Real Analysis

Lecture Notes from Nanjing University
Translated by ChatGPT

February 7, 2026

Contents

Overview

From a narrow perspective, real analysis studies the calculus of real-valued functions of one (or several) real variables. Unlike elementary calculus, we often encounter "pathological" functions. From a broader perspective, most topics in real analysis can be extended to abstract measure spaces. The core subjects are Lebesgue measure theory and Lebesgue integration theory. The Riemann integral is usually regarded as a classical integral, while the Lebesgue integral is viewed as modern integration. It is a foundation of modern analysis (modern partial differential equations, functional analysis, harmonic analysis, ...), and also a foundation of probability theory and stochastic analysis.

Real analysis combines rigorous logical reasoning with rich geometric intuition. At the same time, many exercises are difficult and require deep analytical thinking.

0.1 Limits of Continuous Functions

Let $\{f_n\}$ be a sequence of continuous functions on $[0,1]$, and suppose $\lim\limits_{n\to\infty} f_n(x) = f(x)$ for every $x \in [0,1]$. If this convergence is uniform, then naturally f is also continuous on $[0,1]$.

However, without the assumption of uniform convergence, the situation is completely different. In fact, we can construct a sequence $\{f_n\}$ such that

- $0 \leqslant f_n(x) \leqslant 1, \forall x \in [0,1]$;
- $\{f_n\}$ is monotone decreasing in n;
- the limit function f is not Riemann integrable.

Even so, under the first two assumptions, one can verify that $\int_0^1 f_n(x)\,dx$ converges to a limit. Naturally, we ask: how can we define a new integral so that

$$\int_{[0,1]} f(x)\,dx = \lim_{n\to\infty} \int_0^1 f_n(x)\,dx$$

holds?

Related results will appear in 3.3 and 4.3.

0.2 Length of Curves

In mathematical analysis, we study plane curves and compute their lengths. Let Γ be a continuous curve in the plane, given by the parametric form $\Gamma = \{(x(t), y(t)), a \leqslant t \leqslant b\}$, where $x(t)$ and $y(t)$ are continuous in t. We usually define the length of Γ as the supremum of polygonal lengths obtained by joining finitely many points on Γ in increasing order of t. When this supremum is finite, we call Γ rectifiable. If $x(t)$ and $y(t)$ are continuously differentiable, we have

$$L = \int_a^b \sqrt{(x'(t))^2 + (y'(t))^2}\, dt. \tag{1}$$

For a general curve, we ask:

1. Under what conditions on $x(t)$ and $y(t)$ can we guarantee that Γ is rectifiable?

2. When those conditions hold, is formula (1) valid?

The first question has a complete answer: the curve is rectifiable if and only if $x(t)$ and $y(t)$ are of bounded variation. Then the second question becomes: when $x(t)$ and $y(t)$ are of bounded variation, the integral in (1) is always meaningful. In general, however, the equality does not always hold as stated, though it can be made valid after choosing a new parameter.

For bounded variation functions and the validity of (1) (that is, differentiability of bounded variation functions and integrability of their derivatives), see 5.5.2.

0.3 Differentiation and Integration

The fundamental theorem of calculus states that differentiation and integration are inverse operations. It has two forms:

$$F(b) - F(a) = \int_a^b F'(x)\, dx, \tag{2}$$

$$\frac{d}{dx}\int_0^x f(t)\, dt = f(x). \tag{3}$$

For the first formula, however, we can find continuous functions F that are nowhere differentiable, or functions for which $F'(x)$ exists everywhere but is not Riemann integrable. These issues motivate us to find a broader class of functions F for which (2) remains valid.

For (3), the question is: how can we establish this identity for a broader class of integrable functions? To answer this, we need a covering theory and the notion of absolutely continuous functions; see 5.1 and 5.6.

0.4 Limitations of the Riemann Integral

Definition 0.4.1 (Riemann Integral)**.** *Let $f(x)$ be a bounded function on $[a, b]$, and take a partition*

$$\Delta : a = x_0 < x_1 < x_2 < \cdots < x_n = b,$$

with Riemann sum

$$S(f, \Delta) = \sum_{i=1}^{n} f(\xi_i)(x_i - x_{i-1}),$$

where $\xi_i \in [x_{i-1}, x_i]$. If there exists a constant I such that for every $\varepsilon > 0$, there exists $\delta > 0$ and whenever

$$|\Delta| \stackrel{\text{def.}}{=} \max_{1\leqslant i\leqslant n} \{x_i - x_{i-1}\} < \delta, \tag{4}$$

we have

$$|S(f,\Delta) - I| < \varepsilon,$$

then $f(x)$ is called Riemann integrable on $[a,b]$, and I is called the Riemann integral of $f(x)$ on $[a,b]$, denoted by

$$(R)\int_a^b f(x)\,dx.$$

When no confusion can arise, the (R) on the left is omitted.

Remark 0.4.1. *In topological language, the convergence defined by* (4) *is called convergence of a net.*

The integrability problem is central in Riemann integration. Although the Darboux theorem gives a necessary and sufficient criterion for integrability of $f(x)$ on $[a,b]$, it is still very difficult to classify integrable functions directly from that criterion.

