$$

This is impossible: if $m(E) = 0$, then the sum is 0, whereas if $m(E) > 0$, then the sum is $+\infty$. The contradiction shows that $E \notin \mathcal{M}$. □

Remark 7.42. The above construction uses the axiom of choice. It also explains why outer measure cannot be expected to behave additively on arbitrary subsets of $\mathbb{R}$.

Proposition 7.43 (Consequences for outer measure)**.** *Let $E \subset [0,1]$ be the Vitali set above, and let $E_i = E + r_i$. Then the following statements hold.*

1. *$m^*(E) > 0$.*
2. *Lebesgue outer measure m^* is not countably additive on all subsets of $\mathbb{R}$.*
3. *Lebesgue outer measure m^* is not finitely additive on all subsets of $\mathbb{R}$.*

Proof. (1) Suppose, to the contrary, that $m^*(E) = 0$. Since outer measure is translation invariant, $m^*(E_i) = 0$ for every i. By countable subadditivity,

$$1 = m^*([0,1]) \leq m^*\left(\bigcup_{i=1}^{\infty} E_i\right) \leq \sum_{i=1}^{\infty} m^*(E_i) = 0,$$

a contradiction. Hence $m^*(E) > 0$.

(2) Set

$$A := \bigcup_{i=1}^{\infty} E_i.$$

By Theorem 7.40,

$$1 \leq m^*(A) \leq 3.$$

On the other hand, by translation invariance of outer measure and part (1),

$$m^*(E_i) = m^*(E) > 0 \qquad (i \geq 1),$$

so

$$\sum_{i=1}^{\infty} m^*(E_i) = \sum_{i=1}^{\infty} m^*(E) = +\infty.$$

Therefore

$$m^*\left(\bigcup_{i=1}^{\infty} E_i\right) < \sum_{i=1}^{\infty} m^*(E_i),$$

showing that m^* is not countably additive on all subsets of $\mathbb{R}$.

(3) Suppose, for contradiction, that m^* were finitely additive on all subsets of $\mathbb{R}$. Then for every $n \in \mathbb{N}$,

$$n\, m^*(E) = \sum_{i=1}^{n} m^*(E_i) = m^*\left(\bigcup_{i=1}^{n} E_i\right) \leq m^*\left(\bigcup_{i=1}^{\infty} E_i\right) \leq 3.$$

Hence $m^*(E) \leq 3/n$ for every n, so $m^*(E) = 0$, contradicting part (1). Thus m^* is not finitely additive on arbitrary subsets of $\mathbb{R}$. □

Proposition 7.44 (Measurable subsets of a Vitali set have measure zero)**.** *Let $A \subset E$. If $A \in \mathcal{M}$, then $m(A) = 0$.*

Proof. For each $i \in \mathbb{N}$, set

$$A_i := A + r_i.$$

Since $A \subset E$, the sets A_i are pairwise disjoint and satisfy

$$\bigcup_{i=1}^{\infty} A_i \subset \bigcup_{i=1}^{\infty} E_i \subset [-1, 2].$$

Because A is measurable, each A_i is measurable and

$$m(A_i) = m(A) \qquad (i \geq 1).$$

Therefore, for every $n \in \mathbb{N}$,

$$n\, m(A) = \sum_{i=1}^{n} m(A_i) = m\left(\bigcup_{i=1}^{n} A_i\right) \leq m([-1, 2]) = 3.$$

Hence $m(A) \leq 3/n$ for all n, and letting $n \to \infty$ yields $m(A) = 0$. □

7.6 Measurable functions and simple approximations

Once measurable sets are available, the next task is to identify the corresponding class of functions. The decisive idea is approximation: measurable functions are flexible enough to model complicated behaviour, but still rigid enough to be approximated by simple functions, and that approximation is exactly what makes the Lebesgue integral workable.

7.6.1 Measurable functions

Let $\mathcal{T}_u$ denote the usual topology on $\mathbb{R}$, and let $\mathcal{M}$ denote the σ-algebra of Lebesgue measurable subsets of $\mathbb{R}$. Since every open set in $\mathbb{R}$ is a countable union of open intervals, we have

$$\mathcal{T}_u \subset \mathcal{M}.$$

If $\mathcal{T}_d$ denotes the discrete topology on $\mathbb{R}$, then

$$\mathcal{T}_d = \mathcal{P}(\mathbb{R}),$$

so the measurable sets sit strictly between the usual open sets and the full power set:

$$\mathcal{T}_u \subsetneq \mathcal{M} \subsetneq \mathcal{P}(\mathbb{R}) = \mathcal{T}_d.$$

This observation helps explain why measurability is weaker than continuity but still much stronger than arbitrary set-theoretic behaviour.

In particular, every continuous function $h\colon (\mathbb{R}, \mathcal{T}_u) \to (\mathbb{R}, \mathcal{T}_u)$ is Lebesgue measurable, because for every open set $U \subset \mathbb{R}$ one has $h^{-1}(U) \in \mathcal{T}_u \subset \mathcal{M}$. Thus the class of measurable functions properly contains the continuous functions while still retaining enough structure for approximation and integration.

To accommodate functions such as $x \mapsto 1/x$, whose natural values may include "points at infinity", we enlarge the codomain to the extended real line

$$\overline{\mathbb{R}} := [-\infty, \infty].$$

We equip $\overline{\mathbb{R}}$ with its Borel σ-algebra $\mathcal{B}(\overline{\mathbb{R}})$, generated by the usual open intervals together with the rays $[-\infty, a)$ and $(b, \infty]$ for $a, b \in \mathbb{R}$. Working with $\overline{\mathbb{R}}$ is convenient because pointwise limits of measurable functions may take the values $\pm\infty$, and many natural examples are most cleanly expressed in this language.

Definition 7.45 (Lebesgue measurable function)**.** A function $f : (\mathbb{R}, \mathcal{M}) \to (\overline{\mathbb{R}}, \mathcal{B}(\overline{\mathbb{R}}))$ is called *Lebesgue measurable* if $f^{-1}(B) \in \mathcal{M}$ for every $B \in \mathcal{B}(\overline{\mathbb{R}})$. Equivalently, it suffices to check this on the generating rays:

$$\{x \in \mathbb{R} :\ f(x) > \alpha\} = f^{-1}\big((\alpha, \infty]\big) \in \mathcal{M} \quad \text{for all } \alpha \in \mathbb{R}.$$

Proposition 7.46. *Let $f : (\mathbb{R}, \mathcal{M}) \to \overline{\mathbb{R}}$. The following are equivalent:*

(i) $f^{-1}\big((\alpha, \infty]\big) \in \mathcal{M}$ *for every* $\alpha \in \mathbb{R}$*;*

(ii) $f^{-1}\big([\alpha, \infty]\big) \in \mathcal{M}$ *for every* $\alpha \in \mathbb{R}$*;*

(iii) $f^{-1}\big([-\infty, \alpha)\big) \in \mathcal{M}$ *for every* $\alpha \in \mathbb{R}$*;*

(iv) $f^{-1}\big([-\infty, \alpha]\big) \in \mathcal{M}$ *for every* $\alpha \in \mathbb{R}$*;*

(v) $f^{-1}(\{\infty\}),\ f^{-1}(\{-\infty\}) \in \mathcal{M}$ *and* $f^{-1}((a,b)) \in \mathcal{M}$ *for every* $a, b \in \mathbb{R}$.

Proof. The equivalence of (i)–(iv) follows from elementary set identities.

(i)⇒(ii). For every $\alpha \in \mathbb{R}$,

$$[\alpha, \infty] = \bigcap_{n=1}^{\infty} \left(\alpha - \frac{1}{n}, \infty\right].$$

Hence

$$f^{-1}([\alpha, \infty]) = \bigcap_{n=1}^{\infty} f^{-1}\left(\left(\alpha - \frac{1}{n}, \infty\right]\right) \in \mathcal{M}.$$

(ii)⇒(iii). Since

$$[-\infty, \alpha) = [\alpha, \infty]^c,$$

the conclusion follows by taking complements.

(iii)⇒(iv). For every $\alpha \in \mathbb{R}$,

$$[-\infty, \alpha] = \bigcap_{n=1}^{\infty} \left[-\infty, \alpha + \frac{1}{n}\right),$$

so closure of $\mathcal{M}$ under countable intersections gives (iv).

(iv)⇒(i). Again, this is obtained by taking complements:

$$(\alpha, \infty] = [-\infty, \alpha]^c.$$

Thus (i)–(iv) are equivalent.

To deduce (v), observe that

$$\{\infty\} = \bigcap_{n=1}^{\infty} (n, \infty], \qquad \{-\infty\} = \bigcap_{n=1}^{\infty} [-\infty, -n),$$

and

$$(a, b) = (a, \infty] \cap [-\infty, b).$$

Hence the corresponding preimages are measurable.

Conversely, assume (v). Then for every $\alpha \in \mathbb{R}$,

$$(\alpha, \infty] = \{\infty\} \cup \bigcup_{n=1}^{\infty} (\alpha, n),$$

so

$$f^{-1}((\alpha, \infty]) = f^{-1}(\{\infty\}) \cup \bigcup_{n=1}^{\infty} f^{-1}((\alpha, n)) \in \mathcal{M}.$$

Thus (i) holds, and therefore all five conditions are equivalent. □

Example 7.47. Let $E \in \mathcal{M}$ and define the characteristic function

$$\chi_E(x) = \begin{cases} 1, & x \in E, \\ 0, & x \notin E. \end{cases}$$

Then χ_E is measurable. Indeed, for $\alpha \in \mathbb{R}$,

$$\chi_E^{-1}((\alpha, \infty]) = \begin{cases} \mathbb{R}, & \alpha < 0, \\ E, & 0 \leq \alpha < 1, \\ \varnothing, & \alpha \geq 1, \end{cases}$$

and each of these sets is measurable.

Example 7.48. Every constant function $f : \mathbb{R} \to \overline{\mathbb{R}}$, $f(x) \equiv k$, is measurable. Indeed,

$$f^{-1}((\alpha, \infty]) = \begin{cases} \mathbb{R}, & \alpha < k, \\ \varnothing, & \alpha \geq k, \end{cases}$$

and both sets belong to $\mathcal{M}$. In particular, this covers both the finite case $k \in \mathbb{R}$ and the case $k = \infty$.

A useful practical observation is that one may test measurability using only rational thresholds. Indeed, if $\alpha \in \mathbb{R}$ and $(r_j) \subset \mathbb{Q}$ satisfies $r_j \uparrow \alpha$, then

$$\{x : f(x) \geq \alpha\} = \bigcap_{j=1}^{\infty} \{x : f(x) > r_j\}.$$

Thus, in many arguments, it suffices to verify measurability of sets of the form $\{x : f(x) > r\}$ for $r \in \mathbb{Q}$.

Proposition 7.49 (Stability under algebraic operations and continuous composition)**.** *Let $f, g \colon \mathbb{R} \to \mathbb{R}$ be measurable and let $c \in \mathbb{R}$. Then the following functions are measurable:*

(i) *cf;*

(ii) *$f + g$ and $f - g$;*

(iii) *fg;*

(iv) *$|f|$, f^+, and f^-;*

(v) *$\Phi \circ f$ for every continuous map $\Phi \colon \mathbb{R} \to \mathbb{R}$.*

Proof. For $c = 0$ the claim is immediate, and for $c > 0$ we have

$$\{x : cf(x) > \alpha\} = \{x : f(x) > \alpha/c\}.$$

If $c < 0$, then

$$\{x : cf(x) > \alpha\} = \{x : f(x) < \alpha/c\},$$

which is measurable by the equivalent characterizations of measurable functions. Thus cf is measurable.

To prove measurability of $f + g$, fix $\alpha \in \mathbb{R}$. Then

$$\{x : f(x) + g(x) > \alpha\} = \bigcup_{r \in \mathbb{Q}} \big(\{x : f(x) > r\} \cap \{x : g(x) > \alpha - r\}\big).$$

The inclusion "$\subset$" follows by choosing a rational number strictly between $-g(x) + \alpha$ and $f(x)$, while the reverse inclusion is immediate. Since $\mathbb{Q}$ is countable, the right-hand side is measurable. Replacing g by $-g$ shows that $f - g$ is measurable.

If $\Phi\colon \mathbb{R} \to \mathbb{R}$ is continuous and $U \subset \mathbb{R}$ is open, then $\Phi^{-1}(U)$ is open and hence Borel. Therefore

$$(\Phi \circ f)^{-1}(U) = f^{-1}(\Phi^{-1}(U)) \in \mathcal{M},$$

so $\Phi \circ f$ is measurable. Applying this with $\Phi(t) = |t|$, $\Phi(t) = t^2$, and the continuous maps $t \mapsto \max\{t, 0\}$ and $t \mapsto \max\{-t, 0\}$ proves measurability of $|f|$, f^+, and f^-. Finally,

$$4fg = (f + g)^2 - (f - g)^2,$$

so fg is measurable as well. □

Exercise 7.50. Let $f, g : \mathbb{R} \to \overline{\mathbb{R}}$ be Lebesgue measurable, and assume that $f(x)+g(x) \neq \infty-\infty$ for every $x \in \mathbb{R}$. Prove that $f + g$ is Lebesgue measurable.

A useful route is to analyze the superlevel sets of $f + g$. For each $\alpha \in \mathbb{R}$, show that

$$\{x \in \mathbb{R} : f(x) + g(x) > \alpha\} = \bigcup_{r \in \mathbb{Q}} \big(\{x \in \mathbb{R} : f(x) > r\} \cap \{x \in \mathbb{R} : g(x) > \alpha - r\}\big),$$

and then treat separately the sets on which $f + g = \pm\infty$.

Exercise 7.51. Show that f^2 is Lebesgue measurable whenever f is Lebesgue measurable. For $\alpha > 0$,

$$\{x : f^2(x) > \alpha\} = \{x : f(x) > \sqrt{\alpha}\} \cup \{x : -f(x) > \sqrt{\alpha}\}.$$

Exercise 7.52. Show that if f and g are Lebesgue measurable, then fg is Lebesgue measurable. A useful identity is

$$4fg = (f + g)^2 - (f - g)^2.$$

Definition 7.53. A property P is said to hold *almost everywhere* if the set of points at which P fails has Lebesgue measure zero; that is,

$$m^*\big(\{x \in \mathbb{R} : P \text{ fails at } x\}\big) = 0.$$

Similarly, we write $f = g$ almost everywhere on $\mathbb{R}$ if

$$m^*\big(\{x \in \mathbb{R} : f(x) \neq g(x)\}\big) = 0.$$

Proposition 7.54 (Modification on a null set)**.** *Let $f, g\colon \mathbb{R} \to \overline{\mathbb{R}}$ be functions such that f is measurable and $f = g$ almost everywhere. Then g is measurable.*

Proof. Set

$$N := \{x \in \mathbb{R} : f(x) \neq g(x)\}.$$

Then $m(N) = 0$, so every subset of N is measurable. Fix $\alpha \in \mathbb{R}$. Since $f = g$ on N^c,

$$\{x : g(x) > \alpha\} = \big(\{x : f(x) > \alpha\} \cap N^c\big) \cup \big(\{x : g(x) > \alpha\} \cap N\big).$$

The first set on the right-hand side is measurable because f is measurable, and the second is measurable because it is contained in the null set N. Therefore $\{g > \alpha\} \in \mathcal{M}$ for every $\alpha \in \mathbb{R}$, so g is measurable. □

Example 7.55. Any function that vanishes almost everywhere is measurable, because it agrees almost everywhere with the measurable constant function 0.

Proposition 7.56. *If f and g are Lebesgue measurable, then so are* $\max\{f, g\}$ *and* $\min\{f, g\}$*. More generally, if (f_n) is a sequence of measurable functions, then*

$$\sup_{n\geq 1} f_n, \qquad \inf_{n\geq 1} f_n, \qquad \limsup_{n\to\infty} f_n, \qquad \liminf_{n\to\infty} f_n$$

are measurable. In particular, whenever $f_n \to f$ pointwise, the limit function f is measurable.

Proof. For every $\alpha \in \mathbb{R}$,

$$\{\max\{f, g\} > \alpha\} = \{f > \alpha\} \cup \{g > \alpha\}, \qquad \{\min\{f, g\} > \alpha\} = \{f > \alpha\} \cap \{g > \alpha\}.$$

Hence $\max\{f, g\}$ and $\min\{f, g\}$ are measurable.

Likewise,

$$\Big\{\sup_{n\geq 1} f_n > \alpha\Big\} = \bigcup_{n=1}^{\infty} \{f_n > \alpha\}, \qquad \Big\{\inf_{n\geq 1} f_n > \alpha\Big\} = \bigcap_{n=1}^{\infty} \{f_n > \alpha\},$$

so $\sup_n f_n$ and $\inf_n f_n$ are measurable. Furthermore,

$$\Big\{\limsup_{n\to\infty} f_n > \alpha\Big\} = \bigcap_{m=1}^{\infty} \bigcup_{n\geq m} \{f_n > \alpha\},$$

and

$$\left\{\liminf_{n\to\infty} f_n > \alpha\right\} = \bigcup_{m=1}^{\infty} \bigcap_{n\geq m} \{f_n > \alpha\},$$

so $\limsup f_n$ and $\liminf f_n$ are measurable as well. If $f_n \to f$ pointwise, then $\limsup f_n = \liminf f_n = f$, and hence f is measurable. □

Proposition 7.57. *If $f : \mathbb{R} \to \overline{\mathbb{R}}$ is measurable and $f(x) \neq 0$ for every $x \in \mathbb{R}$, then the reciprocal function $x \mapsto 1/f(x)$ is measurable, where we use the convention $1/(\pm\infty) = 0$.*

Proof. Let $\alpha \in \mathbb{R}$. We consider three cases.

If $\alpha > 0$, then

$$\left\{x : \frac{1}{f(x)} > \alpha\right\} = \{x : 0 < f(x) < 1/\alpha\}.$$

If $\alpha = 0$, then

$$\left\{x : \frac{1}{f(x)} > 0\right\} = \{x : f(x) > 0\}.$$

If $\alpha < 0$, then

$$\left\{x : \frac{1}{f(x)} > \alpha\right\} = \{x : f(x) > 0\} \cup \{x : f(x) < 1/\alpha\}.$$

Each set on the right-hand side is measurable because f is measurable. Therefore all superlevel sets of $1/f$ are measurable, and hence $1/f$ is measurable. □

7.6.2 Simple functions

Definition 7.58 (Simple function)**.** A measurable function $\varphi\colon \mathbb{R} \to \mathbb{R}$ is called a *simple function* if it assumes only finitely many real values. Equivalently, φ can be written in the form

$$\varphi = \sum_{i=1}^{n} \alpha_i \chi_{E_i}, \qquad E_i \in \mathcal{M}, \quad \alpha_i \in \mathbb{R},$$

for some finite family of measurable sets $E_1, \ldots, E_n$. If each $\alpha_i \geq 0$, then φ is called a *nonnegative simple function*. In the integration theory below, nonnegative simple functions therefore take only finitely many *finite* values.

Example 7.59. A basic example is

$$\varphi = 1 \cdot \chi_{[0,1]} + 2 \cdot \chi_{[2,3]}.$$

Theorem 7.60 (Approximation of nonnegative measurable functions by simple functions)**.** *Let*

$$f : (\mathbb{R}, \mathcal{M}) \longrightarrow [0, \infty]$$

be measurable. Then there exists a sequence $(\varphi_n)_{n\geq 1}$ of nonnegative simple functions on $(\mathbb{R}, \mathcal{M})$ such that:

(i) $0 \leq \varphi_n \leq f$ for every n, and $\varphi_n(x) \uparrow f(x)$ for every $x \in \mathbb{R}$;

(ii) if $A \subset \mathbb{R}$ and f is bounded on A, then $\varphi_n \to f$ uniformly on A.

Proof. For each $n \in \mathbb{N}$, divide $[0, 2^n)$ into dyadic intervals of length 2^{-n}:

$$\left[\frac{k}{2^n}, \frac{k+1}{2^n}\right), \qquad k = 0, 1, \ldots, 2^{2n} - 1.$$

Define

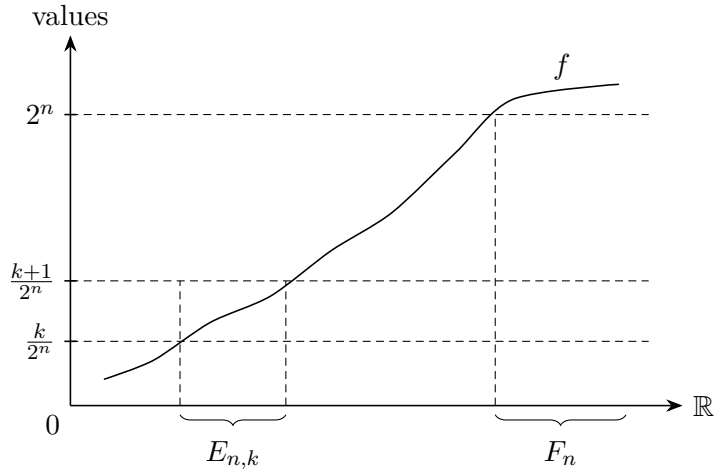
$$F_n := \{x \in \mathbb{R} : f(x) \geq 2^n\}, \qquad E_{n,k} := \left\{x \in \mathbb{R} : \frac{k}{2^n} \leq f(x) < \frac{k+1}{2^n}\right\}.$$

The sets $E_{n,0}, \ldots, E_{n,2^{2n}-1}, F_n$ are measurable, pairwise disjoint, and cover $\mathbb{R}$. Put

$$\varphi_n := \sum_{k=0}^{2^{2n}-1} \frac{k}{2^n} \chi_{E_{n,k}} + 2^n \chi_{F_n}.$$

Then φ_n is nonnegative, simple, and measurable. By construction,

$$0 \leq \varphi_n(x) \leq f(x) \qquad (x \in \mathbb{R}).$$



On each dyadic band, φ_n takes the lower endpoint.

