Measure and Integration

Lecture Notes

Master M1 — 2025–2026

Yaë Ulrich Gaba

*"One must know when to doubt,
when to submit, and when to believe."*
— Blaise Pascal

March 25, 2026

Contents

Preface

Objectives and motivation

Measure theory and integration form one of the fundamental pillars of modern mathematical analysis. Born from the pioneering work of Émile Borel and Henri Lebesgue at the beginning of the 20th century, this theory profoundly transformed our understanding of the concepts of length, area, volume, and more generally the "size" of a set. The Lebesgue integral, which arises from it, considerably extends the Riemann integral and provides the natural framework for probability theory, functional analysis, partial differential equations, and many other branches of mathematics.

This course is aimed at graduate students (Master's level) in pure or applied mathematics. It assumes a solid command of real analysis (sequences, series, continuity, differentiability, Riemann integral) as well as basic notions of general topology. Familiarity with linear algebra and normed vector spaces is also desirable.

Historical context

The Riemann integral, developed in the 1850s, has well-known limitations. Consider the Dirichlet function:

$$f(x) = \mathbf{1}_{\mathbb{Q}}(x) = \begin{cases} 1 & \text{if } x \in \mathbb{Q}, \\ 0 & \text{if } x \notin \mathbb{Q}. \end{cases}$$

This function is not Riemann-integrable, yet intuitively, since $\mathbb{Q}$ is "negligible" in $\mathbb{R}$, one would expect $\int_0^1 f = 0$.

Similarly, if (f_n) is a sequence of Riemann-integrable functions converging pointwise to a function f, it is generally not true that

$$\lim_{n\to\infty} \int_a^b f_n(x)\,dx = \int_a^b f(x)\,dx,$$

even when the right-hand side makes sense. For the interchange of limit and integral to be valid in the Riemann sense, one needs *uniform* convergence, which is a very restrictive condition.

The Lebesgue integral solves both problems elegantly: the Dirichlet function becomes integrable (with integral zero), and the dominated convergence theorem allows the interchange of limit and integral under much weaker hypotheses than uniform convergence.

Guiding ideas

Lebesgue's approach rests on a fundamentally different idea from Riemann's. While Riemann partitions the *domain* of the function into small intervals, Lebesgue partitions the *range* of the function and measures the preimages:

> **Lebesgue's philosophy**
>
> Imagine you need to count a large sum of money composed of coins of different values. Riemann's method takes the coins one by one, in the order they are found. Lebesgue's method first sorts the coins by value, then counts how many there are of each type. The result is the same, but the second method is more flexible and applies to more situations.

For this approach to work, one must be able to "measure" the preimages $f^{-1}([a, b))$. This requires:

1. A theory of **measurable sets** (σ-algebras), which determines which sets can be measured.
2. A theory of **measures**, which assigns a "size" to each of these sets.
3. A definition of **measurable functions**, whose preimages of reasonable sets are measurable.
4. The construction of the **integral** itself, by approximation via step functions (*simple functions*).

Course outline

The course is organized into twelve chapters following this logical progression:

Chapter 1. σ-Algebras and measurable spaces: the fundamental structures. Generated σ-algebras, Borel σ-algebra.

Chapter 2. Measures: definitions, examples, σ-additivity, finite and σ-finite measures, continuity lemma.

Chapter 3. Carathéodory's theorem: construction of measures via outer measures, extension from a pre-measure to a complete measure.

Chapter 4. Lebesgue measure on $\mathbb{R}^n$: construction, invariance properties, non-measurable sets.

Chapter 5. Measurable functions: definitions, stability under algebraic operations and limits, approximation by simple functions.

Chapter 6. Lebesgue integral: three-step construction (simple functions, non-negative functions, integrable functions).

Chapter 7. Convergence theorems: monotone convergence (Beppo Levi), Fatou's lemma, dominated convergence (Lebesgue).

Conventions and notation

- $\mathbb{N} = \{0, 1, 2, \dots\}$ denotes the set of natural numbers (including zero).
- $\mathbb{R}^+ = [0, +\infty)$ and $\overline{\mathbb{R}}^+ = [0, +\infty]$.
- The extended real line is $\overline{\mathbb{R}} = [-\infty, +\infty]$.
- $\mathbf{1}_A$ denotes the indicator (characteristic) function of A.
- $f^+ = \max(f, 0)$, $f^- = \max(-f, 0)$, $f = f^+ - f^-$, $|f| = f^+ + f^-$.
- $A_n \uparrow A$ means (A_n) is increasing and $A = \bigcup_n A_n$.
- $A_n \downarrow A$ means (A_n) is decreasing and $A = \bigcap_n A_n$.
- "a.e." means "almost everywhere" (outside a set of measure zero).
- $(X, \mathcal{A}, \mu)$ denotes a measure space.
- $\mathcal{B}(\mathbb{R}^n)$ denotes the Borel σ-algebra on $\mathbb{R}^n$.
- λ^n denotes the Lebesgue measure on $\mathbb{R}^n$ ($\lambda = \lambda^1$).

Prerequisites

- **Real analysis**: properties of $\mathbb{R}$, numerical sequences and series, continuity, differentiability, Riemann integral, series of functions, uniform convergence.
- **General topology**: metric spaces, open and closed sets, compactness, connectedness.
- **Linear algebra**: vector spaces, norms, Banach spaces (basic notions).
- **Set theory**: set operations, axiom of choice (for the construction of non-measurable sets).

References

- H. L. Royden and P. M. Fitzpatrick, *Real Analysis*, 4th ed., Pearson, 2010.
- W. Rudin, *Real and Complex Analysis*, 3rd ed., McGraw-Hill, 1987.
- G. B. Folland, *Real Analysis: Modern Techniques and Their Applications*, 2nd ed., Wiley, 1999.
- P. Halmos, *Measure Theory*, Springer, 1974.
- R. M. Dudley, *Real Analysis and Probability*, Cambridge University Press, 2002.
- E. M. Stein and R. Shakarchi, *Real Analysis: Measure Theory, Integration, and Hilbert Spaces*, Princeton University Press, 2005.
- R. L. Schilling, *Measures, Integrals and Martingales*, Cambridge University Press, 2005.
- A. N. Kolmogorov and S. V. Fomin, *Introductory Real Analysis*, Dover, 1975.

The Author
March 2026

Chapter 1

σ-Algebras and Measurable Spaces

At the end of the nineteenth century, mathematicians faced a fundamental problem: how to rigorously define the "size" of a set. The Riemann integral, which had served so well for decades, was showing its limits in the face of pathological functions discovered by Dirichlet, Weierstrass, and others. Henri Lebesgue, in his 1902 thesis, proposed a revolution: instead of slicing the x-axis into small intervals, he sliced the y-axis. But for this idea to work, one first had to answer a preliminary question: to which sets can we assign a measure in a consistent way?

The answer — and this is the subject of this first chapter — is that we must restrict ourselves to a collection of sets with good closure properties: a σ-algebra. This concept, which may appear purely technical at first sight, is in reality the keystone of all of measure theory, integration, and by extension modern probability.

1.1 Introduction

The first step in building measure theory is to identify the sets to which we wish to assign a "size." In general, one cannot consistently measure *all* subsets of a given set (we shall see in Chapter 4 that non-measurable sets exist). One must therefore restrict attention to a collection of sets that is rich enough to be useful, yet structured enough for the measure to be well-defined. This collection is a σ-algebra.

1.2 Fundamental definitions

Definition 1.1 (σ-algebra)**.** Let X be a non-empty set. A *σ-algebra* on X is a collection $\mathcal{A} \subset \mathcal{P}(X)$ of subsets of X satisfying:

(i) $X \in \mathcal{A}$;

(ii) if $A \in \mathcal{A}$, then $A^c = X \setminus A \in \mathcal{A}$;

(iii) if $(A_n)_{n\in\mathbb{N}}$ is a sequence of elements of $\mathcal{A}$, then $\bigcup_{n=0}^{\infty} A_n \in \mathcal{A}$.

The pair $(X, \mathcal{A})$ is called a *measurable space*.

Remark 1.2. Condition (iii) concerns *countable* unions. This is the essential difference with the notion of an *algebra* (or Boolean algebra), which only requires closure under *finite* unions.

CHAPTER 1. σ-ALGEBRAS AND MEASURABLE SPACES

Proposition 1.3 (Elementary properties)**.** Let $\mathcal{A}$ be a σ-algebra on X. Then:

1. $\varnothing \in \mathcal{A}$;

2. $\mathcal{A}$ is closed under countable intersections: if $(A_n)_{n\in\mathbb{N}} \subset \mathcal{A}$, then $\bigcap_{n=0}^{\infty} A_n \in \mathcal{A}$;

3. $\mathcal{A}$ is closed under set difference: if $A, B \in \mathcal{A}$, then $A \setminus B \in \mathcal{A}$;

4. $\mathcal{A}$ is closed under finite unions and intersections.

Proof. (1) $\varnothing = X^c \in \mathcal{A}$ by (i) and (ii).

(2) By De Morgan's laws: $\bigcap_n A_n = \left(\bigcup_n A_n^c\right)^c$. Since each $A_n^c \in \mathcal{A}$ by (ii), the union $\bigcup_n A_n^c \in \mathcal{A}$ by (iii), and its complement is in $\mathcal{A}$ by (ii).

(3) $A \setminus B = A \cap B^c$, which is in $\mathcal{A}$ by (ii) and (2).

(4) Special case of (iii) and (2) with $A_n = \varnothing$ for n large enough. $\square$

Definition 1.4 (Algebra)**.** A collection $\mathcal{A}_0 \subset \mathcal{P}(X)$ is an *algebra* on X if it satisfies (i), (ii), and closure under *finite* unions: if $A, B \in \mathcal{A}_0$, then $A \cup B \in \mathcal{A}_0$.

Example 1.5. Here are some fundamental examples of σ-algebras:

1. The *trivial σ-algebra* $\{\varnothing, X\}$ is the smallest σ-algebra on X.

2. The *discrete σ-algebra* $\mathcal{P}(X)$ is the largest σ-algebra on X.

3. For any $A \subset X$, $\mathcal{A} = \{\varnothing, A, A^c, X\}$ is a σ-algebra.

4. If X is uncountable, the collection of countable sets and sets with countable complement forms a σ-algebra, called the *co-countable σ-algebra*.

1.3 Generated σ-algebra

In practice, one does not define a σ-algebra by listing all its elements, but by "generating" it from a smaller collection.

Theorem 1.6 (Generated σ-algebra)**.** *Let $\mathcal{E} \subset \mathcal{P}(X)$ be any collection of subsets of X. There exists a smallest σ-algebra containing $\mathcal{E}$, called the* σ-algebra generated by $\mathcal{E}$ *and denoted $\sigma(\mathcal{E})$. It is given by:*

$$\sigma(\mathcal{E}) = \bigcap_{\substack{\mathcal{A} \text{ is a } \sigma\text{-algebra} \\ \mathcal{E} \subset \mathcal{A}}} \mathcal{A}.$$

Proof. The intersection is well-defined since at least $\mathcal{P}(X)$ is a σ-algebra containing $\mathcal{E}$, so the family is non-empty.

We verify that $\sigma(\mathcal{E})$ is a σ-algebra:

- X belongs to every σ-algebra, so $X \in \sigma(\mathcal{E})$.

- If $A \in \sigma(\mathcal{E})$, then A belongs to every σ-algebra $\mathcal{A} \supset \mathcal{E}$, so A^c does too, hence $A^c \in \sigma(\mathcal{E})$.

- The same argument applies to countable unions.

Finally, $\sigma(\mathcal{E})$ is the *smallest* such σ-algebra since it is contained in every σ-algebra containing $\mathcal{E}$ (by definition as an intersection). □

Remark 1.7. The intersection construction is non-constructive: it does not allow one to describe the elements of $\sigma(\mathcal{E})$ explicitly. In general, $\sigma(\mathcal{E})$ is much larger than $\mathcal{E}$ and its structure is complex.

1.4 The Borel σ-algebra

Definition 1.8 (Borel σ-algebra)**.** Let (X, τ) be a topological space. The *Borel σ-algebra* of X, denoted $\mathcal{B}(X)$, is the σ-algebra generated by the open sets of X:

$$\mathcal{B}(X) = \sigma(\tau).$$

The elements of $\mathcal{B}(X)$ are called *Borel sets*.

Theorem 1.9 (Generators of the Borel σ-algebra of $\mathbb{R}$)**.** *The Borel σ-algebra $\mathcal{B}(\mathbb{R})$ is generated by each of the following collections:*

1. *the open subsets of* $\mathbb{R}$*;*
2. *the closed subsets of* $\mathbb{R}$*;*
3. *the open intervals* (a, b)*,* $a < b$*;*
4. *the half-open intervals* $[a, b)$*,* $a < b$*;*
5. *the half-open intervals* $(a, b]$*,* $a < b$*;*
6. *the rays* $(-\infty, a)$*,* $a \in \mathbb{R}$*;*
7. *the rays* $(-\infty, a]$*,* $a \in \mathbb{R}$*;*
8. *the rays* $(a, +\infty)$*,* $a \in \mathbb{R}$*.*

Proof. Denote by $\mathcal{B}$ the Borel σ-algebra and by σ_i the σ-algebra generated by the i-th collection. It suffices to show circular inclusions.

$\sigma_6 \subset \sigma_7$: $(-\infty, a) = \bigcup_{n=1}^{\infty}(-\infty, a - 1/n]$, so $(-\infty, a) \in \sigma_7$.

$\sigma_7 \subset \sigma_6$: $(-\infty, a] = \bigcap_{n=1}^{\infty}(-\infty, a + 1/n)$, so $(-\infty, a] \in \sigma_6$.

$\sigma_3 \subset \sigma_6$: $(a, b) = (-\infty, b) \cap (a, +\infty)$. Now $(a, +\infty) = (-\infty, a]^c \in \sigma_6$ (since $\sigma_6 = \sigma_7$ as shown).

$\sigma_1 \subset \sigma_3$: every open subset of $\mathbb{R}$ is a countable union of open intervals (Lindelöf property).

Since $\sigma_1 = \mathcal{B}$ by definition and $\sigma_1 \subset \sigma_3 \subset \sigma_6 \subset \mathcal{B}$, we have equality. The remaining cases are analogous. □

Common generators of $\mathcal{B}(\mathbb{R})$

$$\mathcal{B}(\mathbb{R}) = \sigma\big(\{(a, b) : a < b\}\big) = \sigma\big(\{(-\infty, a) : a \in \mathbb{R}\}\big) = \sigma\big(\{(-\infty, a] : a \in \mathbb{R}\}\big).$$

CHAPTER 1. σ-ALGEBRAS AND MEASURABLE SPACES

1.5 Borel σ-algebra on $\mathbb{R}^n$

Definition 1.10. The Borel σ-algebra on $\mathbb{R}^n$ is $\mathcal{B}(\mathbb{R}^n) = \sigma(\tau_n)$, where τ_n is the standard topology on $\mathbb{R}^n$.

Proposition 1.11. We have $\mathcal{B}(\mathbb{R}^n) = \sigma(\mathcal{R}_n)$, where $\mathcal{R}_n$ is the collection of open rectangles $\prod_{i=1}^n (a_i, b_i)$ with $a_i < b_i$ for all i.

Proof. Every open set in $\mathbb{R}^n$ is a countable union of open rectangles with rational vertices (since $\mathbb{Q}^{2n}$ is countable), so $\sigma(\tau_n) \subset \sigma(\mathcal{R}_n)$. Conversely, each open rectangle is open in $\mathbb{R}^n$, so $\sigma(\mathcal{R}_n) \subset \sigma(\tau_n)$. □

1.6 Product σ-algebra

Definition 1.12 (Product σ-algebra)**.** Let $(X_1, \mathcal{A}_1)$ and $(X_2, \mathcal{A}_2)$ be measurable spaces. The *product σ-algebra* is:

$$\mathcal{A}_1 \otimes \mathcal{A}_2 = \sigma\big(\{A_1 \times A_2 : A_1 \in \mathcal{A}_1,\ A_2 \in \mathcal{A}_2\}\big).$$

More generally, for a family $(X_i, \mathcal{A}_i)_{i \in I}$:

$$\bigotimes_{i \in I} \mathcal{A}_i = \sigma\big(\{\pi_i^{-1}(A_i) : i \in I,\ A_i \in \mathcal{A}_i\}\big),$$

where $\pi_i : \prod_{j \in I} X_j \to X_i$ is the canonical projection.

Theorem 1.13. *If X_1 and X_2 are separable metric spaces, then*

$$\mathcal{B}(X_1) \otimes \mathcal{B}(X_2) = \mathcal{B}(X_1 \times X_2).$$

In particular, $\mathcal{B}(\mathbb{R}^n) = \mathcal{B}(\mathbb{R}) \otimes \cdots \otimes \mathcal{B}(\mathbb{R})$ (n times).

Proof. Let U be an open set in $X_1 \times X_2$. For every $(x_1, x_2) \in U$, there exist open sets $V_1 \ni x_1$ and $V_2 \ni x_2$ with $V_1 \times V_2 \subset U$. By separability, one can choose V_1 and V_2 from countable bases $\mathcal{B}_1$ and $\mathcal{B}_2$ of X_1 and X_2. Thus $U = \bigcup_k V_1^{(k)} \times V_2^{(k)}$, a countable union of measurable rectangles. This shows $\mathcal{B}(X_1 \times X_2) \subset \mathcal{B}(X_1) \otimes \mathcal{B}(X_2)$.

Conversely, if $A_1 \in \mathcal{B}(X_1)$, then $\pi_1^{-1}(A_1) = A_1 \times X_2$ is Borel in $X_1 \times X_2$ since π_1 is continuous. Similarly for π_2. The sets $A_1 \times A_2 = \pi_1^{-1}(A_1) \cap \pi_2^{-1}(A_2)$ are Borel, giving the reverse inclusion. □

1.7 Sub-σ-algebra and trace

Definition 1.14 (Trace)**.** Let $(X, \mathcal{A})$ be a measurable space and $Y \subset X$. The *trace* of $\mathcal{A}$ on Y is:

$$\mathcal{A}|_Y = \{A \cap Y : A \in \mathcal{A}\}.$$

This is a σ-algebra on Y.

Proposition 1.15. If $\mathcal{A} = \sigma(\mathcal{E})$, then $\mathcal{A}|_Y = \sigma(\mathcal{E}|_Y)$, where $\mathcal{E}|_Y = \{E \cap Y : E \in \mathcal{E}\}$.

Proof. Set $\mathcal{F} = \{A \subset X : A \cap Y \in \sigma(\mathcal{E}|_Y)\}$. One checks that $\mathcal{F}$ is a σ-algebra containing $\mathcal{E}$, hence $\sigma(\mathcal{E}) \subset \mathcal{F}$, giving $\mathcal{A}|_Y \subset \sigma(\mathcal{E}|_Y)$. The reverse inclusion is immediate. □

1.8 Measurable maps

Definition 1.16 (Measurable map)**.** Let $(X, \mathcal{A})$ and $(Y, \mathcal{B})$ be measurable spaces. A map $f : X \to Y$ is $(\mathcal{A}, \mathcal{B})$-*measurable* (or simply *measurable*) if:

$$\forall B \in \mathcal{B}, \quad f^{-1}(B) \in \mathcal{A}.$$

Proposition 1.17 (Generator criterion for measurability)**.** Let $\mathcal{B} = \sigma(\mathcal{E})$. Then $f : (X, \mathcal{A}) \to (Y, \mathcal{B})$ is measurable if and only if:

$$\forall E \in \mathcal{E}, \quad f^{-1}(E) \in \mathcal{A}.$$

Proof. The forward direction is clear. For the converse, set $\mathcal{C} = \{B \subset Y : f^{-1}(B) \in \mathcal{A}\}$. We verify that $\mathcal{C}$ is a σ-algebra:

- $f^{-1}(Y) = X \in \mathcal{A}$, so $Y \in \mathcal{C}$.
- If $B \in \mathcal{C}$, then $f^{-1}(B^c) = f^{-1}(B)^c \in \mathcal{A}$, so $B^c \in \mathcal{C}$.
- If $(B_n) \subset \mathcal{C}$, then $f^{-1}(\bigcup_n B_n) = \bigcup_n f^{-1}(B_n) \in \mathcal{A}$, so $\bigcup_n B_n \in \mathcal{C}$.