Example 0.4.1 (Dirichlet Function)**.**

$$D(x) = \begin{cases} 1, & x \in \mathbb{Q} \cap [0,1], \\ 0, & x \in [0,1]\backslash\mathbb{Q}, \end{cases} \tag{5}$$

is not Riemann integrable.

In addition, the class of Riemann integrable functions is not complete. If we impose extra conditions, we get the following bounded convergence statement.

Example 0.4.2. *Let $f_n(x)$ be a sequence of Riemann integrable functions on $[a,b]$, and suppose*

(1) *there exists $M > 0$ such that $|f_n(x)| \leqslant M$ for all $x \in [a,b]$;*

(2) *there exists a real-valued function $f(x)$ on $[a,b]$ such that for all $x \in [a,b]$,*

$$\lim_{n\to\infty} f_n(x) = f(x).$$

Question: is $f(x)$ Riemann integrable on $[a,b]$?

Remark 0.4.2. *If $f(x)$ is Riemann integrable, then*

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

Finally, in Riemann integration, the conditions for exchanging limits and integrals, and for iterated integration, are complicated. To overcome these limitations, we need to "design" a new integral: the Lebesgue integral.

0.5 Lebesgue Integral and Lebesgue Measure

Definition 0.5.1 (Lebesgue Integral)**.** *Let $f(x)$ be a real-valued function on $[a,b]$ satisfying*

$$m \leqslant f(x) < M,$$

and let

$$\Delta : m = y_0 < y_1 < y_2 < \cdots < y_n = M.$$

Define

$$E_i = \{x \in [a,b] \mid y_{i-1} \leqslant f(x) < y_i\},$$

and form

$$S_L(f,\Delta) = \sum_{i=1}^{n} \xi_i m(E_i),$$

where $\xi_i \in [y_{i-1}, y_i)$ *is arbitrary. If there exists a constant* I *such that for every* $\varepsilon > 0$*, there exists* $\delta > 0$ *and whenever*

$$|\Delta| \overset{\text{def.}}{=} \max_{1\leqslant i\leqslant n}\{y_i - y_{i-1}\} < \delta,$$

we have

$$|S_L(f,\Delta) - I| < \varepsilon,$$

then $f(x)$ *is called Lebesgue integrable on* $[a,b]$*, and* I *is called the Lebesgue integral of* $f(x)$ *on* $[a,b]$*, denoted by*

$$(L)\int_{[a,b]} f(x)\,dx.$$

When no confusion can arise, the (L) *on the left is omitted.*

The following two questions are essential:

- What is $m(E_i)$? It generalizes the notion of length and is called Lebesgue measure.
- How is $m(E_i)$ defined? What properties does it satisfy? Which sets can be assigned Lebesgue measure?

Example 0.5.1 (Lebesgue Integral of the Dirichlet Function)**.** *Since the Dirichlet function* (5) *has only two isolated values,* 0 *and* 1*, we have*

$$S_L^*(D,\Delta) = m(\mathbb{Q} \cap [0,1]).$$

Then what is the "length" of the set of all rational numbers in $[0,1]$*?*

Proposition 0.5.1 (Relation Between the Lebesgue and Riemann Integrals)**.** *If* $f(x)$ *is Riemann integrable on* $[a,b]$*, then* $f(x)$ *is Lebesgue integrable on* $[a,b]$*, and*

$$(L)\int_{[a,b]} f(x)\,dx = (R)\int_a^b f(x)\,dx.$$

Measure theory is built on the foundation of Cantor's set theory. It began with work of G. Peano and C. Jordan. Following the model of Riemann integration, Jordan established integration on Jordan measurable sets, but this class has a major defect: there exist open sets that are not Jordan measurable. Later, Borel further developed measure theory and established Borel measure on the class generated from open and closed sets by basic operations such as intersections, unions, and differences (the Borel class, a σ-ring).

Lebesgue established measure theory on a larger class of sets. He proved that Lebesgue measurable sets form a σ-ring and clarified the relationship between Borel measurable sets and Lebesgue measurable sets. Further development of measure theory relied on work by Riesz, Caratheodory, and others.

Chapter 1

Sets and Point Sets

G. Cantor (1845–1918) is the founder of set theory. He introduced concepts such as cardinality, accumulation points, open sets, and closed sets, and proved that transcendental numbers are far more numerous than algebraic numbers. Cantor's descriptive definition of a set is: when all objects sharing a certain property are regarded as a whole, that whole is called a set, and those objects are called elements of the set.

Definition 1.0.1 (Paradox)**.** *An argument is called a paradox if it leads to a conclusion opposite to common judgment, while it is difficult to provide a justified refutation.*

Definition 1.0.2 (Fallacy)**.** *A statement is called a fallacy if both the statement and its negation can be proved by seemingly logically equivalent reasoning, and no error in the derivation can be identified.*

Cantor's descriptive definition of sets leads to the famous Russell paradox (1903).

Example 1.0.1 (Russell Paradox)**.** *Let*

$$E = \{x | x \in x\},$$

then

$$E \in E \Leftrightarrow E \notin E.$$

1.1 Set Operations

The basic operations on sets are well known; we give only a simple example.

Example 1.1.1. *Let $f : [a, b] \to \mathbb{R}^1$. Then*

$$[a, b] = \bigcup_{n=0}^{\infty} \{x | |f(x)| \leqslant n\},$$

$$\{x | |f(x)| > 0\} = \bigcup_{n=1}^{\infty} \{x | |f(x)| \geqslant \frac{1}{n}\}.$$

Besides intersection, union, and difference, we also define symmetric difference and limit operations for sets.