Figure 7.2: Dyadic lower approximation of a nonnegative measurable function on $\mathbb{R}$. The set $E_{n,k}$ consists of points where f lies in the kth dyadic band, while F_n collects the part where $f \geq 2^n$.

We prove that (φ_n) is increasing. Fix $x \in \mathbb{R}$. If $x \in E_{n,k}$, then at the next stage the interval

$$\left[\frac{k}{2^n}, \frac{k+1}{2^n}\right)$$

is divided into

$$\left[\frac{2k}{2^{n+1}}, \frac{2k+1}{2^{n+1}}\right) \quad \text{and} \quad \left[\frac{2k+1}{2^{n+1}}, \frac{2k+2}{2^{n+1}}\right).$$

Thus $\varphi_{n+1}(x)$ is either $k/2^n$ or $(2k+1)/2^{n+1}$, and in both cases

$$\varphi_{n+1}(x) \geq \varphi_n(x).$$

If $x \in F_n$, then $f(x) \geq 2^n$. If $x \in F_{n+1}$, then $\varphi_{n+1}(x) = 2^{n+1} \geq 2^n = \varphi_n(x)$. If $x \in F_n \setminus F_{n+1}$, then $2^n \leq f(x) < 2^{n+1}$, so $x \in E_{n+1,j}$ for some $j \geq 2^{2n+1}$. Hence

$$\varphi_{n+1}(x) = \frac{j}{2^{n+1}} \geq \frac{2^{2n+1}}{2^{n+1}} = 2^n = \varphi_n(x).$$

Therefore $\varphi_n \leq \varphi_{n+1}$ on $\mathbb{R}$.

It remains to prove convergence. If $f(x) = \infty$, then $x \in F_n$ for every n, and so $\varphi_n(x) = 2^n \to \infty = f(x)$. If $f(x) < \infty$, choose n_0 such that $f(x) < 2^n$ for all $n \geq n_0$. For such n, the point x belongs to some $E_{n,k}$, and therefore

$$\frac{k}{2^n} \leq f(x) < \frac{k+1}{2^n}.$$

Consequently

$$0 \leq f(x) - \varphi_n(x) < 2^{-n}, \qquad n \geq n_0,$$

which proves $\varphi_n(x) \to f(x)$. Since the sequence is increasing, this gives $\varphi_n(x) \uparrow f(x)$.

Finally, suppose f is bounded on A; say $0 \leq f(x) \leq M$ for $x \in A$. Choose n_0 such that $M < 2^n$ for all $n \geq n_0$. Then the estimate above holds uniformly for all $x \in A$, and hence

$$\sup_{x \in A} |f(x) - \varphi_n(x)| \leq 2^{-n}, \qquad n \geq n_0.$$

Thus $\varphi_n \to f$ uniformly on A. □

Corollary 7.61 (Approximation of real-valued measurable functions by simple functions)**.** *Let $(\mathbb{R}, \mathcal{M})$ be a measurable space, and let $f : \mathbb{R} \to \mathbb{R}$ be measurable. Then there exists a sequence (ψ_n) of simple functions on $(\mathbb{R}, \mathcal{M})$ such that*

$$\psi_n(x) \longrightarrow f(x) \qquad \textit{and} \qquad |\psi_n(x)| \uparrow |f(x)|$$

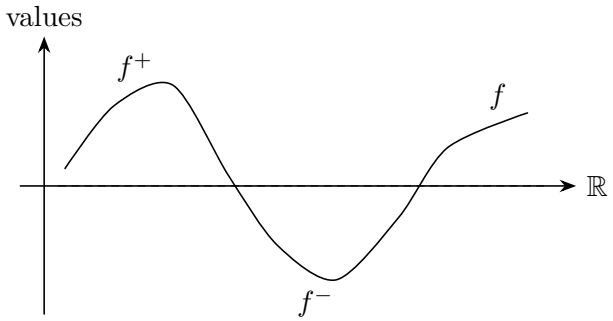
for every $x \in \mathbb{R}$. If f is bounded on $A \subset \mathbb{R}$, the sequence can be chosen so that $\psi_n \to f$ uniformly on A.

Proof. Write

$$f = f^+ - f^-, \qquad f^+ := \max\{f, 0\}, \qquad f^- := \max\{-f, 0\}.$$

By Theorem 7.60, there exist nonnegative simple functions ψ_n^+ and ψ_n^- such that

$$\psi_n^+ \uparrow f^+, \qquad \psi_n^- \uparrow f^-, \qquad 0 \leq \psi_n^+ \leq f^+, \qquad 0 \leq \psi_n^- \leq f^-.$$



$f = f^+ - f^-$; approximate f^+ and f^- separately.

Figure 7.3: Reduction of the real-valued case on $\mathbb{R}$ to the nonnegative case by using the positive and negative parts of f.

Set

$$\psi_n := \psi_n^+ - \psi_n^-.$$

Then ψ_n is simple and $\psi_n(x) \to f(x)$ for every $x \in \mathbb{R}$. Moreover, since $f^+(x)f^-(x) = 0$ and $0 \le \psi_n^+ \le f^+$, $0 \le \psi_n^- \le f^-$, at most one of $\psi_n^+(x)$ and $\psi_n^-(x)$ is nonzero. Therefore

$$|\psi_n(x)| = \psi_n^+(x) + \psi_n^-(x) \uparrow f^+(x) + f^-(x) = |f(x)|.$$

If f is bounded on A, then both f^+ and f^- are bounded on A, and the uniform convergence assertion follows from the uniform part of Theorem 7.60. □

Remark 7.62 (Conventions).

$$0 \cdot \infty = 0, \qquad \infty \cdot 0 = 0.$$

Example: $0 \cdot m(\mathbb{R}) = 0$, $\infty \cdot m(\mathbb{Q}) = 0$.

Remark 7.63 (Conventions). Avoid the indeterminate form $\infty - \infty$.

7.7 Lebesgue integration

This section introduces the integral that measure theory has been preparing all along. The definition proceeds in stages—first for simple functions, then for nonnegative measurable functions, and finally for integrable functions of arbitrary sign—because that layered construction is what makes the later convergence theorems possible.

7.7.1 The Lebesgue integral

Let $\varphi : \mathbb{R} \to [0, \infty]$ be a nonnegative simple measurable function. Thus

$$\varphi = \sum_{j=1}^{m} \alpha_j \chi_{E_j}, \qquad \alpha_j \in [0, \infty), \quad E_j \in \mathcal{M}.$$

We define its Lebesgue integral by

$$\int_{\mathbb{R}} \varphi \, dm := \sum_{j=1}^{m} \alpha_j m(E_j). \tag{7.2}$$

Formula (7.2) is the basic building block of the theory.

Proposition 7.64 (Well-definedness of the integral of a simple function)**.** *The value of* (7.2) *is independent of the chosen representation of φ as a finite linear combination of characteristic functions of measurable sets.*

Proof. Assume that

$$\varphi = \sum_{j=1}^{m} \alpha_j \chi_{E_j} = \sum_{k=1}^{n} \beta_k \chi_{F_k},$$

where the families $(E_j)_{j=1}^m$ and $(F_k)_{k=1}^n$ are measurable and pairwise disjoint within each family. Consider the common refinement

$$G_{jk} := E_j \cap F_k, \qquad 1 \le j \le m, \quad 1 \le k \le n.$$

The sets G_{jk} are measurable and pairwise disjoint, and on each G_{jk} the function φ has both the value α_j and the value β_k; hence $\alpha_j = \beta_k$ whenever $m(G_{jk}) > 0$. Therefore

$$\sum_{j=1}^{m} \alpha_j m(E_j) = \sum_{j=1}^{m} \sum_{k=1}^{n} \alpha_j m(G_{jk}) = \sum_{j=1}^{m} \sum_{k=1}^{n} \beta_k m(G_{jk}) = \sum_{k=1}^{n} \beta_k m(F_k).$$

Thus the definition is independent of the representation. □

Remark 7.65. If $\int_{\mathbb{R}} \varphi \, dm = 0$, then $\varphi = 0$ almost everywhere; if, in addition, the sets in the representation are pairwise disjoint and each coefficient $\alpha_j \ge 0$, then the integral vanishes only when every term is null in measure.

If $E \in \mathcal{M}$, then

$$\varphi|_E = \sum_{j=1}^{m} \alpha_j \chi_{E_j \cap E},$$

so it is natural to set

$$\int_E \varphi \, dm := \sum_{j=1}^{m} \alpha_j m(E_j \cap E).$$

The triple $(\mathbb{R}, \mathcal{M}, m)$ is called the *Lebesgue measure space*. More generally, if $E \in \mathcal{M}$ and

$$\mathcal{M}_E := \{F \cap E : F \in \mathcal{M}\},$$

then $(E, \mathcal{M}_E, m)$ is the induced measure space on E.

Remark 7.66. A simple function can always be written using pairwise disjoint measurable sets. Indeed, by repeatedly decomposing overlaps such as

$$E_1 \cup E_2 = (E_1 \setminus E_2) \cup (E_1 \cap E_2) \cup (E_2 \setminus E_1),$$

one refines an arbitrary representation into a disjoint one.

Example 7.67 (A standard simple-function computation)**.** Let $E_1, \dots, E_n \in \mathcal{M}$ be pairwise disjoint measurable sets of finite measure, let $\alpha_1, \dots, \alpha_n \in \mathbb{C}$, and define

$$f = \sum_{k=1}^{n} \alpha_k \chi_{E_k}.$$

Then for every $1 \le p < \infty$,

$$|f|^p = \sum_{k=1}^{n} |\alpha_k|^p \chi_{E_k},$$

and therefore

$$\int_{\mathbb{R}} |f|^p \, dm = \sum_{k=1}^{n} |\alpha_k|^p \, m(E_k).$$

Proof. Because the sets E_k are pairwise disjoint, at each point x at most one indicator $\chi_{E_k}(x)$ is nonzero. Hence

$$|f(x)|^p = \left| \sum_{k=1}^{n} \alpha_k \chi_{E_k}(x) \right|^p = \sum_{k=1}^{n} |\alpha_k|^p \chi_{E_k}(x).$$

Integrating and using (7.2) gives the stated identity. □

Definition 7.68 (Integral of a nonnegative measurable function)**.** Let $f : \mathbb{R} \to [0, \infty]$ be measurable. Its Lebesgue integral is defined by

$$\int_{\mathbb{R}} f \, dm := \sup \left\{ \int_{\mathbb{R}} \varphi \, dm : \ \varphi \text{ is a nonnegative simple function and } 0 \le \varphi \le f \right\}. \tag{7.3}$$

The value of the integral may be $+\infty$.

Lemma 7.69 (Evaluation by increasing simple approximants)**.** *Let $f : \mathbb{R} \to [0, \infty]$ be measurable, and let (φ_n) be a sequence of nonnegative simple functions such that $\varphi_n \uparrow f$ pointwise. Then*

$$\int_{\mathbb{R}} f \, dm = \lim_{n \to \infty} \int_{\mathbb{R}} \varphi_n \, dm.$$

Proof. Monotonicity for simple functions shows that the sequence $\int \varphi_n \, dm$ is increasing and bounded above by the supremum in (7.3). Hence

$$\lim_{n \to \infty} \int \varphi_n \, dm \le \int f \, dm.$$

Conversely, let ψ be a nonnegative simple function with $\psi \le f$, and fix $0 < \eta < 1$. Set

$$E_n := \{x : \varphi_n(x) \ge (1-\eta)\psi(x)\}.$$

Then $E_n \uparrow \mathbb{R}$. Since the set function $E \mapsto \int_E \psi\, dm$ is a finite or extended measure on $\mathcal{M}$, continuity from below gives

$$\int_{\mathbb{R}} \psi\, dm = \lim_{n\to\infty} \int_{E_n} \psi\, dm.$$

On E_n we have $(1-\eta)\psi \le \varphi_n$, and therefore

$$(1-\eta)\int_{E_n} \psi\, dm \le \int_{\mathbb{R}} \varphi_n\, dm.$$

Letting $n \to \infty$ yields

$$(1-\eta)\int_{\mathbb{R}} \psi\, dm \le \lim_{n\to\infty} \int_{\mathbb{R}} \varphi_n\, dm.$$

Now let $\eta \downarrow 0$ and then take the supremum over all such $\psi \le f$. This gives the reverse inequality. □

Definition 7.70 (The classes L^+ and L^1)**.** We write

$$L^+(\mathbb{R}, \mathcal{M}, m) := \{f : \mathbb{R} \to [0,\infty] : f \text{ is measurable}\}.$$

If $f : \mathbb{R} \to \overline{\mathbb{R}}$ is measurable, write $f = f^+ - f^-$. Whenever at least one of the two quantities $\int_{\mathbb{R}} f^+\, dm$ and $\int_{\mathbb{R}} f^-\, dm$ is finite, we define the signed Lebesgue integral by

$$\int_{\mathbb{R}} f\, dm := \int_{\mathbb{R}} f^+\, dm - \int_{\mathbb{R}} f^-\, dm.$$

We say that f is *integrable* if both $\int_{\mathbb{R}} f^+\, dm$ and $\int_{\mathbb{R}} f^-\, dm$ are finite; equivalently, if $\int_{\mathbb{R}} |f|\, dm < \infty$. The corresponding space is

$$L^1(\mathbb{R}, \mathcal{M}, m) := \left\{ f : \mathbb{R} \to \overline{\mathbb{R}} : f \text{ is measurable and } \int_{\mathbb{R}} |f|\, dm < \infty \right\}.$$

Proposition 7.71 (Basic properties of the simple integral)**.** *Let φ, ψ be nonnegative simple functions on $(\mathbb{R}, \mathcal{M}, m)$.*

(i) For every $c \ge 0$,

$$\int_{\mathbb{R}} c\varphi\, dm = c\int_{\mathbb{R}} \varphi\, dm.$$

(ii)

$$\int_{\mathbb{R}} (\varphi + \psi)\, dm = \int_{\mathbb{R}} \varphi\, dm + \int_{\mathbb{R}} \psi\, dm.$$

(iii) If $\varphi \le \psi$, then

$$\int_{\mathbb{R}} \varphi\, dm \le \int_{\mathbb{R}} \psi\, dm.$$

(iv) For a fixed nonnegative simple function φ, the set function

$$\nu(E) := \int_E \varphi \, dm, \qquad E \in \mathcal{M},$$

is a measure on $(\mathbb{R}, \mathcal{M})$.

Proof. Write

$$\varphi = \sum_{j=1}^{n} \alpha_j \chi_{E_j}, \qquad \psi = \sum_{k=1}^{r} \beta_k \chi_{F_k},$$

where $E_1, \ldots, E_n \in \mathcal{M}$ are pairwise disjoint, $F_1, \ldots, F_r \in \mathcal{M}$ are pairwise disjoint, and, after adjoining a zero coefficient term if necessary,

$$\bigcup_{j=1}^{n} E_j = \bigcup_{k=1}^{r} F_k = \mathbb{R}.$$

For part (i), the function $c\varphi$ has the disjoint representation

$$c\varphi = \sum_{j=1}^{n} c\alpha_j \chi_{E_j}.$$

Thus

$$\int_{\mathbb{R}} c\varphi \, dm = \sum_{j=1}^{n} c\alpha_j m(E_j) = c \sum_{j=1}^{n} \alpha_j m(E_j) = c \int_{\mathbb{R}} \varphi \, dm.$$

For part (ii), use the common measurable refinement

$$G_{jk} := E_j \cap F_k, \qquad 1 \le j \le n, \quad 1 \le k \le r.$$

The sets G_{jk} are pairwise disjoint, measurable, and their union is $\mathbb{R}$. On G_{jk} we have

$$\varphi + \psi = \alpha_j + \beta_k.$$

Therefore

$$\begin{aligned}
\int_{\mathbb{R}} (\varphi + \psi) \, dm &= \sum_{j=1}^{n} \sum_{k=1}^{r} (\alpha_j + \beta_k) m(G_{jk}) \\
&= \sum_{j=1}^{n} \alpha_j \sum_{k=1}^{r} m(E_j \cap F_k) + \sum_{k=1}^{r} \beta_k \sum_{j=1}^{n} m(E_j \cap F_k) \\
&= \sum_{j=1}^{n} \alpha_j m(E_j) + \sum_{k=1}^{r} \beta_k m(F_k) \\
&= \int_{\mathbb{R}} \varphi \, dm + \int_{\mathbb{R}} \psi \, dm.
\end{aligned}$$

For part (iii), assume that $\varphi \le \psi$. On every nonempty set G_{jk}, the functions have the constant values α_j and β_k, respectively; hence $\alpha_j \le \beta_k$ whenever $G_{jk} \neq \varnothing$. Consequently,

$$\begin{aligned}\int_{\mathbb{R}} \varphi \, dm &= \sum_{j=1}^{n} \sum_{k=1}^{r} \alpha_j m(G_{jk}) \\ &\le \sum_{j=1}^{n} \sum_{k=1}^{r} \beta_k m(G_{jk}) \\ &= \int_{\mathbb{R}} \psi \, dm.\end{aligned}$$

It remains to prove part (iv). Fix a nonnegative simple function φ and define

$$\nu(E) := \int_E \varphi \, dm, \qquad E \in \mathcal{M}.$$

Clearly $\nu(\varnothing) = 0$. If $A, B \in \mathcal{M}$ are disjoint, then by part (ii), applied to the simple functions $\varphi\chi_A$ and $\varphi\chi_B$,

$$\nu(A \cup B) = \int_{A\cup B} \varphi \, dm = \int_A \varphi \, dm + \int_B \varphi \, dm = \nu(A) + \nu(B).$$

Now let $A = \bigcup_{k=1}^{\infty} A_k$, where the sets $A_k \in \mathcal{M}$ are pairwise disjoint. Write φ in disjoint form as

$$\varphi = \sum_{i=1}^{s} \lambda_i \chi_{H_i}.$$

Then, using the countable additivity of Lebesgue measure m,

$$\begin{aligned}\nu(A) &= \sum_{i=1}^{s} \lambda_i m(A \cap H_i) \\ &= \sum_{i=1}^{s} \lambda_i \sum_{k=1}^{\infty} m(A_k \cap H_i) \\ &= \sum_{k=1}^{\infty} \sum_{i=1}^{s} \lambda_i m(A_k \cap H_i) \\ &= \sum_{k=1}^{\infty} \nu(A_k).\end{aligned}$$

Hence ν is countably additive, and therefore ν is a measure on $(\mathbb{R}, \mathcal{M})$. □

Proposition 7.72. *If $f, g \in L^+(\mathbb{R}, \mathcal{M}, m)$ and $f \le g$, then*

$$\int_{\mathbb{R}} f \, dm \le \int_{\mathbb{R}} g \, dm.$$

Proof. Let φ be a nonnegative simple function with $\varphi \le f$. Then $\varphi \le g$ as well, and therefore

$$\int_{\mathbb{R}} \varphi \, dm \le \sup_{\psi \le g} \int_{\mathbb{R}} \psi \, dm = \int_{\mathbb{R}} g \, dm.$$

Taking the supremum over all nonnegative simple functions $\varphi \le f$ yields

$$\int_{\mathbb{R}} f \, dm = \sup_{\varphi \le f} \int_{\mathbb{R}} \varphi \, dm \le \int_{\mathbb{R}} g \, dm.$$

□

Proposition 7.73 (Additivity on L^+)**.** *If $f, g \in L^+(\mathbb{R}, \mathcal{M}, m)$, then*

$$\int_{\mathbb{R}} (f+g) \, dm = \int_{\mathbb{R}} f \, dm + \int_{\mathbb{R}} g \, dm.$$

Proof. Choose increasing sequences of nonnegative simple functions (φ_n) and (ψ_n) such that

$$\varphi_n \uparrow f, \qquad \psi_n \uparrow g.$$

Then $\varphi_n + \psi_n$ is a nonnegative simple function for every n, and

$$\varphi_n + \psi_n \uparrow f + g.$$

By additivity for nonnegative simple functions,

$$\int_{\mathbb{R}} (\varphi_n + \psi_n) \, dm = \int_{\mathbb{R}} \varphi_n \, dm + \int_{\mathbb{R}} \psi_n \, dm.$$

Set

$$a_n := \int_{\mathbb{R}} \varphi_n \, dm, \qquad b_n := \int_{\mathbb{R}} \psi_n \, dm.$$

Then (a_n) and (b_n) are increasing sequences with

$$\sup_n a_n = \int_{\mathbb{R}} f \, dm, \qquad \sup_n b_n = \int_{\mathbb{R}} g \, dm.$$

Since the sequences are increasing,

$$\sup_n (a_n + b_n) = \sup_n a_n + \sup_n b_n.$$

By Theorem 7.69,

$$\int_{\mathbb{R}} (f+g) \, dm = \lim_{n\to\infty} \int_{\mathbb{R}} (\varphi_n + \psi_n) \, dm = \lim_{n\to\infty} (a_n + b_n) = \int_{\mathbb{R}} f \, dm + \int_{\mathbb{R}} g \, dm.$$

□

7.7.2 Complex-valued measurable functions and integrals (optional)

Although the core MA224 syllabus is formulated for real-valued functions, complex-valued functions arise naturally in Fourier analysis, differential equations, and probability. It is therefore convenient to record the standard extension.