By hypothesis, $\mathcal{E} \subset \mathcal{C}$, hence $\mathcal{B} = \sigma(\mathcal{E}) \subset \mathcal{C}$, meaning f is measurable. □

Corollary 1.18. *A map $f : \mathbb{R} \to \mathbb{R}$ is Borel measurable (i.e. $(\mathcal{B}(\mathbb{R}), \mathcal{B}(\mathbb{R}))$-measurable) if and only if $f^{-1}((-\infty, a)) \in \mathcal{B}(\mathbb{R})$ for every $a \in \mathbb{R}$.*

Proposition 1.19 (Composition)**.** If $f : (X, \mathcal{A}) \to (Y, \mathcal{B})$ and $g : (Y, \mathcal{B}) \to (Z, \mathcal{C})$ are measurable, then $g \circ f : (X, \mathcal{A}) \to (Z, \mathcal{C})$ is measurable.

Proof. For any $C \in \mathcal{C}$, $(g \circ f)^{-1}(C) = f^{-1}(g^{-1}(C))$. Since g is measurable, $g^{-1}(C) \in \mathcal{B}$, and since f is measurable, $f^{-1}(g^{-1}(C)) \in \mathcal{A}$. □

1.9 Dynkin systems and the π-λ lemma

Dynkin's π-λ lemma is a fundamental tool for proving that two measures agree.

Definition 1.20 (π-system and λ-system)**.**

- A π-*system* on X is a collection $\mathcal{P} \subset \mathcal{P}(X)$ closed under finite intersections.
- A λ-*system* (or *Dynkin system*) on X is a collection $\mathcal{D} \subset \mathcal{P}(X)$ such that:
    (i) $X \in \mathcal{D}$;
    (ii) if $A, B \in \mathcal{D}$ and $A \subset B$, then $B \setminus A \in \mathcal{D}$;
    (iii) if (A_n) is an increasing sequence in $\mathcal{D}$, then $\bigcup_n A_n \in \mathcal{D}$.

Theorem 1.21 (Dynkin's π-λ lemma)**.** *If $\mathcal{P}$ is a π-system and $\mathcal{D}$ is a λ-system with $\mathcal{P} \subset \mathcal{D}$, then $\sigma(\mathcal{P}) \subset \mathcal{D}$.*

Proof. It suffices to show that $d(\mathcal{P})$, the smallest λ-system containing $\mathcal{P}$, is a σ-algebra (for then $\sigma(\mathcal{P}) \subset d(\mathcal{P}) \subset \mathcal{D}$).

For this, it suffices to show that $d(\mathcal{P})$ is closed under finite intersections (since a λ-system closed under finite intersections is a σ-algebra).

For $A \in d(\mathcal{P})$, set $\mathcal{G}_A = \{B \in d(\mathcal{P}) : A \cap B \in d(\mathcal{P})\}$.

CHAPTER 1. σ-ALGEBRAS AND MEASURABLE SPACES

Step 1. If $A \in \mathcal{P}$, then $\mathcal{G}_A$ is a λ-system containing $\mathcal{P}$ (since $\mathcal{P}$ is a π-system), hence $\mathcal{G}_A \supset d(\mathcal{P})$.

Step 2. For any $A \in d(\mathcal{P})$, $\mathcal{G}_A$ is a λ-system. By Step 1, $\mathcal{P} \subset \mathcal{G}_A$ for every $A \in d(\mathcal{P})$, so $\mathcal{G}_A \supset d(\mathcal{P})$.

Thus $d(\mathcal{P})$ is closed under finite intersections, which concludes. □

Using the π-λ lemma

This lemma is very useful for showing that a property true on a π-system extends to the generated σ-algebra. Typically, to show that two measures μ and ν on $\sigma(\mathcal{P})$ agree, it suffices to verify $\mu = \nu$ on the π-system $\mathcal{P}$ (under a σ-finiteness condition).

1.10 Uniqueness of measures

Theorem 1.22 (Uniqueness of measures)**.** *Let μ and ν be two measures on $(X, \sigma(\mathcal{P}))$, where $\mathcal{P}$ is a π-system. Suppose there exists a sequence $(E_n)_{n\geq 1} \subset \mathcal{P}$ with $E_n \uparrow X$ and $\mu(E_n) = \nu(E_n) < \infty$ for all n. If $\mu = \nu$ on $\mathcal{P}$, then $\mu = \nu$ on $\sigma(\mathcal{P})$.*

Proof. Fix n and set $\mu_n(\cdot) = \mu(\cdot \cap E_n)$ and $\nu_n(\cdot) = \nu(\cdot \cap E_n)$. These are finite measures. The collection

$$\mathcal{D}_n = \{A \in \sigma(\mathcal{P}) : \mu_n(A) = \nu_n(A)\}$$

is a λ-system (direct verification). By hypothesis, $\mathcal{P} \subset \mathcal{D}_n$. By the π-λ lemma, $\sigma(\mathcal{P}) \subset \mathcal{D}_n$, so $\mu_n = \nu_n$ on $\sigma(\mathcal{P})$.

For every $A \in \sigma(\mathcal{P})$, $A \cap E_n \uparrow A$, so by monotone continuity: $\mu(A) = \lim_n \mu_n(A) = \lim_n \nu_n(A) = \nu(A)$. □

1.11 Exercises

Exercise 1.1. Let X be an uncountable set. Show that the collection

$$\mathcal{A} = \{A \subset X : A \text{ is countable or } A^c \text{ is countable}\}$$

is a σ-algebra on X.

Exercise 1.2. Show that $\mathcal{B}(\mathbb{R})$ is generated by the open intervals with rational endpoints.

Exercise 1.3. Let $f : \mathbb{R} \to \mathbb{R}$ be a continuous function. Show that f is Borel measurable.

Exercise 1.4. Show that $\mathcal{B}(\mathbb{R}) \neq \mathcal{P}(\mathbb{R})$. *Hint*: use a cardinality argument ($|\mathcal{B}(\mathbb{R})| = |\mathbb{R}| = \mathfrak{c}$ while $|\mathcal{P}(\mathbb{R})| = 2^{\mathfrak{c}}$).

Exercise 1.5. Let $(X, \mathcal{A})$ be a measurable space and $f, g : X \to \mathbb{R}$ two measurable functions. Show that $\{x \in X : f(x) < g(x)\} \in \mathcal{A}$.

Exercise 1.6 (Atoms of a σ-algebra)**.** Let $(X, \mathcal{A})$ be a measurable space. For $x \in X$, define the *atom* of x as $[x] = \bigcap_{A \in \mathcal{A},\, x \in A} A$.

1. Show that if $\mathcal{A}$ is generated by a countable partition, then $[x] \in \mathcal{A}$ for every x.

2. Give an example where $[x] \notin \mathcal{A}$.

Chapter 2

Measures

Now that we have σ-algebras — the collections of sets to which we wish to assign a size — it is time to define that "size" itself. This is the role of the notion of *measure*, which simultaneously generalizes length, area, volume, and probability. This unification, due to Lebesgue and refined by Carathéodory, is one of the great achievements of modern mathematics.

2.1 Definition and first properties

Definition 2.1 (Measure)**.** Let $(X, \mathcal{A})$ be a measurable space. A *measure* on $(X, \mathcal{A})$ is a function $\mu : \mathcal{A} \to [0, +\infty]$ satisfying:

(i) $\mu(\varnothing) = 0$;

(ii) (σ-additivity) for every sequence $(A_n)_{n\in\mathbb{N}}$ of *pairwise disjoint* elements of $\mathcal{A}$:

$$\mu\left(\bigsqcup_{n=0}^{\infty} A_n\right) = \sum_{n=0}^{\infty} \mu(A_n).$$

The triple $(X, \mathcal{A}, \mu)$ is called a *measure space*.

Definition 2.2 (Types of measures)**.** Let $(X, \mathcal{A}, \mu)$ be a measure space.

- μ is *finite* if $\mu(X) < \infty$.
- μ is a *probability measure* if $\mu(X) = 1$.
- μ is *σ-finite* if there exists a sequence $(X_n)_{n\in\mathbb{N}} \subset \mathcal{A}$ with $X = \bigcup_n X_n$ and $\mu(X_n) < \infty$ for all n.
- μ is *complete* if for every $N \in \mathcal{A}$ with $\mu(N) = 0$ and every $A \subset N$, we have $A \in \mathcal{A}$.

Proposition 2.3 (Monotonicity and subadditivity)**.** Let $(X, \mathcal{A}, \mu)$ be a measure space.

1. **Monotonicity**: if $A \subset B$ with $A, B \in \mathcal{A}$, then $\mu(A) \leq \mu(B)$.
2. **Excision**: if $A \subset B$ and $\mu(A) < \infty$, then $\mu(B \setminus A) = \mu(B) - \mu(A)$.
3. **Countable subadditivity**: for any sequence $(A_n) \subset \mathcal{A}$, $\mu(\bigcup_n A_n) \leq \sum_n \mu(A_n)$.

Proof. (1) Write $B = A \sqcup (B \setminus A)$, so $\mu(B) = \mu(A) + \mu(B \setminus A) \geq \mu(A)$.

(2) Follows directly from the decomposition $B = A \sqcup (B \setminus A)$ and the finiteness of $\mu(A)$.

(3) Set $B_0 = A_0$ and $B_n = A_n \setminus \bigcup_{k=0}^{n-1} A_k$ for $n \geq 1$. Then (B_n) is a disjoint sequence with $\bigcup_n B_n = \bigcup_n A_n$ and $B_n \subset A_n$. Therefore

$$\mu\left(\bigcup_n A_n\right) = \mu\left(\bigsqcup_n B_n\right) = \sum_n \mu(B_n) \leq \sum_n \mu(A_n). \qquad \square$$

2.2 Continuity of measures

Theorem 2.4 (Monotone continuity)**.** *Let $(X, \mathcal{A}, \mu)$ be a measure space.*

1. ***Continuity from below**: if $A_n \uparrow A$ (increasing sequence), then $\mu(A_n) \uparrow \mu(A)$.*
2. ***Continuity from above**: if $A_n \downarrow A$ (decreasing sequence) and $\mu(A_1) < \infty$, then $\mu(A_n) \downarrow \mu(A)$.*

Proof. (1) Set $B_0 = A_0$ and $B_n = A_n \setminus A_{n-1}$ for $n \geq 1$. Then (B_n) is disjoint with $\bigsqcup_{k=0}^{n} B_k = A_n$ and $\bigsqcup_{k=0}^{\infty} B_k = A$. By σ-additivity:

$$\mu(A) = \sum_{k=0}^{\infty} \mu(B_k) = \lim_{n\to\infty} \sum_{k=0}^{n} \mu(B_k) = \lim_{n\to\infty} \mu(A_n).$$

(2) Set $C_n = A_1 \setminus A_n$. Then $C_n \uparrow A_1 \setminus A$. By (1), $\mu(C_n) \uparrow \mu(A_1 \setminus A)$. Since $\mu(A_1) < \infty$: $\mu(C_n) = \mu(A_1) - \mu(A_n)$ and $\mu(A_1 \setminus A) = \mu(A_1) - \mu(A)$, giving $\mu(A_n) \downarrow \mu(A)$. $\square$

Finiteness condition

The hypothesis $\mu(A_1) < \infty$ in (2) is essential. Counterexample: $A_n = [n, +\infty)$ with Lebesgue measure. Then $A_n \downarrow \varnothing$ but $\mu(A_n) = +\infty$ for all n.

Lemma 2.5 (Borel–Cantelli lemma)**.** *Let $(X, \mathcal{A}, \mu)$ be a measure space and $(A_n) \subset \mathcal{A}$. If $\sum_{n=0}^{\infty} \mu(A_n) < \infty$, then*

$$\mu\left(\limsup_{n\to\infty} A_n\right) = 0,$$

where $\limsup_n A_n = \bigcap_{n=0}^{\infty} \bigcup_{k=n}^{\infty} A_k$ is the set of points belonging to infinitely many A_n.

Proof. Set $B_n = \bigcup_{k=n}^{\infty} A_k$. Then $B_n \downarrow \limsup_n A_n$ and

$$\mu(B_n) \leq \sum_{k=n}^{\infty} \mu(A_k) \to 0 \quad (n \to \infty)$$

since the tail of a convergent series tends to zero. Since $\mu(B_0) \leq \sum_k \mu(A_k) < \infty$, continuity from above gives $\mu(\limsup_n A_n) = \lim_n \mu(B_n) = 0$. $\square$

2.3 Fundamental examples

Example 2.6 (Counting measure)**.** Let X be any set and $\mathcal{A} = \mathcal{P}(X)$. The *counting measure* is defined by:

$$\mu(A) = \begin{cases} |A| & \text{if } A \text{ is finite,} \\ +\infty & \text{otherwise.} \end{cases}$$

It is a measure on $(X, \mathcal{P}(X))$, σ-finite if and only if X is countable.

Example 2.7 (Dirac measure)**.** Let $x_0 \in X$. The *Dirac measure* (or point mass) at x_0 is the probability measure:

$$\delta_{x_0}(A) = \begin{cases} 1 & \text{if } x_0 \in A, \\ 0 & \text{if } x_0 \notin A. \end{cases}$$

Example 2.8 (Linear combinations)**.** If μ_1, μ_2 are measures on $(X, \mathcal{A})$ and $\alpha, \beta \geq 0$, then $\alpha\mu_1 + \beta\mu_2$ is a measure. More generally, if $(x_n)_{n\in\mathbb{N}} \subset X$ and $(a_n)_{n\in\mathbb{N}} \subset [0, +\infty)$, then $\mu = \sum_n a_n \delta_{x_n}$ is a measure.

2.4 Pre-measures and extension

Definition 2.9 (Pre-measure)**.** Let $\mathcal{A}_0$ be an algebra on X. A function $\mu_0 : \mathcal{A}_0 \to [0, +\infty]$ is a *pre-measure* if:

(i) $\mu_0(\varnothing) = 0$;

(ii) for every disjoint sequence $(A_n) \subset \mathcal{A}_0$ such that $\bigsqcup_n A_n \in \mathcal{A}_0$: $\mu_0(\bigsqcup_n A_n) = \sum_n \mu_0(A_n)$.

Remark 2.10. The difference from a measure is that $\mathcal{A}_0$ is only an algebra, and σ-additivity is required only when the union falls in $\mathcal{A}_0$.

Proposition 2.11 (Criterion for σ-additivity)**.** Let μ_0 be a finitely additive set function on an algebra $\mathcal{A}_0$ with $\mu_0(\varnothing) = 0$. The following are equivalent:

1. μ_0 is a pre-measure (σ-additive).

2. μ_0 is continuous from below: if $(A_n) \subset \mathcal{A}_0$, $A_n \uparrow A \in \mathcal{A}_0$, then $\mu_0(A_n) \uparrow \mu_0(A)$.

3. μ_0 is continuous at $\varnothing$: if $(A_n) \subset \mathcal{A}_0$, $A_n \downarrow \varnothing$, then $\mu_0(A_n) \downarrow 0$.

Proof. (1) $\Rightarrow$ (2): same proof as Theorem 2.4(1).

(2) $\Rightarrow$ (3): if $A_n \downarrow \varnothing$ and $\mu_0(A_1) < \infty$, apply the same technique as Theorem 2.4(2). If $\mu_0(A_1) = +\infty$, the result is trivial.

(3) $\Rightarrow$ (1): let (A_n) be disjoint with $A = \bigsqcup_n A_n \in \mathcal{A}_0$. Set $B_N = A \setminus \bigsqcup_{n=0}^{N} A_n$. By finite additivity, $\mu_0(A) = \sum_{n=0}^{N} \mu_0(A_n) + \mu_0(B_N)$. Since $B_N \downarrow \varnothing$, we get $\mu_0(B_N) \to 0$, hence $\mu_0(A) = \sum_n \mu_0(A_n)$. □

2.5 Extension theorem

Theorem 2.12 (Carathéodory extension — preliminary statement)**.** *Let μ_0 be a pre-measure on an algebra $\mathcal{A}_0$. Then μ_0 extends to a measure μ on $\sigma(\mathcal{A}_0)$. If μ_0 is σ-finite, this extension is unique.*

The complete proof will be given in Chapter 3, but we sketch the construction here.

Outer measure construction

One defines the *outer measure* associated with μ_0 by:

$$\mu^*(A) = \inf\left\{\sum_{n=0}^{\infty} \mu_0(A_n) : (A_n) \subset \mathcal{A}_0,\ A \subset \bigcup_{n=0}^{\infty} A_n\right\}$$

for every $A \subset X$. Then one restricts μ^* to the *Carathéodory-measurable* sets.

2.6 Completion of a measure

Definition 2.13 (Completion)**.** Let $(X, \mathcal{A}, \mu)$ be a measure space. The *completion* of μ is the measure $\overline{\mu}$ defined on the σ-algebra:

$$\overline{\mathcal{A}} = \{A \cup N : A \in \mathcal{A},\ N \subset M \text{ for some } M \in \mathcal{A} \text{ with } \mu(M) = 0\}$$

by $\overline{\mu}(A \cup N) = \mu(A)$.

Proposition 2.14. $\overline{\mathcal{A}}$ is a σ-algebra and $\overline{\mu}$ is a complete measure extending μ.

Proof. We verify that $\overline{\mathcal{A}}$ is a σ-algebra.

Closure under complementation: if $E = A \cup N$ with $N \subset M$, $\mu(M) = 0$, then $E^c = A^c \cap N^c = (A^c \cap M^c) \cup (A^c \cap M \setminus N)$. Now $A^c \cap M^c \in \mathcal{A}$ and $A^c \cap M \setminus N \subset M$, so $E^c \in \overline{\mathcal{A}}$.

Closure under countable unions: if $E_n = A_n \cup N_n$ with $N_n \subset M_n$, $\mu(M_n) = 0$, then $\bigcup_n E_n = \bigcup_n A_n \cup \bigcup_n N_n$ with $\bigcup_n N_n \subset \bigcup_n M_n$ and $\mu(\bigcup_n M_n) \leq \sum_n \mu(M_n) = 0$.

Well-definedness of $\overline{\mu}$ (independence of the decomposition) is verified directly, and σ-additivity follows from that of μ. □

2.7 Image measure

Definition 2.15 (Image measure)**.** Let $(X, \mathcal{A}, \mu)$ be a measure space and $f : X \to Y$ a measurable map to $(Y, \mathcal{B})$. The *image measure* (or *push-forward*) of μ under f is the measure $f_*\mu$ on $(Y, \mathcal{B})$ defined by:

$$(f_*\mu)(B) = \mu(f^{-1}(B)), \quad B \in \mathcal{B}.$$

Proposition 2.16. $f_*\mu$ is indeed a measure on $(Y, \mathcal{B})$.

Proof. $(f_*\mu)(\varnothing) = \mu(f^{-1}(\varnothing)) = \mu(\varnothing) = 0$. For a disjoint sequence (B_n) in $\mathcal{B}$, $(f^{-1}(B_n))$ is disjoint in $\mathcal{A}$ and $f^{-1}(\bigsqcup_n B_n) = \bigsqcup_n f^{-1}(B_n)$, so $(f_*\mu)(\bigsqcup_n B_n) = \mu(\bigsqcup_n f^{-1}(B_n)) = \sum_n \mu(f^{-1}(B_n)) = \sum_n (f_*\mu)(B_n)$. □

2.8 Exercises

Exercise 2.1. Let $(X, \mathcal{A}, \mu)$ be a measure space and $(A_n) \subset \mathcal{A}$. Show that $\mu(\liminf_n A_n) \leq \liminf_n \mu(A_n)$ (Fatou's lemma for sets).