Definition 1.1.1 (Symmetric Difference)**.** *Let A, B be sets. Define*

$$C = (A \backslash B) \cup (B \backslash A),$$

called the symmetric difference of A and B, denoted by $A \triangle B$.

To motivate the definition of set limits, let us recall limsup and liminf for sequences. Let $\{a_n\}$ be a real sequence and define

$$b_k = \sup_{i \geqslant k} a_i,$$

then $\{b_k\}$ is monotone decreasing and hence convergent. We define

$$\limsup_{n \to \infty} a_n = \inf_k b_k = \inf_k \sup_{i \geqslant k} a_i.$$

Similarly, we define upper and lower limits of a sequence of sets.

Definition 1.1.2 (Limit Set)**.** *Let $\{A_k\}$ be a sequence of sets. The set*

$$\bigcap_{j=1}^{\infty} \bigcup_{k=j}^{\infty} A_k$$

is called the upper limit set of $\{A_k\}$, denoted by $\overline{\lim_{k \to \infty}} A_k$ or $\limsup_{k \to \infty} A_k$. The set

$$\bigcup_{j=1}^{\infty} \bigcap_{k=j}^{\infty} A_k$$

is called the lower limit set of $\{A_k\}$, denoted by $\underline{\lim_{k \to \infty}} A_k$ or $\liminf_{k \to \infty} A_k$. If

$$\overline{\lim_{k \to \infty}} A_k = \underline{\lim_{k \to \infty}} A_k,$$

then this common set is called the limit set of $\{A_k\}$, denoted by $\lim_{k \to \infty} A_k$.

Remark 1.1.1. *In fact, set inclusion forms a partial order. For any partially ordered set, one can define suprema and infima, and therefore limsup and liminf analogously. Here, $\bigcup_{k=j}^{\infty} A_k$ is the supremum of $\{A_k\}_{k=j}^{\infty}$.*

Now let us look at a basic example.

Example 1.1.2. *Let*

$$A_k = \begin{cases} [-1, 1] \cup [1, 2], & k \text{ even;} \\ [-1, 1] \cup [-2, -1], & k \text{ odd.} \end{cases}$$

Then

$$\overline{\lim_{k \to \infty}} A_k = [-2, 2], \ \underline{\lim_{k \to \infty}} A_k = [-1, 1].$$

Think: 1. For an arbitrary sequence of sets $\{A_k\}$, do the upper and lower limit sets always exist?

2. Assume $\lim_{n \to \infty} x_n = x$ and $\lim_{n \to \infty} r_n = r$. Study limsup and liminf for the family of n-dimensional balls $\{B(x_n, r_n)\}$.

Definition 1.1.3 (Monotonicity). *Let $\{A_k\}$ be a sequence of sets. If for every $k \in \mathbb{N}$,*

$$A_k \subseteq A_{k-1}, (A_k \supseteq A_{k-1}),$$

then $\{A_k\}$ is called monotone increasing (decreasing), and both are called monotone sequences of sets.

Theorem 1.1.1. *Let $\{A_k\}$ be a monotone sequence of sets. Then*

$$\overline{\lim_{k\to\infty}} A_k = \underline{\lim_{k\to\infty}} A_k.$$

Proof. Consider the monotone increasing case. First, for fixed $j \in \mathbb{N}$,

$$\bigcap_{k=j}^{\infty} A_k \subseteq \bigcup_{k=l}^{\infty} A_k, l = 1, 2, \cdots .$$

Hence

$$\bigcap_{k=j}^{\infty} A_k \subseteq \bigcap_{l=l}^{\infty} \bigcup_{k=l}^{\infty} A_k.$$

Since j is arbitrary,

$$\bigcup_{j=1}^{\infty} \bigcap_{k=j}^{\infty} A_k \subseteq \bigcap_{l=l}^{\infty} \bigcup_{k=l}^{\infty} A_k.$$

Conversely, for any $x \in \bigcap\limits_{l=1}^{\infty} \bigcup\limits_{k=l}^{\infty} A_k$,

$$x \in \bigcup_{k=l}^{\infty} A_k, l = 1, 2, \cdots .$$

So there exists some k_i such that $x \in A_{k_i}$. Because $\{A_k\}$ is monotone increasing, for $k > k_i$ we have $x \in A_k$, hence

$$x \in \bigcap_{k=k_i}^{\infty} A_k,$$

which implies

$$x \in \bigcup_{j=1}^{\infty} \bigcap_{k=j}^{\infty} A_k.$$

□

Recall the monotone convergence theorem in classical analysis.

Proposition 1.1.1. *Let $\{A_k\}$ be a sequence of subsets of E. Then*

(1) $E \backslash \Big(\overline{\lim_{k\to\infty}} A_k \Big) = \underline{\lim_{k\to\infty}} (E \backslash A_k)$,

(2) $E \backslash \Big(\underline{\lim_{k\to\infty}} A_k \Big) = \overline{\lim_{k\to\infty}} (E \backslash A_k)$.