Definition 7.74. A function $f : \mathbb{R} \to \mathbb{C}$ is called *Lebesgue measurable* if its real and imaginary parts

$$u := \Re f, \qquad v := \Im f$$

are Lebesgue measurable real-valued functions. The function f is called *Lebesgue integrable* if

$$\int_{\mathbb{R}} |f| \, dm < \infty.$$

Proposition 7.75. *Let $f = u + iv : \mathbb{R} \to \mathbb{C}$ be measurable. Then the following hold.*

(i) *$|f|$, $\Re f$, and $\Im f$ are measurable.*

(ii) *If $\int_{\mathbb{R}} |f| \, dm < \infty$, then u and v belong to $L^1(\mathbb{R})$.*

(iii) *For integrable f, the complex Lebesgue integral*

$$\int_{\mathbb{R}} f \, dm := \int_{\mathbb{R}} u \, dm + i \int_{\mathbb{R}} v \, dm$$

is well-defined and satisfies

$$\left| \int_{\mathbb{R}} f \, dm \right| \leq \int_{\mathbb{R}} |f| \, dm.$$

(iv) *If f and g are integrable and $\alpha, \beta \in \mathbb{C}$, then*

$$\int_{\mathbb{R}} (\alpha f + \beta g) \, dm = \alpha \int_{\mathbb{R}} f \, dm + \beta \int_{\mathbb{R}} g \, dm.$$

Proof. Since u and v are measurable, so is

$$|f| = \sqrt{u^2 + v^2},$$

because sums, products, and compositions with continuous functions preserve measurability. Moreover,

$$|u| \leq |f|, \qquad |v| \leq |f|.$$

Hence $\int_{\mathbb{R}} |f| \, dm < \infty$ implies

$$\int_{\mathbb{R}} |u| \, dm \leq \int_{\mathbb{R}} |f| \, dm < \infty, \qquad \int_{\mathbb{R}} |v| \, dm \leq \int_{\mathbb{R}} |f| \, dm < \infty.$$

Thus $u, v \in L^1(\mathbb{R})$, so both real integrals defining $\int f \, dm$ exist.

To prove the estimate, write

$$I := \int_{\mathbb{R}} f \, dm = \int_{\mathbb{R}} u \, dm + i \int_{\mathbb{R}} v \, dm.$$

If $I = 0$, the estimate is immediate. Otherwise choose $\theta \in \mathbb{R}$ such that $e^{-i\theta} I = |I|$. Then

$$|I| = \Re(e^{-i\theta} I) = \Re\Big(\int_{\mathbb{R}} e^{-i\theta} f \, dm \Big) = \int_{\mathbb{R}} \Re(e^{-i\theta} f(x)) \, dm(x),$$

where the last step uses linearity for real and imaginary parts. Since

$$\Re(e^{-i\theta} f(x)) \le |f(x)| \qquad \text{for all } x,$$

we obtain

$$|I| \le \int_{\mathbb{R}} |f| \, dm.$$

Finally, linearity follows immediately from the linearity of the real Lebesgue integral applied to the real and imaginary parts. □

Remark 7.76. This extension is standard in many widely used analysis texts and is included here as an optional supplement. It does not alter the core syllabus, but it does make the present notes more useful as a broader reference.

7.8 Convergence theorems and L^1 consequences

The decisive advantage of the Lebesgue integral is its compatibility with limiting processes. The theorems in this section explain when one may pass a limit through an integral, and they supply the basic toolkit from which much of modern analysis is built.

7.8.1 Monotone convergence theorem

Theorem 7.77 (Monotone Convergence Theorem)**.** *Let (f_n) be a sequence of nonnegative measurable functions on $\mathbb{R}$ such that*

$$0 \le f_1 \le f_2 \le \cdots \qquad \text{and} \qquad f_n(x) \uparrow f(x) \quad \text{for every } x \in \mathbb{R}.$$

Then f is nonnegative and measurable, and

$$\int_{\mathbb{R}} f \, dm = \lim_{n \to \infty} \int_{\mathbb{R}} f_n \, dm. \tag{7.4}$$

Both sides are allowed to have the value $+\infty$.

Proof. Since $f_n \le f$ for every n, monotonicity gives

$$\int_{\mathbb{R}} f_n \, dm \le \int_{\mathbb{R}} f \, dm,$$

so

$$\limsup_{n\to\infty} \int_{\mathbb{R}} f_n \, dm \le \int_{\mathbb{R}} f \, dm.$$

To prove the reverse inequality, fix $\varepsilon \in (0,1)$ and let φ be a nonnegative simple function with $\varphi \le f$. Define

$$E_n := \{x \in \mathbb{R} : f_n(x) \ge (1-\varepsilon)\varphi(x)\}.$$

Because $f_n \uparrow f$ and $\varphi \le f$, the sets E_n are measurable, satisfy $E_n \subset E_{n+1}$, and obey

$$\bigcup_{n=1}^{\infty} E_n = \mathbb{R}.$$

Let

$$\nu(E) := \int_E \varphi \, dm, \qquad E \in \mathcal{M}.$$

By the basic properties of the integral of nonnegative simple functions, ν is a measure. Since $E_n \uparrow \mathbb{R}$, continuity from below yields

$$\nu(E_n) \uparrow \nu(\mathbb{R}) = \int_{\mathbb{R}} \varphi \, dm.$$

On each E_n we have $(1-\varepsilon)\varphi \le f_n$, so

$$(1-\varepsilon)\nu(E_n) = \int_{E_n} (1-\varepsilon)\varphi \, dm \le \int_{E_n} f_n \, dm \le \int_{\mathbb{R}} f_n \, dm.$$

Passing to the limit in n gives

$$(1-\varepsilon) \int_{\mathbb{R}} \varphi \, dm \le \liminf_{n\to\infty} \int_{\mathbb{R}} f_n \, dm.$$

Since this holds for every nonnegative simple function $\varphi \le f$, taking the supremum over such φ yields

$$(1-\varepsilon) \int_{\mathbb{R}} f \, dm \le \liminf_{n\to\infty} \int_{\mathbb{R}} f_n \, dm.$$

Finally let $\varepsilon \downarrow 0$ to obtain

$$\int_{\mathbb{R}} f \, dm \le \liminf_{n\to\infty} \int_{\mathbb{R}} f_n \, dm.$$

Together with the first inequality, this proves (7.4). □

Remark 7.78. The monotonicity assumption in the Monotone Convergence Theorem is essential.

Example 7.79. $f_n = \dfrac{1}{n}\chi_{[0,n]} \to 0$.

$$\int_{\mathbb{R}} f_n \, dm = 1 \ne 0 = \int_{\mathbb{R}} \lim f_n \, dm.$$

Exercise 7.80. Verify the Monotone Convergence Theorem for the following sequences $f_n : \mathbb{R} \to [0, \infty]$:

(i) $f_n = \chi_{(n,n+1)}$.

(ii) $f_n = n\chi_{(0,\frac{1}{n})}$.

Remark 7.81. The Monotone Convergence Theorem is the first genuinely measure-theoretic convergence result. It is the mechanism behind many of the structural properties established earlier and is the starting point for Fatou's lemma and the Dominated Convergence Theorem.

Proposition 7.82. *Let $E \in \mathcal{M}$ and let $f \in L^+(\mathbb{R}, \mathcal{M}, m)$. If*

$$\int_E f\, dm = 0,$$

then $f = 0$ almost everywhere on E.

Proof. For each $n \geq 1$ let

$$E_n := \left\{ x \in E : f(x) > \frac{1}{n} \right\}.$$

Since $\frac{1}{n}\chi_{E_n} \leq f$, monotonicity gives

$$\frac{1}{n}\, m(E_n) = \int_E \frac{1}{n}\chi_{E_n}\, dm \leq \int_E f\, dm = 0.$$

Thus $m(E_n) = 0$ for every n, and therefore

$$\{x \in E : f(x) > 0\} = \bigcup_{n=1}^{\infty} E_n$$

has measure zero. Hence $f = 0$ almost everywhere on E. □

Corollary 7.83 (Monotone convergence almost everywhere)**.** *Let $f_n, f \in L^+(\mathbb{R}, \mathcal{M}, m)$ and suppose that $f_n \uparrow f$ pointwise almost everywhere on $\mathbb{R}$. Then*

$$\int_{\mathbb{R}} f\, dm = \lim_{n\to\infty} \int_{\mathbb{R}} f_n\, dm.$$

Proof. Let

$$A := \{x \in \mathbb{R} : f_n(x) \uparrow f(x)\}.$$

By assumption, $m(A^c) = 0$. Define

$$g_n := \chi_A f_n, \qquad g := \chi_A f.$$

Then $g_n \uparrow g$ pointwise on all of $\mathbb{R}$, so the Monotone Convergence Theorem gives

$$\int_{\mathbb{R}} g\, dm = \lim_{n\to\infty} \int_{\mathbb{R}} g_n\, dm.$$

Since A^c has measure zero and the functions are nonnegative,

$$\int_{\mathbb{R}} g\,dm = \int_{\mathbb{R}} f\,dm \qquad \text{and} \qquad \int_{\mathbb{R}} g_n\,dm = \int_{\mathbb{R}} f_n\,dm \quad (n \geq 1).$$

Substituting these identities yields the conclusion. □

Theorem 7.84 (Vanishing integral criterion)**.** *Let $f \in L^+(\mathbb{R}, \mathcal{M}, m)$. Then*

$$\int_{\mathbb{R}} f\,dm = 0 \quad \iff \quad f = 0 \text{ almost everywhere on } \mathbb{R}.$$

Proof. If $f = 0$ almost everywhere, then every nonnegative simple function $\varphi \leq f$ also vanishes almost everywhere, and therefore

$$\int_{\mathbb{R}} \varphi\,dm = 0.$$

Taking the supremum over all such simple functions gives

$$\int_{\mathbb{R}} f\,dm = 0.$$

Conversely, assume that

$$\int_{\mathbb{R}} f\,dm = 0.$$

For each $n \geq 1$, set

$$E_n := \left\{ x \in \mathbb{R} : f(x) > \frac{1}{n} \right\}.$$

Since $\frac{1}{n}\chi_{E_n} \leq f$, monotonicity of the integral implies

$$\frac{1}{n}\, m(E_n) = \int_{\mathbb{R}} \frac{1}{n}\chi_{E_n}\,dm \leq \int_{\mathbb{R}} f\,dm = 0.$$

Hence $m(E_n) = 0$ for every n. Now

$$\{x \in \mathbb{R} : f(x) > 0\} = \bigcup_{n=1}^{\infty} E_n,$$

so countable subadditivity gives

$$m(\{x : f(x) > 0\}) = 0.$$

That is, $f = 0$ almost everywhere. □

7.8.2 Fatou's lemma

Lemma 7.85 (Fatou's lemma)**.** *Let (f_n) be a sequence in $L^+(\mathbb{R}, \mathcal{M}, m)$. Then*

$$\int_{\mathbb{R}} \liminf_{n\to\infty} f_n\,dm \leq \liminf_{n\to\infty} \int_{\mathbb{R}} f_n\,dm.$$

Proof. For each $k \geq 1$, define

$$g_k := \inf_{n \geq k} f_n.$$

Then $0 \leq g_k \leq f_n$ for every $n \geq k$, so

$$\int_{\mathbb{R}} g_k \, dm \leq \inf_{n \geq k} \int_{\mathbb{R}} f_n \, dm.$$

Moreover, the sequence (g_k) is increasing and

$$g_k \uparrow \sup_{k \geq 1} \inf_{n \geq k} f_n = \liminf_{n \to \infty} f_n.$$

By the Monotone Convergence Theorem,

$$\int_{\mathbb{R}} \liminf_{n \to \infty} f_n \, dm = \lim_{k \to \infty} \int_{\mathbb{R}} g_k \, dm \leq \lim_{k \to \infty} \inf_{n \geq k} \int_{\mathbb{R}} f_n \, dm = \liminf_{n \to \infty} \int_{\mathbb{R}} f_n \, dm.$$

□

Remark 7.86. Strict inequality can occur in Fatou's lemma.

For example, if

$$f_n = \frac{1}{n} \chi_{[0,n]},$$

then $f_n \to 0$ pointwise, while

$$\int_{\mathbb{R}} \liminf_{n \to \infty} f_n \, dm = 0 < 1 = \liminf_{n \to \infty} \int_{\mathbb{R}} f_n \, dm.$$

Remark 7.87. Fatou's lemma need not hold for functions that are not nonnegative.

Example 7.88. Let

$$f_n = -\frac{1}{n} \chi_{[n,2n]}, \qquad n \geq 1.$$

Then $f_n(x) \to 0$ for every $x \in \mathbb{R}$, so

$$\int_{\mathbb{R}} \liminf_{n \to \infty} f_n \, dm = 0.$$

On the other hand,

$$\int_{\mathbb{R}} f_n \, dm = -1 \qquad \text{for every } n,$$

and hence

$$0 = \int_{\mathbb{R}} \liminf_{n \to \infty} f_n \, dm > -1 = \liminf_{n \to \infty} \int_{\mathbb{R}} f_n \, dm.$$

7.8.3 Signed integrals and the space L^1

Let $f : (\mathbb{R}, \mathcal{M}, m) \to \overline{\mathbb{R}}$ be measurable. Recall that

$$f = f^+ - f^-, \qquad |f| = f^+ + f^-.$$

Thus f is integrable precisely when both f^+ and f^- have finite integral, or equivalently when

$$\int_{\mathbb{R}} |f|\, dm < \infty.$$

In that case,

$$\int_{\mathbb{R}} f\, dm = \int_{\mathbb{R}} f^+\, dm - \int_{\mathbb{R}} f^-\, dm.$$

We therefore write

$$L^1(\mathbb{R}, \mathcal{M}, m) := \left\{ f : \mathbb{R} \to \overline{\mathbb{R}} : f \text{ is measurable and } \int_{\mathbb{R}} |f|\, dm < \infty \right\}.$$

We also use the equivalent notations $L^1(\mathbb{R})$ and $L^1(\mathbb{R}, m)$ when no ambiguity can arise.

A standard and important observation is that

$$\int_{\mathbb{R}} |f|\, dm = 0 \quad \Longleftrightarrow \quad f = 0 \text{ almost everywhere.}$$

Accordingly, one usually identifies functions that agree almost everywhere; with this convention, $L^1(\mathbb{R}, \mathcal{M}, m)$ becomes a normed linear space with norm

$$\|f\|_1 := \int_{\mathbb{R}} |f|\, dm.$$

Proposition 7.89. *If $f \in L^1(\mathbb{R}, \mathcal{M}, m)$, then*

$$\left| \int_{\mathbb{R}} f\, dm \right| \le \int_{\mathbb{R}} |f|\, dm.$$

Proof. Write $f = f^+ - f^-$ with $f^+, f^- \ge 0$. Then

$$\begin{aligned} \left| \int_{\mathbb{R}} f\, dm \right| &= \left| \int_{\mathbb{R}} f^+\, dm - \int_{\mathbb{R}} f^-\, dm \right| \\ &\le \int_{\mathbb{R}} f^+\, dm + \int_{\mathbb{R}} f^-\, dm \\ &= \int_{\mathbb{R}} |f|\, dm. \end{aligned}$$

□

7.8.4 Chebyshev's inequality

Proposition 7.90 (Chebyshev's inequality)**.** *Let $f \in L^1(\mathbb{R}, \mathcal{M}, m)$ and let $\alpha > 0$. Then*

$$m(\{x \in \mathbb{R} : |f(x)| \geq \alpha\}) \leq \frac{1}{\alpha} \|f\|_1.$$

Proof. Set $E_\alpha := \{x \in \mathbb{R} : |f(x)| \geq \alpha\}$. Since $\alpha \chi_{E_\alpha} \leq |f|$, monotonicity of the integral gives

$$\alpha \, m(E_\alpha) = \int_{\mathbb{R}} \alpha \chi_{E_\alpha} \, dm \leq \int_{\mathbb{R}} |f| \, dm = \|f\|_1.$$

Dividing by α yields the stated estimate. □

Corollary 7.91. *If $f \in L^1(\mathbb{R}, \mathcal{M}, m)$, then*

$$m(\{x \in \mathbb{R} : |f(x)| = \infty\}) = 0.$$

In other words, every L^1-function is finite almost everywhere.

Proof. For each $n \in \mathbb{N}$, set $E_n := \{x \in \mathbb{R} : |f(x)| \geq n\}$. Then

$$\{x \in \mathbb{R} : |f(x)| = \infty\} = \bigcap_{n=1}^{\infty} E_n.$$

By Chebyshev's inequality,

$$m(E_n) \leq \frac{1}{n} \|f\|_1 \qquad (n \geq 1).$$

Since $\{x : |f(x)| = \infty\} \subset E_n$ for every n, monotonicity gives

$$m(\{x : |f(x)| = \infty\}) \leq m(E_n) \leq \frac{1}{n} \|f\|_1.$$

Letting $n \to \infty$ yields the claim. □

7.8.5 Dominated convergence theorem

Theorem 7.92 (Dominated Convergence Theorem)**.** *Let $f_n : (\mathbb{R}, \mathcal{M}, m) \to \overline{\mathbb{R}}$ be a sequence of measurable functions such that*

(i) $f_n(x) \to f(x)$ pointwise for all $x \in \mathbb{R}$;

(ii) $|f_n| \leq g$ for some $g \in L^1(\mathbb{R}, \mathcal{M}, m)$.

Then

$$\int_{\mathbb{R}} f \, dm = \lim_{n \to \infty} \int_{\mathbb{R}} f_n \, dm.$$

Proof. Since $f_n \to f$ pointwise and $|f_n| \leq g$, we have $|f| \leq g$ pointwise. Hence $f \in L^1(\mathbb{R}, \mathcal{M}, m)$.

Now $g + f_n \geq 0$ and $g - f_n \geq 0$ for every n, and

$$g + f_n \to g + f, \qquad g - f_n \to g - f \qquad \text{pointwise on } \mathbb{R}.$$

Applying Fatou's lemma to $(g + f_n)$ gives

$$\int_{\mathbb{R}} (g + f)\, dm \leq \liminf_{n\to\infty} \int_{\mathbb{R}} (g + f_n)\, dm,$$

so

$$\int_{\mathbb{R}} f\, dm \leq \liminf_{n\to\infty} \int_{\mathbb{R}} f_n\, dm.$$

Applying Fatou's lemma to $(g - f_n)$ gives

$$\int_{\mathbb{R}} (g - f)\, dm \leq \liminf_{n\to\infty} \int_{\mathbb{R}} (g - f_n)\, dm,$$

which is equivalent to

$$\limsup_{n\to\infty} \int_{\mathbb{R}} f_n\, dm \leq \int_{\mathbb{R}} f\, dm.$$

Combining the two inequalities, we obtain

$$\limsup_{n\to\infty} \int_{\mathbb{R}} f_n\, dm \leq \int_{\mathbb{R}} f\, dm \leq \liminf_{n\to\infty} \int_{\mathbb{R}} f_n\, dm.$$

Therefore the limit exists and equals $\int_{\mathbb{R}} f\, dm$. □

Exercise 7.93. For each of the following sequences $f_n : (\mathbb{R}, \mathcal{M}, m) \to \mathbb{R}$, determine whether the hypotheses of the Dominated Convergence Theorem are satisfied. If they are, compute the limiting integral; if they are not, identify which hypothesis fails.

(i) $f_n = n\chi_{[0,1/n]}$.

(ii) $f_n = \frac{1}{n}\chi_{[n,n+1]}$.

(iii) $f_n = \chi_{[n,n+1]}$.

Remark 7.94 (Hint). In part (i), $f_n \to 0$ pointwise but $\int_{\mathbb{R}} f_n\, dm = 1$ for every n, so no integrable dominating function can exist.

Exercise 7.95 (Pointwise limits and interchange of limit and integral)**.** Consider the following sequences $f_n : \mathbb{R} \to \mathbb{R}$:

(i) $f_n(t) = \dfrac{1}{n} e^{-t^2}$;

(ii) $f_n(t) = \dfrac{1}{1 + nt^2}$;

(iii) $f_n(t) = e^{-t^2/n}\, \chi_{(0,n)}(t)$;

(iv) $f_n(t) = e^{-t^2/n}$.

In each case determine the pointwise limit f, decide whether $f \in L^1(\mathbb{R}, m)$, and determine whether

$$\lim_{n\to\infty} \int_{\mathbb{R}} f_n\, dm = \int_{\mathbb{R}} \lim_{n\to\infty} f_n\, dm.$$

This exercise is meant to separate three issues that are often confused: pointwise convergence, integrability of the pointwise limit, and the validity of passing a limit through an integral.

7.8.6 Bounded convergence theorem

Theorem 7.96 (Bounded Convergence Theorem)**.** *Let $E \in \mathcal{M}$ satisfy $m(E) < \infty$. If $f_n, f : (E, \mathcal{M}_E, m) \to \overline{\mathbb{R}}$ are such that*

(i) $|f_n(x)| \leq M$ for every $n \in \mathbb{N}$ and every $x \in E$;

(ii) $f_n(x) \to f(x)$ for every $x \in E$.