Exercise 2.2. Let μ be a measure on $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ such that $\mu(K) < \infty$ for every compact K. Show that μ is σ-finite.

Exercise 2.3. Show that the counting measure on $(\mathbb{R}, \mathcal{P}(\mathbb{R}))$ is not σ-finite.

Exercise 2.4. Let $(X, \mathcal{A}, \mu)$ be a measure space with $\mu(X) < \infty$ and $(A_n) \subset \mathcal{A}$ a sequence with $\mu(A_n) \geq \alpha > 0$ for all n. Show that there exists $x \in X$ belonging to infinitely many A_n.

Exercise 2.5. Let μ be a finite measure on $(\mathbb{N}, \mathcal{P}(\mathbb{N}))$. Show that μ is determined by the values $\mu(\{n\})$, $n \in \mathbb{N}$, and that $\mu = \sum_{n=0}^{\infty} \mu(\{n\})\delta_n$.

Exercise 2.6 (Lebesgue–Stieltjes measure)**.** Let $F : \mathbb{R} \to \mathbb{R}$ be an increasing right-continuous function. Define $\mu_0((a, b]) = F(b) - F(a)$ on the algebra $\mathcal{A}_0$ of finite disjoint unions of half-open intervals $(a, b]$.

1. Show that μ_0 is a pre-measure on $\mathcal{A}_0$.
2. Deduce the existence of a unique measure μ_F on $\mathcal{B}(\mathbb{R})$ such that $\mu_F((a, b]) = F(b) - F(a)$.

Chapter 3

Carathéodory Extension Theorem

3.1 Introduction

How does one *construct* a measure? The question may seem circular: we want to measure sets, but to rigorously define what "measuring" means, we first need a theory. It was Constantin Carathéodory who, in 1914, cut this Gordian knot with a construction of remarkable elegance. His idea: start with a cruder notion, the *outer measure*, which has the defect of not being additive in general, then select among all sets those on which this outer measure behaves well. The "measurable" sets in Carathéodory's sense automatically form a σ-algebra, and the outer measure restricted to this σ-algebra is a genuine measure.

This construction, as elegant as it is abstract, is the universal machine that produces the Lebesgue measure, the Hausdorff measure, and many others. It is the cornerstone on which the entire edifice of modern integration rests.

3.2 Outer measures

Definition 3.1 (Outer measure)**.** Let X be a set. An *outer measure* on X is a function $\mu^* : \mathcal{P}(X) \to [0, +\infty]$ satisfying:

(i) $\mu^*(\varnothing) = 0$;

(ii) (monotonicity) if $A \subset B$, then $\mu^*(A) \leq \mu^*(B)$;

(iii) (countable subadditivity) for every sequence $(A_n)_{n\in\mathbb{N}} \subset \mathcal{P}(X)$:

$$\mu^*\left(\bigcup_{n=0}^{\infty} A_n\right) \leq \sum_{n=0}^{\infty} \mu^*(A_n).$$

Remark 3.2. An outer measure is defined on *all* subsets of X, but it is generally not σ-additive. Carathéodory's idea is to identify the sets on which it is additive.

3.3 Carathéodory-measurable sets

Definition 3.3 (Carathéodory measurability)**.** Let μ^* be an outer measure on X. A set $A \subset X$ is μ^**-measurable* (or *Carathéodory-measurable*) if:

$$\forall\, E \subset X, \quad \mu^*(E) = \mu^*(E \cap A) + \mu^*(E \cap A^c).$$

We denote by $\mathcal{M}^*$ the collection of all μ^*-measurable sets.

Interpreting the Carathéodory criterion

The condition says that A "splits" every set E properly: the outer measure of E equals the sum of the outer measures of the two pieces $E \cap A$ and $E \cap A^c$. By subadditivity, the inequality $\mu^*(E) \leq \mu^*(E \cap A) + \mu^*(E \cap A^c)$ always holds. So it suffices to verify the reverse inequality: $\mu^*(E) \geq \mu^*(E \cap A) + \mu^*(E \cap A^c)$.

3.4 The main theorem

Theorem 3.4 (Carathéodory). *Let μ^* be an outer measure on X. Then:*

1. *$\mathcal{M}^*$ is a σ-algebra;*
2. *the restriction $\mu = \mu^*|_{\mathcal{M}^*}$ is a complete measure on $(X, \mathcal{M}^*)$.*

Proof. The proof proceeds in several steps.

Step 1: $\mathcal{M}^*$ is an algebra.

Clearly $X \in \mathcal{M}^*$ (take $A = X$: $\mu^*(E) = \mu^*(E) + \mu^*(\varnothing) = \mu^*(E)$). If $A \in \mathcal{M}^*$, then $A^c \in \mathcal{M}^*$ since the Carathéodory condition is symmetric in A and A^c.

Let us show closure under finite unions. Let $A, B \in \mathcal{M}^*$. For every $E \subset X$:

$$\begin{aligned}\mu^*(E) &= \mu^*(E \cap A) + \mu^*(E \cap A^c) \\ &= \mu^*(E \cap A \cap B) + \mu^*(E \cap A \cap B^c) + \mu^*(E \cap A^c \cap B) + \mu^*(E \cap A^c \cap B^c).\end{aligned}$$

Now $E \cap (A \cup B) = (E \cap A \cap B) \cup (E \cap A \cap B^c) \cup (E \cap A^c \cap B)$ and $E \cap (A \cup B)^c = E \cap A^c \cap B^c$. By subadditivity:

$$\mu^*(E \cap (A \cup B)) \leq \mu^*(E \cap A \cap B) + \mu^*(E \cap A \cap B^c) + \mu^*(E \cap A^c \cap B).$$

Therefore $\mu^*(E) \geq \mu^*(E \cap (A \cup B)) + \mu^*(E \cap (A \cup B)^c)$, and the reverse inequality follows from subadditivity. Thus $A \cup B \in \mathcal{M}^*$.

Step 2: Finite additivity.

If $A, B \in \mathcal{M}^*$ are disjoint, taking $E = A \cup B$: $\mu^*(A \cup B) = \mu^*((A \cup B) \cap A) + \mu^*((A \cup B) \cap A^c) = \mu^*(A) + \mu^*(B)$.

By induction, for $A_1, \ldots, A_n \in \mathcal{M}^*$ pairwise disjoint: $\mu^*(\bigsqcup_{k=1}^n A_k) = \sum_{k=1}^n \mu^*(A_k)$.

Step 3: σ-additivity and closure under countable unions.

Let $(A_n)_{n \geq 1} \subset \mathcal{M}^*$ be a disjoint sequence. Set $S_n = \bigsqcup_{k=1}^n A_k$ and $A = \bigsqcup_{k=1}^\infty A_k$. By Step 1, $S_n \in \mathcal{M}^*$, so for every E:

$$\mu^*(E) = \mu^*(E \cap S_n) + \mu^*(E \cap S_n^c) \geq \sum_{k=1}^n \mu^*(E \cap A_k) + \mu^*(E \cap A^c)$$

since $S_n^c \supset A^c$. Letting $n \to \infty$:

$$\mu^*(E) \geq \sum_{k=1}^\infty \mu^*(E \cap A_k) + \mu^*(E \cap A^c) \geq \mu^*(E \cap A) + \mu^*(E \cap A^c)$$

by subadditivity of μ^*. The reverse inequality always holds, so $A \in \mathcal{M}^*$.

Taking $E = A$: $\mu^*(A) = \sum_{k=1}^{\infty} \mu^*(A_k) + \mu^*(\varnothing) = \sum_k \mu^*(A_k)$, giving σ-additivity.

Step 4: Completeness.

Let $N \in \mathcal{M}^*$ with $\mu^*(N) = 0$ and $A \subset N$. For every $E \subset X$: $\mu^*(E \cap A) \leq \mu^*(N) = 0$ and $\mu^*(E \cap A^c) \leq \mu^*(E)$, so $\mu^*(E \cap A) + \mu^*(E \cap A^c) \leq \mu^*(E)$. The reverse inequality is subadditivity, so $A \in \mathcal{M}^*$. □

3.5 Application to pre-measures

Theorem 3.5 (Carathéodory extension). *Let μ_0 be a pre-measure on an algebra $\mathcal{A}_0 \subset \mathcal{P}(X)$. Define:*

$$\mu^*(A) = \inf\left\{\sum_{n=0}^{\infty} \mu_0(A_n) : (A_n) \subset \mathcal{A}_0,\, A \subset \bigcup_{n=0}^{\infty} A_n\right\}, \quad A \subset X.$$

Then:

1. *μ^* is an outer measure on X.*
2. *$\mathcal{A}_0 \subset \mathcal{M}^*$ (every element of the algebra is Carathéodory-measurable).*
3. *$\mu^*|_{\mathcal{A}_0} = \mu_0$ (the extension agrees with the pre-measure on the algebra).*
4. *In particular, $\mu = \mu^*|_{\sigma(\mathcal{A}_0)}$ is a measure on $\sigma(\mathcal{A}_0)$ extending μ_0.*

Proof. Part 1. $\mu^*(\varnothing) = 0$ (take $A_n = \varnothing$). Monotonicity is clear. For countable subadditivity, let (B_k) be a sequence and $\varepsilon > 0$. For each k, choose $(A_n^{(k)}) \subset \mathcal{A}_0$ with $B_k \subset \bigcup_n A_n^{(k)}$ and $\sum_n \mu_0(A_n^{(k)}) \leq \mu^*(B_k) + \varepsilon/2^{k+1}$. Then $\bigcup_k B_k \subset \bigcup_{k,n} A_n^{(k)}$ and

$$\mu^*\left(\bigcup_k B_k\right) \leq \sum_{k,n} \mu_0(A_n^{(k)}) \leq \sum_k \mu^*(B_k) + \varepsilon.$$

Letting $\varepsilon \to 0$ gives countable subadditivity.

Part 2. Let $A \in \mathcal{A}_0$ and $E \subset X$. It suffices to show $\mu^*(E) \geq \mu^*(E \cap A) + \mu^*(E \cap A^c)$. Let $\varepsilon > 0$ and $(A_n) \subset \mathcal{A}_0$ with $E \subset \bigcup_n A_n$ and $\sum_n \mu_0(A_n) \leq \mu^*(E) + \varepsilon$. Since $\mathcal{A}_0$ is an algebra, $A_n \cap A$ and $A_n \cap A^c$ are in $\mathcal{A}_0$ and $\mu_0(A_n) = \mu_0(A_n \cap A) + \mu_0(A_n \cap A^c)$ (finite additivity). Thus:

$$\begin{aligned}\mu^*(E) + \varepsilon &\geq \sum_n \mu_0(A_n \cap A) + \sum_n \mu_0(A_n \cap A^c) \\ &\geq \mu^*(E \cap A) + \mu^*(E \cap A^c).\end{aligned}$$

Part 3. For $A \in \mathcal{A}_0$, $\mu^*(A) \leq \mu_0(A)$ (take $A_0 = A$, $A_n = \varnothing$). For the reverse, let $(A_n) \subset \mathcal{A}_0$ with $A \subset \bigcup_n A_n$. Set $B_N = A \cap \left(\bigcup_{n=0}^N A_n\right) \in \mathcal{A}_0$. Then $B_N \uparrow A$ and $B_N \subset \bigcup_{n=0}^N A_n$. By finite subadditivity: $\mu_0(B_N) \leq \sum_{n=0}^N \mu_0(A_n)$. By continuity from below (Proposition 2.11): $\mu_0(A) = \lim_N \mu_0(B_N) \leq \sum_n \mu_0(A_n)$. Taking the infimum: $\mu_0(A) \leq \mu^*(A)$. □

3.6 Uniqueness of the extension

Theorem 3.6 (Uniqueness)**.** *If the pre-measure μ_0 is σ-finite on $\mathcal{A}_0$, then the extension μ to $\sigma(\mathcal{A}_0)$ is unique.*

Proof. By the uniqueness theorem for measures (Theorem 1.22), applied to the π-system $\mathcal{A}_0$ (an algebra is a π-system). The σ-finiteness of μ_0 provides the required exhausting sequence. $\square$

Summary of the Carathéodory construction

$$\text{Pre-measure } \mu_0 \text{ on } \mathcal{A}_0 \xrightarrow{\text{outer meas.}} \mu^* \text{ on } \mathcal{P}(X) \xrightarrow{\text{restrict}} \mu \text{ on } \mathcal{M}^* \supset \sigma(\mathcal{A}_0).$$

3.7 Carathéodory criterion for metric outer measures

Definition 3.7 (Metric outer measure)**.** Let (X, d) be a metric space. An outer measure μ^* on X is *metric* if:

$$d(A, B) > 0 \implies \mu^*(A \cup B) = \mu^*(A) + \mu^*(B),$$

where $d(A, B) = \inf\{d(a, b) : a \in A,\ b \in B\}$.

Theorem 3.8. *If μ^* is a metric outer measure on (X, d), then every Borel set is μ^*-measurable: $\mathcal{B}(X) \subset \mathcal{M}^*$.*

Proof. It suffices to show that every closed set F is μ^*-measurable. Let $E \subset X$ and set $E_n = \{x \in E \cap F^c : d(x, F) \geq 1/n\}$. Then $E_n \uparrow E \cap F^c$ and $d(E \cap F, E_n) \geq 1/n > 0$, so $\mu^*(E \cap F \cup E_n) = \mu^*(E \cap F) + \mu^*(E_n)$ by the metric property.

By monotonicity, $\mu^*(E) \geq \mu^*(E \cap F \cup E_n) = \mu^*(E \cap F) + \mu^*(E_n)$. It remains to show $\mu^*(E_n) \to \mu^*(E \cap F^c)$.

Set $D_n = E_{n+1} \setminus E_n$. For n and m with $m \geq n + 2$, $d(D_n, D_m) > 0$, so the sets D_n of fixed parity are at positive distance from each other. By the metric property and countable subadditivity:

$$\mu^*(E \cap F^c) \leq \mu^*(E_n) + \sum_{k=n}^{\infty} \mu^*(D_k).$$

If $\sum_k \mu^*(D_k) < \infty$, the tail tends to zero and $\mu^*(E_n) \to \mu^*(E \cap F^c)$. If the series diverges, one shows that $\mu^*(E) = +\infty$ and the inequality is trivial. $\square$

3.8 Application: Hausdorff measure

Definition 3.9 (Hausdorff measure)**.** Let (X, d) be a metric space and $s \geq 0$. For $\delta > 0$, define:

$$\mathcal{H}^s_\delta(A) = \inf\left\{ \sum_{n=1}^{\infty} (\operatorname{diam} U_n)^s : A \subset \bigcup_{n=1}^{\infty} U_n,\ \operatorname{diam} U_n \leq \delta \right\}.$$

The *s-dimensional Hausdorff measure* is:

$$\mathcal{H}^s(A) = \lim_{\delta \to 0} \mathcal{H}^s_\delta(A) = \sup_{\delta > 0} \mathcal{H}^s_\delta(A).$$

Proposition 3.10. $\mathcal{H}^s$ is a metric outer measure. Consequently, it is a (Borel) measure on $(X, \mathcal{B}(X))$.

3.9 Exercises

Exercise 3.1. Verify in detail that the set $\mathcal{M}^*$ in Carathéodory's theorem is closed under countable intersections.

Exercise 3.2. Let μ^* be an outer measure on X. Show that if $\mu^*(A) = 0$, then $A \in \mathcal{M}^*$.

Exercise 3.3. Show that the Hausdorff measure $\mathcal{H}^n$ on $\mathbb{R}^n$ equals (up to a multiplicative constant) the Lebesgue measure.

Exercise 3.4. Let μ_0 be a *finite* pre-measure on an algebra $\mathcal{A}_0$. Show that for every $A \in \sigma(\mathcal{A}_0)$ and every $\varepsilon > 0$, there exists $B \in \mathcal{A}_0$ such that $\mu(A \triangle B) < \varepsilon$, where $A \triangle B = (A \setminus B) \cup (B \setminus A)$.

Exercise 3.5. Let X be an uncountable set and $\mu^*(A) = \begin{cases} 0 & \text{if } A \text{ is countable,} \\ 1 & \text{otherwise.} \end{cases}$ Show that μ^* is an outer measure and determine $\mathcal{M}^*$.

CHAPTER 3. CARATHÉODORY EXTENSION THEOREM

Chapter 4

Lebesgue Measure on $\mathbb{R}^n$

We finally arrive at the object that all the preceding machinery has been preparing: the *Lebesgue measure*. It is the measure that gives rigorous meaning to the notions of length, area, and volume, where Riemann's construction fell short. Henri Lebesgue, in his 1902 thesis, showed that one can measure far more pathological sets than simple unions of intervals—and that the resulting integral possesses limit-passage properties incomparably superior to Riemann's. The construction we present here applies Carathéodory's theorem to the volume content of rectangles, and produces the most fundamental measure in all of analysis.

4.1 Construction

Lebesgue measure is the unique measure on $\mathbb{R}^n$ that extends the notion of volume of rectangles. We construct it via Carathéodory's theorem.

Definition 4.1 (Rectangle and volume)**.** A *rectangle* (or *box*) in $\mathbb{R}^n$ is a set of the form $I = \prod_{k=1}^n [a_k, b_k)$ with $a_k \leq b_k$. Its *volume* is:

$$\mathrm{vol}(I) = \prod_{k=1}^{n} (b_k - a_k).$$

Definition 4.2 (Algebra of disjoint rectangles)**.** Let $\mathcal{A}_0$ denote the collection of finite disjoint unions of half-open rectangles in $\mathbb{R}^n$. This is an algebra.

Proposition 4.3. The function $\mu_0 : \mathcal{A}_0 \to [0, +\infty]$ defined by $\mu_0(\bigsqcup_{k=1}^m I_k) = \sum_{k=1}^m \mathrm{vol}(I_k)$ is a σ-finite pre-measure on $\mathcal{A}_0$.

Proof. Finite additivity is immediate by definition. The σ-additivity follows from continuity at $\varnothing$ (Proposition 2.11). Let (A_m) be a decreasing sequence in $\mathcal{A}_0$ with $A_m \downarrow \varnothing$. We show $\mu_0(A_m) \to 0$ by a compactness argument.

Suppose for contradiction that $\mu_0(A_m) \geq \alpha > 0$ for all m. Each A_m is a finite union of rectangles. By slightly shrinking each rectangle $[a_k, b_k)$ to $[a_k + \varepsilon, b_k - \varepsilon]$, we obtain a compact set $K_m \subset A_m$ with $\mu_0(A_m) - \mu_0(K_m)$ small. We can ensure $\mu_0(K_m) \geq \alpha/2 > 0$, so $K_m \neq \varnothing$. Since $K_1 \supset K_2 \supset \dots$ are nested non-empty compact sets, $\bigcap_m K_m \neq \varnothing$, contradicting $\bigcap_m A_m = \varnothing$.

σ-finiteness follows from $\mathbb{R}^n = \bigcup_{k=1}^\infty [-k, k)^n$. $\square$

Theorem 4.4 (Lebesgue measure). *There exists a unique measure λ^n on $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$ such that:*

$$\lambda^n\left(\prod_{k=1}^n [a_k, b_k)\right) = \prod_{k=1}^n (b_k - a_k).$$

This measure is called Lebesgue measure *on $\mathbb{R}^n$.*

Proof. Direct application of the Carathéodory extension theorem (Theorem 3.5) to the pre-measure μ_0. Uniqueness follows from σ-finiteness. □

Remark 4.5. We often write $\lambda = \lambda^1$ for Lebesgue measure on $\mathbb{R}$. The *complete Lebesgue measure* is the completion of λ^n with respect to the Lebesgue σ-algebra $\mathcal{L}^n \supset \mathcal{B}(\mathbb{R}^n)$.