Proof. For (1), by De Morgan's law,

$$\begin{aligned} E\backslash\Big(\overline{\lim_{k\to\infty}} A_k\Big) &= E\backslash\Big(\bigcap_{j=1}^{\infty}\bigcup_{k=j}^{\infty} A_k\Big) \\ &= \bigcup_{j=1}^{\infty}(E\backslash(\bigcup_{k=j}^{\infty} A_k)) \\ &= \bigcup_{j=1}^{\infty}\bigcap_{k=j}^{\infty}(E\backslash A_k) \\ &= \underline{\lim_{k\to\infty}}(E\backslash A_k). \end{aligned}$$

□

Theorem 1.1.2. *Let $\{A_k\}$ be a sequence of sets. Then*

(1) $\overline{\lim_{k\to\infty}} A_k = \{x|\forall j\in\mathbb{N},\exists k\geqslant j, s.t. x\in A_k\}$, *i.e., there exists a subsequence $\{A_{k_i}\}$ such that $x\in A_{k_i}$.*

(2) $\underline{\lim_{k\to\infty}} A_k = \{x|\exists j_0\in\mathbb{N}, s.t.\forall k\geqslant j_0, x\in A_k\}$, *i.e., from some index j_0 onward, x belongs to all A_k.*

Proof. This descriptive formulation follows directly from the definition of limsup/liminf. □

Example 1.1.3 (Structure of Non-convergence Points)**.** *Let $\{f_n(x)\}$ and $f(x)$ be real-valued functions on $\mathbb{R}^1$. Denote by D the set of points where $\{f_n(x)\}$ does not converge to $f(x)$. Then*

$$D=\bigcup_{k=1}^{\infty}\Big(\bigcap_{N=1}^{\infty}\bigcup_{n=N}^{\infty}\{x||f_n(x)-f(x)|\geqslant\frac{1}{k}\}\Big).$$

Remark 1.1.2. *We will use this in the proof of Egorov's theorem; see Theorem 3.3.2.*

At the end of this section, we define Cartesian products. Let X, Y be nonempty sets. The Cartesian product is

$$X\times Y=\{(x,y)|x\in X, y\in Y\}.$$

The definition of infinite products is more subtle and is formulated via maps. Let $\{X_\alpha\}_{\alpha\in I}$ be a family of sets. Define

$$\prod_{\alpha\in I} X_\alpha=\Big\{x\Big|x: I\to\bigcup_{\alpha\in I} X_\alpha, s.t., x(\alpha)\in X_\alpha,\forall\alpha\in I\Big\}.$$

Conversely, a map can be viewed as an element of a product space:

$$\{f|f: X\to Y\}\cong Y^X.$$

As a simple example, when $X=\{1,2,\cdots,n\}$, elements of $\{X\to Y\}$ correspond one-to-one to elements of $Y\times Y\times\cdots\times Y$.

Exercises 1.1

[Unless otherwise noted, page and problem numbers in exercises refer to the corresponding page/problem in Zhou Mingqiang's *Real Analysis*, 2nd edition.]

1. Prove Theorem 1.1.2.
2. Characterize convergence of a function sequence $\{f_n(x)\}$ to $f(x)$.
3. P11 1,2,3
4. P63 1,2

1.2 Mappings and Cardinality

1.2.1 Mappings

Sets are a central object of study in mathematics. For a given set, one may introduce a topology to obtain a topological space, or an algebraic structure (usually natural) to obtain a group, ring, or field, or study both structures simultaneously. These are studies of individual sets in relative isolation. A more effective approach is to classify a family of sets by grouping together sets with the same properties (equivalent sets). A key way to realize this classification is to build relations between sets, and one major tool is mappings between sets.

Definition 1.2.1 (Mapping)**.** *Let X, Y be sets. If there is a rule f such that for each $x \in X$ there exists a unique $y \in Y$ corresponding to x, then this relation is called a mapping, also called a function or transformation.*

Let $f : X \to Y$ be a mapping, with $A \subset X$, $B \subset Y$. Define

$$f(A) = \{f(x) | x \in A\}$$

as the image of A under f, and

$$f^{-1}(B) = \{x | f(x) \in B\}$$

as the preimage of B under f. Question: is $f^{-1}(B)$ unique?

Definition 1.2.2. *Let $f : X \to Y$ be a mapping. If for all $x_1, x_2 \in X$ with $x_1 \neq x_2$, one has*

$$f(x_1) \neq f(x_2),$$

then f is called injective (one-to-one). If $f(X) = Y$, then f is called surjective (onto). If f is both injective and surjective, then f is called bijective.

Let $f : X \to Y$ be bijective. Define $g : Y \to X$ by

$$g(y) = x,$$

where $y = f(x)$. Then g is called the inverse mapping of f, denoted by f^{-1}. Clearly, $f^{-1}(f(x)) = x$ and $f(f^{-1}(y)) = y$.