Then

$$\int_E f\, dm = \lim_{n\to\infty} \int_E f_n\, dm.$$

Proof. Since $|f_n| \leq M\chi_E$ on E and

$$\int_E M\, dm = Mm(E) < \infty,$$

the function $M\chi_E$ is integrable. Thus the sequence (f_n) is dominated by an integrable function. The Dominated Convergence Theorem therefore yields

$$\int_E f\, dm = \lim_{n\to\infty} \int_E f_n\, dm.$$

□

7.9 Relation with the Riemann integral

A new theory becomes more convincing when it clarifies, rather than discards, what came before it. The point of this section is to show that Lebesgue integration extends Riemann integration on compact intervals and to explain precisely how the older theory sits inside the newer one.

7.9.1 Riemann and Lebesgue integration on compact intervals

Theorem 7.97 (Lebesgue's criterion for Riemann integrability)**.** *Let $f : [a, b] \to \mathbb{R}$ be bounded. Then f is Riemann integrable on $[a, b]$ if and only if the set of points at which f is discontinuous has Lebesgue measure zero. Equivalently, f is Riemann integrable if and only if it is continuous almost everywhere on $[a, b]$.*

Proof. For an interval $J \subset [a, b]$, write

$$\operatorname{osc}(f; J) := \sup_J f - \inf_J f.$$

For $x \in [a, b]$, define the oscillation of f at x by

$$\omega_f(x) := \inf_{\delta>0} \operatorname{osc}\big(f; [a, b] \cap (x - \delta, x + \delta)\big).$$

Then f is continuous at x if and only if $\omega_f(x) = 0$. Hence the discontinuity set is

$$D_f = \bigcup_{r=1}^{\infty} D_r, \qquad D_r := \{x \in [a,b] : \omega_f(x) \geq 1/r\}.$$

Assume first that f is Riemann integrable. Fix $r \in \mathbb{N}$ and $\varepsilon > 0$. Choose a partition

$$P : a = x_0 < x_1 < \cdots < x_N = b$$

such that $U(P,f) - L(P,f) < \varepsilon/r$. If $J_i = [x_{i-1}, x_i]$ and

$$I_r := \{i : \operatorname{osc}(f; J_i) \geq 1/r\},$$

then

$$\frac{1}{r} \sum_{i \in I_r} (x_i - x_{i-1}) \leq \sum_{i=1}^{N} \operatorname{osc}(f; J_i)(x_i - x_{i-1}) = U(P,f) - L(P,f) < \frac{\varepsilon}{r}.$$

Thus the total length of the intervals indexed by I_r is less than ε. Apart from the finitely many partition points, every point of D_r must lie in one of these intervals. Since finite sets have measure zero and ε is arbitrary, $m(D_r) = 0$. Therefore

$$m(D_f) \leq \sum_{r=1}^{\infty} m(D_r) = 0.$$

Conversely, suppose $m(D_f) = 0$. Let $|f| \leq M$ on $[a,b]$. It is enough, by Darboux's criterion, to show that for every $\varepsilon > 0$ there is a partition P with

$$U(P,f) - L(P,f) < \varepsilon.$$

Choose $\eta > 0$ so small that $\eta(b-a) < \varepsilon/2$. Since $D_\eta := \{x : \omega_f(x) \geq \eta\}$ is a closed subset of the null set D_f, it is compact and has measure zero. Hence there is a finite union G of open intervals such that

$$D_\eta \subset G, \qquad \ell(G) < \frac{\varepsilon}{4M+1},$$

where $\ell(G)$ denotes the sum of the lengths of the intervals in the union. On the compact set $K = [a,b] \setminus G$ we have $\omega_f(x) < \eta$ at every point. Therefore K may be covered by finitely many open intervals on each of which the oscillation of f is less than η. Choose a partition refining the endpoints of the intervals forming G and fine enough on K so that each remaining subinterval lies inside one of these oscillation intervals. Then the subintervals contained in G contribute at most $2M\ell(G)$ to $U(P,f) - L(P,f)$, while all other subintervals contribute at most $\eta(b-a)$. Consequently,

$$U(P,f) - L(P,f) \leq 2M\ell(G) + \eta(b-a) < \varepsilon.$$

Thus f is Riemann integrable. □

Remark 7.98. Every Riemann integrable function on $[a, b]$ is also Lebesgue integrable. A useful heuristic is that the upper and lower step-function approximations arising from partitions already resemble the simple-function approximations used in Lebesgue integration. The rigorous statement and proof are given in Theorem 7.99.

Theorem 7.99 (Riemann integrability implies Lebesgue integrability)**.** *Let $f \in \mathcal{R}[a, b]$ be Riemann integrable on $[a, b]$. Then f defines an element of $L^1([a, b], m)$ and*

$$\int_{[a,b]} f \, dm = \int_a^b f(x) \, dx, \tag{7.5}$$

where the right-hand side is the Riemann integral.

Proof. Set $I = [a, b]$. Since f is Riemann integrable, it is bounded; choose $m_0, M_0 \in \mathbb{R}$ such that

$$m_0 \leq f(x) \leq M_0, \qquad x \in I. \tag{7.6}$$

Choose a sequence of partitions (P_n) such that P_{n+1} refines P_n and

$$\lim_{n\to\infty} U(P_n, f) = \lim_{n\to\infty} L(P_n, f) = \int_a^b f(x) \, dx. \tag{7.7}$$

Writing $P_n = \{a = t_{0,n} < t_{1,n} < \cdots < t_{N_n,n} = b\}$, let

$$I_{k,n} := [t_{k-1,n}, t_{k,n}) \quad (1 \leq k < N_n), \qquad I_{N_n,n} := [t_{N_n-1,n}, b].$$

Define the lower and upper step functions

$$\psi_n(x) := \sum_{k=1}^{N_n} \Big(\inf_{t\in I_{k,n}} f(t) \Big) \chi_{I_{k,n}}(x),$$
$$\varphi_n(x) := \sum_{k=1}^{N_n} \Big(\sup_{t\in I_{k,n}} f(t) \Big) \chi_{I_{k,n}}(x).$$

Then

$$\psi_n \leq f \leq \varphi_n \quad \text{on } I,$$

and refinement gives

$$\psi_1 \leq \psi_2 \leq \cdots \leq \psi_n \leq \cdots \leq \varphi_n \leq \cdots \leq \varphi_2 \leq \varphi_1.$$

Moreover,

$$m_0 \leq \psi_n \leq f \leq \varphi_n \leq M_0 \qquad \text{on } I.$$

Let

$$\psi := \sup_{n\geq 1} \psi_n, \qquad \varphi := \inf_{n\geq 1} \varphi_n.$$

Then ψ and φ are measurable and $\psi \le f \le \varphi$ on I. Since $\psi_n - m_0 \uparrow \psi - m_0$ and these functions are nonnegative, the Monotone Convergence Theorem gives

$$\int_I \psi \, dm = \lim_{n\to\infty} \int_I \psi_n \, dm = \lim_{n\to\infty} L(P_n, f).$$

Similarly, $M_0 - \varphi_n \uparrow M_0 - \varphi$, and hence

$$\int_I \varphi \, dm = \lim_{n\to\infty} \int_I \varphi_n \, dm = \lim_{n\to\infty} U(P_n, f).$$

Using (7.7), we obtain

$$\int_I (\varphi - \psi) \, dm = 0.$$

Since $\varphi - \psi \ge 0$, it follows that $\varphi = \psi$ almost everywhere on I. Because $\psi \le f \le \varphi$, we have $f = \psi$ almost everywhere on I. Therefore f is measurable by Theorem 7.54. The bound (7.6) gives

$$|f| \le \max\{|m_0|, |M_0|\}\chi_I,$$

so $f \in L^1(I, m)$. Finally,

$$\int_I f \, dm = \int_I \psi \, dm = \lim_{n\to\infty} L(P_n, f) = \int_a^b f(x) \, dx.$$

This proves (7.5). □

Remark 7.100. One has the strict inclusion $\mathcal{R}[a,b] \subsetneq L^1([a,b], m)$. For instance, the Dirichlet function $f = \chi_{(\mathbb{R}\setminus\mathbb{Q})\cap[0,1]}$ belongs to $L^1([0,1], m)$ with $\int_{[0,1]} f \, dm = 1$, but is not Riemann integrable since $L(P, f) = 0$ and $U(P, f) = 1$ for every partition P.

Example 7.101 (Incompleteness of Riemann integrable functions in the L^1-norm)**.** On the pointwise space of Riemann integrable functions, the formula

$$\|f\|_1 := \int_a^b |f(x)| \, dx$$

is a seminorm rather than a genuine norm, because functions that differ only on a finite set may have distance zero. After quotienting by equality almost everywhere, however, it becomes the usual L^1-norm. In that normed sense the Riemann integrable functions form a proper, non-complete subspace of $L^1([a,b], m)$.

For example, define $f : [0,1] \to \mathbb{R}$ by

$$f(x) = \begin{cases} x^{-1/2}, & 0 < x \le 1, \\ 0, & x = 0. \end{cases}$$

Then $f \notin \mathcal{R}[0,1]$, since it is unbounded near 0, but $f \in L^1([0,1], m)$ and

$$\int_{[0,1]} f\, dm = \lim_{n\to\infty} \int_{1/n}^{1} \frac{dx}{\sqrt{x}} = 2.$$

Set

$$f_n := f\chi_{[1/n,1]}, \qquad n \in \mathbb{N}.$$

Each f_n is bounded and Riemann integrable on $[0,1]$. If $m > n$, then

$$\|f_m - f_n\|_1 = \int_{1/m}^{1/n} \frac{dx}{\sqrt{x}} = 2\left(\frac{1}{\sqrt{n}} - \frac{1}{\sqrt{m}}\right),$$

which tends to 0 as $m, n \to \infty$. Thus (f_n) is Cauchy in the L^1-norm, but its L^1-limit is f, which is not Riemann integrable. Hence $\mathcal{R}[0,1]$ is not complete in the L^1-norm.

7.9.2 Improper Riemann integration

The preceding subsection concerns Riemann integration on compact intervals. Many important integrals in analysis, however, are taken over unbounded intervals, where the appropriate Riemann notion is defined by a limiting process.

Let $f : [a, \infty) \to \mathbb{R}$ be such that $f \in \mathcal{R}[a,b]$ for every $b > a$. We say that f is *improperly Riemann integrable* on $[a,\infty)$ if the finite limit

$$\lim_{b\to\infty} \int_a^b f(x)\, dx$$

exists. In that case we write

$$\int_a^\infty f(x)\, dx := \lim_{b\to\infty} \int_a^b f(x)\, dx.$$

The same definition applies to intervals of the form $(-\infty, a]$ and, by splitting at a finite point, to the whole real line.

Remark 7.102. Improper Riemann integrability does *not* by itself imply Lebesgue integrability. The missing hypothesis is absolute integrability. Thus conditional convergence of an improper integral is a genuinely Riemannian phenomenon and need not produce a function in L^1.

Example 7.103. The classical integral

$$\int_0^\infty \frac{\sin x}{x}\, dx$$

exists as an improper Riemann integral, but the function $x \mapsto \sin x/x$ does not belong to $L^1([0,\infty), m)$.

Indeed, define

$$g(x) = \begin{cases} \dfrac{\sin x}{x}, & x \neq 0, \\ 1, & x = 0. \end{cases}$$

Then $g \in \mathcal{R}[0,1]$. For $1 \leq A < B$, integration by parts gives

$$\int_A^B \frac{\sin x}{x}\,dx = \left[-\frac{\cos x}{x}\right]_A^B - \int_A^B \frac{\cos x}{x^2}\,dx.$$

The boundary term has a finite limit as $B \to \infty$, and the second integral converges absolutely because

$$\int_A^\infty \frac{dx}{x^2} < \infty.$$

Hence $\int_1^\infty (\sin x/x)\,dx$ converges, so the improper integral exists on $[0,\infty)$.

On the other hand,

$$\int_{[0,\infty)} \left|\frac{\sin x}{x}\right| dm \geq \sum_{n=1}^{\infty} \int_{n\pi+\pi/6}^{n\pi+5\pi/6} \frac{|\sin x|}{x}\,dx \geq c \sum_{n=1}^{\infty} \frac{1}{n+1} = \infty$$

for a suitable constant $c > 0$, because $|\sin x| \geq 1/2$ on each interval $[n\pi + \pi/6, n\pi + 5\pi/6]$. Thus $g \notin L^1([0,\infty), m)$.

Example 7.104. The function $x \mapsto \sin x$ on $[0,\infty)$ is not improperly Riemann integrable, since

$$\int_0^b \sin x\,dx = 1 - \cos b$$

does not converge as $b \to \infty$.

7.9.3 Absolute convergence and Lebesgue integrability on unbounded intervals

On unbounded intervals, ordinary improper Riemann convergence is too weak to capture the measure-theoretic notion of integrability. Absolute convergence is the decisive strengthening: for locally Riemann integrable functions, it is precisely the condition that restores compatibility with the Lebesgue integral.

Theorem 7.105. *Let $f : [a,\infty) \to \mathbb{R}$ be such that $f \in \mathcal{R}[a,b]$ for every $b > a$. Then the following are equivalent:*

(i) $f \in L^1([a,\infty), m)$*;*

(ii) $|f|$ *is improperly Riemann integrable on* $[a,\infty)$*.*

In this case,

$$\int_{[a,\infty)} f\,dm = \lim_{b\to\infty} \int_a^b f(x)\,dx.$$

Proof. Assume first that $f \in L^1([a,\infty), m)$. Let (b_n) be any sequence with $b_n > a$ and $b_n \to \infty$, and set

$$f_n := f\chi_{[a,b_n]}.$$

Then $f_n \to f$ pointwise almost everywhere and $|f_n| \le |f| \in L^1([a,\infty), m)$. By the Dominated Convergence Theorem,

$$\int_{[a,\infty)} f\,dm = \lim_{n\to\infty} \int_{[a,\infty)} f_n\,dm.$$

By Theorem 7.99,

$$\int_{[a,\infty)} f_n\,dm = \int_a^{b_n} f(x)\,dx.$$

Since this holds for every sequence $b_n \to \infty$, the improper Riemann integral of f exists and equals the Lebesgue integral. Applying the same argument to $|f|$ shows that $|f|$ is improperly Riemann integrable.

Conversely, assume that $|f|$ is improperly Riemann integrable on $[a,\infty)$. For $n \in \mathbb{N}$ sufficiently large, set

$$g_n := |f|\chi_{[a,a+n]}.$$

Then $g_n \uparrow |f|$ pointwise. By Theorem 7.99,

$$\int g_n\,dm = \int_a^{a+n} |f(x)|\,dx.$$

The right-hand side is bounded above by the finite improper integral of $|f|$. Hence the Monotone Convergence Theorem yields

$$\int_{[a,\infty)} |f|\,dm = \lim_{n\to\infty} \int_a^{a+n} |f(x)|\,dx < \infty.$$

Thus $f \in L^1([a,\infty), m)$. The identity for $\int f\,dm$ follows from the first part of the proof. □

Corollary 7.106 (A useful boundedness criterion on $[a,\infty)$)**.** *Let $f : [a,\infty) \to \mathbb{R}$ be such that*

(i) $f \in \mathcal{R}[a,b]$ for every finite $b > a$;

(ii) there exists a constant $M < \infty$, independent of b, such that

$$\int_a^b |f(x)|\,dx \le M, \qquad b > a.$$

Then $f \in L^1([a,\infty), m)$ and

$$\int_{[a,\infty)} f\,dm = \lim_{b\to\infty} \int_{[a,b]} f\,dm = \lim_{b\to\infty} \int_a^b f(x)\,dx.$$

Proof. For $n \in \mathbb{N}$ with $n > a$, set $f_n := f\chi_{[a,n]}$. Since $|f_n| \uparrow |f|$ pointwise and

$$\int_{[a,\infty)} |f_n|\, dm = \int_a^n |f(x)|\, dx \le M,$$

the Monotone Convergence Theorem gives

$$\int_{[a,\infty)} |f|\, dm = \lim_{n\to\infty} \int_a^n |f(x)|\, dx \le M.$$

Thus $f \in L^1([a,\infty), m)$. Since $|f_n| \le |f|$ and $f_n \to f$ pointwise, the Dominated Convergence Theorem gives

$$\int_{[a,\infty)} f\, dm = \lim_{n\to\infty} \int_{[a,\infty)} f_n\, dm = \lim_{n\to\infty} \int_a^n f(x)\, dx.$$

The same conclusion for an arbitrary parameter $b \to \infty$ follows by applying the argument to every sequence $b_j \to \infty$. □

Example 7.107. For $f(x) = \dfrac{1}{1+x^2}$ on $\mathbb{R}$, one has

$$\int_{\mathbb{R}} f\, dm = \lim_{\substack{A\to-\infty \\ B\to\infty}} \int_A^B \frac{dx}{1+x^2} = \pi.$$

In particular, $f \in L^1(\mathbb{R}, m)$.

Theorem 7.108. *Let $f : (0,a] \to \mathbb{R}$ be such that $f \in \mathcal{R}[\varepsilon, a]$ for every $\varepsilon > 0$. If*

$$\sup_{0<\varepsilon<a} \int_\varepsilon^a |f(x)|\, dx < \infty,$$

then $f \in L^1([0,a], m)$ after redefining $f(0)$ arbitrarily, and

$$\int_{[0,a]} f\, dm = \lim_{\varepsilon\downarrow 0} \int_\varepsilon^a f(x)\, dx.$$

Proof. Define $f(0) := 0$ and, for $n \in \mathbb{N}$ with $1/n < a$, set

$$f_n := f\chi_{[1/n,a]}.$$

Then $f_n \to f$ pointwise on $[0,a]$ and $|f_n| \uparrow |f|$. By Theorem 7.99,

$$\int_{[0,a]} |f_n|\, dm = \int_{1/n}^a |f(x)|\, dx.$$

The hypothesis implies that the right-hand side is bounded, so the Monotone Convergence Theorem gives

$$\int_{[0,a]} |f|\, dm = \lim_{n\to\infty} \int_{1/n}^a |f(x)|\, dx < \infty.$$

Hence $f \in L^1([0,a], m)$. Since $|f_n| \le |f| \in L^1([0,a], m)$ and $f_n \to f$ pointwise, the Dominated Convergence Theorem yields

$$\int_{[0,a]} f\, dm = \lim_{n\to\infty} \int_{[0,a]} f_n\, dm = \lim_{n\to\infty} \int_{1/n}^{a} f(x)\, dx.$$

This proves the one-sided improper-integral formula. The full limit as $\varepsilon \downarrow 0$ follows by the same argument along arbitrary sequences $\varepsilon_n \downarrow 0$. □

7.9.4 L^p spaces

Definition 7.109. Let $(X, \mathcal{A}, \mu)$ be a measure space and let $1 \le p < \infty$. The space $L^p(X, \mu)$ consists of equivalence classes of measurable scalar-valued functions f on X such that

$$\int_X |f|^p\, d\mu < \infty,$$

where two functions are identified if they agree μ-almost everywhere. The L^p-norm is

$$\|f\|_p := \left(\int_X |f|^p\, d\mu \right)^{1/p}.$$

For $p = \infty$, the essential supremum norm is

$$\|f\|_\infty := \inf\{ M \ge 0 : \ |f| \le M \ \mu\text{-almost everywhere} \},$$

and $L^\infty(X, \mu)$ is the space of essentially bounded measurable functions modulo equality almost everywhere. In all cases the notation L^p refers to equivalence classes, although representatives are often written as ordinary functions when no confusion can arise.

Theorem 7.110 (Hölder's inequality)**.** *Let $1 < p < \infty$ and let q be the conjugate exponent, i.e. $\frac{1}{p} + \frac{1}{q} = 1$. If $f \in L^p(X, \mu)$ and $g \in L^q(X, \mu)$, then $fg \in L^1(X, \mu)$ and*

$$\int_X |fg|\, d\mu \le \|f\|_p \, \|g\|_q. \tag{7.8}$$

Proof. If $\|f\|_p = 0$ or $\|g\|_q = 0$, then $fg = 0$ almost everywhere and there is nothing to prove. For $x \in X$, Young's inequality applied to

$$a(x) := \frac{|f(x)|}{\|f\|_p}, \qquad b(x) := \frac{|g(x)|}{\|g\|_q}$$

yields

$$\frac{|f(x)g(x)|}{\|f\|_p \, \|g\|_q} \le \frac{|f(x)|^p}{p\, \|f\|_p^p} + \frac{|g(x)|^q}{q\, \|g\|_q^q}.$$

Integrating gives

$$\frac{1}{\|f\|_p \, \|g\|_q} \int_X |fg| \, d\mu \leq \frac{1}{p} \frac{\int_X |f|^p \, d\mu}{\|f\|_p^p} + \frac{1}{q} \frac{\int_X |g|^q \, d\mu}{\|g\|_q^q} = \frac{1}{p} + \frac{1}{q} = 1.$$

This proves (7.8). □

Theorem 7.111 (Minkowski's inequality)**.** *Let $1 \leq p \leq \infty$ and let $f, g \in L^p(X, \mu)$. Then $f + g \in L^p(X, \mu)$ and*

$$\|f + g\|_p \leq \|f\|_p + \|g\|_p. \tag{7.9}$$

In particular, for $1 \leq p \leq \infty$, the map $f \mapsto \|f\|_p$ defines a norm on $L^p(X, \mu)$.