4.2 Fundamental properties

Theorem 4.6 (Regularity). *For every $A \in \mathcal{B}(\mathbb{R}^n)$:*

1. *(Outer regularity) $\lambda^n(A) = \inf\{\lambda^n(U) : U \supset A,\ U \text{ open}\}$.*

2. *(Inner regularity) $\lambda^n(A) = \sup\{\lambda^n(K) : K \subset A,\ K \text{ compact}\}$.*

Proof. (1) Let $\varepsilon > 0$. By definition of the outer measure, there exists a sequence of rectangles (I_k) with $A \subset \bigcup_k I_k$ and $\sum_k \operatorname{vol}(I_k) \leq \lambda^n(A) + \varepsilon$. Replacing each $[a_j, b_j)$ by $(a_j - \delta, b_j + \delta)$ with δ small enough, we obtain an open set $U \supset A$ with $\lambda^n(U) \leq \lambda^n(A) + 2\varepsilon$.

(2) For $\lambda^n(A) < \infty$, use the approximation $A = \bigcup_m (A \cap [-m, m]^n)$ and outer regularity applied to the complement. □

Theorem 4.7 (Translation invariance). *For every $A \in \mathcal{B}(\mathbb{R}^n)$ and every $x \in \mathbb{R}^n$:*

$$\lambda^n(A + x) = \lambda^n(A), \quad \textit{where } A + x = \{a + x : a \in A\}.$$

Proof. The measure $\nu(A) = \lambda^n(A + x)$ is a measure on $\mathcal{B}(\mathbb{R}^n)$ that agrees with λ^n on rectangles (since volume is translation-invariant). By uniqueness, $\nu = \lambda^n$. □

Theorem 4.8 (Scaling behavior). *For every $A \in \mathcal{B}(\mathbb{R}^n)$ and every $c \in \mathbb{R} \setminus \{0\}$:*

$$\lambda^n(cA) = |c|^n \lambda^n(A).$$

More generally, if $T : \mathbb{R}^n \to \mathbb{R}^n$ is an invertible linear map:

$$\lambda^n(T(A)) = |\det T| \cdot \lambda^n(A).$$

Proof. For scaling, the measure $\nu(A) = \lambda^n(cA)/|c|^n$ agrees with λ^n on rectangles and is σ-finite, so $\nu = \lambda^n$ by uniqueness.

For the general case, reduce to shears and scalings (Gaussian elimination). □

4.3 Null sets

Proposition 4.9. The following sets have Lebesgue measure zero:

1. every countable set;
2. every affine subspace of dimension $< n$ in $\mathbb{R}^n$;
3. the Cantor set $\mathcal{C} \subset [0,1]$.

Proof. (1) A singleton $\{x\}$ is contained in $[x_1, x_1+\varepsilon) \times \ldots \times [x_n, x_n+\varepsilon)$ of volume ε^n, for every $\varepsilon > 0$. So $\lambda^n(\{x\}) = 0$. By countable subadditivity, every countable set has measure zero.

(2) By isometry invariance, it suffices to treat the hyperplane $\mathbb{R}^{n-1} \times \{0\}$. We have $\mathbb{R}^{n-1} \times \{0\} \subset \mathbb{R}^{n-1} \times [-\varepsilon, \varepsilon)$ for all $\varepsilon > 0$, and the latter has finite measure tending to zero.

(3) The Cantor set is the intersection $\mathcal{C} = \bigcap_n C_n$ where C_n is a union of 2^n intervals of length 3^{-n}. So $\lambda(\mathcal{C}) \le \lambda(C_n) = (2/3)^n \to 0$. □

4.4 Non-measurable sets

Theorem 4.10 (Vitali, 1905)**.** *There exists a subset of* $[0,1]$ *that is not Lebesgue-measurable.*

Proof. Define the equivalence relation $x \sim y \iff x - y \in \mathbb{Q}$ on $[0,1)$. By the axiom of choice, let $V \subset [0,1)$ contain exactly one representative from each equivalence class.

For $r \in \mathbb{Q} \cap [0,1)$, define $V_r = \{x + r \pmod 1 : x \in V\}$ (translation modulo 1). The sets V_r are pairwise disjoint and $[0,1) = \bigsqcup_{r \in \mathbb{Q}\cap[0,1)} V_r$.

If V were measurable, by translation invariance (modulo 1), all V_r would have the same measure $\alpha = \lambda(V)$. By σ-additivity:

$$1 = \lambda([0,1)) = \sum_{r \in \mathbb{Q}\cap[0,1)} \lambda(V_r) = \sum_{r \in \mathbb{Q}\cap[0,1)} \alpha.$$

If $\alpha = 0$, the sum equals $0 \neq 1$. If $\alpha > 0$, the sum equals $+\infty \neq 1$. Contradiction. □

Role of the axiom of choice

Vitali's construction uses the axiom of choice. Indeed, Solovay (1970) showed that it is consistent with ZF (without the axiom of choice) that every subset of $\mathbb{R}$ is Lebesgue-measurable.

4.5 The Cantor set

Definition 4.11 (Cantor set)**.** The Cantor set $\mathcal{C}$ is defined by:

$$\mathcal{C} = \bigcap_{n=0}^{\infty} C_n, \quad \text{where } C_0 = [0,1], \quad C_{n+1} = \frac{C_n}{3} \cup \left(\frac{2}{3} + \frac{C_n}{3}\right).$$

Equivalently, $\mathcal{C} = \{x \in [0,1] : x = \sum_{k=1}^{\infty} a_k 3^{-k},\ a_k \in \{0,2\}\}$.

 CHAPTER 4. LEBESGUE MEASURE ON $\mathbb{R}^n$

Proposition 4.12 (Properties of the Cantor set)**.** 1. $\mathcal{C}$ is compact, uncountable, perfect (closed with no isolated points), and totally disconnected.

2. $\lambda(\mathcal{C}) = 0$.

3. $\mathcal{C}$ is in bijection with $[0, 1]$ (via the Cantor function).

4. $\mathcal{C}$ contains non-Borel subsets (since $|\mathcal{P}(\mathcal{C})| = 2^{\mathfrak{c}} > \mathfrak{c} = |\mathcal{B}(\mathbb{R})|$).

4.6 Characterization of Lebesgue-measurable sets

Theorem 4.13. *A set $A \subset \mathbb{R}^n$ is Lebesgue-measurable if and only if there exist a Borel set B and a null set N such that $A = B \cup N$. In other words, $\mathcal{L}^n = \overline{\mathcal{B}(\mathbb{R}^n)}$ (the completion of the Borel σ-algebra).*

Theorem 4.14 (Characterization via G_δ and F_σ)**.** *Let $A \subset \mathbb{R}^n$ be Lebesgue-measurable with $\lambda^n(A) < \infty$. Then:*

1. *there exists a G_δ set (countable intersection of open sets) $G \supset A$ with $\lambda^n(G \setminus A) = 0$;*

2. *there exists an F_σ set (countable union of closed sets) $F \subset A$ with $\lambda^n(A \setminus F) = 0$.*

4.7 Exercises

Exercise 4.1. Show that $\lambda(\mathbb{Q}) = 0$ directly from the definition of Lebesgue measure (by constructing an explicit covering).

Exercise 4.2. Let $A \subset \mathbb{R}$ be measurable with $\lambda(A) > 0$. Show that the difference set $A - A = \{a - b : a, b \in A\}$ contains an interval centered at zero. (*Steinhaus' theorem.*)

Exercise 4.3 (Generalized Cantor set)**.** For $\alpha \in (0, 1)$, construct a closed set $\mathcal{C}_\alpha \subset [0, 1]$, homeomorphic to the Cantor set, with $\lambda(\mathcal{C}_\alpha) = 1 - \alpha$.

Exercise 4.4. Let $f : \mathbb{R} \to \mathbb{R}$ be Lipschitz with constant L. Show that for every $A \in \mathcal{B}(\mathbb{R})$, $\lambda(f(A)) \leq L \cdot \lambda(A)$.

Exercise 4.5. Show that Lebesgue measure λ is the only (complete) measure on $\mathbb{R}$ that is translation-invariant and satisfies $\lambda([0, 1]) = 1$.

Exercise 4.6. Show that there exists a Lebesgue-measurable set that is not Borel. *Hint*: use the surjection $\mathcal{C} \to [0, 1]$ and the fact that $\mathcal{C}$ has measure zero.

Chapter 5

Measurable Functions

Before we can integrate, we must know which functions are allowed to be integrated. This is the question of *measurability*: a function is measurable if the preimages of "reasonable" sets are themselves measurable. This condition, which may seem technical, is in fact very natural: it guarantees that one can unambiguously define the probability that $f(X)$ falls in a given interval, or the integral of f with respect to a measure. Continuous functions are always measurable, but the class of measurable functions is vastly larger—and it is precisely this flexibility that makes the Lebesgue integral so powerful.

5.1 Definitions and measurability criteria

Definition 5.1 (Measurable function)**.** Let $(X, \mathcal{A})$ be a measurable space. A function $f : X \to \overline{\mathbb{R}} = [-\infty, +\infty]$ is $\mathcal{A}$*-measurable* (or simply *measurable*) if:

$$\forall\, a \in \mathbb{R}, \quad f^{-1}([-\infty, a)) = \{x \in X : f(x) < a\} \in \mathcal{A}.$$

Proposition 5.2 (Equivalent criteria)**.** For $f : X \to \overline{\mathbb{R}}$, the following are equivalent:

1. $\{f < a\} \in \mathcal{A}$ for all $a \in \mathbb{R}$;
2. $\{f \leq a\} \in \mathcal{A}$ for all $a \in \mathbb{R}$;
3. $\{f > a\} \in \mathcal{A}$ for all $a \in \mathbb{R}$;
4. $\{f \geq a\} \in \mathcal{A}$ for all $a \in \mathbb{R}$.

Moreover, if any of these holds, then $\{f = a\} \in \mathcal{A}$, $\{f = +\infty\} \in \mathcal{A}$, and $\{f = -\infty\} \in \mathcal{A}$.

Proof. (1) ⇒ (2): $\{f \leq a\} = \bigcap_{n=1}^{\infty} \{f < a + 1/n\}$.

(2) ⇒ (3): $\{f > a\} = \{f \leq a\}^c$.

(3) ⇒ (4): $\{f \geq a\} = \bigcap_{n=1}^{\infty} \{f > a - 1/n\}$.

(4) ⇒ (1): $\{f < a\} = \{f \geq a\}^c$.

Finally, $\{f = a\} = \{f \leq a\} \cap \{f \geq a\}$, $\{f = +\infty\} = \bigcap_n \{f > n\}$, $\{f = -\infty\} = \bigcap_n \{f < -n\}$. □

Example 5.3.

1. Every constant function is measurable.
2. The indicator function $\mathbf{1}_A$ is measurable if and only if $A \in \mathcal{A}$.
3. If X is a topological space and $\mathcal{A} = \mathcal{B}(X)$, every continuous function $f : X \to \mathbb{R}$ is measurable.
4. Every monotone function $f : \mathbb{R} \to \mathbb{R}$ is Borel measurable.

5.2 Operations on measurable functions

Theorem 5.4 (Algebraic stability). *Let $f, g : X \to \mathbb{R}$ be measurable functions and $\alpha \in \mathbb{R}$. Then:*

1. *αf is measurable;*
2. *$f + g$ is measurable (when well-defined);*
3. *fg is measurable;*
4. *$|f|$ is measurable;*
5. *f/g is measurable (where $g \neq 0$);*
6. *f^+, f^-, $\max(f, g)$, $\min(f, g)$ are measurable.*

Proof. (2) The key observation is:

$$\{f + g < a\} = \bigcup_{r \in \mathbb{Q}} (\{f < r\} \cap \{g < a - r\}).$$

This is a countable union of measurable sets, hence measurable.

(3) Use the identity $fg = \frac{1}{4}[(f + g)^2 - (f - g)^2]$ and the fact that h^2 is measurable when h is (since $\{h^2 < a\} = \{-\sqrt{a} < h < \sqrt{a}\}$ for $a > 0$).

(4) $\{|f| < a\} = \{-a < f < a\} = \{f > -a\} \cap \{f < a\}$.

(6) $\max(f, g) = \frac{1}{2}(f + g + |f - g|)$ and $\min(f, g) = \frac{1}{2}(f + g - |f - g|)$. □

5.3 Limits of measurable functions

Theorem 5.5 (Stability under limits). *Let $(f_n)_{n \in \mathbb{N}}$ be a sequence of measurable functions $X \to \overline{\mathbb{R}}$. Then the following are measurable:*

1. *$\sup_n f_n$ and $\inf_n f_n$;*
2. *$\limsup_n f_n$ and $\liminf_n f_n$;*
3. *if $f = \lim_n f_n$ exists (pointwise), then f is measurable.*

Proof. (1) $\{\sup_n f_n > a\} = \bigcup_n \{f_n > a\} \in \mathcal{A}$. Similarly, $\{\inf_n f_n < a\} = \bigcup_n \{f_n < a\} \in \mathcal{A}$.

(2) $\limsup_n f_n = \inf_n \sup_{k \geq n} f_k$ and $\liminf_n f_n = \sup_n \inf_{k \geq n} f_k$, hence measurable by (1).

(3) If $f = \lim_n f_n$, then $f = \limsup_n f_n = \liminf_n f_n$. □

Contrast with continuity

A pointwise limit of continuous functions need not be continuous (e.g., $f_n(x) = x^n$ on $[0, 1]$). But a pointwise limit of measurable functions is *always* measurable. This is a major advantage of measurability over continuity.

5.4 Simple functions

Definition 5.6 (Simple function)**.** A function $\varphi : X \to \mathbb{R}$ is *simple* if it takes only finitely many values. It can be written as:

$$\varphi = \sum_{k=1}^{m} c_k \mathbf{1}_{A_k},$$

where $c_1, \ldots, c_m$ are the distinct values of φ and $A_k = \varphi^{-1}(\{c_k\})$ are the corresponding preimages. φ is measurable if and only if each $A_k \in \mathcal{A}$.

Theorem 5.7 (Approximation by simple functions)**.** *Let $f : X \to [0, +\infty]$ be a measurable function. There exists a sequence $(\varphi_n)_{n \geq 1}$ of measurable simple functions such that:*

1. $0 \leq \varphi_1 \leq \varphi_2 \leq \ldots$;
2. $\varphi_n(x) \uparrow f(x)$ *for every* $x \in X$;
3. *if f is bounded, the convergence is uniform.*

More generally, every measurable function $f : X \to \overline{\mathbb{R}}$ is the pointwise limit of a sequence of measurable simple functions.

Proof. For $n \geq 1$ and $x \in X$, set:

$$\varphi_n(x) = \begin{cases} k \cdot 2^{-n} & \text{if } k \cdot 2^{-n} \leq f(x) < (k+1) \cdot 2^{-n}, \quad k = 0, 1, \ldots, n \cdot 2^n - 1, \\ n & \text{if } f(x) \geq n. \end{cases}$$

This function is simple (taking at most $n \cdot 2^n + 1$ values) and measurable (each preimage is a measurable set of the form $\{k \cdot 2^{-n} \leq f < (k+1) \cdot 2^{-n}\}$).

Monotonicity: for every x, $\varphi_n(x) \leq \varphi_{n+1}(x)$ since the partition at step $n+1$ refines that at step n.

Convergence: if $f(x) < \infty$, for $n > f(x)$, we have $0 \leq f(x) - \varphi_n(x) < 2^{-n} \to 0$. If $f(x) = +\infty$, $\varphi_n(x) = n \to +\infty$.

Uniform convergence when f is bounded: if $f \leq M$, for $n > M$, $\sup_x |f(x) - \varphi_n(x)| \leq 2^{-n}$.

For general f, write $f = f^+ - f^-$ and approximate f^+ and f^- separately. ☐

Approximation construction

$$\varphi_n = \sum_{k=0}^{n2^n - 1} \frac{k}{2^n} \mathbf{1}_{\{k/2^n \leq f < (k+1)/2^n\}} + n \, \mathbf{1}_{\{f \geq n\}}.$$

5.5 Measurable functions and Borel functions

Proposition 5.8. Let $(X, \mathcal{A})$ be a measurable space. If $f : X \to \mathbb{R}$ is measurable and $g : \mathbb{R} \to \mathbb{R}$ is Borel measurable, then $g \circ f$ is measurable.

Proof. For every $B \in \mathcal{B}(\mathbb{R})$, $(g \circ f)^{-1}(B) = f^{-1}(g^{-1}(B))$. Since g is Borel, $g^{-1}(B) \in \mathcal{B}(\mathbb{R})$. Since f is $(\mathcal{A}, \mathcal{B}(\mathbb{R}))$-measurable, $f^{-1}(g^{-1}(B)) \in \mathcal{A}$. ☐

Corollary 5.9. *If f is measurable, then e^f, $\sin(f)$, f^p $(p > 0)$, $\log(|f|)$ (where $f \neq 0$) are measurable.*

CHAPTER 5. MEASURABLE FUNCTIONS

5.6 Almost everywhere convergence and convergence in measure

Definition 5.10 (Almost everywhere convergence)**.** We say $f_n \to f$ *μ-almost everywhere* (μ-a.e.) if

$$\mu(\{x \in X : f_n(x) \not\to f(x)\}) = 0.$$

Definition 5.11 (Convergence in measure)**.** We say $f_n \to f$ *in measure* if for every $\varepsilon > 0$:

$$\mu(\{x \in X : |f_n(x) - f(x)| > \varepsilon\}) \to 0 \quad (n \to \infty).$$

Theorem 5.12. *If $\mu(X) < \infty$ and $f_n \to f$ a.e., then $f_n \to f$ in measure.*

Proof. Set $A_n(\varepsilon) = \{x : |f_n(x) - f(x)| > \varepsilon\}$ and $B_N(\varepsilon) = \bigcup_{n \geq N} A_n(\varepsilon)$. Then $B_N(\varepsilon) \downarrow \{x : |f_n(x) - f(x)| > \varepsilon$ for infinitely many $n\} \subset \{f_n \not\to f\}$.

By hypothesis, $\mu(\{f_n \not\to f\}) = 0$, so $\mu(B_N(\varepsilon)) \to 0$. Since $A_N(\varepsilon) \subset B_N(\varepsilon)$, $\mu(A_N(\varepsilon)) \to 0$. □

Theorem 5.13 (Partial converse)**.** *If $f_n \to f$ in measure, there exists a subsequence (f_{n_k}) such that $f_{n_k} \to f$ a.e.*

Proof. For each k, choose n_k such that $\mu(\{|f_{n_k} - f| > 2^{-k}\}) < 2^{-k}$. By the Borel–Cantelli lemma (Lemma 2.5), $\mu(\limsup_k \{|f_{n_k} - f| > 2^{-k}\}) = 0$. Outside this set, $|f_{n_k} - f| \leq 2^{-k}$ for k large enough, so $f_{n_k} \to f$. □

Theorem 5.14 (Egorov)**.** *Let $\mu(X) < \infty$ and $f_n \to f$ μ-a.e. For every $\delta > 0$, there exists $A \in \mathcal{A}$ with $\mu(A^c) < \delta$ and $f_n \to f$ uniformly on A.*

Proof. For $m, n \in \mathbb{N}^*$, set $E_{m,n} = \bigcup_{k \geq n} \{|f_k - f| \geq 1/m\}$. For fixed m, $E_{m,n} \downarrow E_m$ where $E_m \subset \{f_n \not\to f\}$, so $\mu(E_m) = 0$. Thus $\mu(E_{m,n}) \to 0$ as $n \to \infty$.

Choose n_m so that $\mu(E_{m,n_m}) < \delta/2^m$. Set $A = X \setminus \bigcup_m E_{m,n_m}$. Then $\mu(A^c) \leq \sum_m \delta/2^m = \delta$. On A, for every m and every $k \geq n_m$, $|f_k - f| < 1/m$, which is uniform convergence. □

5.7 Lusin's theorem

Theorem 5.15 (Lusin)**.** *Let $f : \mathbb{R}^n \to \mathbb{R}$ be Lebesgue-measurable and $\varepsilon > 0$. There exists a closed set $F \subset \mathbb{R}^n$ with $\lambda^n(F^c) < \varepsilon$ and $f|_F$ continuous.*

Interpretation of Lusin's theorem

"Every measurable function is nearly continuous." More precisely, one can make a measurable function continuous by removing a set of arbitrarily small measure.