Proposition 1.2.1. *Let $f : X \to Y$ be a mapping. Let $\{A_i\}_{i \in I}$ and $\{B_j\}_{j \in J}$ be families of subsets of X and Y, respectively. Then*

(1) $f(\bigcup_{i \in I} A_i) = \bigcup_{i \in I} f(A_i)$*;*

(2) $f(\bigcap_{i\in I} A_i) \subseteq \bigcap_{i\in I} f(A_i)$;
(3) $f^{-1}(\bigcup_{i\in I} A_i) = \bigcup_{i\in I} f^{-1}(A_i)$;
(4) $f^{-1}(\bigcap_{i\in I} A_i) = \bigcap_{i\in I} f^{-1}(A_i)$;
(5) *If* $B_1, B_2 \subset Y$ *and* $B_1 \cap B_2 = \varnothing$, *then*

$$f^{-1}(B_1) \cap f^{-1}(B_2) = \varnothing,$$

and further

$$f^{-1}(B_1^c) \cap (f^{-1}(B_1))^c = \varnothing.$$

Proof. For (5), use $X = f^{-1}(B_1 \cup B_1^c) = f^{-1}(B_1) \cup f^{-1}(B_1^c)$. □

Remark 1.2.1. *This shows that a bijection* f *does not necessarily preserve set operations, while its inverse* f^{-1} *preserves almost all of them. This is one reason for the following definitions:*

1. *A continuous function is defined as a map whose inverse image of every open set is open.*

2. *A measurable function is defined as a map whose inverse image of every Borel set is Borel.*

Definition 1.2.3 (Composite Mapping)**.** *Let* $f : X \to Y$ *and* $g : Y \to Z$ *be mappings. Define* $h : X \to Z$ *by*

$$h(x) = g(f(x)), \forall x \in X,$$

called the composition of f *and* g*, denoted by* $g \circ f$.

1.2.2 Characteristic Functions and Power Sets

Definition 1.2.4 (Power Set)**.** *Let* X *be a set. The set*

$$\mathcal{P}(X) = \{A | A \subset X\}$$

is called the power set of X.

Definition 1.2.5 (Characteristic Function)**.** *Let* X *be a set. For any subset* $A \subset X$*, define*

$$\chi_A(x) = \begin{cases} 1, & x \in A; \\ 0, & x \notin A. \end{cases}$$

This map is called the characteristic function of A.

Let X be a set and define

$$f : \mathcal{P}(X) \to \chi(X)$$

$$A \mapsto \chi_A,$$

where $\chi(X)$ denotes the set of all characteristic functions on X. Then f is bijective.

Some basic properties of characteristic functions are listed below; their proofs are straightforward.

Proposition 1.2.2. *Let A, B be sets. Then*

(1) *If $A \cap B = \varnothing$, then $\chi_{A\cup B} = \chi_A + \chi_B$;*

(2) $\chi_{A\cup B} = \chi_{A\backslash B} + \chi_{B\backslash A} + \chi_{A\cap B} = \chi_A + \chi_B - \chi_{A\cap B}$*;*

(3) $\chi_{A\cap B} = \chi_A \times \chi_B$*;*

(4) $\chi_{A\backslash B} = \chi_A - \chi_{A\cap B} = \chi_A(1 - \chi_B)$*;*

(5) $\chi_{A\triangle B} = |\chi_A - \chi_B| = \begin{cases} \chi_A - \chi_B, & x \in A\backslash B; \\ \chi_B - \chi_A, & x \in B\backslash A. \end{cases}$

These identities connect operations on sets with operations on functions. Furthermore, if one introduces an order structure on the function space and considers lattice operations

$$\chi_A \vee \chi_B = \chi_{A\cup B}, \chi_A \wedge \chi_B = \chi_{A\cap B},$$

one obtains an isomorphism between the two structures in the lattice sense.

Think: Prove

$$\limsup_{n\to\infty} \chi_{A_n} = \chi_{\limsup\limits_{n\to\infty} A_n}, \liminf_{n\to\infty} \chi_{A_n} = \chi_{\liminf\limits_{n\to\infty} A_n}.$$

In fact, since the definitions of $\limsup$ and $\liminf$ only involve order structure and χ gives an isomorphism of the two order relations, the conclusion should hold.

Proposition 1.2.3 (Fixed Point Problem for Monotone Set Maps)**.** *Let X be nonempty and $f : \mathcal{P}(X) \to \mathcal{P}(X)$ satisfy*

$$A \subset B \Rightarrow f(A) \subset f(B).$$

Then there exists $T \in \mathcal{P}(X)$ such that $f(T) = T$.

Proof. Let

$$S = \{A | A \in \mathcal{P}(x), A \subset f(A)\}.$$

Since $\varnothing \subset f(\varnothing)$, $S \neq \varnothing$. Define

$$T = \bigcup_{A\in S} A,$$

then $T \in \mathcal{P}(X)$. Now prove $f(T) = T$. First, for any $A \in S$, $A \subset T$, so $f(A) \subset f(T)$. Since $A \subset f(A)$, we get $A \subset f(T)$. By arbitrariness of A, $T \subset f(T)$. On the other hand,

$$T \subset f(T) \Rightarrow f(T) \subset f^2(T),$$

so $f(T) \in S$. By maximality of T, we have $f(T) \subset T$. □

Example 1.2.1. *Let $f : X \to X$ be a mapping. Define $F : \mathcal{P}(X) \to \mathcal{P}(X)$ by*

$$F(A) = f(A).$$

Call F the map induced by f. Clearly F is monotone and has the trivial fixed point $\varnothing$. Question: does F have nontrivial fixed points?

1.2.3 Cardinality

One major question in set theory is how to describe and compare the number of elements in two sets; this is essentially a classification problem. For a finite set A with n elements, we say A has cardinality n, written $\overline{\overline{A}} = n$ or $\mathrm{Card}A = n$. What about infinite sets?