Proof. For $p = \infty$, the pointwise estimate $|f + g| \leq |f| + |g|$ implies immediately that

$$\|f + g\|_\infty \leq \|f\|_\infty + \|g\|_\infty.$$

For $p = 1$, we have

$$\|f + g\|_1 = \int_X |f + g| \, d\mu \leq \int_X |f| \, d\mu + \int_X |g| \, d\mu = \|f\|_1 + \|g\|_1.$$

Now assume $1 < p < \infty$ and let $q = \frac{p}{p-1}$. If $\|f + g\|_p = 0$, there is nothing to prove. Otherwise,

$$\begin{aligned} \|f + g\|_p^p &= \int_X |f + g| \, |f + g|^{p-1} \, d\mu \\ &\leq \int_X |f| \, |f + g|^{p-1} \, d\mu + \int_X |g| \, |f + g|^{p-1} \, d\mu. \end{aligned}$$

By Hölder's inequality,

$$\int_X |f| \, |f + g|^{p-1} \, d\mu \leq \|f\|_p \Big(\int_X |f + g|^{(p-1)q} \, d\mu \Big)^{1/q} = \|f\|_p \, \|f + g\|_p^{p-1},$$

and similarly for g. Hence

$$\|f + g\|_p^p \leq (\|f\|_p + \|g\|_p) \, \|f + g\|_p^{p-1}.$$

Dividing by $\|f + g\|_p^{p-1}$ yields (7.9). The norm axioms now follow from the definition and Minkowski's inequality. □

Theorem 7.112 (Completeness of L^p)**.** *For each $1 \leq p \leq \infty$, the normed space $(L^p(X, \mu), \|\cdot\|_p)$ is complete.*

Proof. We treat first the case $1 \leq p < \infty$. Let (f_n) be a Cauchy sequence in $L^p(X, \mu)$. Choose a subsequence (f_{n_k}) such that

$$\|f_{n_{k+1}} - f_{n_k}\|_p < 2^{-k}, \qquad k \geq 1.$$

Set $g_k := |f_{n_{k+1}} - f_{n_k}|$ and define

$$G_N := \sum_{k=1}^{N} g_k, \qquad G := \sum_{k=1}^{\infty} g_k \quad \text{(pointwise, possibly } = \infty).$$

By Minkowski's inequality (7.9) applied to the nonnegative functions g_k,

$$\|G_N\|_p \le \sum_{k=1}^{N} \|g_k\|_p \le \sum_{k=1}^{\infty} 2^{-k} = 1.$$

Since $G_N \uparrow G$, the Monotone Convergence Theorem yields $\|G\|_p \le 1$, hence $G \in L^p(X,\mu)$ and $G(x) < \infty$ for μ-almost every x. For such x, the series $\sum_{k\ge1} g_k(x)$ converges, so $(f_{n_k}(x))$ is a Cauchy sequence in $\mathbb{R}$ and therefore converges to a finite limit. Define $f(x) := \lim_{k\to\infty} f_{n_k}(x)$ for μ-almost every x (and set $f(x) = 0$ on the remaining null set). Then f is measurable. Fix $K \ge 1$. For $k \ge K$, the telescoping identity gives

$$\begin{aligned} |f_{n_k}(x) - f_{n_K}(x)| &\le \sum_{j=K}^{k-1} g_j(x) \\ &\le \sum_{j=K}^{\infty} g_j(x) \\ &=: H_K(x), \end{aligned}$$

where $H_K \downarrow 0$ pointwise almost everywhere and $H_K \le G \in L^p$. Letting $k \to \infty$ and using $f_{n_k} \to f$ pointwise almost everywhere yields

$$|f(x) - f_{n_K}(x)| \le H_K(x) \qquad \text{for } \mu\text{-almost every } x.$$

By the Dominated Convergence Theorem (dominated by G^p), we obtain

$$\begin{aligned} \|f - f_{n_K}\|_p^p &= \int_X |f - f_{n_K}|^p \, d\mu \\ &\le \int_X H_K^p \, d\mu \xrightarrow[K\to\infty]{} 0. \end{aligned}$$

Hence $f_{n_K} \to f$ in L^p. Since (f_n) is Cauchy, it follows that $f_n \to f$ in L^p as well, proving completeness for $1 \le p < \infty$. For $p = \infty$, let (f_n) be Cauchy in L^∞ and choose a subsequence (f_{n_k}) with

$$\|f_{n_{k+1}} - f_{n_k}\|_\infty < 2^{-k}.$$

For each k there exists a measurable null set N_k such that

$$|f_{n_{k+1}}(x) - f_{n_k}(x)| \le 2^{-k} \qquad (x \in X \setminus N_k).$$

Set $N := \bigcup_{k\ge 1} N_k$, so $\mu(N) = 0$. Then for every $x \in X \setminus N$ and every $K < L$,

$$|f_{n_L}(x) - f_{n_K}(x)| \le \sum_{k=K}^{L-1} |f_{n_{k+1}}(x) - f_{n_k}(x)| \le \sum_{k=K}^{L-1} 2^{-k}.$$

Thus $(f_{n_k}(x))$ is Cauchy for every $x \in X \setminus N$, hence converges there to a finite limit. Define f to be this limit on $X \setminus N$ and set $f = 0$ on N. Then f is measurable.

Passing to the limit as $L \to \infty$ in the previous estimate yields, for $x \in X \setminus N$,

$$|f(x) - f_{n_K}(x)| \le \sum_{k=K}^{\infty} 2^{-k}.$$

Therefore

$$\|f - f_{n_K}\|_\infty \le \sum_{k=K}^{\infty} 2^{-k} \xrightarrow[K\to\infty]{} 0.$$

Hence $f_{n_K} \to f$ in L^∞, and the Cauchy property of (f_n) implies $f_n \to f$ in L^∞ as well. □

7.10 Further optional topics

The main chapter already provides a complete first course in measure and integration, but a few further results help indicate how quickly the subject expands. Each topic below points toward a substantial branch of analysis: finer modes of convergence, approximation by regular functions, product measures, and changes of measure.

7.10.1 Egoroff's theorem and Lusin's theorem (optional)

Theorem 7.113 (Egoroff)**.** *Let $(X, \mathcal{A}, \mu)$ be a measure space with $\mu(X) < \infty$. If (f_n) is a sequence of measurable functions such that $f_n \to f$ μ-almost everywhere, then for every $\varepsilon > 0$ there exists a measurable set $E \subseteq X$ with $\mu(X \setminus E) < \varepsilon$ such that $f_n \to f$ uniformly on E.*

Proof. Let

$$N := \{x \in X : f_n(x) \not\to f(x)\}.$$

By hypothesis, $\mu(N) = 0$. For $k, m \in \mathbb{N}$ define

$$E_{k,m} := \bigcup_{n\ge m} \Big\{x \in X : \ |f_n(x) - f(x)| > \frac{1}{k}\Big\}.$$

For each fixed k, the sets $E_{k,m}$ decrease as $m \to \infty$, and

$$\bigcap_{m=1}^{\infty} E_{k,m} \subset N.$$

Since $\mu(X) < \infty$, continuity from above gives $\mu(E_{k,m}) \downarrow 0$. Choose m_k such that

$$\mu(E_{k,m_k}) < \varepsilon 2^{-(k+1)}.$$

Set

$$E := X \setminus \left(N \cup \bigcup_{k=1}^{\infty} E_{k,m_k} \right).$$

Then

$$\mu(X \setminus E) \leq \mu(N) + \sum_{k=1}^{\infty} \mu(E_{k,m_k}) < \varepsilon.$$

If $x \in E$, then for every k and every $n \geq m_k$ one has

$$|f_n(x) - f(x)| \leq \frac{1}{k}.$$

Given $\eta > 0$, choose k with $1/k < \eta$. Then $|f_n(x) - f(x)| < \eta$ for every $x \in E$ and every $n \geq m_k$. Hence $f_n \to f$ uniformly on E. □

Theorem 7.114 (Lusin)**.** *Let $f\colon [a,b] \to \mathbb{R}$ be measurable and finite almost everywhere, and let $\varepsilon > 0$. Then there exists a continuous function $g\colon [a,b] \to \mathbb{R}$ and a measurable set $E \subseteq [a,b]$ such that $m([a,b] \setminus E) < \varepsilon$ and $f = g$ on E.*

Proof. We first treat the case in which f is bounded.

Approximation claim for simple functions. Let $s\colon [a,b] \to \mathbb{R}$ be a simple measurable function and let $\eta > 0$. Write

$$s = \sum_{j=1}^{N} c_j \chi_{A_j},$$

where the measurable sets $A_1, \ldots, A_N$ are pairwise disjoint and satisfy $\bigcup_{j=1}^{N} A_j = [a,b]$. For each j, inner regularity on the finite-measure interval $[a,b]$ yields a compact set $K_j \subset A_j$ such that

$$m(A_j \setminus K_j) < \frac{\eta}{N}.$$

Set

$$K := \bigcup_{j=1}^{N} K_j.$$

Then K is compact, and

$$m([a,b] \setminus K) \leq \sum_{j=1}^{N} m(A_j \setminus K_j) < \eta.$$

Define $h\colon K \to \mathbb{R}$ by $h|_{K_j} = c_j$. Because the sets K_j are pairwise disjoint and closed, the function h is continuous on the closed set K. A continuous function on a closed subset of the compact

interval $[a,b]$ extends continuously to all of $[a,b]$; using linear interpolation on the complementary intervals of K, one may choose an extension $\widetilde{h} \in C([a,b])$ with

$$\|\widetilde{h}\|_\infty \le \max_{1\le j\le N} |c_j| = \|s\|_\infty.$$

By construction, $\widetilde{h} = s$ on K.

Choose simple functions $(s_n)_{n\ge 1}$ such that

$$|s_n(x) - f(x)| \le 2^{-n} \qquad (x \in [a,b],\ n \ge 1).$$

This is possible because f is bounded and measurable, so the simple-function approximation theorem gives uniform approximation on bounded sets. Define

$$u_1 := s_1, \qquad u_n := s_n - s_{n-1} \quad (n \ge 2).$$

Then each u_n is a simple measurable function and

$$\|u_n\|_\infty \le 2^{-n} + 2^{-(n-1)} \le 2^{-(n-2)} \qquad (n \ge 2),$$

while $\|u_1\|_\infty \le \|f\|_\infty + 1$. Apply the approximation claim to u_n with $\eta = \varepsilon 2^{-(n+1)}$. We obtain continuous functions $v_n \in C([a,b])$ and compact sets $K_n \subset [a,b]$ such that

$$u_n = v_n \text{ on } K_n, \qquad m([a,b] \setminus K_n) < \varepsilon 2^{-(n+1)},$$

and

$$\|v_1\|_\infty \le \|f\|_\infty + 1, \qquad \|v_n\|_\infty \le 2^{-(n-2)} \quad (n \ge 2).$$

Hence the series $\sum_{n=1}^{\infty} v_n$ converges uniformly on $[a,b]$ by the Weierstrass M-test. Let

$$g := \sum_{n=1}^{\infty} v_n \in C([a,b]), \qquad E := \bigcap_{n=1}^{\infty} K_n.$$

Then

$$m([a,b] \setminus E) \le \sum_{n=1}^{\infty} m([a,b] \setminus K_n) < \sum_{n=1}^{\infty} \varepsilon 2^{-(n+1)} < \varepsilon.$$

If $x \in E$, then $v_n(x) = u_n(x)$ for every n, so for every N,

$$\sum_{n=1}^{N} v_n(x) = \sum_{n=1}^{N} u_n(x) = s_N(x).$$

Letting $N \to \infty$, the left-hand side converges to $g(x)$ and the right-hand side converges to $f(x)$. Thus $g(x) = f(x)$ for every $x \in E$. This proves Lusin's theorem for bounded measurable functions.

For the general case, define

$$A_M := \{x \in [a,b] : |f(x)| > M\} \qquad (M \in \mathbb{N}).$$

Since f is finite almost everywhere, the sets $(A_M)_{M\geq 1}$ decrease to the null set

$$N := \{x \in [a,b] : |f(x)| = \infty\}.$$

Because $m([a,b]) < \infty$, continuity from above gives $m(A_M) \downarrow m(N) = 0$. Choose M so large that

$$m(A_M) < \frac{\varepsilon}{2}.$$

Let

$$T_M(t) := \max\{-M, \min\{t, M\}\}, \qquad h := T_M \circ f.$$

Then h is bounded and measurable on $[a,b]$, and $h = f$ on $[a,b] \setminus A_M$. Applying the bounded case to h with $\varepsilon/2$ in place of ε, we obtain a continuous function $g\colon [a,b] \to \mathbb{R}$ and a measurable set $F \subset [a,b]$ such that

$$m([a,b] \setminus F) < \frac{\varepsilon}{2} \qquad \text{and} \qquad g = h \text{ on } F.$$

Set

$$E := F \cap ([a,b] \setminus A_M).$$

Then

$$m([a,b] \setminus E) \leq m([a,b] \setminus F) + m(A_M) < \varepsilon,$$

and on E we have $g = h = f$. This completes the proof. □

Remark 7.115. Lusin's theorem expresses the principle that measurable functions are "nearly continuous" on sets of arbitrarily large measure. Together with Egoroff's theorem, it forms one of the standard bridges between pointwise measurability and the approximation of measurable objects by classical continuous ones.

7.10.2 Product measures and Fubini–Tonelli (optional)

Definition 7.116 (Product σ-algebra)**.** Let $(X, \mathcal{A})$ and $(Y, \mathcal{B})$ be measurable spaces. The *product σ-algebra* $\mathcal{A} \otimes \mathcal{B}$ is the σ-algebra on $X \times Y$ generated by measurable rectangles $A \times B$ with $A \in \mathcal{A}$ and $B \in \mathcal{B}$.

Theorem 7.117 (Tonelli)**.** *Let $(X, \mathcal{A}, \mu)$ and $(Y, \mathcal{B}, \nu)$ be σ-finite measure spaces, and let $f\colon X \times Y \to [0, \infty]$ be measurable with respect to $\mathcal{A} \otimes \mathcal{B}$. Then the iterated integrals exist (possibly $= \infty$) and*

$$\int_{X\times Y} f \, d(\mu \times \nu) = \int_X \Big(\int_Y f(x,y)\, d\nu(y) \Big)\, d\mu(x) = \int_Y \Big(\int_X f(x,y)\, d\mu(x) \Big)\, d\nu(y).$$

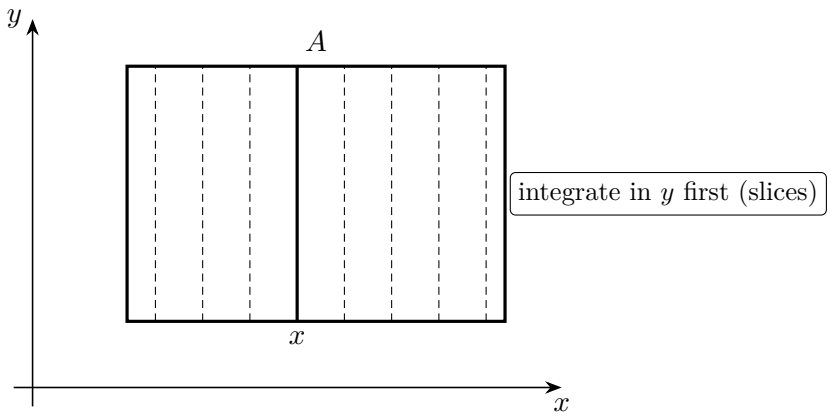


Figure 7.4: Slicing a planar region: intuition behind iterated integrals in Fubini's theorem.

Theorem 7.118 (Fubini)**.** *Under the hypotheses of Tonelli, if $f \in L^1(X \times Y, \mu \times \nu)$, then for μ-almost every x the function $y \mapsto f(x,y)$ is integrable on Y, for ν-almost every y the function $x \mapsto f(x,y)$ is integrable on X, and the equalities of iterated integrals in Theorem 7.117 hold with finite values.*

Remark 7.119. Both theorems are proved by first checking the statement for indicator functions of rectangles, extending to nonnegative simple functions, and then using monotone convergence (Tonelli) and dominated convergence applied to the positive and negative parts (Fubini).

7.10.3 Radon–Nikodym theorem (optional)

Theorem 7.120 (Radon–Nikodym)**.** *Let $(X, \mathcal{A}, \mu)$ be a σ-finite measure space and let ν be a σ-finite measure on $(X, \mathcal{A})$ that is absolutely continuous with respect to μ (denoted $\nu \ll \mu$). Then there exists a measurable function $h \geq 0$ such that*

$$\nu(A) = \int_A h \, d\mu, \qquad A \in \mathcal{A}.$$

The function h is unique up to equality μ-almost everywhere and is denoted $h = \frac{d\nu}{d\mu}$.

Part V

Problems and Revision

Problem Set 1: The Real Line, Sequences, and Countability

The first cluster consolidates foundational set-theoretic arguments; the second develops subsequences, liminf/limsup, and completeness; the final cluster emphasizes structural equivalences and explicit constructions.

A. Countability and cardinality

1. (Countability warm-up.) Show that $\mathbb{N} \times \mathbb{N}$ is countable. Deduce that $\mathbb{Z}$ and $\mathbb{Q}$ are countable.
2. Prove that a countable union of countable sets is countable. (In particular, $\mathbb{Q}^n$ is countable for each fixed $n \in \mathbb{N}$.)
3. (Cantor diagonal argument.) Prove that $(0, 1)$ is uncountable. Conclude that $\mathbb{R}$ is uncountable.

B. Sequences, liminf/limsup, and completeness

4. Let (x_n) be a bounded sequence of real numbers. Define

$$a_n := \inf_{k \geq n} x_k, \qquad b_n := \sup_{k \geq n} x_k.$$

   Show that (a_n) is increasing, (b_n) is decreasing, and that

$$\lim_{n \to \infty} a_n = \liminf_{n \to \infty} x_n, \qquad \lim_{n \to \infty} b_n = \limsup_{n \to \infty} x_n.$$

5. Prove that $\liminf x_n \leq \limsup x_n$ for every real sequence (x_n), and that $\lim x_n$ exists iff $\liminf x_n = \limsup x_n$.
6. (Bolzano–Weierstrass.) Prove that every bounded sequence in $\mathbb{R}$ has a convergent subsequence.
7. (Monotone convergence.) If (x_n) is monotone and bounded, prove that (x_n) converges and identify its limit using sup or inf.
8. Give an example of a sequence (x_n) with $\limsup x_n = 1$ and $\liminf x_n = -1$ but which does *not* converge.
9. Let (x_n) be a sequence in $\mathbb{R}$ such that $|x_{n+1} - x_n| \leq 2^{-n}$ for all n. Show that (x_n) is Cauchy and hence convergent.
10. Prove that every Cauchy sequence in $\mathbb{R}$ converges (completeness of $\mathbb{R}$).
11. (Subsequence selection.) Show that every real sequence has a monotone subsequence.

C. Enrichment and structural equivalences

12. (Algebraic numbers.) Show that the set of algebraic numbers is countable. Conclude that the set of transcendental numbers is uncountable.
13. Construct an explicit bijection between $(0, 1)$ and $\mathbb{R}$. (Hint: use a tangent or a rational map and verify it is bijective.)
14. (Completeness equivalences.) Show that the following are equivalent:
    (i) every nonempty set $S \subset \mathbb{R}$ that is bounded above has $\sup S \in \mathbb{R}$;
    (ii) every nested sequence of closed bounded intervals has nonempty intersection;
    (iii) every Cauchy sequence in $\mathbb{R}$ converges.

Problem Set 2: Metric Spaces and Topology

Begin with the language of open, closed, and boundary points; then move to testing candidate metrics and sequence behavior; finish with continuity and image-closure questions.

A. Topological vocabulary

1. (Topological basics.) Let (X, d) be a metric space and $A \subset X$. Prove that A is open iff for every $x \in A$ there exists $\varepsilon > 0$ such that $B_\varepsilon(x) \subset A$. Prove that A is closed iff it contains all of its limit points.
2. (Closure, interior, boundary.) For a subset $A \subset X$ define A°, $\overline{A}$, and ∂A. Show that $\partial A = \overline{A} \setminus A^\circ$ and that $\overline{A} = A \cup \partial A$.
3. (Bases.) Show that the family of open balls in a metric space forms a base for the topology (every open set is a union of open balls).