5.8 Exercises

Exercise 5.1. Show that if $f : \mathbb{R} \to \mathbb{R}$ is monotone, then f is Borel measurable.

Exercise 5.2. Let $f : X \to \mathbb{R}$ be measurable and $g : X \to \mathbb{R}$ with $f = g$ μ-a.e. Show that g need not be measurable (give a counterexample). However, show that if μ is complete, then g is measurable.

Exercise 5.3. Show that convergence in measure is metrizable when $\mu(X) < \infty$ via the distance $d(f,g) = \int_X \frac{|f-g|}{1+|f-g|}\, d\mu$.

Exercise 5.4. Construct a sequence (f_n) on $([0,1],\lambda)$ that converges in measure to 0 but converges at no point. (*Typewriter functions.*)

Exercise 5.5. Let (f_n) be a sequence of measurable non-negative functions. Show that $\{x : \sum_n f_n(x) < \infty\}$ is a measurable set.

CHAPTER 5. MEASURABLE FUNCTIONS

Chapter 6

Lebesgue Integral

We arrive at the heart of the edifice: the *Lebesgue integral*. While Riemann sliced the x-axis into small intervals and summed the areas of rectangles, Lebesgue had the ingenious idea of slicing the y-axis instead: group the points where the function takes similar values, measure these level sets, and sum. This approach, presented in his 1902 thesis, produces an integral that coincides with Riemann's when the latter exists, but applies to an incomparably larger class of functions—and above all, possesses convergence theorems (Fatou, dominated convergence, monotone convergence) of unmatched power.

6.1 Integral of simple functions

Definition 6.1 (Integral of a non-negative simple function)**.** Let $(X, \mathcal{A}, \mu)$ be a measure space and $\varphi = \sum_{k=1}^{m} c_k \mathbf{1}_{A_k}$ a non-negative measurable simple function ($c_k \geq 0$). We define:

$$\int_X \varphi \, d\mu = \sum_{k=1}^{m} c_k \, \mu(A_k),$$

with the convention $0 \cdot (+\infty) = 0$.

Proposition 6.2 (Properties of the integral of simple functions)**.** Let φ, ψ be non-negative measurable simple functions and $\alpha, \beta \geq 0$. Then:

1. (Linearity) $\int (\alpha\varphi + \beta\psi) \, d\mu = \alpha \int \varphi \, d\mu + \beta \int \psi \, d\mu$.
2. (Monotonicity) If $\varphi \leq \psi$, then $\int \varphi \, d\mu \leq \int \psi \, d\mu$.
3. If $\varphi = 0$ μ-a.e., then $\int \varphi \, d\mu = 0$.
4. For any $A \in \mathcal{A}$, $\int_A \varphi \, d\mu := \int_X \varphi \cdot \mathbf{1}_A \, d\mu$.

Proof. (1) Consider the common partition generated by the level sets of φ and ψ. If $\varphi = \sum_i a_i \mathbf{1}_{B_i}$ and $\psi = \sum_j b_j \mathbf{1}_{C_j}$ with $\{B_i\}$ and $\{C_j\}$ partitions, then the sets $B_i \cap C_j$ form a finer partition and

$$\alpha\varphi + \beta\psi = \sum_{i,j} (\alpha a_i + \beta b_j) \mathbf{1}_{B_i \cap C_j}.$$

The equality follows by direct computation.

(2) Write $\psi - \varphi \geq 0$ and use linearity. □

6.2 Integral of non-negative measurable functions

Definition 6.3 (Integral of a non-negative function)**.** Let $f : X \to [0, +\infty]$ be measurable. We define:

$$\int_X f\,d\mu = \sup\left\{\int_X \varphi\,d\mu : \varphi \text{ simple measurable}, 0 \le \varphi \le f\right\}.$$

Proposition 6.4 (Fundamental properties)**.** Let $f, g : X \to [0, +\infty]$ be measurable.

1. (Monotonicity) If $f \le g$, then $\int f\,d\mu \le \int g\,d\mu$.
2. (Homogeneity) For $c \ge 0$, $\int cf\,d\mu = c\int f\,d\mu$.
3. $\int f\,d\mu = 0$ if and only if $f = 0$ μ-a.e.
4. If $\mu(A) = 0$, then $\int_A f\,d\mu = 0$.

Proof. (3) If $f = 0$ a.e., every simple function $0 \le \varphi \le f$ satisfies $\varphi = 0$ a.e., so $\int \varphi\,d\mu = 0$.

Conversely, if $\int f\,d\mu = 0$, set $A_n = \{f > 1/n\}$. Then $\frac{1}{n}\mathbf{1}_{A_n} \le f$, so $\frac{1}{n}\mu(A_n) \le \int f\,d\mu = 0$, giving $\mu(A_n) = 0$ for all n. Since $\{f > 0\} = \bigcup_n A_n$, we get $\mu(\{f > 0\}) = 0$. □

Theorem 6.5 (Chebyshev / Markov inequality)**.** *Let $f : X \to [0, +\infty]$ be measurable. For every $a > 0$:*

$$\mu(\{f \ge a\}) \le \frac{1}{a}\int_X f\,d\mu.$$

Proof. $a \cdot \mathbf{1}_{\{f \ge a\}} \le f$, so $a \cdot \mu(\{f \ge a\}) = \int a \cdot \mathbf{1}_{\{f \ge a\}}\,d\mu \le \int f\,d\mu$. □

6.3 Additivity of the integral (non-negative functions)

Theorem 6.6. *Let $f, g : X \to [0, +\infty]$ be measurable. Then:*

$$\int_X (f + g)\,d\mu = \int_X f\,d\mu + \int_X g\,d\mu.$$

This result is a consequence of the monotone convergence theorem (Chapter 7). It can also be proved directly by approximation with simple functions.

6.4 Integral of integrable functions

Definition 6.7 (Integrable function)**.** A measurable function $f : X \to \overline{\mathbb{R}}$ is *integrable* (or *μ-integrable*, or *summable*) if:

$$\int_X |f|\,d\mu < \infty.$$

In this case, we define:

$$\int_X f\,d\mu = \int_X f^+\,d\mu - \int_X f^-\,d\mu,$$

where $f^+ = \max(f, 0)$ and $f^- = \max(-f, 0)$. We write $L^1(X, \mathcal{A}, \mu)$ (or $L^1(\mu)$) for the space of integrable functions.

The form $\infty - \infty$ is forbidden

The integrability condition $\int |f|\, d\mu < \infty$ guarantees that both $\int f^+\, d\mu$ and $\int f^-\, d\mu$ are finite, thus avoiding the indeterminate form $\infty - \infty$.

Theorem 6.8 (Properties of the integral). *Let $f, g \in L^1(\mu)$ and $\alpha, \beta \in \mathbb{R}$.*

1. *(Linearity) $\int (\alpha f + \beta g)\, d\mu = \alpha \int f\, d\mu + \beta \int g\, d\mu$.*
2. *(Triangle inequality) $\left|\int f\, d\mu\right| \le \int |f|\, d\mu$.*
3. *(Monotonicity) If $f \le g$ a.e., then $\int f\, d\mu \le \int g\, d\mu$.*
4. *If $f = g$ a.e., then $\int f\, d\mu = \int g\, d\mu$.*
5. *If $f \ge 0$ a.e. and $\int f\, d\mu = 0$, then $f = 0$ a.e.*

Proof. (2) We have $-|f| \le f \le |f|$, so by monotonicity $-\int |f|\, d\mu \le \int f\, d\mu \le \int |f|\, d\mu$. □

6.5 Comparison with the Riemann integral

Theorem 6.9. *Let $f : [a, b] \to \mathbb{R}$ be a bounded function.*

1. *If f is Riemann-integrable, then f is Lebesgue-measurable and the two integrals agree:*
$$(\mathcal{R})\int_a^b f(x)\, dx = \int_{[a,b]} f\, d\lambda.$$
2. *f is Riemann-integrable if and only if its set of discontinuities has Lebesgue measure zero (Lebesgue's criterion for Riemann integrability).*

Proof idea. (1) The lower and upper Darboux sums are integrals of simple functions with respect to λ. Taking the limit over partitions yields equality.

(2) Let D be the set of discontinuities of f. One shows that the upper and lower Darboux sums converge to the same value if and only if $\lambda(D) = 0$. □

Lebesgue integral — Summary

For $f : X \to [0, +\infty]$ measurable:
$$\int_X f\, d\mu = \sup\left\{\sum_{k=1}^m c_k \mu(A_k) : 0 \le \sum_k c_k \mathbf{1}_{A_k} \le f\right\}.$$

For general integrable f:
$$\int_X f\, d\mu = \int_X f^+\, d\mu - \int_X f^-\, d\mu.$$

6.6 The measure ν_f associated with an integrable function

Proposition 6.10. If $f \in L^1(\mu)$, $f \geq 0$, the set function $\nu_f(A) = \int_A f \, d\mu$ is a finite measure on $(X, \mathcal{A})$, absolutely continuous with respect to μ.

Proposition 6.11 (Absolute continuity of the integral)**.** Let $f \in L^1(\mu)$. For every $\varepsilon > 0$, there exists $\delta > 0$ such that:

$$\mu(A) < \delta \implies \int_A |f| \, d\mu < \varepsilon.$$

Proof. Set $f_n = \min(|f|, n)$. Then $f_n \uparrow |f|$ and by monotone convergence, $\int f_n \, d\mu \uparrow \int |f| \, d\mu < \infty$. Fix n such that $\int (|f| - f_n) \, d\mu < \varepsilon/2$. Then for $\mu(A) < \delta$ with $\delta = \varepsilon/(2n)$:

$$\int_A |f| \, d\mu \leq \int_A f_n \, d\mu + \int_A (|f| - f_n) \, d\mu \leq n \cdot \mu(A) + \varepsilon/2 < \varepsilon. \qquad \square$$

6.7 Exercises

Exercise 6.1. Compute $\int_{[0,1]} \mathbf{1}_{\mathbb{Q}} \, d\lambda$ and $\int_{[0,1]} \mathbf{1}_{\mathbb{R}\setminus\mathbb{Q}} \, d\lambda$.

Exercise 6.2. Show that if $f \in L^1(\mu)$, then $\{f \neq 0\}$ is σ-finite (for μ).

Exercise 6.3. Let $f : [0, \infty) \to [0, \infty)$ be measurable. Show that

$$\int_0^\infty f(x) \, dx = \int_0^\infty \lambda(\{f > t\}) \, dt.$$

(*Cavalieri's principle / layer cake formula.*)

Exercise 6.4. Let $f \in L^1(\mathbb{R}, \lambda)$. Show that $\lim_{n\to\infty} \int_{\mathbb{R}} f(x) \sin(nx) \, dx = 0$ (*Riemann–Lebesgue lemma*).

Exercise 6.5. Let $f \geq 0$ be measurable on $(X, \mathcal{A}, \mu)$. Show that $\nu(A) = \int_A f \, d\mu$ defines a measure on $(X, \mathcal{A})$ and that for every measurable $g \geq 0$: $\int_X g \, d\nu = \int_X g f \, d\mu$.

Exercise 6.6. Let $f : \mathbb{R} \to \mathbb{R}$ be Riemann-integrable on every $[a, b]$. Show that f is Lebesgue-measurable. Is f necessarily in $L^1(\mathbb{R}, \lambda)$?

Chapter 7

Convergence Theorems

7.1 Introduction

Here is the question every analyst eventually faces: when can one interchange a limit and an integral? The Riemann integral offers a spartan answer: under uniform convergence, and that is all. But in practice — Fourier series, probability theory, statistical mechanics — uniform convergence is a rare luxury. This is precisely where the Lebesgue integral reveals its full power. Between 1902 and 1910, Lebesgue, then Beppo Levi and Pierre Fatou, established three convergence theorems that form the beating heart of modern integration. Beppo Levi's monotone convergence theorem (1906) handles increasing sequences; Fatou's lemma (1906) provides a universal inequality; and Lebesgue's dominated convergence theorem crowns the edifice by allowing the limit-integral exchange whenever an integrable function "dominates" the sequence. These three results are, without exaggeration, the most frequently used tools in all of modern analysis.

7.2 Monotone convergence theorem (Beppo Levi)

Theorem 7.1 (Monotone convergence)**.** *Let* $(f_n)_{n\in\mathbb{N}}$ *be an* increasing *sequence of non-negative measurable functions:* $0 \le f_1 \le f_2 \le \dots$ *Let* $f = \lim_n f_n = \sup_n f_n$*. Then:*

$$\int_X f\,d\mu = \lim_{n\to\infty} \int_X f_n\,d\mu = \sup_{n\in\mathbb{N}} \int_X f_n\,d\mu.$$

Proof. Set $I = \lim_n \int f_n\,d\mu$ (the limit exists in $[0,+\infty]$ since the sequence $(\int f_n\,d\mu)$ is increasing by monotonicity of the integral).

Inequality $\int f\,d\mu \ge I$: since $f_n \le f$ for all n, $\int f_n\,d\mu \le \int f\,d\mu$, so $I \le \int f\,d\mu$.

Inequality $\int f\,d\mu \le I$: let φ be a measurable simple function with $0 \le \varphi \le f$ and $\alpha \in (0,1)$. Set $E_n = \{x : f_n(x) \ge \alpha\varphi(x)\}$. Then $E_n \uparrow X$ (since for every x, $f_n(x) \uparrow f(x) \ge \varphi(x) > \alpha\varphi(x)$ eventually). We have:

$$\int f_n\,d\mu \ge \int_{E_n} f_n\,d\mu \ge \alpha \int_{E_n} \varphi\,d\mu.$$

Letting $n \to \infty$, by monotone continuity of the measure $A \mapsto \int_A \varphi\,d\mu$:

$$I \ge \alpha \int_X \varphi\,d\mu.$$

CHAPTER 7. CONVERGENCE THEOREMS

Letting $\alpha \to 1^-$: $I \geq \int_X \varphi \, d\mu$. Taking the supremum over all simple functions $\varphi \leq f$: $I \geq \int f \, d\mu$. □

Monotone convergence

If $0 \leq f_n \uparrow f$, then $\displaystyle\int f_n \, d\mu \uparrow \int f \, d\mu$.

Corollary 7.2 (Series of non-negative functions). *Let $(g_n)_{n\in\mathbb{N}}$ be a sequence of non-negative measurable functions. Then:*

$$\int_X \sum_{n=0}^{\infty} g_n \, d\mu = \sum_{n=0}^{\infty} \int_X g_n \, d\mu.$$

Proof. Apply the monotone convergence theorem to $f_N = \sum_{n=0}^{N} g_n$. □

Corollary 7.3 (Additivity of the integral). *For measurable $f, g : X \to [0, +\infty]$: $\int (f + g) \, d\mu = \int f \, d\mu + \int g \, d\mu$.*

Proof. Let (φ_n) and (ψ_n) be increasing sequences of non-negative simple functions with $\varphi_n \uparrow f$ and $\psi_n \uparrow g$. Then $\varphi_n + \psi_n \uparrow f + g$ and $\int (\varphi_n + \psi_n) \, d\mu = \int \varphi_n \, d\mu + \int \psi_n \, d\mu$. Apply the monotone convergence theorem to both sides. □

7.3 Fatou's lemma

Theorem 7.4 (Fatou's lemma). *Let $(f_n)_{n\in\mathbb{N}}$ be a sequence of non-negative measurable functions. Then:*

$$\int_X \liminf_{n\to\infty} f_n \, d\mu \leq \liminf_{n\to\infty} \int_X f_n \, d\mu.$$

Proof. Set $g_n = \inf_{k\geq n} f_k$. Then g_n is non-negative measurable, $g_n \leq f_n$, and $g_n \uparrow \liminf_n f_n$. By the monotone convergence theorem:

$$\int \liminf_n f_n \, d\mu = \lim_n \int g_n \, d\mu \leq \liminf_n \int f_n \, d\mu,$$

the last inequality following from $\int g_n \, d\mu \leq \int f_n \, d\mu$. □

The inequality can be strict

Take $f_n = n \cdot \mathbf{1}_{(0,1/n)}$ on $(\mathbb{R}, \lambda)$. Then $f_n \to 0$ a.e. but $\int f_n \, d\lambda = 1$ for all n. So $0 = \int \liminf f_n \, d\lambda < \liminf \int f_n \, d\lambda = 1$.

Corollary 7.5 (Reverse Fatou lemma). *If (f_n) is a sequence of non-negative measurable functions and there exists $g \in L^1(\mu)$ with $f_n \leq g$ a.e. for all n, then:*

$$\limsup_{n\to\infty} \int_X f_n \, d\mu \leq \int_X \limsup_{n\to\infty} f_n \, d\mu.$$

Proof. Apply Fatou's lemma to the sequence $g - f_n \geq 0$. □

7.4 Dominated convergence theorem (Lebesgue)

Theorem 7.6 (Dominated convergence)**.** *Let (f_n) be a sequence of measurable functions $X \to \overline{\mathbb{R}}$ such that:*

(i) $f_n \to f$ μ-a.e. (pointwise convergence almost everywhere);

*(ii) there exists $g \in L^1(\mu)$, $g \geq 0$, with $|f_n| \leq g$ μ-a.e. for all n (*domination*).*

Then $f \in L^1(\mu)$, each $f_n \in L^1(\mu)$, and:

$$\lim_{n\to\infty} \int_X f_n \, d\mu = \int_X f \, d\mu.$$

Moreover, $\int_X |f_n - f| \, d\mu \to 0$.

Proof. Since $|f_n| \leq g$ a.e. and $f_n \to f$ a.e., we have $|f| \leq g$ a.e., so $f \in L^1(\mu)$.

Consider the functions $g + f_n \geq 0$ and $g - f_n \geq 0$. By Fatou's lemma:

$$\int (g+f)\, d\mu \leq \liminf_n \int (g+f_n)\, d\mu = \int g\, d\mu + \liminf_n \int f_n \, d\mu,$$
$$\int (g-f)\, d\mu \leq \liminf_n \int (g-f_n)\, d\mu = \int g\, d\mu - \limsup_n \int f_n \, d\mu.$$

Since $\int g \, d\mu < \infty$, we deduce:

$$\int f\, d\mu \leq \liminf_n \int f_n \, d\mu \leq \limsup_n \int f_n \, d\mu \leq \int f \, d\mu.$$

For the last assertion, apply dominated convergence to $|f_n - f| \leq 2g$ and $|f_n - f| \to 0$ a.e. □

Dominated convergence theorem

$$f_n \to f \text{ a.e.}, \quad |f_n| \leq g \in L^1 \implies \int f_n \, d\mu \to \int f \, d\mu.$$

7.5 Applications

Corollary 7.7 (Continuity under the integral sign)**.** *Let $f : X \times I \to \mathbb{R}$ where $I \subset \mathbb{R}$ is an interval. Suppose:*

1. *for every $t \in I$, $x \mapsto f(x,t)$ is measurable;*
2. *for μ-a.e. x, $t \mapsto f(x,t)$ is continuous at t_0;*
3. *there exists $g \in L^1(\mu)$ with $|f(x,t)| \leq g(x)$ for all $t \in I$ and μ-a.e. x.*

Then $F(t) = \int_X f(x,t)\, d\mu(x)$ is continuous at t_0.