- The algebraic viewpoint classifies by isomorphism.
- The topological viewpoint classifies by homotopy/homeomorphism.
- The differential-geometric viewpoint classifies by diffeomorphism.
- The set-theoretic viewpoint classifies by equipotence.

Definition 1.2.6 (Equipotence)**.** *Let A, B be sets. If there exists a bijection from A to B, then A and B are called equipotent, denoted $A \sim B$.*

Ignoring the Russell paradox issue, equipotence is an equivalence relation between sets.[1] If $A \sim B$, then A and B have the same cardinal number, written

$$\overline{\overline{A}} = \overline{\overline{B}}.$$

Thus cardinality is a concept on equivalence classes of sets.

Intuitively, if A is equipotent to a subset of B, define

$$\overline{\overline{A}} \leqslant \overline{\overline{B}}.$$

Conversely, if B is equipotent to a subset of A, define

$$\overline{\overline{A}} \geqslant \overline{\overline{B}}.$$

Then we ask:

$$\left.\begin{array}{l} \overline{\overline{A}} \leqslant \overline{\overline{B}} \\ \overline{\overline{A}} \geqslant \overline{\overline{B}} \end{array}\right\} \Rightarrow \overline{\overline{A}} = \overline{\overline{B}}?$$

Theorem 1.2.1 (Cantor–Bernstein)**.** *If X is equipotent to a subset of Y, and Y is equipotent to a subset of X, then $X \sim Y$.*

Using the following lemmas, we obtain a proof of Theorem 1.2.1.

Lemma 1.2.1. *Let $X_1 \subset X_2$, $Y_1 \subset Y_2$ be sets, and $\varphi : X \to Y$ a bijection, with*

$$X_i \overset{\varphi}{\sim} Y_i, i = 1, 2.$$

Then

$$(X_2 \backslash X_1) \overset{\varphi}{\sim} (Y_2 \backslash Y_1).$$

Proof. Omitted. ☐

[1] An equivalence relation can be viewed as the diagonal subset in the product space of two sets.

Lemma 1.2.2. *Let $\{A_i\}_{i\in I}$ and $\{B_i\}_{i\in I}$ be two families indexed by the same set I. Assume members of $\{A_i\}$ are pairwise disjoint and members of $\{B_i\}$ are pairwise disjoint, and*

$$A_i \sim B_i.$$

Then

$$\bigcup_{i\in I} A_i \sim \bigcup_{i\in I} B_i.$$

Proof. Omitted. □

Remark 1.2.2. *If the two families use different index sets that are themselves equipotent, the lemma still holds.*

Proof. (Proof of Theorem 1.2.1). By hypothesis, there exist $Y_1 \subset Y$ and an injection/surjection $f : X \to Y_1$; similarly there exist $X_1 \subset X$ and $g : Y \to X_1$ bijective onto X_1. Write $Y_1 = f(X)$, $X_1 = g(Y)$, $X_2 = g \circ f(X)$. Then

$$X \xrightarrow{f} Y_1 \xrightarrow{g} X_2,$$

where $X_2 = g(Y_1) \subset X_1$. Let $\varphi = g \circ f$. Then

$$X \overset{\varphi}{\sim} X_2 \subset X_1.$$

Set $X_3 = \varphi(X_1) \subset X_2$. Then

$$X_1 \subset X \Rightarrow X_1 \overset{\varphi}{\sim} X_3 \subset X_2.$$

By Lemma 1.2.1,

$$(X_0 \backslash X_1) \overset{\varphi}{\sim} (X_2 \backslash X_3).$$

Further, define $X_{n+2} = \varphi(X_n)$ with $X_n \subset X_{n-1}$ and $X_0 = X$. Then

(1) $X = X_0 \supset X_1 \supset X_2 \supset \cdots \supset X_n \supset \cdots$

(2) $X \overset{\varphi}{\sim} X_2 \overset{\varphi}{\sim} X_4 \overset{\varphi}{\sim} \cdots \overset{\varphi}{\sim} X_{2n} \overset{\varphi}{\sim} \cdots$

(3) $X_1 \overset{\varphi}{\sim} X_3 \overset{\varphi}{\sim} X_5 \overset{\varphi}{\sim} \cdots \overset{\varphi}{\sim} X_{2n+1} \overset{\varphi}{\sim} \cdots$

(4) X and X_1 decompose as

$$X = \Big(\bigcup_{n=1}^{\infty} (X_{n-1} \backslash X_n)\Big) \cup \Big(\bigcup_{n=1}^{\infty} X_n\Big).$$

$$X_1 = \Big(\bigcup_{n=1}^{\infty} (X_n \backslash X_{n+1})\Big) \cup \Big(\bigcup_{n=1}^{\infty} X_n\Big).$$

(5) For every $n \in \mathbb{N}$,

$$(X_{2n} \backslash X_{2n+1}) \overset{\varphi}{\sim} (X_{2n+2} \backslash X_{2n+3}),$$
$$(X_{2n+1} \backslash X_{2n+2}) \overset{\mathrm{id}}{\sim} (X_{2n+1} \backslash X_{2n+2}). \tag{1.1}$$

Now reorder the decompositions in (4) as

$$X = \Big(\bigcup_{n=1}^{\infty} X_n\Big) \cup (X_0 \backslash X_1) \cup (X_1 \backslash X_2) \cup (X_2 \backslash X_3) \cup (X_3 \backslash X_4) \cup \cdots,$$

$$X_1 = \Big(\bigcup_{n=1}^{\infty} X_n\Big) \cup (X_2 \backslash X_3) \cup (X_1 \backslash X_2) \cup (X_4 \backslash X_5) \cup (X_3 \backslash X_4) \cup \cdots,$$

then by (5) and Lemma 1.2.2,

$$X \sim X_1 \sim Y.$$

In (1.1), id denotes the identity map. □

There is another proof of Theorem 1.2.1. We list it below, first requiring one lemma.