B. Metrics, openness, and sequences

4. For each of the following choices of X and d, decide whether d defines a metric on X.
   (a) $X = \mathbb{R}$ and $d(x, y) = \frac{|x-y|}{1+|xy|}$ for all $x, y \in \mathbb{R}$.
   (b) $X = \mathbb{R}$ and $d(x, y) = \min\{\sqrt{|x-y|}, |x-y|^2\}$ for all $x, y \in \mathbb{R}$.
   (c) $X = \mathbb{R}$ and $d(x, y) = |x-y|^p$ for all $x, y \in \mathbb{R}$ $(0 < p < 1)$.
   (d) $X = \mathbb{R}$ and

$$d(x,y) = \begin{cases} 1 + |x-y|, & \text{if exactly one of } x \text{ and } y \text{ is positive,} \\ |x-y|, & \text{otherwise,} \end{cases} \qquad x, y \in \mathbb{R}.$$

   (e) $X = \mathbb{R}^2$ and $d(x, y) = (|x_1 - y_1| + |x_2 - y_2|^{\frac{1}{2}})^{\frac{1}{2}}$ for all $x = (x_1, x_2), y = (y_1, y_2) \in \mathbb{R}^2$.
   (f) $X = \mathbb{R}^n$ and $d(x, y) = [(x_1-y_1)^2 + \frac{1}{2}(x_2-y_2)^2 + \cdots + \frac{1}{n}(x_n-y_n)^2]^{\frac{1}{2}}$ for all $x = (x_1, \ldots, x_n)$, $y = (y_1, \ldots, y_n) \in \mathbb{R}^n$.
   (g) $X = \mathbb{C}$ and

$$d(z,w) = \begin{cases} \min\{|z| + |w|,\, |z-1| + |w-1|\}, & \text{if } z \neq w, \\ 0, & \text{if } z = w, \end{cases} \qquad z, w \in \mathbb{C}.$$

(h) $X = \mathbb{C}$ and

$$d(z,w) = \begin{cases} |z-w|, & \text{if } \dfrac{z}{|z|} = \dfrac{w}{|w|}, \\ |z|+|w|, & \text{otherwise,} \end{cases} \qquad z, w \in \mathbb{C}.$$

(i) $X = \mathbb{C}$ and $d(z,w) = \frac{2|z-w|}{\sqrt{1+|z|^2}\sqrt{1+|w|^2}}$ for all $z, w \in \mathbb{C}$.

(j) $X =$ The class of all finite subsets of a nonempty set and $d(A,B) =$ The number of elements of the set $A \triangle B$ (the symmetric difference of A and B).

(k) $X = C[0,1]$ and $d(f,g) = (\int_0^1 |f(t) - g(t)|^2\, dt)^{\frac{1}{2}}$ for all $f, g \in C[0,1]$.

5. Determine whether the following sets are open and/or closed in $\mathbb{R}^2$ (with the usual metric).
   (a) $\{(x,y) \in \mathbb{R}^2 : xy > 0\}$
   (b) $\{(x,x) : x \in \mathbb{R}\}$
   (c) $(0,1) \times \{0\}$
   (d) $\{(x,y) \in \mathbb{R}^2 : 0 < x < y\}$
   (e) $\{(x,y) \in \mathbb{R}^2 : x + y < 1\}$
   (f) $\{(x,y) \in \mathbb{R}^2 : y \in \mathbb{Z}\}$
6. Let $A = \{(x,y,z) \in \mathbb{R}^3 : x^2 + y^2 < 1\}$ and $B = \{(x,y,z) \in \mathbb{R}^3 : z = 0\}$. Examine whether $A \cap B$ is a closed/an open subset of $\mathbb{R}^3$ with respect to the usual metric on $\mathbb{R}^3$.
7. Determine whether a finite subset of a metric space is open and/or closed.
8. For all $x, y \in \mathbb{R}$, let $d_1(x,y) = |x-y|$, $d_2(x,y) = \min\{1, |x-y|\}$ and $d_3(x,y) = \frac{|x-y|}{1+|x-y|}$. If G is an open set in any one of the three metric spaces $(\mathbb{R}, d_i)$ $(i = 1,2,3)$, show that G is also open in the other two metric spaces.
9. Let (x_n) and (y_n) be Cauchy sequences in a metric space (X,d). Show that the sequence $(d(x_n, y_n))$ is convergent.

C. Continuity and closure under maps

10. Let (x_n) be a sequence in a metric space X such that each of the subsequences (x_{2n}), (x_{2n-1}) and (x_{3n}) converges in X. Show that (x_n) converges in X.
11. Let F_n be a sequence of closed sets in $\mathbb{R}$ such that $F_n \subset (n, n+1]$ and $F_n \cap F_m = \emptyset$, whenever $m \neq n$. Show that $F = \bigcup_{n=1}^{\infty} F_n$ is a closed set in $\mathbb{R}$.
12. Let $f : (X,d) \to \mathbb{R}$ be a continuous function. Show that $\{x \in X : f(x) \neq 0\}$ is an open set in the metric space (X,d).
13. Let $f : X \to Y$ be continuous, where X and Y are metric spaces. If $x \in X$ is a limit point of $A \subset X$, then show that $f(x) \in f(A)$ or $f(x)$ is a limit point of $f(A)$.
14. Let $f : [0,1] \to \mathbb{R}$ be continuous. Show that $\overline{f(A)} \subset f(\overline{A})$ for all $A \subset [0,1]$.

Problem Set 3: Completeness, Compactness, and Fixed Points

The architecture mirrors the chapter itself: completion and Cauchy methods first, compactness next, and fixed-point arguments last.

A. Completeness and completion

1. (Completion.) Give an explicit construction of the completion of a metric space (X, d) via Cauchy sequences, and prove that the completion is unique up to isometry.
2. (Total boundedness.) Recall that X is totally bounded if for every $\varepsilon > 0$ there exist finitely many balls of radius ε covering X. Show that in a metric space, compactness implies total boundedness.
3. Show that a metric space is compact iff it is complete and totally bounded.
4. Let (x_n) be a sequence in a complete metric space (X, d) such that $\sum_{n=1}^{\infty} d(x_n, x_{n+1}) < \infty$. Show that (x_n) converges in (X, d).
5. Show that each of the following metric spaces is not complete.
   (a) $(\mathbb{N}, d)$, where $d(m, n) = |\frac{1}{m} - \frac{1}{n}|$ for all $m, n \in \mathbb{N}$
   (b) $((0, \infty), d)$, where $d(x, y) = |\frac{1}{x} - \frac{1}{y}|$ for all $x, y \in (0, \infty)$
   (c) $(\mathbb{R}, d)$, where $d(x, y) = |\frac{x}{1+|x|} - \frac{y}{1+|y|}|$ for all $x, y \in \mathbb{R}$
   (d) $(\mathbb{R}, d)$, where $d(x, y) = |e^x - e^y|$ for all $x, y \in \mathbb{R}$
6. Examine whether the following metric spaces are complete.
   (a) $([0, 1), d)$, where $d(x, y) = |\frac{x}{1-x} - \frac{y}{1-y}|$ for all $x, y \in [0, 1)$
   (b) $((-1, 1), d)$, where $d(x, y) = |\tan \frac{\pi x}{2} - \tan \frac{\pi y}{2}|$ for all $x, y \in (-1, 1)$
7. For $X(\neq \emptyset) \subset \mathbb{R}$, let $d(x, y) = \frac{|x-y|}{1+|x-y|}$ for all $x, y \in X$. Determine whether the metric space (X, d) is complete, where X is
   (a) $[0, 1] \cap \mathbb{Q}$.
   (b) $[-1, 0] \cup [1, \infty)$.
   (c) $\{n^2 : n \in \mathbb{N}\}$.

B. Compactness and its consequences

8. Let $\alpha > 0$ and let $f : (\mathbb{R}^n, \|\cdot\|_2) \to (\mathbb{R}^m, \|\cdot\|_2)$ be continuous such that $\|f(\mathbf{x}) - f(\mathbf{y})\|_2 \geq \alpha\|\mathbf{x} - \mathbf{y}\|_2$ for all $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$. Show that the range of f is a closed subset of $(\mathbb{R}^m, \|\cdot\|_2)$.

9. Let $A(\neq \emptyset) \subset \mathbb{R}^n$ be such that every continuous function $f : A \to \mathbb{R}$ is bounded. Show that A is a compact subset of $\mathbb{R}^n$.
10. Let $f : \mathbb{R}^2 \to \mathbb{R}$ be continuous such that $\lim\limits_{\|\mathbf{x}\|_2 \to \infty} f(\mathbf{x}) = 1$. Show that f is bounded on $\mathbb{R}^2$.
11. Let X, Y be metric spaces and let X be compact. If $f : X \to Y$ is a bijective continuous function, then show that $f^{-1} : Y \to X$ is continuous.
12. Let (X, d) be a compact metric space and let $f : X \to X$ be such that $d(f(x), f(y)) < d(x, y)$ for all $x, y \in X$ with $x \neq y$. Show that f has a fixed point in X.

C. Fixed points and applications

13. Let X be a complete metric space. Let $f : X \to X$ be a contraction and let $g : X \to X$ be such that $f \circ g = g \circ f$. Show that g has a fixed point in X. Is it necessary that the fixed point of g in X is unique? Justify.
14. Let (X, d) be a complete metric space and $f : X \to X$ be such that $f^m : X \to X$ is a contraction for some $m \in \mathbb{N}$. Show that f has a unique fixed point in X.
15. Let $f : \mathbb{R}^n \to \mathbb{R}^n$ be a contraction and $g(\mathbf{x}) = \mathbf{x} - f(\mathbf{x})$ for all $\mathbf{x} \in \mathbb{R}^n$. Show that $g : \mathbb{R}^n \to \mathbb{R}^n$ is injective and surjective. Also, show that both g and $g^{-1} : \mathbb{R}^n \to \mathbb{R}^n$ are continuous.
16. Using the contraction mapping theorem, show that the equation $4x^5 - 2x^2 - 4x + 1 = 0$ has exactly one root in $\left(0, \frac{1}{2}\right)$.
17. Using the contraction mapping theorem, show that the initial value problem $\frac{dy}{dx} = x + \sin(x^2 y)$ for $x \in [0, 1]$, $y(0) = \frac{1}{2}$, has a unique solution.
18. Let $a_1 = 1$ and $a_{n+1} = \sqrt[3]{1 + a_n^2}$. Using fixed point theory, prove that the sequence $\{a_n\}$ is convergent, and that its limit satisfies the equation $x^3 - x^2 - 1 = 0$.
19. Using fixed point theory, determine all functions $f \in C[0, 1]$ such that $f(x) = \int_0^x (x - y) f(y) dy$.
20. Let $B = \{\mathbf{x} \in \mathbb{R}^n : \|\mathbf{x}\|_2 < 1\}$ and let $f : (B, d_2) \to (B, d_2)$ be continuous such that $\|f(\mathbf{x})\|_2 < \|\mathbf{x}\|_2$ for all $\mathbf{x} \in B \setminus \{\mathbf{0}\}$. Let $\mathbf{x}_1 \in B \setminus \{\mathbf{0}\}$ and let $\mathbf{x}_{k+1} = f(\mathbf{x}_k)$ for all $k \in \mathbb{N}$. Show that $\mathbf{x}_k \to \mathbf{0}$.

Problem Set 4: Normed Linear Spaces and Fundamental Inequalities

This set moves from norm axioms and model spaces to comparison estimates for ℓ^p-type norms, and then to topological and completeness questions for normed spaces.

A. Norms, Banach spaces, and operators

1. (Young $\Rightarrow$ Hölder $\Rightarrow$ Minkowski.) Prove Young's inequality: if $p, q > 1$ with $\frac{1}{p} + \frac{1}{q} = 1$, then for all $a, b \geq 0$,
$$ab \leq \frac{a^p}{p} + \frac{b^q}{q}.$$
Deduce Hölder's inequality and then Minkowski's inequality for $\|\cdot\|_p$ on $\mathbb{R}^n$.

2. Let $1 \leq p \leq \infty$. Prove that $(\ell^p, \|\cdot\|_p)$ is a Banach space.

3. Let $T : \mathbb{R}^n \to \mathbb{R}^m$ be linear. Show that T is continuous iff it is bounded, and compute the operator norm of the matrix $A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ acting on $(\mathbb{R}^2, \|\cdot\|_2)$.

4. Let X be a nonzero normed vector space. Show that $\{x \in X : \|x\| < 1\}$ is not closed in X and $\{x \in X : \|x\| \leq 1\}$ is not open in X.

5. Let X be a normed vector space and let $Y (\neq X)$ be a subspace of X. Show that Y is not open in X.

6. For each of the following definitions, determine whether $\|\cdot\|$ is a norm on $\mathbb{R}^2$, where for each $(x, y) \in \mathbb{R}^2$,
   (a) $\|(x, y)\| = (\sqrt{|x|} + \sqrt{|y|})^2$.
   (b) $\|(x, y)\| = \sqrt{\frac{x^2}{9} + \frac{y^2}{4}}$.
   (c) $\|(x, y)\| = \begin{cases} \sqrt{x^2 + y^2} & \text{if } xy \geq 0, \\ \max\{|x|, |y|\} & \text{if } xy < 0. \end{cases}$

7. Determine whether $\|\cdot\|$ is a norm on $C[0, 1]$, where for each $f \in C[0, 1]$,
   (a) $\|f\| = \min\{\|f\|_\infty, 2\|f\|_1\}$. (b) $\|f\| = \sup\{t\,|f(t)| : t \in [0, 1]\}$.

8. Let X be a vector space over $\mathbb{R}$ and let $p : X \to [0, \infty)$ satisfy the following three conditions:
   (i) $p(x) = 0$ iff $x = 0$ $(x \in X)$.
   (ii) $p(\alpha x) = |\alpha| p(x)$ for all $x \in X$ and for all $\alpha \in \mathbb{R}$.
   (iii) $\{x \in X : p(x) \leq 1\}$ is a convex subset of X.
   Show that p is a norm on X.
   (A subset S of X is called convex if $(1 - t)x + ty \in S$ for all $x, y \in S$ and for all $t \in [0, 1]$.)

9. Let d be a metric on a real vector space X satisfying the following two conditions:
   (i) $d(x+z, y+z) = d(x,y)$ for all $x, y, z \in X$,
   (ii) $d(\alpha x, \alpha y) = |\alpha| d(x,y)$ for all $x, y \in X$ and for all $\alpha \in \mathbb{R}$.
   Show that there exists a norm $\|\cdot\|$ on X such that $d(x,y) = \|x-y\|$ for all $x, y \in X$.
10. Let X be a normed vector space and let $x \in X$. Show that $\|x\| = \inf\left\{\frac{1}{|\alpha|} : \alpha \in \mathbb{R} \setminus \{0\},\ \|\alpha x\| \leq 1\right\}$.

B. Norm inequalities and p-norm asymptotics

11. Let X be a normed vector space and let $x, y \in X$. Show that $\|x\| \leq \max\{\|x+y\|, \|x-y\|\}$.
12. Show that
   (a) if $\mathbf{x} \in \mathbb{R}^n$, then $\lim_{p\to\infty} \|\mathbf{x}\|_p = \|\mathbf{x}\|_\infty$.
   (b) if $x \in \ell^q$ for some $1 \leq q < \infty$, then $\lim_{p\to\infty} \|x\|_p = \|x\|_\infty$.
   (c) if $f \in C[a,b]$, then $\lim_{p\to\infty} \|f\|_p = \|f\|_\infty$.
13. If $1 \leq p < q \leq \infty$, then show that $\|x\|_q \leq \|x\|_p$ for all $x \in \ell^p$.

C. Topology and completeness in normed spaces

14. Let $(X, \|\cdot\|)$ be a nonzero normed vector space. Consider the metrics d_1, d_2 and d_3 on X:

$$d_1(x,y) := \min\{1, \|x-y\|\},$$
$$d_2(x,y) := \frac{\|x-y\|}{1+\|x-y\|},$$
$$d_3(x,y) := \begin{cases} 1+\|x-y\| & \text{if } x \neq y, \\ 0 & \text{if } x = y, \end{cases}$$

   for all $x, y \in X$. Prove that none of d_1, d_2 and d_3 is induced by any norm on X.
15. Let $\mathbb{R}^\infty$ be the real vector space of all sequences in $\mathbb{R}$, where addition and scalar multiplication are defined componentwise. Let $d((x_n),(y_n)) = \sum_{n=1}^{\infty} \frac{1}{2^n} \cdot \frac{|x_n - y_n|}{1+|x_n-y_n|}$ for all $(x_n), (y_n) \in \mathbb{R}^\infty$. Show that d is a metric on $\mathbb{R}^\infty$ but that no norm on $\mathbb{R}^\infty$ induces d.
16. Let X be a normed vector space containing more than one point, let $x, y \in X$ and let $\varepsilon, \delta > 0$. If $B_\varepsilon[x] = B_\delta[y]$, show that $x = y$ and $\varepsilon = \delta$. Does the result remain true if X is assumed to be a metric space? Justify.
17. Show that $A = \{f \in C[0,1] : \|f\|_1 < 1\}$ is an unbounded subset of the normed linear space $(C[0,1], \|\cdot\|_\infty)$.
18. For $f \in C^1[0,1]$, define $\|f\| = \|f\|_1 + \|f\|_\infty$. Determine whether $(C^1[0,1], \|\cdot\|)$ is a complete normed linear space.

Problem Set 5: Sequences of Functions and Uniform Convergence

Work first with concrete convergence tests, then with the core preservation theorems, and finally with stability questions for algebraic operations and series.

A. Concrete convergence models

1. Determine the pointwise and uniform limits (if any) of the following sequences on the indicated domains:
   (a) $f_n(x) = x^n$ on $[0,1]$;
   (b) $g_n(x) = \frac{x}{1+nx}$ on $[0,1]$;
   (c) $h_n(x) = \frac{\sin(nx)}{n}$ on $\mathbb{R}$.

B. Core theorems and compactness

2. (Weierstrass M-test.) Let (f_n) be functions on X and suppose $|f_n(x)| \le M_n$ for all $x \in X$, where $\sum_{n=1}^{\infty} M_n < \infty$. Prove that $\sum_{n=1}^{\infty} f_n$ converges uniformly and absolutely on X.
3. Let $f_n \to f$ uniformly on $[a,b]$ and suppose each f_n is Riemann integrable. Prove that $\int_a^b f_n \to \int_a^b f$ and that f is Riemann integrable.
4. (Term-by-term differentiation.) Suppose (f_n) are differentiable on $[a,b]$, $f_n(x_0)$ converges for some $x_0 \in [a,b]$, and $f_n' \to g$ uniformly on $[a,b]$. Prove that $f_n \to f$ uniformly on $[a,b]$ for some differentiable f and that $f' = g$.
5. (Arzelà–Ascoli, compact version.) Let $\mathcal{F} \subset C([0,1])$ be uniformly bounded and equicontinuous. Prove that every sequence in $\mathcal{F}$ has a uniformly convergent subsequence.

C. Uniform convergence in practice

6. Determine whether the sequence (f_n) converges in $(C[0,1], d_\infty)$, where for all $n \in \mathbb{N}$ and for all $t \in [0,1]$,
   (a) $f_n(t) = \frac{nt^2}{1+nt}$.
   (b) $f_n(t) = 1 + t + \frac{t^2}{2!} + \cdots + \frac{t^n}{n!}$.
   (c)
$$f_n(t) = \begin{cases} nt, & \text{if } 0 \le t \le \dfrac{1}{n}, \\ \dfrac{1}{nt}, & \text{if } \dfrac{1}{n} < t \le 1. \end{cases}$$

(d)

$$f_n(t) = \begin{cases} nt, & \text{if } 0 \le t \le \dfrac{1}{n}, \\ \dfrac{n}{n-1}(1-t), & \text{if } \dfrac{1}{n} < t \le 1. \end{cases}$$

7. Let $f : [0,1] \to \mathbb{R}$ be continuous and $f(1) = 0$. If $f_n(x) = f(x)x^n$ for all $x \in [0,1]$ and for all $n \in \mathbb{N}$, then examine whether the sequence (f_n) converges uniformly on $[0,1]$.
8. Let X be a metric space and let $f, g, f_n, g_n : X \to \mathbb{R}$ $(n \in \mathbb{N})$ be such that $f_n \to f$ uniformly on X and $g_n \to g$ uniformly on X.
   (a) Show that $f_n + g_n \to f + g$ uniformly on X.
   (b) Is it necessary that $f_n.g_n \to f.g$ uniformly on X ? Justify.
   (c) If f and g are bounded on X, then show that $f_n.g_n \to f.g$ uniformly on X.
9. Let X be a metric space and let (f_n) be a sequence of real-valued bounded functions on X. If $f : X \to \mathbb{R}$ is such that $f_n \to f$ uniformly on X, then show that f is bounded on X. Does this result hold if $f_n \to f$ pointwise on X ? Justify.
10. Let (f_n) be a uniformly convergent sequence of real-valued bounded functions defined on a metric space X. Show that there exists $M > 0$ such that $|f_n(x)| \le M$ for all $x \in X$ and for all $n \in \mathbb{N}$.
11. Determine whether the series $\sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{(2n-1)!} x^{2n-1}$ is uniformly convergent in $\mathbb{R}$.

Problem Set 6: Functions of Several Variables

The problem flow starts with diagnostics on limits and continuity, advances through differentiability and the chain rule, and culminates in local invertibility, implicit maps, and Taylor expansion.