Proof. If $t_n \to t_0$, then $f(x,t_n) \to f(x,t_0)$ a.e. with $|f(x,t_n)| \leq g(x)$. By dominated convergence, $F(t_n) \to F(t_0)$. □

Corollary 7.8 (Differentiation under the integral sign). *Let $f : X \times I \to \mathbb{R}$ where I is an open interval. Suppose:*

1. *for every $t \in I$, $x \mapsto f(x,t)$ is integrable;*
2. *for μ-a.e. x, $t \mapsto f(x,t)$ is differentiable on I;*
3. *there exists $g \in L^1(\mu)$ with $\left|\frac{\partial f}{\partial t}(x,t)\right| \leq g(x)$ for all $t \in I$ and μ-a.e. x.*

Then $F(t) = \int_X f(x,t)\, d\mu(x)$ is differentiable on I and:

$$F'(t) = \int_X \frac{\partial f}{\partial t}(x,t)\, d\mu(x).$$

Proof. For $h \neq 0$, $\frac{F(t+h)-F(t)}{h} = \int \frac{f(x,t+h)-f(x,t)}{h}\, d\mu(x)$. By the mean value theorem, $\left|\frac{f(x,t+h)-f(x,t)}{h}\right| \leq g(x)$. Letting $h \to 0$, dominated convergence gives the result. □

7.6 Convergence in L^1

Theorem 7.9 (Characterization of L^1 convergence). *Let $f_n, f \in L^1(\mu)$. The following are equivalent:*

1. *$\int |f_n - f|\, d\mu \to 0$ (L^1 convergence).*
2. *$f_n \to f$ in measure and the family (f_n) is uniformly integrable.*

Definition 7.10 (Uniform integrability). A family $(f_i)_{i \in I}$ of integrable functions is *uniformly integrable* if:

$$\lim_{M\to\infty} \sup_{i\in I} \int_{\{|f_i|>M\}} |f_i|\, d\mu = 0.$$

Theorem 7.11 (Vitali). *Let $\mu(X) < \infty$. If (f_n) is uniformly integrable and $f_n \to f$ in measure, then $f \in L^1(\mu)$ and $\int |f_n - f|\, d\mu \to 0$.*

7.7 Exercises

Exercise 7.1. Compute $\lim_{n\to\infty} \int_0^\infty \frac{n\sin(x/n)}{x(1+x^2)}\, dx$, justifying the interchange of limit and integral.

Exercise 7.2. Show that $F(\alpha) = \int_0^\infty \frac{e^{-\alpha x}}{1+x^2}\, dx$ is C^∞ on $(0,\infty)$ and compute $F'(\alpha)$.

Exercise 7.3. Let $f \in L^1(\mathbb{R})$. Show that $F(t) = \int_{\mathbb{R}} e^{itx} f(x)\, dx$ is continuous on $\mathbb{R}$ (this is the Fourier transform of f).

Exercise 7.4. Give an example of a sequence (f_n) in $L^1([0,1],\lambda)$ with $f_n \to 0$ a.e. but $\int f_n\, d\lambda \not\to 0$. Why does the dominated convergence theorem not apply?

Exercise 7.5. Compute $\lim_{n\to\infty} \int_0^1 (1+x/n)^n e^{-2x}\, dx$.

Exercise 7.6. Let $f \in L^1(\mathbb{R})$. Show that $\lim_{h\to 0} \int_{\mathbb{R}} |f(x+h) - f(x)|\, dx = 0$ (continuity of translation in L^1).

Chapter 8

L^p Spaces

The L^p spaces are the most important Banach spaces in analysis. Born from the Lebesgue integral, they provide the natural framework for Fourier analysis, partial differential equations, probability theory, and machine learning. Frigyes Riesz, in the 1910s, was the first to study these spaces systematically, establishing with Ernst Fischer the celebrated completeness theorem for L^2. The Hölder and Minkowski inequalities structure the theory; the duality $(L^p)^* = L^q$ (for $1/p + 1/q = 1$) crowns it. As for L^2, endowed with its inner product, it is the only one that is a Hilbert space, and it is upon L^2 that all of spectral theory rests.

8.1 Definitions

Definition 8.1 (L^p space)**.** Let $(X, \mathcal{A}, \mu)$ be a measure space and $1 \leq p < \infty$. The space $\mathcal{L}^p(\mu)$ is the set of measurable functions $f : X \to \mathbb{R}$ (or $\mathbb{C}$) such that:

$$\int_X |f|^p \, d\mu < \infty.$$

The L^p *norm* (or *semi-norm* on $\mathcal{L}^p$) is:

$$\|f\|_p = \left(\int_X |f|^p \, d\mu \right)^{1/p}.$$

The space $L^p(\mu) = \mathcal{L}^p(\mu)/\sim$ is the quotient by the equivalence relation $f \sim g \iff f = g$ μ-a.e.

Definition 8.2 (L^∞ space)**.** The space $L^\infty(\mu)$ is the set of equivalence classes of measurable essentially bounded functions, equipped with the norm:

$$\|f\|_\infty = \inf\{M \geq 0 : |f| \leq M \ \mu\text{-a.e.}\} = \operatorname{ess\,sup} |f| .$$

Remark 8.3. One works with equivalence classes (modulo a.e. equality) so that $\|\cdot\|_p$ is a genuine norm (not merely a semi-norm).

8.2 Hölder's inequality

Definition 8.4 (Conjugate exponents)**.** Two reals $p, q \in [1, +\infty]$ are *conjugate* (or *dual*) if:

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

By convention, $q = \infty$ if $p = 1$ and $q = 1$ if $p = \infty$.

Lemma 8.5 (Young's inequality)**.** *For $a, b \geq 0$ and conjugate $p, q > 1$:*

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

Proof. The function $t \mapsto \log t$ is concave, so for $\alpha = 1/p$ and $\beta = 1/q$:

$$\log(\alpha a^p + \beta b^q) \geq \alpha \log(a^p) + \beta \log(b^q) = \log(a) + \log(b) = \log(ab).$$

Exponentiating: $ab \leq \frac{a^p}{p} + \frac{b^q}{q}$. □

Theorem 8.6 (Hölder's inequality)**.** *Let $p, q \in [1, +\infty]$ be conjugate. If $f \in L^p(\mu)$ and $g \in L^q(\mu)$, then $fg \in L^1(\mu)$ and:*

$$\|fg\|_1 = \int_X |fg| \, d\mu \leq \|f\|_p \cdot \|g\|_q \, .$$

Proof. The case $p = 1$, $q = \infty$ is immediate: $|fg| \leq |f| \cdot \|g\|_\infty$ a.e.

For $1 < p < \infty$, we may assume $\|f\|_p = \|g\|_q = 1$ (otherwise normalize). By Young's inequality applied to $a = |f(x)|$ and $b = |g(x)|$:

$$|f(x)g(x)| \leq \frac{|f(x)|^p}{p} + \frac{|g(x)|^q}{q}.$$

Integrating:

$$\int |fg| \, d\mu \leq \frac{\|f\|_p^p}{p} + \frac{\|g\|_q^q}{q} = \frac{1}{p} + \frac{1}{q} = 1 = \|f\|_p \cdot \|g\|_q \, .$$ □

Corollary 8.7 (Cauchy–Schwarz inequality)**.** *For $p = q = 2$: $\int |fg| \, d\mu \leq \|f\|_2 \cdot \|g\|_2$.*

8.3 Minkowski's inequality

Theorem 8.8 (Minkowski's inequality)**.** *Let $1 \leq p \leq \infty$. For $f, g \in L^p(\mu)$:*

$$\|f + g\|_p \leq \|f\|_p + \|g\|_p \, .$$

That is, $\|\cdot\|_p$ is a norm on $L^p(\mu)$.

Proof. The case $p = 1$ is the triangle inequality. The case $p = \infty$ is clear.

For $1 < p < \infty$:

$$\begin{aligned} \int |f + g|^p \, d\mu &\leq \int |f + g|^{p-1} \left(|f| + |g| \right) d\mu \\ &= \int |f + g|^{p-1} |f| \, d\mu + \int |f + g|^{p-1} |g| \, d\mu. \end{aligned}$$

Apply Hölder to each term with $q = p/(p-1)$:

$$\int |f + g|^{p-1} |f| \, d\mu \leq \left(\int |f + g|^p \, d\mu \right)^{1/q} \|f\|_p \, .$$

Thus $\|f + g\|_p^p \leq \|f + g\|_p^{p/q} (\|f\|_p + \|g\|_p)$. Dividing by $\|f + g\|_p^{p/q}$ (if nonzero) and noting $p - p/q = 1$: $\|f + g\|_p \leq \|f\|_p + \|g\|_p$. □

Fundamental inequalities

Hölder: $\|fg\|_1 \le \|f\|_p \|g\|_q$, Minkowski: $\|f+g\|_p \le \|f\|_p + \|g\|_p$.

8.4 Completeness: L^p is a Banach space

Theorem 8.9 (Riesz–Fischer). *For every $1 \le p \le \infty$, the space $(L^p(\mu), \|\cdot\|_p)$ is a Banach space (complete normed vector space).*

Proof. Let (f_n) be a Cauchy sequence in L^p. It suffices to show it has a convergent subsequence (every Cauchy sequence with a convergent subsequence converges).

Choose n_k so that $\left\|f_{n_{k+1}} - f_{n_k}\right\|_p < 2^{-k}$. Set:

$$g_N = \sum_{k=1}^{N} \left|f_{n_{k+1}} - f_{n_k}\right|, \qquad g = \sum_{k=1}^{\infty} \left|f_{n_{k+1}} - f_{n_k}\right|.$$

By Minkowski, $\|g_N\|_p \le \sum_{k=1}^N 2^{-k} \le 1$. By monotone convergence, $\|g\|_p \le 1$, so $g < \infty$ a.e. This means the telescoping series $f_{n_1} + \sum_{k=1}^\infty (f_{n_{k+1}} - f_{n_k})$ converges absolutely a.e. Let f be its sum (a.e.). Then $f_{n_k} \to f$ a.e. and $|f_{n_k} - f| \le 2g \in L^p$, so by dominated convergence, $\|f_{n_k} - f\|_p \to 0$. □

8.5 Inclusion relations between L^p spaces

Proposition 8.10. If $\mu(X) < \infty$ and $1 \le p \le q \le \infty$, then $L^q(\mu) \subset L^p(\mu)$ and:

$$\|f\|_p \le \mu(X)^{1/p - 1/q} \|f\|_q.$$

Proof. Apply Hölder with $r = q/p \ge 1$ to $|f|^p$ and 1: $\int |f|^p \, d\mu \le \left(\int |f|^q \, d\mu\right)^{p/q} \mu(X)^{1-p/q}$. □

No inclusion in general

If μ is not finite, there is generally no inclusion between the L^p spaces. For example, on $(\mathbb{R}, \lambda)$: $f(x) = x^{-1/2}\mathbf{1}_{(0,1)} \in L^1 \setminus L^2$ and $g(x) = (1+x)^{-1} \in L^2(\mathbb{R}) \setminus L^1(\mathbb{R})$.

8.6 Density and separability

Theorem 8.11. *For $1 \le p < \infty$:*

1. *Integrable simple functions are dense in $L^p(\mu)$.*
2. *If $X = \mathbb{R}^n$ and $\mu = \lambda^n$, the continuous functions with compact support $C_c(\mathbb{R}^n)$ are dense in $L^p(\mathbb{R}^n)$.*
3. *$L^p(\mathbb{R}^n)$ is separable for $1 \le p < \infty$.*

Remark 8.12. L^∞ is generally *not* separable.

8.7 Duality of L^p spaces

Theorem 8.13 (L^p–L^q duality)**.** *Let $1 \leq p < \infty$ and q the conjugate exponent. For every continuous linear functional $\ell : L^p(\mu) \to \mathbb{R}$, there exists a unique $g \in L^q(\mu)$ such that:*

$$\ell(f) = \int_X fg\,d\mu, \quad \forall f \in L^p(\mu).$$

Moreover, $\|\ell\| = \|g\|_q$. In other words, $(L^p(\mu))^* \cong L^q(\mu)$ (isometrically).

Remark 8.14. This theorem requires μ to be σ-finite. For $p = 1$, $(L^1)^* \cong L^\infty$. For $p = \infty$, $(L^\infty)^* \supsetneq L^1$ in general.

8.8 L^2 as a Hilbert space

Definition 8.15 (L^2 inner product)**.** The space $L^2(\mu)$ is equipped with the inner product:

$$\langle f, g\rangle = \int_X f\overline{g}\,d\mu.$$

It is complete for the associated norm $\|f\|_2 = \sqrt{\langle f, f\rangle}$, hence it is a Hilbert space.

Theorem 8.16 (Orthogonal projection)**.** *If V is a closed subspace of $L^2(\mu)$ and $f \in L^2(\mu)$, there exists a unique $\hat{f} \in V$ such that:*

$$\left\|f - \hat{f}\right\|_2 = \inf_{g\in V} \|f - g\|_2 .$$

Moreover, $f - \hat{f} \perp V$ (i.e. $\langle f - \hat{f}, g\rangle = 0$ for all $g \in V$).

8.9 Complementary inequalities

Proposition 8.17 (Jensen's inequality)**.** Let $(X, \mathcal{A}, \mu)$ be a probability space ($\mu(X) = 1$), $f \in L^1(\mu)$, and $\varphi : \mathbb{R} \to \mathbb{R}$ convex. If $\varphi \circ f$ is integrable, then:

$$\varphi\left(\int_X f\,d\mu\right) \leq \int_X \varphi \circ f\,d\mu.$$

Proposition 8.18 (Interpolation inequality)**.** If $f \in L^{p_1} \cap L^{p_2}$ with $1 \leq p_1 < p_2 \leq \infty$, then $f \in L^p$ for every $p \in [p_1, p_2]$ and:

$$\|f\|_p \leq \|f\|_{p_1}^{\theta} \|f\|_{p_2}^{1-\theta}, \quad \text{where } \frac{1}{p} = \frac{\theta}{p_1} + \frac{1-\theta}{p_2}.$$

8.10 Exercises

Exercise 8.1. Show that for $1 \leq p < q < \infty$, $\ell^p \subset \ell^q$ (where $\ell^p = L^p(\mathbb{N}, \mathcal{P}(\mathbb{N}), \#)$ with $\#$ the counting measure). This is the reverse of the inclusion for finite measure spaces.

Exercise 8.2. Show that $L^2([0, 1])$ is a separable Hilbert space and that $(e_n)_{n\in\mathbb{Z}}$ defined by $e_n(x) = e^{2\pi inx}$ is an orthonormal basis.

Exercise 8.3. Prove the integral form of Minkowski's inequality: if $f \geq 0$ is measurable on $X \times Y$ and $1 \leq p < \infty$, then

$$\left(\int_X \left(\int_Y f(x,y)\,d\nu(y)\right)^p d\mu(x)\right)^{1/p} \leq \int_Y \left(\int_X f(x,y)^p\,d\mu(x)\right)^{1/p} d\nu(y).$$

Exercise 8.4. Let $(f_n) \subset L^p(\mu)$ with $\sum_n \|f_n\|_p < \infty$. Show that $\sum_n f_n$ converges a.e. and in L^p.

Exercise 8.5. Show that if $f_n \to f$ in L^p and $f_n \to g$ a.e., then $f = g$ a.e.

CHAPTER 8. L^p SPACES

Chapter 9

Signed Measures and Hahn-Jordan Decomposition

9.1 Introduction

Until now, measures have been positive quantities: area, volume, probability. But physics and economics abound with quantities that change sign: an electric charge can be positive or negative, a financial balance can be profit or loss, the difference of two probability measures is no longer a positive measure. Johann Radon and Otton Nikodym, in the 1910s–1930s, showed that every signed measure decomposes canonically into a positive part and a negative part (Jordan decomposition), concentrated on disjoint sets (Hahn decomposition). The Radon–Nikodym theorem, which expresses an absolutely continuous measure with respect to another as a weighted integral, crowns this theory and provides the rigorous definition of the probability density function.

Why signed measures?

Consider an electric charge distribution on a surface: some regions carry positive charge, others negative. The total "charge measure" of a set can be positive, negative, or zero. This is exactly what a signed measure models. Moreover, the difference of two (positive) finite measures is always a signed measure.

9.2 Definitions and first examples

Definition 9.1 (Signed measure)**.** Let $(X, \mathcal{A})$ be a measurable space. A *signed measure* on $(X, \mathcal{A})$ is a function $\nu : \mathcal{A} \to [-\infty, +\infty]$ satisfying:

(i) $\nu(\varnothing) = 0$;

(ii) ν assumes *at most* one of the values $+\infty$ and $-\infty$;

(iii) (σ*-additivity*) for every sequence $(A_n)_{n\geq 1}$ of mutually disjoint sets in $\mathcal{A}$:

$$\nu\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} \nu(A_n),$$

where the series converges absolutely if the left-hand side is finite.

Remark 9.2. Condition (ii) is essential to avoid indeterminate expressions of the form $+\infty - \infty$. Every positive measure is a signed measure (special case).

Example 9.3.

1. If μ and λ are positive measures on $(X, \mathcal{A})$ with λ finite, then $\nu = \mu - \lambda$ is a signed measure.
2. If $f \in L^1(\mu)$, then $\nu(A) = \int_A f \, d\mu$ is a finite signed measure.

Definition 9.4 (Positive, negative, and null sets)**.** Let ν be a signed measure on $(X, \mathcal{A})$.

- A set $A \in \mathcal{A}$ is *positive* for ν if $\nu(B) \geq 0$ for every $B \in \mathcal{A}$ with $B \subset A$.
- A set A is *negative* for ν if $\nu(B) \leq 0$ for every $B \in \mathcal{A}$ with $B \subset A$.
- A set A is *null* for ν if $\nu(B) = 0$ for every $B \in \mathcal{A}$ with $B \subset A$.

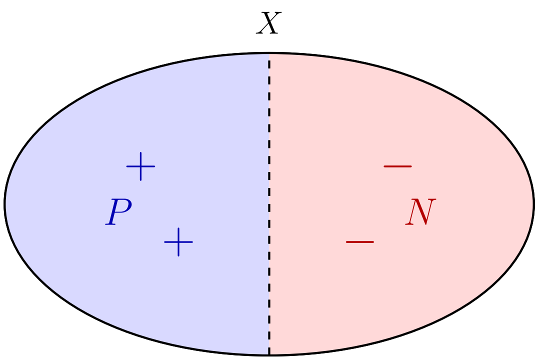
Positive for $\nu \neq \nu(A) \geq 0$

A set A is positive for ν if *all* its measurable subsets have nonnegative signed measure. The fact that $\nu(A) \geq 0$ alone does not suffice: there may exist $B \subset A$ with $\nu(B) < 0$.

9.3 Hahn decomposition

Lemma 9.5. *Let ν be a signed measure on $(X, \mathcal{A})$ and let $A \in \mathcal{A}$ with $-\infty < \nu(A) < 0$. Then there exists a negative set $B \subset A$ with $\nu(B) \leq \nu(A)$.*

Proof. If A is already negative, take $B = A$. Otherwise, there exists $A_1 \subset A$ with $\nu(A_1) > 0$. Let n_1 be the smallest integer such that some measurable subset of A has ν-value $\geq 1/n_1$. Iterate: in $A \setminus A_1$, find A_2 with $\nu(A_2) \geq 1/n_2$, etc. Set $B = A \setminus \bigcup_k A_k$. Then $\nu(B) = \nu(A) - \sum_k \nu(A_k) \leq \nu(A)$, and $\sum_k 1/n_k < \infty$, so $1/n_k \to 0$. If $C \subset B$ had $\nu(C) > 0$, we could find n with $\nu(C) \geq 1/n$, contradicting the choice of n_k when $1/n_k < 1/n$. Hence B is negative. □



Hahn decomposition: $X = P \cup N, \quad \nu|_P \geq 0, \quad \nu|_N \leq 0$

Theorem 9.6 (Hahn Decomposition)**.** *Let ν be a signed measure on $(X, \mathcal{A})$. There exists a partition $X = P \cup N$ with $P, N \in \mathcal{A}$, $P \cap N = \varnothing$, such that P is a positive set and N is a negative set for ν.*

Moreover, this decomposition is essentially unique*: if $X = P' \cup N'$ is another Hahn decomposition, then $P \triangle P'$ (and $N \triangle N'$) is a null set for ν.*

Proof. Assume ν does not take the value $+\infty$ (the symmetric case is analogous). Let $\lambda = \inf\{\nu(A) : A \in \mathcal{A}\} \in [-\infty, 0]$. For each $n \geq 1$, choose A_n with $\nu(A_n) < \lambda + 1/n$. By Lemma 9.5, there exists a negative set $B_n \subset A_n$ with $\nu(B_n) \leq \nu(A_n)$.