Lemma 1.2.3 (Decomposition Theorem Under Set Mappings)**.** *Let X, Y be sets. If $f : X \to Y$ and $g : Y \to X$, then there are decompositions*

$$X = A \cup A_1, A \cap A_1 = \varnothing,$$

$$Y = B \cup B_1, A \cap B_1 = \varnothing,$$

such that

$$f(A) = B, g(B_1) = A_1.$$

Proof. Without loss of generality assume

$$f(X) \subsetneqq Y, g(Y) \subsetneqq X.$$

Call a subset $E \subset X$ *separated* if

$$E \cap g(Y \backslash f(E)) = \varnothing.$$

Let Γ be the family of all separated sets. Since $\varnothing \in \Gamma$, Γ is nonempty. We claim Γ is closed under arbitrary unions. Indeed, if A is a union of members of Γ, then

$$\begin{aligned} & E \in \Gamma, (Y \backslash f(A)) \subset (Y \backslash f(E)) \\ \Rightarrow \quad & \forall E \in \Gamma, E \cap g(Y \backslash f(A)) = \varnothing \\ \Rightarrow \quad & A \cap g(Y \backslash f(A)) = \varnothing \\ \Rightarrow \quad & A \text{ is separated.} \end{aligned}$$

With this property, define[2]

$$A = \bigcup_{E \in \Gamma} E.$$

Then:

(1) By definition, A is maximal among separated sets under inclusion.

(2) Set $B = f(A)$, $B_1 = Y \backslash B$, $A_1 = g(B_1)$. We prove

$$A \cap A_1 = \varnothing, \text{and} A_1 = X \backslash A.$$

First,

$$A \cap g(Y \backslash f(A)) = \varnothing \Rightarrow A \cap A_1 = \varnothing.$$

Next prove $A_1 = X \backslash A$ by contradiction. Suppose $\exists x_0 \in X \backslash A$ and $x_0 \notin A_1$. Let $A^* = A \cup \{x_0\}$. Then

$$\begin{aligned} & A^* = A \cup \{x_0\} \\ \Rightarrow \quad & f(A) \subset f(A^*) \\ \Rightarrow \quad & Y \backslash f(A^*) \subset Y \backslash f(A) \\ \Rightarrow \quad & \Big(g(Y \backslash f(A^*)) \subset g(Y \backslash f(A))\Big) + \Big(A \cap g(Y \backslash f(A)) = \varnothing\Big) \\ \Rightarrow \quad & A \cap g(Y \backslash f(A^*)) = \varnothing. \end{aligned}$$

[2] This avoids explicitly invoking the well-ordering principle.

Since $x_0 \notin A_1$, we have $x_0 \notin g(Y \backslash f(A^*)) \subset g(Y \backslash f(A))$. Hence $A^* \in \Gamma$, but $A \subsetneqq A^*$, contradicting maximality. □

Using this lemma, Theorem 1.2.1 can be proved (left as an exercise).

Let us also see a corollary.

Corollary 1.2.1. *Let $C \subset A \subset B$ and $B \sim C$. Then $A \sim B$.*

Proof. Take $f : B \to C \subsetneqq A$ and $g : A \to A \subsetneqq B$. □

Theorem 1.2.1 not only solves

$$\left.\begin{array}{l} \overline{\overline{A}} \leqslant \overline{\overline{B}} \\ \overline{\overline{A}} \geqslant \overline{\overline{B}} \end{array}\right\} \Rightarrow \overline{\overline{A}} = \overline{\overline{B}}?$$

but also provides a convenient tool for proving equipotence. In the next part, we focus on the following cardinalities:

1. finite sets;
2. countable sets;
3. $\overline{\overline{[0,1]}}$;
4. $\overline{\overline{\mathcal{P}(X)}} = \overline{\overline{2^X}} > \overline{\overline{X}}$.

Definition 1.2.7 (Finite Set)**.** *Let A be nonempty. Write $M_n = \{1, 2, \cdots, n\}$. If $A \sim M_n$ for some n, then A is called finite and n is its cardinality. If $A \nsim M_n$ for every natural number n, then A is called infinite.*

Definition 1.2.8 (Countable Set)**.** *A set equipotent to $\mathbb{N}$ is called countable. The cardinality of a countable set is denoted by $\aleph_0$.*

The next two examples are standard countability examples.