A. Diagnostics: limits, continuity, and first principles

1. For each of the following statements, determine whether it is TRUE or FALSE, and provide a complete justification.
   (a) There exists a continuous function $f : \mathbb{R} \to \mathbb{R}^2$ such that $f(\cos n) = (n, \frac{1}{n})$ for all $n \in \mathbb{N}$.
   (b) There exists a non-constant continuous function $f : \mathbb{R}^2 \to \mathbb{R}$ such that $f(x, y) = 5$ for all $(x, y) \in \mathbb{R}^2$ with $x^2 + y^2 < 1$.
   (c) There exists a one-to-one continuous function from $\{(x, y) \in \mathbb{R}^2 : x^2 + y^2 < 1\}$ onto $\mathbb{R}^2$.
   (d) There exists a continuous function from $\{(x, y) \in \mathbb{R}^2 : x^2 + y^2 \leq 1\}$ onto $\mathbb{R}^2$.
   (e) If $f : \mathbb{R}^2 \to \mathbb{R}$ is continuous such that $f_x(0, 0)$ exists, then $f_y(0, 0)$ must exist.
   (f) There exists a function $f : \mathbb{R}^2 \to \mathbb{R}^2$ which is differentiable only at $(1, 0)$.
   (g) Let $f : \mathbb{R}^2 \to \mathbb{R}$ be such that $f_x(0, 0) = 0$. Then there exists some $\delta > 0$ such that $f(x, 0)$ is continuous on $(-\delta, \delta)$.
   (h) If $f : \mathbb{R}^2 \to \mathbb{R}$ is continuous such that all the directional derivatives of f at $(0, 0)$ exist, then f must be differentiable at $(0, 0)$.
   (i) If $f : \mathbb{R}^2 \to \mathbb{R}^2$ is differentiable with $f(0, 0) = (1, 1)$ and $[f'(0, 0)] = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$, then there cannot exist a differentiable function $g : \mathbb{R}^2 \to \mathbb{R}^2$ with $g(1, 1) = (0, 0)$ and $(f \circ g)(x, y) = (y, x)$ for all $(x, y) \in \mathbb{R}^2$.
   (j) A continuously differentiable function $f : \mathbb{R}^2 \to \mathbb{R}^2$ cannot be one-to-one and onto if $\det[f'(x, y)] = 0$ for some $(x, y) \in \mathbb{R}^2$.
   (k) If $f : \mathbb{R}^2 \to \mathbb{R}^2$ is a differentiable function with a differentiable inverse $f^{-1} : \mathbb{R}^2 \to \mathbb{R}^2$, then $f'(0) : \mathbb{R}^2 \to \mathbb{R}^2$ must be one-to-one and onto.
   (l) The equation $\sin(xyz) = z$ defines x implicitly as a differentiable function of y and z locally around the point $(x, y, z) = (\frac{\pi}{2}, 1, 1)$.
   (m) The equation $\sin(xyz) = z$ defines z implicitly as a differentiable function of x and y locally around the point $(x, y, z) = (\frac{\pi}{2}, 1, 1)$.

2. Let $\alpha \in (0,1)$ and let $\mathbf{x}_n = (n^3\alpha^n, \frac{1}{n}[n\alpha])$ for all $n \in \mathbb{N}$. (For each $x \in \mathbb{R}$, $[x]$ denotes the greatest integer not exceeding x.) Determine whether the sequence $(\mathbf{x}_n)$ converges in $\mathbb{R}^2$. Also, find $\lim_{n\to\infty} \mathbf{x}_n$ if it exists.
3. Determine whether each of the following limits exists; if it does, compute its value.

   (a) $\lim_{(x,y)\to(0,0)} \frac{x^3y}{x^4+y^2}$, (b) $\lim_{(x,y)\to(0,0)} \frac{x^3-y^3}{x^2+y^2}$, (c) $\lim_{(x,y)\to(0,0)} \frac{x^2y^2}{x^2y^2+(x^2-y^2)^2}$,

   (d) $\lim_{(x,y)\to(0,0)} \frac{|x|}{y^2}e^{-|x|/y^2}$, (e) $\lim_{(x,y)\to(0,0)} \frac{1-\cos(x^2+y^2)}{(x^2+y^2)^2}$, (f) $\lim_{(x,y)\to(0,0)} \frac{\sqrt{x^2y^2+1}-1}{x^2+y^2}$.

4. Study the continuity of $f : \mathbb{R}^2 \to \mathbb{R}$ at $(0,0)$, where for all $(x,y) \in \mathbb{R}^2$,

   (a)
   $$f(x,y) = \begin{cases} xy, & \text{if } xy \geq 0, \\ -xy, & \text{if } xy < 0. \end{cases}$$

   (b)
   $$f(x,y) = \begin{cases} 1, & \text{if } x > 0 \text{ and } 0 < y < x^2, \\ 0, & \text{otherwise.} \end{cases}$$

   (c)
   $$f(x,y) = \begin{cases} \frac{x^3}{x^2+y^2}, & \text{if } (x,y) \neq (0,0), \\ 0, & \text{if } (x,y) = (0,0). \end{cases}$$

5. Determine all the points of $\mathbb{R}^2$ where $f : \mathbb{R}^2 \to \mathbb{R}$ is continuous, if for all $(x,y) \in \mathbb{R}^2$,

   (a) $f(x,y) = \begin{cases} \frac{xy}{x-y} & \text{if } x \neq y, \\ 0 & \text{if } x = y. \end{cases}$

   (b) $f(x,y) = \begin{cases} \frac{xy}{x^2-y^2} & \text{if } x^2 \neq y^2, \\ 0 & \text{if } x^2 = y^2. \end{cases}$

   (c) $f(x,y) = \begin{cases} xy & \text{if } xy \in \mathbb{Q}, \\ -xy & \text{if } xy \in \mathbb{R} \setminus \mathbb{Q}. \end{cases}$

6. Let Ω be an open subset of $\mathbb{R}^n$ and let $f : \Omega \to \mathbb{R}^m$ and $g : \Omega \to \mathbb{R}^m$ be continuous at $\mathbf{x}_0 \in \Omega$. If for each $\varepsilon > 0$, there exist $\mathbf{x}, \mathbf{y} \in B_\varepsilon(\mathbf{x}_0)$ such that $f(\mathbf{x}) = g(\mathbf{y})$, then show that $f(\mathbf{x}_0) = g(\mathbf{x}_0)$.
7. Let $A(\neq \emptyset) \subset \mathbb{R}^n$ be such that every continuous function $f : A \to \mathbb{R}$ is bounded. Show that A is a closed and bounded subset of $\mathbb{R}^n$.
8. Let $f : \mathbb{R}^2 \to \mathbb{R}$ be continuous at $(x_0, y_0) \in \mathbb{R}^2$ and let $f(x_0, y_0) \neq 0$. Show that there exists $\delta > 0$ such that $f(x,y) \neq 0$ for all $(x,y) \in \mathbb{R}^2$ satisfying $(x-x_0)^2 + (y-y_0)^2 < \delta$.
9. Let $T : \mathbb{R}^n \to \mathbb{R}^n$ be linear and let $f(\mathbf{x}) = T(\mathbf{x}) \cdot \mathbf{x}$ for all $\mathbf{x} \in \mathbb{R}^n$. Find $f'(\mathbf{x})$, where $\mathbf{x} \in \mathbb{R}^n$.

B. Differentiability and directional behavior

10. Study the differentiability of f at $\mathbf{0}$, where
    (a) $f : \mathbb{R}^n \to \mathbb{R}$ satisfies $|f(\mathbf{x})| \leq \|\mathbf{x}\|_2^2$ for all $\mathbf{x} \in \mathbb{R}^n$.
    (b) $f : \mathbb{R}^2 \to \mathbb{R}$ is defined by $f(x, y) = \sqrt{|xy|}$ for all $(x, y) \in \mathbb{R}^2$.
    (c) $f : \mathbb{R}^2 \to \mathbb{R}$ is defined by $f(x, y) = ||x| - |y|| - |x| - |y|$ for all $(x, y) \in \mathbb{R}^2$.
    (d) $f : \mathbb{R}^2 \to \mathbb{R}$ is defined by $f(x, y) = \begin{cases} \frac{y}{|y|}\sqrt{x^2 + y^2} & \text{if } y \neq 0, \\ 0 & \text{if } y = 0. \end{cases}$
    (e) $f : \mathbb{R}^2 \to \mathbb{R}$ is defined by $f(x, y) = \begin{cases} \frac{x^3}{x^2+y^2} & \text{if } (x, y) \neq (0, 0), \\ 0 & \text{if } (x, y) = (0, 0). \end{cases}$
    (f) $f : \mathbb{R}^2 \to \mathbb{R}$ is defined by $f(x, y) = \begin{cases} \frac{\sin(x^2y^2)}{x^2+y^2} & \text{if } (x, y) \neq (0, 0), \\ 0 & \text{if } (x, y) = (0, 0). \end{cases}$
    (g) $f : \mathbb{R}^2 \to \mathbb{R}^2$ is defined by $f(x, y) = \begin{cases} (\sin^2 x + x^2 \sin \frac{1}{x}, y^2) & \text{if } x \neq 0, \\ (0, y^2) & \text{if } x = 0. \end{cases}$
    (h) $f : \mathbb{R}^n \to \mathbb{R}$ is defined by $f(\mathbf{x}) = \|\mathbf{x}\|_2$ for all $\mathbf{x} \in \mathbb{R}^n$.
    (i) $f : \mathbb{R}^n \to \mathbb{R}^n$ is defined by $f(\mathbf{x}) = \|\mathbf{x}\|_2\mathbf{x}$ for all $\mathbf{x} \in \mathbb{R}^n$.
11. Let $f(x, y) = (x^2 + y^3, x^3 + y^2, 2x^2y^2)$ for all $(x, y) \in \mathbb{R}^2$. Examine whether $f : \mathbb{R}^2 \to \mathbb{R}^3$ is differentiable at $(1, 2)$ and find $f'(1, 2)$ if f is differentiable at $(1, 2)$.
12. Determine all the points of $\mathbb{R}^2$ where $f : \mathbb{R}^2 \to \mathbb{R}$ is differentiable, if for all $(x, y) \in \mathbb{R}^2$,
    (a) $f(x, y) = \begin{cases} x^2 + y^2 & \text{if both } x, y \in \mathbb{Q}, \\ 0 & \text{otherwise.} \end{cases}$
    (b) $f(x, y) = \begin{cases} x^{4/3} \sin(\frac{y}{x}) & \text{if } x \neq 0, \\ 0 & \text{if } x = 0. \end{cases}$
13. Let $f : \mathbb{R}^2 \to \mathbb{R}$ be defined by $f(x, y) = \begin{cases} (x^2 + y^2) \sin(\frac{1}{\sqrt{x^2+y^2}}) & \text{if } (x, y) \neq (0, 0), \\ 0 & \text{if } (x, y) = (0, 0). \end{cases}$
    Show that f is differentiable at $(0, 0)$ although neither f_x nor f_y is continuous at $(0, 0)$.
14. Let $f : \mathbb{R}^2 \to \mathbb{R}$ be defined by $f(x, y) = \begin{cases} \frac{x^2y(x-y)}{x^2+y^2} & \text{if } (x, y) \neq (0, 0), \\ 0 & \text{if } (x, y) = (0, 0). \end{cases}$
    Examine whether $f_{xy}(0, 0) = f_{yx}(0, 0)$.
15. Let $f : \mathbb{R}^2 \to \mathbb{R}$ be defined by $f(x, y) = \begin{cases} \frac{xy(x^2-y^2)}{x^2+y^2} & \text{if } (x, y) \neq (0, 0), \\ 0 & \text{if } (x, y) = (0, 0). \end{cases}$
    Determine all the points of $\mathbb{R}^2$ where f_{xy} and f_{yx} are continuous.
16. Let Ω be a nonempty open subset of $\mathbb{R}^n$. Let $f : \Omega \to \mathbb{R}$ be differentiable at $\mathbf{x}_0 \in \Omega$, let $f(\mathbf{x}_0) = 0$ and let $g : \Omega \to \mathbb{R}$ be continuous at $\mathbf{x}_0$. Prove that $fg : \Omega \to \mathbb{R}$, defined by $(fg)(\mathbf{x}) = f(\mathbf{x})g(\mathbf{x})$ for all $\mathbf{x} \in \Omega$, is differentiable at $\mathbf{x}_0$.
17. Let Ω be a nonempty open subset of $\mathbb{R}^n$ and let $g : \Omega \to \mathbb{R}^n$ be continuous at $\mathbf{x}_0 \in \Omega$. If $f : \Omega \to \mathbb{R}$ is such that $f(\mathbf{x}) - f(\mathbf{x}_0) = g(\mathbf{x}) \cdot (\mathbf{x} - \mathbf{x}_0)$ for all $\mathbf{x} \in \Omega$, then show that f is differentiable at $\mathbf{x}_0$.

18. The directional derivatives of a differentiable function $f : \mathbb{R}^2 \to \mathbb{R}$ at $(0,0)$ in the directions of $(1,2)$ and $(2,1)$ are 1 and 2 respectively. Find $f_x(0,0)$ and $f_y(0,0)$.
19. Let $A \in GL(\mathbb{R}^n)$ and $\alpha \geq 2$. If $f : \mathbb{R}^n \to \mathbb{R}^n$ satisfies $\|f(x)\| \leq k\|x\|^\alpha$, for some $k > 0$. Prove or disprove that the map $g = f + A$ is continuously differentiable at $\mathbf{0}$ and g is invertible in the neighborhood of $\mathbf{0}$.
20. Find all $\mathbf{v} \in \mathbb{R}^2$ for which the directional derivative $f'_{\mathbf{v}}(0,0)$ exists, where for all $(x,y) \in \mathbb{R}^2$,
    (a) $f(x,y) = \sqrt{|x^2 - y^2|}$.
    (b) $f(x,y) = \begin{cases} \frac{xy}{x^2+y^2} & \text{if } (x,y) \neq (0,0), \\ 0 & \text{if } (x,y) = (0,0). \end{cases}$
    (c) $f(x,y) = \begin{cases} \frac{x}{y} & \text{if } y \neq 0, \\ 0 & \text{if } y = 0. \end{cases}$
    (d) $f(x,y) = ||x| - |y|| - |x| - |y|$.

C. Chain rule, homogeneity, and linear approximation

21. Let $f(x,y,z) = (x^3y + y^2z, xyz)$ and $g(x,y) = (x^2y, xy, x - 2y, x^2 + 3y)$ for all $x, y, z \in \mathbb{R}$. Using the chain rule, compute $(g \circ f)'(\mathbf{a})$, where $\mathbf{a} = (1, 2, -3)$.
22. Let $f : \mathbb{R}^2 \to \mathbb{R}$ be differentiable such that $f(1,1) = 1$, $f_x(1,1) = 2$ and $f_y(1,1) = 5$. If $g(x) = f(x, f(x,x))$ for all $x \in \mathbb{R}$, determine $g'(1)$.
23. Let $\varphi : \mathbb{R} \to \mathbb{R}$ and $\psi : \mathbb{R} \to \mathbb{R}$ be differentiable. Show that $f : \mathbb{R}^2 \to \mathbb{R}$, defined by $f(x,y) = \varphi(x) + \psi(y)$ for all $(x,y) \in \mathbb{R}^2$, is differentiable.
24. Prove that a differentiable function $f : \mathbb{R}^n \setminus \{\mathbf{0}\} \to \mathbb{R}^m$ is homogeneous of degree $\alpha \in \mathbb{R}$ (i.e., $f(t\mathbf{x}) = t^\alpha f(\mathbf{x})$ for all $\mathbf{x} \in \mathbb{R}^n \setminus \{\mathbf{0}\}$ and for all $t > 0$) if and only if $f'(\mathbf{x})(\mathbf{x}) = \alpha f(\mathbf{x})$ for all $\mathbf{x} \in \mathbb{R}^n \setminus \{\mathbf{0}\}$.
25. Suppose $f : \mathbb{R}^n \to \mathbb{R}^n$ is satisfying $f(rx) = r^{\frac{3}{2}} f(x)$ for all $(x,r) \in \mathbb{R}^n \times (0,\infty)$. Is f differentiable at $\mathbf{0}$?
26. Let $\mathbf{x}_0 \in \mathbb{R}^n$ and $r > 0$. If $L : \mathbb{R}^n \to \mathbb{R}^n$ is a linear function such that $L(\mathbf{x}) = \mathbf{0}$ for all $\mathbf{x} \in \mathbb{R}^n$ satisfying $\|\mathbf{x} - \mathbf{x}_0\|_2 < r$, then show that $L(\mathbf{x}) = \mathbf{0}$ for all $\mathbf{x} \in \mathbb{R}^n$.
27. Let $\mathbf{x}_0 \in \mathbb{R}^n$ and $r > 0$. If $L : \mathbb{R}^n \to \mathbb{R}^m$ is a linear function such that $L(\mathbf{x}) = \mathbf{0}$ for all $\mathbf{x} \in \mathbb{R}^n$ satisfying $\|\mathbf{x} - \mathbf{x}_0\|_2 = r$, then show that $L(\mathbf{x}) = \mathbf{0}$ for all $\mathbf{x} \in \mathbb{R}^n$.
28. Let $f : \mathbb{R}^n \to \mathbb{R}^m$ be differentiable at $\mathbf{0}$ and let $f(\alpha\mathbf{x}) = \alpha f(\mathbf{x})$ for all $\mathbf{x} \in \mathbb{R}^n$ and for all $\alpha \in \mathbb{R}$. Show that f is a linear function.
29. Let $f(x,y) = (x + y^2, xy)$ for all $(x,y) \in \mathbb{R}^2$. Using the definition of differentiability, show that $f : \mathbb{R}^2 \to \mathbb{R}^2$ is differentiable and also find $f'(x_0, y_0)$, where $(x_0, y_0) \in \mathbb{R}^2$.

D. Mean-value arguments and local invertibility

30. Let $f(x,y) = (x^5 + 1, y^5 + y + 1)$ for all $(x,y) \in \mathbb{R}^2$. Show that $f : \mathbb{R}^2 \to \mathbb{R}^2$ is invertible. Also, show that $f^{-1} : \mathbb{R}^2 \to \mathbb{R}^2$ is differentiable at $(0,1)$ and find $(f^{-1})'(0,1)$.

31. Let $T : \mathbb{R}^n \to \mathbb{R}^n$ be linear and let $f(\mathbf{x}) = T(\mathbf{x}) \cdot \mathbf{x}$ for all $\mathbf{x} \in \mathbb{R}^n$. Find $f'(\mathbf{x})$, where $\mathbf{x} \in \mathbb{R}^n$.
32. Let $L : \mathbb{R}^n \to \mathbb{R}^n$ be a linear function and $f(\mathbf{x}) = \langle L(\mathbf{x}), \mathbf{x}\rangle \mathbf{x}$ for all $\mathbf{x} \in \mathbb{R}^n$. Find $f'(\mathbf{x})$, where $\mathbf{x} \in \mathbb{R}^n$.
33. Let $f : \mathbb{R}^2 \to \mathbb{R}$ be continuously differentiable such that $f_x(a, b) = f_y(a, b)$ for all $(a, b) \in \mathbb{R}^2$ and $f(a, 0) > 0$ for all $a \in \mathbb{R}$. Show that $f(a, b) > 0$ for all $(a, b) \in \mathbb{R}^2$.
34. Let Ω be an open convex subset of $\mathbb{R}^n$. Let $L : \mathbb{R}^n \to \mathbb{R}^m$ be a linear function and let $f : \Omega \to \mathbb{R}^m$ be differentiable with $f'(\mathbf{x}) = L$ for all $\mathbf{x} \in \Omega$. Show that there exists $\mathbf{v} \in \mathbb{R}^m$ such that $f(\mathbf{x}) = L(\mathbf{x}) + \mathbf{v}$ for all $\mathbf{x} \in \Omega$.
35. Let Ω be an open subset of $\mathbb{R}^n$ such that $\mathbf{a}, \mathbf{b} \in \Omega$ and $S = \{(1 - t)\mathbf{a} + t\mathbf{b} : t \in [0, 1]\} \subset \Omega$. If $f : \Omega \to \mathbb{R}^m$ is differentiable at each point of S, then show that there exists a linear map $L : \mathbb{R}^n \to \mathbb{R}^m$ such that $f(\mathbf{b}) - f(\mathbf{a}) = L(\mathbf{b} - \mathbf{a})$.
36. Let $\alpha > 0$ and let $f : \mathbb{R}^2 \to \mathbb{R}^2$ satisfy $\|f(\mathbf{x}) - f(\mathbf{y})\|_2 \leq \alpha\|\mathbf{x} - \mathbf{y}\|_2^2$ for all $\mathbf{x}, \mathbf{y} \in \mathbb{R}^2$. Show that f is a constant function.
37. Let $f : \mathbb{R}^n \to \mathbb{R}^m$ be differentiable and let $\mathbf{y} \in \mathbb{R}^m$. If $g(\mathbf{x}) = f(\mathbf{x}) \cdot \mathbf{y}$ for all $\mathbf{x} \in \mathbb{R}^n$, then show that $g : \mathbb{R}^n \to \mathbb{R}$ is differentiable.
38. Let $\Omega \subset \mathbb{R}^n$ and let $\mathbf{x}_0 \in \Omega^0$. If $f : \Omega \to \mathbb{R}$ is differentiable at $\mathbf{x}_0$ and $f(\mathbf{x}) \leq f(\mathbf{x}_0)$ for all $\mathbf{x} \in \Omega$, then determine (with justification) $f'(\mathbf{x}_0)$.
39. Let $f : \mathbb{R}^n \to \mathbb{R}^m$ be differentiable at $\mathbf{0}$ and $f(\mathbf{0}) = \mathbf{0}$. Show that there exist $\alpha > 0$ and $r > 0$ such that $\|f(\mathbf{x})\|_2 \leq \alpha\|\mathbf{x}\|_2$ for all $\mathbf{x} \in \mathbb{R}^n$ with $\|\mathbf{x}\|_2 < r$.
40. Let $f(x, y) = (2ye^{2x}, xe^y)$ for all $(x, y) \in \mathbb{R}^2$. Show that there exist open sets U and V in $\mathbb{R}^2$ containing $(0, 1)$ and $(2, 0)$ respectively such that $f : U \to V$ is one-to-one and onto.
41. Determine all the points of $\mathbb{R}^2$ where $f : \mathbb{R}^2 \to \mathbb{R}^2$ is locally invertible, if for all $(x, y) \in \mathbb{R}^2$,
    (a) $f(x, y) = (x^2 + y^2, xy)$.
    (b) $f(x, y) = (x^2 + xy + y^2, xy)$.
    (c) $f(x, y) = (\cos x + \cos y, \sin x + \sin y)$.
42. Determine all the points of $\mathbb{R}^3$ where $f : \mathbb{R}^3 \to \mathbb{R}^3$ is locally invertible, if for all $(x, y, z) \in \mathbb{R}^3$,
    (a) $f(x, y, z) = (x + y, xy + z, y + z)$.
    (b) $f(x, y, z) = (x - xy, xy - xyz, xyz)$.
    (c) $f(x, y, z) = (x^2 + y^2, y^2 + z^2, z^2 + x^2)$.
43. Let $f(x, y) = (3x - y^2, 2x + y, xy + y^3)$ and $g(x, y) = (2ye^{2x}, xe^y)$ for all $(x, y) \in \mathbb{R}^2$. Examine whether $(f \circ g^{-1})'(2, 0)$ exists (with a meaningful interpretation of g^{-1}) and find $(f \circ g^{-1})'(2, 0)$ if it exists.
44. For $n \geq 2$, let $B = \{\mathbf{x} \in \mathbb{R}^n : \|\mathbf{x}\|_2 < 1\}$ and let $f(\mathbf{x}) = \|\mathbf{x}\|_2^2\mathbf{x}$ for all $\mathbf{x} \in B$. Show that $f : B \to B$ is differentiable and invertible but that $f^{-1} : B \to B$ is not differentiable at $\mathbf{0}$.