Set $N = \bigcup_n B_n$. As a countable union of negative sets, N is negative. Moreover, $\nu(N) \leq \nu(B_n) < \lambda + 1/n$ for all n, so $\nu(N) = \lambda$. Set $P = X \setminus N$. If P were not positive, there would exist $A \subset P$ with $\nu(A) < 0$, yielding a negative subset $B \subset A$ with $\nu(N \cup B) = \nu(N) + \nu(B) < \lambda$, a contradiction.

For essential uniqueness: $P \cap N' \subset P$ (positive) and $P \cap N' \subset N'$ (negative), so $P \cap N'$ is null. Similarly $P' \cap N$ is null. □

9.4 Jordan decomposition and total variation

Definition 9.7 (Jordan decomposition)**.** Let ν be a signed measure and $X = P \cup N$ a Hahn decomposition. The *positive and negative parts* of ν are:

$$\nu^+(A) = \nu(A \cap P), \qquad \nu^-(A) = -\nu(A \cap N),$$

for all $A \in \mathcal{A}$. Then $\nu = \nu^+ - \nu^-$.

Theorem 9.8 (Jordan Decomposition)**.** *The measures ν^+ and ν^- are positive measures on $(X, \mathcal{A})$, at least one of them is finite, and they do not depend on the choice of Hahn decomposition. Moreover, $\nu^+ \perp \nu^-$ (mutual singularity).*

Definition 9.9 (Total variation)**.** The *total variation* of ν is the positive measure:

$$|\nu| = \nu^+ + \nu^-.$$

The *total variation of ν on X* is $\|\nu\| = |\nu|(X)$.

Proposition 9.10 (Properties of total variation)**.** For any signed measure ν and any $A \in \mathcal{A}$:

1. $|\nu(A)| \leq |\nu|(A)$;
2. $|\nu|(A) = \sup\left\{\sum_{i=1}^n |\nu(A_i)| : A = \bigsqcup_{i=1}^n A_i,\ A_i \in \mathcal{A}\right\}$;
3. ν is a finite signed measure if and only if $|\nu|(X) < \infty$.

Proof. (1) $|\nu(A)| = |\nu^+(A) - \nu^-(A)| \leq \nu^+(A) + \nu^-(A) = |\nu|(A)$.

(2) For any partition $A = \bigsqcup_i A_i$, $\sum_i |\nu(A_i)| \leq \sum_i |\nu|(A_i) = |\nu|(A)$. The reverse inequality follows by taking the partition $A = (A \cap P) \sqcup (A \cap N)$.

(3) Immediate since $\|\nu\| = \nu^+(X) + \nu^-(X)$. □

Summary: Jordan decomposition

For any signed measure ν:

$$\nu = \nu^+ - \nu^-, \qquad |\nu| = \nu^+ + \nu^-,$$
$$\nu^+ = \frac{|\nu| + \nu}{2}, \qquad \nu^- = \frac{|\nu| - \nu}{2}.$$

9.5 Mutual singularity

Definition 9.11 (Mutually singular measures)**.** Two signed measures μ and ν on $(X, \mathcal{A})$ are *mutually singular*, written $\mu \perp \nu$, if there exists a partition $X = A \cup B$ with $A, B \in \mathcal{A}$ such that $|\mu|(B) = 0$ and $|\nu|(A) = 0$.

Example 9.12.

1. Lebesgue measure λ on $\mathbb{R}$ and the Dirac measure δ_0 are mutually singular: $\lambda(\{0\}) = 0$ and $\delta_0(\mathbb{R} \setminus \{0\}) = 0$.

2. Lebesgue measure and the Cantor–Vitali measure on $[0, 1]$ are mutually singular.

9.6 The Banach space of finite signed measures

Proposition 9.13. The set $\mathcal{M}(X, \mathcal{A})$ of finite signed measures on $(X, \mathcal{A})$ is a real vector space, and $\|\nu\| = |\nu|(X)$ is a norm. Moreover, $(\mathcal{M}(X, \mathcal{A}), \|\cdot\|)$ is a Banach space.

9.7 Exercises

Exercise 9.1. Let ν be a finite signed measure on $(X, \mathcal{A})$. Show that ν is bounded, i.e. $\sup_{A \in \mathcal{A}} |\nu(A)| < \infty$.

Exercise 9.2. Let $f \in L^1(\mathbb{R}, \lambda)$ where λ is Lebesgue measure. Show that $\nu(A) = \int_A f \, d\lambda$ defines a signed measure and compute ν^+, ν^-, $|\nu|$ in terms of f. *Hint*: show that $\nu^+(A) = \int_A f^+ \, d\lambda$ and $\nu^-(A) = \int_A f^- \, d\lambda$.

Exercise 9.3. Show that two positive measures μ and ν are mutually singular if and only if $\inf_{A \in \mathcal{A}} [\mu(A) + \nu(A^c)] = 0$.

Exercise 9.4. Let ν be a signed measure and (A_n) an increasing sequence in $\mathcal{A}$. Show that $\nu(\bigcup_n A_n) = \lim_n \nu(A_n)$ (monotone continuity for signed measures).

Exercise 9.5 (Explicit Hahn decomposition)**.** Let ν be the signed measure on $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ defined by $\nu(A) = \int_A (x^2 - 1) e^{-x^2} \, d\lambda(x)$. Find a Hahn decomposition of $\mathbb{R}$ for ν and deduce ν^+, ν^-.

Exercise 9.6. Let ν be a finite signed measure. Show that for every $\varepsilon > 0$, the collection $\{A \in \mathcal{A} : |\nu(A)| > \varepsilon\}$ is finite. *Hint*: use $|\nu|(X) < \infty$.

Chapter 10

Radon-Nikodym Theorem

10.1 Introduction

This chapter establishes one of the most important results in measure theory: the Radon-Nikodym theorem, which characterizes measures that are absolutely continuous with respect to a given measure. This theorem is fundamental in probability (densities, conditional expectation) and functional analysis.

Radon-Nikodym: changing the reference measure

If ν is absolutely continuous with respect to μ, it means that ν only "sees" what μ sees: sets negligible for μ are also negligible for ν. The Radon-Nikodym theorem then asserts that ν is obtained by "weighting" μ with a density: $d\nu = f\,d\mu$. This generalizes the fact that any probability on $\mathbb{R}$ with a density can be written as $\mathbb{P}(A) = \int_A f(x)\,dx$.

10.2 Absolute continuity

Definition 10.1 (Absolute continuity)**.** Let μ be a positive measure and ν a signed measure on $(X, \mathcal{A})$. We say that ν is *absolutely continuous* with respect to μ, written $\nu \ll \mu$, if:

$$\forall A \in \mathcal{A}, \quad \mu(A) = 0 \implies \nu(A) = 0.$$

Proposition 10.2 (ε-δ characterization)**.** If μ is a positive measure and ν is a *finite* signed measure, then $\nu \ll \mu$ if and only if:

$$\forall \varepsilon > 0, \exists\, \delta > 0, \forall A \in \mathcal{A}, \quad \mu(A) < \delta \implies |\nu(A)| < \varepsilon.$$

Proof. ($\Leftarrow$) Clear: if $\mu(A) = 0$, then $\mu(A) < \delta$ for every δ, giving $|\nu(A)| < \varepsilon$ for every ε, hence $\nu(A) = 0$.

($\Rightarrow$) By contradiction. Suppose $\nu \ll \mu$ but the ε-δ condition fails. There exist $\varepsilon_0 > 0$ and a sequence (A_n) with $\mu(A_n) < 2^{-n}$ and $|\nu(A_n)| \geq \varepsilon_0$. Set $B_n = \bigcup_{k \geq n} A_k$. Then $\mu(B_n) \leq 2^{1-n}$ and $B = \bigcap_n B_n$ satisfies $\mu(B) = 0$. By decreasing continuity, $|\nu|(B) = \lim_n |\nu|(B_n) \geq \varepsilon_0 > 0$, contradicting $\nu \ll \mu$. $\square$

Example 10.3. 1. If $f \geq 0$ is measurable and $\nu(A) = \int_A f\,d\mu$, then $\nu \ll \mu$.

 CHAPTER 10. RADON-NIKODYM THEOREM

2. The Dirac measure δ_0 is not absolutely continuous with respect to Lebesgue measure λ: $\lambda(\{0\}) = 0$ but $\delta_0(\{0\}) = 1$.

3. If $\nu \ll \mu$ and $\mu \ll \nu$, we say μ and ν are *equivalent*, written $\mu \sim \nu$.

10.3 The Radon-Nikodym theorem

Theorem 10.4 (Radon-Nikodym)**.** *Let $(X, \mathcal{A}, \mu)$ be a σ-finite measure space and let ν be a σ-finite signed measure on $(X, \mathcal{A})$ with $\nu \ll \mu$. Then there exists a measurable function $f : X \to [-\infty, +\infty]$, integrable with respect to μ (if ν is finite), such that:*

$$\nu(A) = \int_A f \, d\mu, \qquad \forall A \in \mathcal{A}.$$

Moreover, f is unique μ-almost everywhere. If ν is a positive measure, then $f \geq 0$ μ-a.e.

Proof (case μ, ν finite positive measures). We present Von Neumann's proof. Set $\lambda = \mu + \nu$, a finite positive measure. For every $g \in L^2(\lambda)$, the linear functional $\varphi(g) = \int g \, d\nu$ is continuous on $L^2(\lambda)$, since by Cauchy–Schwarz:

$$|\varphi(g)| \leq \int |g| \, d\lambda \leq \left(\textstyle\int g^2 \, d\lambda\right)^{1/2} \lambda(X)^{1/2}.$$

By the Riesz representation theorem for Hilbert spaces, there exists $h \in L^2(\lambda)$ such that:

$$\int g \, d\nu = \int gh \, d\lambda, \qquad \forall g \in L^2(\lambda).$$

Taking $g = \mathbb{1}_A$ gives $\nu(A) = \int_A h \, d\lambda$ for all $A \in \mathcal{A}$. Since $0 \leq \nu(A) \leq \lambda(A)$, we deduce $0 \leq h \leq 1$ λ-a.e.

We can write $\nu(A) = \int_A h \, d\mu + \int_A h \, d\nu$, hence $\int_A (1-h) \, d\nu = \int_A h \, d\mu$. Setting $E_0 = \{h = 1\}$, one checks $\mu(E_0) = 0$. Define $f = h/(1-h)$ on E_0^c and $f = 0$ on E_0. The monotone convergence theorem applied to $g_n = (1 + h + \cdots + h^n)\mathbb{1}_A$ yields:

$$\nu(A) = \int_A f \, d\mu.$$

Uniqueness of f μ-a.e. follows from the fact that if $\int_A f \, d\mu = \int_A g \, d\mu$ for all A, then $f = g$ μ-a.e. □

Definition 10.5 (Radon-Nikodym derivative)**.** The function f from Theorem 10.4 is called the *Radon-Nikodym derivative* of ν with respect to μ and is denoted:

$$f = \frac{d\nu}{d\mu}.$$

Proposition 10.6 (Properties of the Radon-Nikodym derivative)**.** Let μ, ν, λ be σ-finite measures.

1. **Chain rule:** if $\nu \ll \mu \ll \lambda$, then $\dfrac{d\nu}{d\lambda} = \dfrac{d\nu}{d\mu} \cdot \dfrac{d\mu}{d\lambda}$ λ-a.e.

2. **Inverse:** if $\nu \ll \mu$ and $\mu \ll \nu$, then $\dfrac{d\mu}{d\nu} = \left(\dfrac{d\nu}{d\mu}\right)^{-1}$ μ-a.e.

3. **Linearity:** if $\nu_1, \nu_2 \ll \mu$, then $\dfrac{d(\alpha\nu_1 + \beta\nu_2)}{d\mu} = \alpha\dfrac{d\nu_1}{d\mu} + \beta\dfrac{d\nu_2}{d\mu}$.

Key Radon-Nikodym formulas

If $\nu \ll \mu$ with $f = \frac{d\nu}{d\mu}$, then for every measurable $g \geq 0$ or $g \in L^1(\nu)$:

$$\int g \, d\nu = \int g \cdot f \, d\mu = \int g \cdot \frac{d\nu}{d\mu} \, d\mu.$$

Chain rule: $\dfrac{d\nu}{d\lambda} = \dfrac{d\nu}{d\mu} \cdot \dfrac{d\mu}{d\lambda}$.

10.4 Lebesgue decomposition theorem

Theorem 10.7 (Lebesgue Decomposition). *Let $(X, \mathcal{A}, \mu)$ be a σ-finite measure space and let ν be a σ-finite signed measure. Then there exists a unique pair of signed measures (ν_a, ν_s) such that:*

$$\nu = \nu_a + \nu_s, \qquad \nu_a \ll \mu, \qquad \nu_s \perp \mu.$$

Moreover, $\nu_a(A) = \int_A f \, d\mu$ for some measurable f (the Radon-Nikodym derivative of ν_a with respect to μ).

Proof sketch. It suffices to treat the case where ν is a finite positive measure and μ is finite. For each $\lambda > 0$, the signed measure $\nu - \lambda\mu$ admits a Hahn decomposition: $X = P_\lambda \cup N_\lambda$ with P_λ positive and N_λ negative for $\nu - \lambda\mu$. One may choose P_λ increasing in λ (replacing P_λ by $\bigcup_{\lambda' \leq \lambda} P_{\lambda'}$ if needed). Set $\lambda_0 = \sup\{\lambda > 0 : \nu(N_\lambda) > 0\}$ and take a sequence $\lambda_n \nearrow \lambda_0$. The set $N = \bigcap_n N_{\lambda_n}$ satisfies $\mu(N) = 0$: indeed, for every n and every $A \subseteq P_{\lambda_n}^c$, $\nu(A) \leq \lambda_n \mu(A)$, and passing to the limit gives $\mu(N) = 0$. Set $\nu_s(\cdot) = \nu(\cdot \cap N)$ and $\nu_a(\cdot) = \nu(\cdot \cap N^c)$. Since $\mu(N) = 0$, we have $\nu_s \perp \mu$. To show $\nu_a \ll \mu$: if $\mu(A) = 0$ with $A \subseteq N^c$, then $A \subseteq P_{\lambda_n}$ for n large enough, and $\nu(A) \leq \lambda_n \mu(A) = 0$. The Radon-Nikodym theorem applied to ν_a provides the density f. See Rudin, *Real and Complex Analysis*, Theorem 6.10. ☐

10.5 Applications to probability

Example 10.8 (Probability density). Let X be a real random variable on $(\Omega, \mathcal{F}, \mathbb{P})$ whose law $\mathbb{P}_X$ is absolutely continuous with respect to Lebesgue measure λ. Then there exists $f_X \geq 0$ with $\int f_X \, d\lambda = 1$ such that:

$$\mathbb{P}(X \in A) = \int_A f_X(x) \, dx, \qquad \forall A \in \mathcal{B}(\mathbb{R}).$$

The function $f_X = \frac{d\mathbb{P}_X}{d\lambda}$ is the *density* of X.

Example 10.9 (Conditional expectation). Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, $\mathcal{G} \subset \mathcal{F}$ a sub-σ-algebra, and $Y \in L^1(\mathbb{P})$. The signed measure $\nu(A) = \int_A Y \, d\mathbb{P}$ for $A \in \mathcal{G}$ is absolutely continuous with respect to $\mathbb{P}|_{\mathcal{G}}$. By Radon-Nikodym, there exists a unique (a.s.) $\mathcal{G}$-measurable function $Z \in L^1(\mathbb{P})$ with:

$$\int_A Z \, d\mathbb{P} = \int_A Y \, d\mathbb{P}, \qquad \forall A \in \mathcal{G}.$$

This function Z is the *conditional expectation* $\mathbb{E}[Y|\mathcal{G}]$.

Example 10.10 (Likelihood ratio)**.** Let $\mathbb{P}$ and $\mathbb{Q}$ be two probability measures on $(\Omega, \mathcal{F})$ with $\mathbb{Q} \ll \mathbb{P}$. The *likelihood ratio* is $L = \frac{d\mathbb{Q}}{d\mathbb{P}}$, and for every $\mathbb{Q}$-integrable random variable Y:

$$\mathbb{E}_{\mathbb{Q}}[Y] = \mathbb{E}_{\mathbb{P}}[Y \cdot L].$$

This is the foundation of change of measure in finance (Girsanov's theorem) and statistics (Neyman–Pearson test).

10.6 Exercises

Exercise 10.1. Let μ be Lebesgue measure on $[0, 1]$ and $\nu(A) = \int_A 2x\,dx$. Verify that $\nu \ll \mu$, identify $\frac{d\nu}{d\mu}$, and compute $\int_{[0,1]} x^2\,d\nu$.

Exercise 10.2. Show that if $\nu \ll \mu$ and $\nu \perp \mu$, then $\nu = 0$.

Exercise 10.3. Let μ and ν be positive σ-finite measures with $\nu \ll \mu$. Show that $|\nu|(A) = \int_A \left| \frac{d\nu}{d\mu} \right| d\mu$.

Exercise 10.4 (Chain rule)**.** Let λ be Lebesgue measure, $\mu(A) = \int_A e^{-x} \mathbb{1}_{[0,\infty)}(x)\,d\lambda(x)$, and $\nu(A) = \int_A xe^{-x} \mathbb{1}_{[0,\infty)}(x)\,d\lambda(x)$. Compute $\frac{d\nu}{d\mu}$ and $\frac{d\mu}{d\lambda}$. Verify the chain rule.

Exercise 10.5. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and X a random variable with $\mathbb{E}[|X|] < \infty$. Let $\mathcal{G} = \sigma(Y)$ where Y is a discrete random variable taking values $y_1, y_2, \ldots$ Show that:

$$\mathbb{E}[X|\mathcal{G}](\omega) = \sum_k \frac{\mathbb{E}[X \mathbb{1}_{\{Y=y_k\}}]}{\mathbb{P}(Y = y_k)} \mathbb{1}_{\{Y=y_k\}}(\omega).$$

Exercise 10.6. Prove the uniqueness of the Lebesgue decomposition: if $\nu = \nu_a + \nu_s$ and $\nu = \nu'_a + \nu'_s$ with $\nu_a, \nu'_a \ll \mu$ and $\nu_s, \nu'_s \perp \mu$, then $\nu_a = \nu'_a$ and $\nu_s = \nu'_s$.

Chapter 11

Product Measures and Fubini-Tonelli

11.1 Introduction

Computing multiple integrals is one of the most frequent operations in analysis and probability. The theorems of Fubini and Tonelli provide conditions under which a double integral can be computed as an iterated integral, in either order.

Why Fubini-Tonelli?

Computing $\iint f(x,y)\,d(x,y)$ directly on a product is often difficult. Fubini and Tonelli allow us to reduce this to successive integrations: first integrate in y for each fixed x (or vice versa). Tonelli handles the case of nonnegative functions (no integrability hypothesis), while Fubini applies to integrable functions.

11.2 Product σ-algebra

Definition 11.1 (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 the *measurable rectangles*:

$$\mathcal{A} \otimes \mathcal{B} = \sigma\big(\{A \times B : A \in \mathcal{A},\ B \in \mathcal{B}\}\big).$$

Definition 11.2 (Sections)**.** For $E \subset X \times Y$ and $x \in X$, $y \in Y$, the *sections* are:

$$E_x = \{y \in Y : (x,y) \in E\}, \qquad E^y = \{x \in X : (x,y) \in E\}.$$

For a function $f : X \times Y \to \overline{\mathbb{R}}$, the sections are: $f_x(y) = f(x,y)$ and $f^y(x) = f(x,y)$.