E. Implicit-function problems

45. Using implicit function theorem, show that the system of equations

$$x^3(y^3+z^3)=0,$$
$$(x-y)^3-z^2=7,$$

can be solved locally near the point $(1,-1,1)$ for y and z as a differentiable function of x.

46. Show that the system of equations

$$x^2-y\cos(uv)+z^2=0,$$
$$x^2+y^2-\sin(uv)+2z^2=2,$$
$$xy-\sin u\cos v+z=0,$$

implicitly defines (x,y,z) as a differentiable function of (u,v) near $x=1$, $y=1$, $z=0$, $u=\frac{\pi}{2}$ and $v=0$.

47. Using implicit function theorem, show that in a neighbourhood of any point $(x_0,y_0,u_0,v_0)\in\mathbb{R}^4$ which satisfies the equations

$$x-e^u\cos v=0,$$
$$v-e^y\sin x=0,$$

there exists a unique solution $(u,v)=\varphi(x,y)$ satisfying $\det[\varphi'(x,y)]=v/x$.

48. Show that in a neighbourhood of any point $(x_0,y_0,z_0)\in\mathbb{R}^3$ which satisfies the equations

$$x^4+(x+z)y^3-3=0,$$
$$x^4+(2x+3z)y^3-6=0,$$

there is a unique continuous solution $y=\varphi_1(x)$, $z=\varphi_2(x)$ of these equations.

49. Show that around the point $(0,1,1)$, the equation $xy-z\log y+e^{xz}=1$ can be solved locally as $y=f(x,z)$ but cannot be solved locally as $z=g(x,y)$.

50. Show that the system of equations

$$3x+y-z+u^2=0,$$
$$x-y+2z+u=0,$$
$$2x+2y-3z+2u=0,$$

can be solved locally for x,y,u in terms of z; for x,z,u in terms of y; for y,z,u in terms of x, but not for x,y,z in terms of u.

51. Show that the system of equations

$$x^2 + y^2 - u^2 - v = 0,$$
$$x^2 + 2y^2 + 3u^2 + 4v^2 = 1,$$

defines (u, v) implicitly as a differentiable function of (x, y) locally around the point $(x, y, u, v) = (\frac{1}{2}, 0, \frac{1}{2}, 0)$ but does not define (x, y) implicitly as a differentiable function of (u, v) locally around the same point.

52. Show that there are points $(x, y, z, u, v, w) \in \mathbb{R}^6$ which satisfy the equations

$$x^2 + u + e^v = 0,$$
$$y^2 + v + e^w = 0,$$
$$z^2 + w + e^u = 0.$$

Prove that in a neighbourhood of such a point there exist unique differentiable solutions $u = \varphi_1(x, y, z)$, $v = \varphi_2(x, y, z)$, $w = \varphi_3(x, y, z)$. If $\varphi = (\varphi_1, \varphi_2, \varphi_3)$, find $\varphi'(x, y, z)$.

F. Taylor expansion and synthesis

53. Find the 3rd order Taylor polynomial of $f(x, y, z) = x^2y + z$ about the point $(1, 2, 1)$.

54. Find the 4th order Taylor polynomial of $g(x, y) = e^{x-2y}/(1 + x^2 - y)$ about the point $(0, 0)$.

55. Let $f : \mathbb{R}^2 \to \mathbb{R}$ be continuously differentiable. Show that f is not one-to-one.

Problem Set 7: Measure and Integration

The arrangement proceeds from outer measure and null sets to measurability criteria, then to measurable functions, and finally to concrete integral computations and convergence questions.

A. Outer measure, null sets, and first consequences

1. State TRUE or FALSE giving proper justification for each of the following statements.
   (a) There exists an unbounded subset A of $\mathbb{R}$ such that $m^*(A) = 5$.
   (b) If $A \subset \mathbb{R}$ such that $m^*(A) = 0$, then there must exist a sequence (x_n) in $\mathbb{R} \setminus A$ such that $x_n \to 1$ as $n \to \infty$.
   (c) If A is an uncountable subset of $\mathbb{R}$ such that $A \cap (a, b) \neq \emptyset$ for all $a, b \in \mathbb{R}$ with $a < b$, then it is necessary that $m^*(A) > 0$.
   (d) There exists an open subset A of $\mathbb{R}$ such that $[\frac{1}{2}, \frac{3}{4}] \subset A$ and $m(A) = \frac{1}{4}$.
   (e) There exists an open subset A of $\mathbb{R}$ such that $m(A) < \frac{1}{5}$ but $A \cap (a, b) \neq \emptyset$ for all $a, b \in \mathbb{R}$ with $a < b$.
   (f) If A and B are open subsets of $\mathbb{R}$ such that $A \subsetneq B$, then it is necessary that $m(A) < m(B)$.
   (g) A subset E of $\mathbb{R}$ is Lebesgue measurable iff $m^*(A \cup B) = m^*(A) + m^*(B)$ for each $A \subset E$ and for each $B \subset \mathbb{R} \setminus E$.
   (h) If $f : \mathbb{R} \to \mathbb{R}$ is continuous a.e. on $\mathbb{R}$, then there must exist a continuous function $g : \mathbb{R} \to \mathbb{R}$ such that $f = g$ a.e. on $\mathbb{R}$.
   (i) If $g : \mathbb{R} \to \mathbb{R}$ is continuous and if $f : \mathbb{R} \to \mathbb{R}$ is such that $f = g$ a.e. on $\mathbb{R}$, then f must be continuous a.e. on $\mathbb{R}$.
   (j) If $f : \mathbb{R} \to \mathbb{R}$ and $g : \mathbb{R} \to \mathbb{R}$ are continuous such that $f = g$ a.e. on $\mathbb{R}$, then it is necessary that $f(x) = g(x)$ for all $x \in \mathbb{R}$.
2. Let $f : [0, 2) \to \mathbb{R}$ be defined by $f(x) = \begin{cases} x^2 & \text{if } 0 \leq x \leq 1, \\ 3 - x & \text{if } 1 < x < 2. \end{cases}$
   Find $m^*(A)$, where $A = f^{-1}((\frac{9}{16}, \frac{5}{4})) = \{x \in [0, 2) : f(x) \in (\frac{9}{16}, \frac{5}{4})\}$.
3. Let $B \subset A \subset \mathbb{R}$ such that $m^*(B) = 0$. Show that $m^*(A \setminus B) = m^*(A)$.
4. Let $A, B, C \subset \mathbb{R}$ such that $m^*(A \triangle B) = 0$ and $m^*(B \triangle C) = 0$. Show that $m^*(A \triangle C) = 0$.
5. Let $A \subset \mathbb{R}$ such that $m^*(A) > 0$. Show that there exists $B \subset A$ such that B is bounded and $m^*(B) > 0$.
6. If $A \subset \mathbb{R}$, then show that $m^*(A) = \inf\{m(G) : A \subset G,\ G \text{ is an open set in } \mathbb{R}\}$.

7. Let $E = \{x \in [0,1] :$ The decimal representation of x does not contain the digit $5\}$. Show that $m(E) = 0$.
8. Let $A_n \subset \mathbb{R}$ for $n = 1, 2, \ldots$ such that $\sum_{n=1}^{\infty} m^*(A_n) < \infty$.
   If $E = \{x \in \mathbb{R} : x \in A_n$ for infinitely many $n\}$, then show that $m(E) = 0$.
9. If G is a nonempty open subset of $\mathbb{R}$, then show that $m(G) > 0$.

B. Measurable sets and structural criteria

10. Show that a subset E of $\mathbb{R}$ is Lebesgue measurable iff $m^*(I) = m^*(I \cap E) + m^*(I \setminus E)$ for every bounded open interval I of $\mathbb{R}$.
11. Let $A \subset E \subset B \subset \mathbb{R}$ such that A, B are Lebesgue measurable and $m(A) = m(B) < \infty$. Show that E is Lebesgue measurable.
    More generally, let $A \subset B \subset \mathbb{R}$ such that A is Lebesgue measurable and $m^*(B) = m(A) < \infty$. Show that B is Lebesgue measurable.
12. Let $A, B \subset \mathbb{R}$ such that $m^*(A) = 0$ and $A \cup B$ is Lebesgue measurable. Show that B is Lebesgue measurable.
13. Let $A, B \subset \mathbb{R}$ such that A is Lebesgue measurable and $m^*(A \triangle B) = 0$. Show that B is Lebesgue measurable.
14. Let $A \subset \mathbb{R}$ such that $A \cap B$ is Lebesgue measurable for every bounded subset B of $\mathbb{R}$. Show that A is Lebesgue measurable.
15. If E is a Lebesgue measurable subset of $\mathbb{R}$ and if $x \in \mathbb{R}$, then show that $E + x$ is Lebesgue measurable.
16. Let A be a countable subset of $\mathbb{R}$ and let $B \subset \mathbb{R}$ such that $m^*(B) = 0$. Show that $m^*(A+B) = 0$.
17. If E and F are Lebesgue measurable subsets of $\mathbb{R}$, then show that $m(E \cup F) + m(E \cap F) = m(E) + m(F)$.
    More generally, let E be a Lebesgue measurable subset of $\mathbb{R}$ and let $A \subset \mathbb{R}$. Show that $m^*(E \cap A) + m^*(E \cup A) = m^*(E) + m^*(A)$.
18. Let I and J be disjoint open intervals in $\mathbb{R}$ and let $A \subset I$, $B \subset J$. Show that $m^*(A \cup B) = m^*(A) + m^*(B)$.
19. Let $A \subset [0,1]$ be Lebesgue measurable with $m(A) = 1$. If $B \subset [0,1]$, then show that $m^*(A \cap B) = m^*(B)$.

C. Further consequences of measurability

20. Let $E_i \subset (0,1)$ $(i = 1, \ldots, n)$ be Lebesgue measurable sets such that $\sum_{i=1}^{n} m(E_i) > n - 1$. Show that $m(\cap_{i=1}^{n} E_i) > 0$.

21. If $A \subset \mathbb{R}$, then show that there exists a Lebesgue measurable subset E of $\mathbb{R}$ such that $m^*(A) = m(E)$.
22. Let $A \subset \mathbb{R}$ such that $m^*(A) > 0$. Show that there exist $x, y \in A$ such that $x - y \in \mathbb{R} \setminus \mathbb{Q}$.
23. Let A and B be Lebesgue measurable subsets of $(0, 1)$ such that $m(A) > \frac{1}{2}$ and $m(B) > \frac{1}{2}$. Prove that there exist $a \in A$ and $b \in B$ such that $a + b = 1$.
24. Let A be an unbounded Lebesgue measurable subset of $\mathbb{R}$ such that $m(A) < \infty$. Show that for each $\varepsilon > 0$, there exists a bounded Lebesgue measurable set B in $\mathbb{R}$ such that $B \subset A$ and $m(A \setminus B) < \varepsilon$.
25. Show that the Borel σ-algebra on $\mathbb{R}$ is generated by the class $\{(-\infty, x] : x \in \mathbb{Q}\}$.
26. Let $A \subset \mathbb{R}$ such that $m^*(A) = 0$. Show that $m^*(\{x^2 : x \in A\}) = 0$.
27. If A is subset of $[1, 2]$ such that $m^*(A) = 0$, then Show that $\{x + x^2 : x \in A\}$ is measurable.
28. For $n \in \mathbb{N}$, let $A = \bigcup_{n=1}^{\infty} \left[n, n + \frac{1}{n^{3/2}}\right]$. Show that $m(A) < \infty$ and $m(A^2) = \infty$, where $A^2 = \{x^2 : x \in A\}$.
29. Let $A, B \subset \mathbb{R}$ such that $A \cup B$ is Lebesgue measurable and $m(A \cup B) = m^*(A) + m^*(B) < \infty$. Show that both A and B are Lebesgue measurable.

D. Measurable functions

30. Let E be a Lebesgue measurable subset of $\mathbb{R}$. Show that $f : E \to \mathbb{R}$ is Lebesgue measurable if and only if $\{x \in E : f(x) > r\}$ is Lebesgue measurable for each $r \in \mathbb{Q}$.
31. Let E be a Lebesgue measurable subset of $\mathbb{R}$ and let $f : E \to \mathbb{R}$ be a Lebesgue measurable function. For each $x \in E$, let $g(x) = \begin{cases} f(x) & \text{if } |f(x)| \leq 5, \\ 0 & \text{if } |f(x)| > 5. \end{cases}$
Show that $g : E \to \mathbb{R}$ is Lebesgue measurable.
32. Let E be a Lebesgue measurable subset of $\mathbb{R}$ and let $f : E \to \mathbb{R}$ be a Lebesgue measurable function. For each $x \in E$, let $g(x) = \begin{cases} 0 & \text{if } f(x) \in \mathbb{Q}, \\ 1 & \text{if } f(x) \in \mathbb{R} \setminus \mathbb{Q}. \end{cases}$
Show that $g : E \to \mathbb{R}$ is Lebesgue measurable.
33. Let E be a Lebesgue measurable subset of $\mathbb{R}$ and let $f : E \to \mathbb{R}$ be a Lebesgue measurable function. For each $x \in E$, let $g(x) = \begin{cases} -2 & \text{if } f(x) < -2, \\ f(x) & \text{if } -2 \leq f(x) \leq 3, \\ 3 & \text{if } f(x) > 3. \end{cases}$
Show that $g : E \to \mathbb{R}$ is Lebesgue measurable.
34. Let E be a Lebesgue measurable subset of $\mathbb{R}$ and let $f : E \to \mathbb{R}$ be a Lebesgue measurable function. If $g : \mathbb{R} \to \mathbb{R}$ is continuous, then show that $g \circ f$ is Lebesgue measurable.
35. Does there exist a continuous function $f : \mathbb{R} \to \mathbb{R}$ such that $f = \chi_{[0,1]}$ a.e. on $\mathbb{R}$? Justify.

36. Let E be a Lebesgue measurable subset of $\mathbb{R}$ and let $f : E \to \mathbb{R}$, $g : E \to \mathbb{R}$ be Lebesgue measurable functions. If G is an open subset of $\mathbb{R}^2$, then show that $\{x \in E : (f(x), g(x)) \in G\}$ is a Lebesgue measurable subset of $\mathbb{R}$.
37. Let $f : \mathbb{R} \to [0, \infty]$ be such that $m^*\{x \in \mathbb{R} : f(x) \geq 2^n\} < \frac{1}{2^n}$, whenever $n \in \mathbb{N}$. Show that $\{x \in \mathbb{R} : f(x) = \infty\}$ is Lebesgue measurable.
38. Let $f : [a, b] \to \mathbb{R}$ be a differentiable function. Show that $f' : [a, b] \to \mathbb{R}$ is Lebesgue measurable.

E. Integration and explicit computations

39. Let $\varphi : (\mathbb{R}, \mathcal{M}, m) \to [0, \infty]$ be a Lebesgue measurable simple function. Define a set function $\nu : M \to [0, \infty]$ by $\nu(E) = \int\limits_E \varphi dm$. Show that $\nu(\bigcup\limits_{n=1}^{\infty} E_n) = \sum\limits_{n=1}^{\infty} \nu(E_n)$, whenever E_n is a sequence of pairwise disjoint sets in M.
40. For each $x \in [0, 1]$, let $f(x) = \begin{cases} \frac{1}{n} & \text{if } x = \frac{m}{n} \text{ for some } m, n \in \mathbb{N} \text{ with } \gcd(m, n) = 1, \\ 0 & \text{otherwise.} \end{cases}$
Evaluate the Lebesgue integral $\int\limits_{[0,1]} f$.
41. For each $x \in [0, 1]$, let $f(x) = \begin{cases} x^2 & \text{if } x = \frac{1}{2^n} \text{ for some } n \in \mathbb{N}, \\ x^3 & \text{if } x = \frac{1}{3^n} \text{ for some } n \in \mathbb{N}, \\ x^4 & \text{otherwise.} \end{cases}$
Evaluate the Lebesgue integral $\int\limits_{[0,1]} f$.
42. Let $f(x) = \begin{cases} \sin(\pi x) & \text{if } x \in [0, \frac{1}{2}] \setminus C, \\ \cos(\pi x) & \text{if } x \in (\frac{1}{2}, 1] \setminus C, \\ x^2 & \text{if } x \in C. \end{cases}$
(C denotes the Cantor set.) Evaluate the Lebesgue integral $\int\limits_{[0,1]} f$.
43. Evaluate the Lebesgue integral $\int\limits_{[0,\infty)} e^{-[x]}\, dx$.
44. Let $f(x) = \begin{cases} e^{[x]} & \text{if } x \in \mathbb{Q}, \\ e^{-[x]} & \text{if } x \in \mathbb{R} \setminus \mathbb{Q}. \end{cases}$
Evaluate the Lebesgue integral $\int\limits_{(0,\infty)} f$.
45. Let $f(x) = \begin{cases} e^{|x|} & \text{if } x \in \mathbb{Q}, \\ e^{-|x|} & \text{if } x \in \mathbb{R} \setminus \mathbb{Q}. \end{cases}$
Evaluate the Lebesgue integral $\int\limits_{\mathbb{R}} f$.
46. Evaluate the Lebesgue integral $\int\limits_{(0,1]} \frac{1}{\sqrt[3]{x}}\, dx$.

47. Let $f(x) = \begin{cases} \frac{1}{\sqrt{x}} & \text{if } 0 < x \leq 1, \\ \frac{1}{x} & \text{if } x > 1. \end{cases}$
    Evaluate the Lebesgue integral $\int\limits_{(0,\infty)} f$.

48. Evaluate the following:
    (a) $\lim\limits_{n\to\infty} \int\limits_{-2}^{2} \frac{x^{2n}}{1+x^{2n}}\, dx$
    (b) $\lim\limits_{n\to\infty} \int\limits_{[0,1]} \frac{1+nx}{(1+x)^n}\, dx$
    (c) $\int\limits_0^1 (\sum\limits_{n=1}^{\infty} \frac{x^n}{n})\, dx$
    (d) $\lim\limits_{n\to\infty} \int\limits_1^{\infty} \frac{1}{1+x^{2n}}\, dx$
    (e) $\sum\limits_{n=0}^{\infty} \int\limits_0^1 \frac{x^2}{(1+x^2)^n}\, dx$

Selected References

[1] T. M. Apostol, *Mathematical Analysis*, 2nd ed., Addison–Wesley, 1974.

[2] R. G. Bartle and D. R. Sherbert, *Introduction to Real Analysis*, 4th ed., Wiley, 2011.

[3] R. C. Buck, *Advanced Calculus*, Waveland Press, 2003.

[4] M. Capinski and E. Kopp, *Measure, Integral and Probability*, 2nd ed., Springer, 2007.

[5] N. L. Carothers, *Real Analysis*, Cambridge University Press, 2000.

[6] G. de Barra, *Measure Theory and Integration*, New Age International, 1981.

[7] G. B. Folland, *Real Analysis: Modern Techniques and Their Applications*, 2nd ed., Wiley, 1999.

[8] J. E. Marsden and M. J. Hoffman, *Elementary Classical Analysis*, 2nd ed., W. H. Freeman, 1993.

[9] H. L. Royden and P. M. Fitzpatrick, *Real Analysis*, 4th ed., Pearson, 2010.

[10] W. Rudin, *Principles of Mathematical Analysis*, 3rd ed., McGraw–Hill, 1976.

[11] E. M. Stein and R. Shakarchi, *Real Analysis: Measure Theory, Integration, and Hilbert Spaces*, Princeton University Press, 2005.