Proposition 11.3 (Measurability of sections)**.** Let $E \in \mathcal{A} \otimes \mathcal{B}$. Then for every $x \in X$, the section $E_x \in \mathcal{B}$, and for every $y \in Y$, the section $E^y \in \mathcal{A}$.

If $f : (X \times Y, \mathcal{A} \otimes \mathcal{B}) \to (\overline{\mathbb{R}}, \mathcal{B}(\overline{\mathbb{R}}))$ is measurable, then f_x is $\mathcal{B}$-measurable and f^y is $\mathcal{A}$-measurable.

Proof. The collection $\mathcal{C} = \{E \in \mathcal{A} \otimes \mathcal{B} : E_x \in \mathcal{B} \text{ for all } x\}$ is a σ-algebra containing the measurable rectangles (since $(A \times B)_x = B$ if $x \in A$ and $\varnothing$ otherwise), hence $\mathcal{C} = \mathcal{A} \otimes \mathcal{B}$. $\square$

11.3 Product measure

Theorem 11.4 (Existence and uniqueness of the product measure)**.** *Let $(X, \mathcal{A}, \mu)$ and $(Y, \mathcal{B}, \nu)$ be σ-finite measure spaces. There exists a unique measure $\mu \otimes \nu$ on $(X \times Y, \mathcal{A} \otimes \mathcal{B})$ such that:*

$$(\mu \otimes \nu)(A \times B) = \mu(A) \cdot \nu(B), \qquad \forall A \in \mathcal{A}, \forall B \in \mathcal{B}.$$

Moreover, for every $E \in \mathcal{A} \otimes \mathcal{B}$:

$$(\mu \otimes \nu)(E) = \int_X \nu(E_x)\, d\mu(x) = \int_Y \mu(E^y)\, d\nu(y).$$

Sketch of proof. Define $\rho(E) = \int_X \nu(E_x)\, d\mu(x)$. The measurability of $x \mapsto \nu(E_x)$ is verified by a monotone class argument. The σ-additivity of ρ follows from the monotone convergence theorem. On rectangles, $\rho(A \times B) = \mu(A)\nu(B)$. Uniqueness follows from the uniqueness of measures on a π-system (measurable rectangles form a π-system) and σ-finiteness. □

σ-finiteness hypothesis

σ-finiteness is essential. Without it, the product measure may not be unique, and the Fubini and Tonelli theorems may fail.

11.4 Tonelli's theorem

Theorem 11.5 (Tonelli)**.** *Let $(X, \mathcal{A}, \mu)$ and $(Y, \mathcal{B}, \nu)$ be σ-finite measure spaces and let $f : X \times Y \to [0, +\infty]$ be $(\mathcal{A} \otimes \mathcal{B})$-measurable. Then:*

(i) for μ-a.e. x, the function $y \mapsto f(x, y)$ is $\mathcal{B}$-measurable and $x \mapsto \int_Y f(x, y)\, d\nu(y)$ is $\mathcal{A}$-measurable;

(ii) the same holds with the roles of x and y swapped;

(iii) we have:

$$\int_{X\times Y} f\, d(\mu \otimes \nu) = \int_X \left(\int_Y f(x, y)\, d\nu(y) \right) d\mu(x) = \int_Y \left(\int_X f(x, y)\, d\mu(x) \right) d\nu(y).$$

Proof. For $f = \mathbb{1}_E$ with $E \in \mathcal{A} \otimes \mathcal{B}$, the result is Theorem 11.4. By linearity, it extends to nonnegative simple functions. The general case follows from the monotone convergence theorem applied to an increasing sequence of simple functions converging to f. □

11.5 Fubini's theorem

Theorem 11.6 (Fubini)**.** *Let $(X, \mathcal{A}, \mu)$ and $(Y, \mathcal{B}, \nu)$ be σ-finite measure spaces and let $f \in L^1(X \times Y, \mu \otimes \nu)$. Then:*

(i) for μ-a.e. x, $f_x \in L^1(Y, \nu)$;

(ii) the function $x \mapsto \int_Y f(x, y)\, d\nu(y)$ (defined a.e.) is in $L^1(X, \mu)$;

(iii) we have:

$$\int_{X\times Y} f\,d(\mu\otimes\nu) = \int_X \left(\int_Y f(x,y)\,d\nu(y)\right) d\mu(x) = \int_Y \left(\int_X f(x,y)\,d\mu(x)\right) d\nu(y).$$

Proof. Write $f = f^+ - f^-$. By Tonelli, $\int f^+\,d(\mu\otimes\nu)$ and $\int f^-\,d(\mu\otimes\nu)$ are both finite (since $f \in L^1$). Apply Tonelli to f^+ and f^- separately to obtain the iterated integrals for each. In particular, $\int_Y f^+(x,y)\,d\nu(y) < \infty$ and $\int_Y f^-(x,y)\,d\nu(y) < \infty$ for μ-a.e. x, so $f_x \in L^1(\nu)$ a.e. Conclude by taking the difference. □

Fubini-Tonelli: user guide

1. **Tonelli:** $f \geq 0$ measurable $\implies$ one may freely swap $\int\int$ (even if $= +\infty$).
2. **Fubini:** $f \in L^1(\mu\otimes\nu)$ $\implies$ one may swap.
3. **Strategy:** to check $f \in L^1$, first apply Tonelli to $|f|$.

11.6 Applications

11.6.1 Convolution

Definition 11.7 (Convolution)**.** Let $f, g \in L^1(\mathbb{R}^n, \lambda)$ where λ is Lebesgue measure. The *convolution* of f and g is:

$$(f * g)(x) = \int_{\mathbb{R}^n} f(x-y)\,g(y)\,dy.$$

Proposition 11.8. If $f, g \in L^1(\mathbb{R}^n)$, then $f * g$ is defined a.e., $f * g \in L^1(\mathbb{R}^n)$, and $\|f * g\|_{L^1} \leq \|f\|_{L^1}\,\|g\|_{L^1}$.

Proof. By Tonelli, $\int_{\mathbb{R}^n}\int_{\mathbb{R}^n} |f(x-y)|\,|g(y)|\,dy\,dx = \|f\|_{L^1}\,\|g\|_{L^1} < \infty$ (by translation invariance). By Fubini, the inner integral is finite for a.e. x, and the result follows. □

11.6.2 Change of variable formula

Theorem 11.9 (Change of variables)**.** *Let $\varphi : U \to V$ be a C^1-diffeomorphism between open subsets of $\mathbb{R}^n$. For every measurable $f : V \to [0, +\infty]$ (or $f \in L^1(V)$):*

$$\int_V f(y)\,dy = \int_U f(\varphi(x))\,|\det D\varphi(x)|\,dx.$$

Example 11.10 (Polar coordinates)**.** For $\varphi(r,\theta) = (r\cos\theta, r\sin\theta)$, we have $|\det D\varphi| = r$, giving:

$$\iint_{\mathbb{R}^2} f(x,y)\,dx\,dy = \int_0^\infty \int_0^{2\pi} f(r\cos\theta, r\sin\theta)\,r\,d\theta\,dr.$$

Example 11.11 (Gaussian integral)**.** Using Tonelli and polar coordinates:

$$\left(\int_{-\infty}^\infty e^{-x^2}\,dx\right)^2 = \int_{\mathbb{R}^2} e^{-(x^2+y^2)}\,dx\,dy = \int_0^\infty 2\pi r\,e^{-r^2}\,dr = \pi.$$

Hence $\int_{-\infty}^\infty e^{-x^2}\,dx = \sqrt{\pi}$.

11.7 Complements: infinite products

Theorem 11.12 (Countable product of probability measures)**.** *Let* $(X_n, \mathcal{A}_n, \mu_n)_{n\geq 1}$ *be a sequence of probability spaces. There exists a unique probability* $\mathbb{P}$ *on* $\left(\prod_{n\geq 1} X_n, \bigotimes_{n\geq 1} \mathcal{A}_n\right)$ *such that:*

$$\mathbb{P}\left(\prod_{n=1}^{N} A_n \times \prod_{n>N} X_n\right) = \prod_{n=1}^{N} \mu_n(A_n)$$

for all N *and all choices of* $A_n \in \mathcal{A}_n$*. This is the Ionescu-Tulcea theorem (or Kolmogorov extension theorem in the product case).*

11.8 Exercises

Exercise 11.1. Compute $\int_0^1 \int_0^1 \frac{x^2-y^2}{(x^2+y^2)^2}\, dx\, dy$ and $\int_0^1 \int_0^1 \frac{x^2-y^2}{(x^2+y^2)^2}\, dy\, dx$. Comment on the result. *Hint*: $f \notin L^1$.

Exercise 11.2. Let $f : [0,1] \times [0,1] \to \mathbb{R}$ be defined by $f(x,y) = \frac{xy}{(x^2+y^2)^2}$. Verify that $f \notin L^1([0,1]^2)$ and that the iterated integrals give different results.

Exercise 11.3. Using Fubini, show that for $f \in L^1([0,\infty), \lambda)$:

$$\int_0^\infty \frac{1}{x} \int_0^x f(t)\, dt\, dx = \int_0^\infty f(t) \int_t^\infty \frac{1}{x}\, dx\, dt$$

and explain why this equality is formal (both sides may diverge).

Exercise 11.4. Show that $\int_0^\infty \frac{e^{-ax}-e^{-bx}}{x}\, dx = \ln(b/a)$ for $0 < a < b$, using Fubini. *Hint*: write $\frac{e^{-ax}-e^{-bx}}{x} = \int_a^b e^{-tx}\, dt$.

Exercise 11.5 (Convolution and probability)**.** Let X and Y be independent real random variables with densities f_X and f_Y respectively. Show that $Z = X + Y$ has density $f_Z = f_X * f_Y$. Apply to the case $X, Y \sim \mathcal{N}(0,1)$.

Exercise 11.6. Using polar coordinates in dimension n, compute the volume of the unit ball $B_n = \{x \in \mathbb{R}^n : \|x\| \leq 1\}$ and show that:

$$\lambda_n(B_n) = \frac{\pi^{n/2}}{\Gamma(n/2+1)}.$$

Chapter 12

Radon Measures and Riesz Representation

12.1 Introduction

This chapter lies at the interface of measure theory and topology. We study *Radon measures* on locally compact spaces and establish the celebrated Riesz representation theorem, which identifies positive linear functionals on the space of compactly supported continuous functions with Radon measures.

Linking topology and measure

Radon measures are measures that "respect" the topology: they are determined by their values on compact sets and open sets. The Riesz theorem says that every "reasonable" way of integrating compactly supported continuous functions comes from such a measure. This is the fundamental bridge between functional analysis and measure theory.

12.2 Locally compact spaces

Definition 12.1 (Locally compact space)**.** A Hausdorff topological space X is *locally compact* if every point has a compact neighborhood, i.e. for every $x \in X$, there exist an open set U and a compact set K with $x \in U \subset K$.

Example 12.2.

1. $\mathbb{R}^n$ is locally compact.
2. Every compact space is locally compact.
3. Every open subset of a locally compact space is locally compact.
4. An infinite-dimensional Banach space is not locally compact.

Definition 12.3 (Support and compactly supported functions)**.** For a continuous function $f : X \to \mathbb{R}$, the *support* of f is $\text{supp}(f) = \overline{\{x \in X : f(x) \neq 0\}}$. We denote by $C_c(X)$ the space of continuous functions with compact support.

Lemma 12.4 (Urysohn for locally compact spaces)**.** *Let X be locally compact Hausdorff, K compact and U open with $K \subset U$. Then there exists $f \in C_c(X)$ with $0 \leq f \leq 1$, $f|_K = 1$, and $\text{supp}(f) \subset U$.*

Proposition 12.5 (Partition of unity)**.** Let X be locally compact Hausdorff and let K be a compact set covered by open sets $U_1, \ldots, U_n$. There exist $\varphi_1, \ldots, \varphi_n \in C_c(X)$ with $0 \leq \varphi_i \leq 1$, $\operatorname{supp}(\varphi_i) \subset U_i$, and $\sum_{i=1}^{n} \varphi_i = 1$ on K.

12.3 Radon measures

Definition 12.6 (Regular Borel measure)**.** Let X be a locally compact Hausdorff space. A positive Borel measure μ on X is called:

- *outer regular* if for every $B \in \mathcal{B}(X)$: $\mu(B) = \inf\{\mu(U) : U \text{ open}, B \subset U\}$;
- *inner regular* if for every open set U: $\mu(U) = \sup\{\mu(K) : K \text{ compact}, K \subset U\}$;
- *regular* if it is both outer and inner regular.

Definition 12.7 (Radon measure)**.** A *Radon measure* on a locally compact Hausdorff space X is a Borel measure μ on X that is:

(i) *locally finite*: for every $x \in X$, there exists an open neighborhood U of x with $\mu(U) < \infty$;

(ii) *inner regular on open sets*: for every open set U, $\mu(U) = \sup\{\mu(K) : K \text{ compact}, K \subset U\}$;

(iii) *outer regular*: for every Borel set B, $\mu(B) = \inf\{\mu(U) : U \text{ open}, B \subset U\}$.

Example 12.8.

1. Lebesgue measure on $\mathbb{R}^n$ is a Radon measure.
2. Every Dirac measure δ_a is a Radon measure.
3. The counting measure on a discrete space is a Radon measure.
4. The counting measure on $\mathbb{R}$ is *not* a Radon measure (it is not locally finite).

Proposition 12.9. Every finite Radon measure on a locally compact Hausdorff space is regular (inner regular on *all* Borel sets, not just open sets).

12.4 Positive linear functionals

Definition 12.10 (Positive linear functional)**.** A linear functional $\Lambda : C_c(X) \to \mathbb{R}$ is called *positive* if:

$$f \geq 0 \implies \Lambda(f) \geq 0.$$

Remark 12.11. Every positive linear functional on $C_c(X)$ is automatically continuous for the topology of uniform convergence on compact sets (by positivity, $|\Lambda(f)| \leq \Lambda(\|f\|_\infty \cdot \mathbb{1}_K)$ where K contains the support of f).

12.5 The Riesz representation theorem

Theorem 12.12 (Riesz-Markov-Kakutani)**.** *Let X be a locally compact Hausdorff space and let $\Lambda : C_c(X) \to \mathbb{R}$ be a positive linear functional. Then there exists a unique Radon measure μ on X such that:*

$$\Lambda(f) = \int_X f \, d\mu, \qquad \forall f \in C_c(X).$$

Sketch of proof. The proof proceeds in several steps:

Step 1. Define μ on open sets. For U open, set:

$$\mu(U) = \sup\{\Lambda(f) : f \in C_c(X),\, 0 \leq f \leq 1,\, \text{supp}(f) \subset U\}.$$

Step 2. Extend to Borel sets by outer regularity:

$$\mu(B) = \inf\{\mu(U) : U \text{ open},\, B \subset U\}.$$

Step 3. Verify that μ is a measure. Use Urysohn's lemma and partitions of unity to verify σ-additivity.

Step 4. Verify that $\Lambda(f) = \int f \, d\mu$. Approximate f by simple functions and use the regularity of μ.

Step 5. Uniqueness. If μ' is another Radon measure with $\Lambda(f) = \int f \, d\mu'$, then μ and μ' agree on open sets (by Urysohn's lemma), hence on all Borel sets (by regularity). □

Conditions of the theorem

The space X must be locally compact and Hausdorff. The functional must be defined on $C_c(X)$ (compactly supported functions), not $C_b(X)$ (bounded functions). For more general spaces, different versions of the theorem are needed.

12.6 Support of a measure

Definition 12.13 (Support of a measure)**.** Let μ be a Radon measure on a locally compact Hausdorff space X. The *support* of μ is the smallest closed set F such that $\mu(X \setminus F) = 0$:

$$\text{supp}(\mu) = X \setminus \bigcup\{U \text{ open} : \mu(U) = 0\}.$$

Proposition 12.14. Let μ be a Radon measure on X. Then:

1. $x \in \text{supp}(\mu)$ if and only if $\mu(U) > 0$ for every open neighborhood U of x.
2. $\text{supp}(\mu) = \varnothing$ if and only if $\mu = 0$.
3. If $f \in C_c(X)$ and $\text{supp}(f) \cap \text{supp}(\mu) = \varnothing$, then $\int f \, d\mu = 0$.

Proof. (1) $x \notin \text{supp}(\mu)$ means x belongs to an open set U with $\mu(U) = 0$, i.e. some open neighborhood of x has measure zero.

(2) $\text{supp}(\mu) = \varnothing$ means $\mu(X) = 0$ by inner regularity (every compact K is contained in $X \setminus \text{supp}(\mu)$, a union of open sets of measure zero).

(3) $\text{supp}(f)$ is compact and contained in the open set $X \setminus \text{supp}(\mu)$, which has measure zero. □

Example 12.15. 1. $\mathrm{supp}(\delta_a) = \{a\}$.

2. $\mathrm{supp}(\lambda|_{[0,1]}) = [0,1]$ where λ is Lebesgue measure.

3. The Cantor measure has the Cantor set as its support.

12.7 Riesz theorem for signed measures

Theorem 12.16. *Let X be locally compact Hausdorff. Every continuous linear functional $\Lambda : C_0(X) \to \mathbb{R}$ (where $C_0(X)$ is the space of continuous functions vanishing at infinity, with the uniform norm) can be uniquely written as:*

$$\Lambda(f) = \int_X f \, d\mu$$

where μ is a regular signed Radon measure with $\|\Lambda\| = |\mu|(X)$.

Riesz correspondence

$$\left(C_c(X)\right)^*_+ \longleftrightarrow \{\text{positive Radon measures on } X\}$$

$$\left(C_0(X)\right)^* \longleftrightarrow \{\text{regular signed Radon measures on } X\}$$

12.8 Exercises

Exercise 12.1. Show that Lebesgue measure on $\mathbb{R}$ is a Radon measure. *Hint*: verify the three conditions in Definition 12.7.

Exercise 12.2. Let $X = \mathbb{R}$ and $\Lambda(f) = f(0)$ for $f \in C_c(\mathbb{R})$. Identify the Radon measure μ such that $\Lambda(f) = \int f \, d\mu$.

Exercise 12.3. Let μ be a Radon measure on $\mathbb{R}^n$ and K a compact set with $\mu(K) = 0$. Show that for every $\varepsilon > 0$, there exists an open set $U \supset K$ with $\mu(U) < \varepsilon$.

Exercise 12.4. Let X be locally compact Hausdorff and μ a finite Radon measure. Show that for every $B \in \mathcal{B}(X)$ and every $\varepsilon > 0$, there exist a compact set $K \subset B$ and an open set $U \supset B$ with $\mu(U \setminus K) < \varepsilon$.

Exercise 12.5 (Support and density)**.** Let $f : \mathbb{R} \to [0, +\infty)$ be continuous and $\mu(A) = \int_A f \, d\lambda$. Show that $\mathrm{supp}(\mu) = \overline{\{x \in \mathbb{R} : f(x) > 0\}}$.

Exercise 12.6. Let X be locally compact Hausdorff and let μ, ν be two Radon measures such that $\int f \, d\mu = \int f \, d\nu$ for every $f \in C_c(X)$. Show that $\mu = \nu$. Is this statement true if $C_c(X)$ is replaced by a proper subspace?

Bibliography

[1] Rudin, W., *Real and Complex Analysis*, 3rd ed., McGraw-Hill, 1987.

[2] Folland, G.B., *Real Analysis*, 2nd ed., Wiley, 1999.

[3] Brezis, H., *Functional Analysis, Sobolev Spaces and PDE*, Springer, 2011.

[4] Halmos, P.R., *Measure Theory*, Springer, 1974.

[5] Bogachev, V.I., *Measure Theory*, Springer, 2007.