Example 1.2.2. *Any family of pairwise disjoint open intervals in $\mathbb{R}^1$ is at most countable.*

Proof. It suffices to build a one-to-one correspondence with a subset of $\mathbb{Q}$. □

Remark 1.2.3. *This property will be used in describing the structure of open sets in $\mathbb{R}^1$.*

Example 1.2.3. *The set of discontinuities of a monotone function f is at most countable.*

Proof. Omitted. □

Theorem 1.2.2. *Every infinite set A contains a countable subset.*

Example 1.2.4 (Characterization of Function Discontinuity Points)**.** *Let f be real-valued. Define the left jump at x by*

$$\omega_-(f, x) = \sup_{y<x} \big|f(y-0) - f(x)\big|,$$

and similarly the right jump by

$$\omega_+(f, x) = \sup_{y>x} \big|f(y+0) - f(x)\big|.$$

Then

$$f \text{ is discontinuous at } x \Leftrightarrow \omega(f,x) = \omega_-(f,x) + \omega_+(f,x) > 0.$$

Now let

$$\Omega_k = \{x \in \mathbb{R}^1 | \omega(f,x) \geqslant \frac{1}{k}\},$$

then the set of all discontinuity points of f on $\mathbb{R}^1$ is

$$\Omega^{(+)} = \bigcup_{k=1}^{\infty} \Omega_k.$$

Theorem 1.2.3. *For any set A,*

$$\overline{\overline{\mathcal{P}(A)}} = \overline{\overline{2^A}} > \overline{\overline{A}}.$$

Proof. First define $\varphi : A \to \mathcal{P}(A)$ by

$$x \mapsto \{x\},$$

so clearly $\overline{\overline{\mathcal{P}(A)}} \geqslant \overline{\overline{A}}$. We prove strict inequality by contradiction. Assume $\overline{\overline{\mathcal{P}(A)}} = \overline{\overline{A}}$. Then there is a bijection $\psi : A \to \mathcal{P}(A)$. Let

$$B = \{x \in A | x \notin \psi(x)\}.$$

Then $B \neq \varnothing$. Since ψ is onto, there exists $x^* \in A$ with

$$\psi(x^*) = B.$$

Thus

$$x^* \in B \Leftrightarrow x^* \notin B,$$

a contradiction. Hence $\overline{\overline{\mathcal{P}(A)}} \neq \overline{\overline{A}}$, and therefore $\overline{\overline{\mathcal{P}(A)}} > \overline{\overline{A}}$. □

Now let us look at two examples on cardinality.

Example 1.2.5. *The cardinality of the set $\mathcal{F}$ of all real-valued functions on $\mathbb{R}^1$ is $2^{\aleph}$.*

Proof. For any $A \subset \mathbb{R}^1$, define

$$\varphi(A) = \chi_A(x).$$

Then φ is a bijection between $\mathcal{P}(A)$ and all characteristic functions, so $\overline{\overline{\mathcal{F}}} \geqslant 2^{\aleph}$. On the other hand, for each $f \in \mathcal{F}$ define

$$g(f) = \text{Graph}(f) = \big\{(x, f(x)) | x \in \mathbb{R}^1\big\} \in \mathcal{P}(\mathbb{R}^2).$$

Hence $\overline{\overline{\mathcal{F}}} \leqslant \overline{\overline{\mathcal{P}(\mathbb{R}^2)}} = 2^{\aleph}$. Therefore $\overline{\overline{\mathcal{F}}} = 2^{\aleph}$. □

Remark 1.2.4. *If one regards maps as elements of infinite product spaces, then $\mathcal{F} \sim \mathbb{R}^{\mathbb{R}}$. Hence*

$$\text{Card}\mathcal{F} = 2^{\aleph}.$$

Example 1.2.6. *The cardinality of $C(\mathbb{R}^1)$, the set of all continuous functions on $\mathbb{R}^1$, is $\aleph$.*

Proof. Write $\mathbb{Q} = \{r_1, r_2, \cdots, r_n, \cdots\}$. For each $f \in C(\mathbb{R}^1)$ define

$$f \mapsto (f(r_1), f(r_2), \cdots, f(r_n), \cdots) \in \{\{x_n\} | x_n \in \mathbb{R}^1\}.$$

By continuity, this map is injective. Conversely, if $f \neq g$, then $\exists x_0 \in \mathbb{R}^1$ such that $f(x_0) \neq g(x_0)$. By density of $\mathbb{Q}$, there is $r_{k_n} \to x_0$ with

$$\lim_{n\to\infty} f(r_{k_n}) \neq \lim_{n\to\infty} g(r_{k_n}),$$

so

$$\exists k_{n_0}, s.t. f(k_{n_0}) \neq g(k_{n_0}).$$

Hence the map is also onto and therefore bijective. Thus $\overline{\overline{C(\mathbb{R}^1)}} \leqslant \aleph$.[3] Since constant functions form a subset of $C(\mathbb{R}^1)$, clearly $\overline{\overline{C(\mathbb{R}^1)}} \geqslant \aleph$. Therefore $\overline{\overline{C(\mathbb{R}^1)}} = \aleph$. □

Exercises 1.2

1. Let $f(x)$ be continuous on $[a, b]$. Then for any $\varepsilon > 0$, there exist finitely many pairwise disjoint intervals $\{[a_i, b_i)\}_{i=1}^n$ and $\xi_i \in \mathbb{R}$, $i = 1, 2, \cdots, n$, such that

$$\left| f(x) - \sum_{i=1}^{n} \xi_i \chi_{[a_i,b_i)}(x) \right| < \varepsilon, \forall x \in [a, b].$$

2. Prove that the definition of cardinality for finite sets is well-defined.

3. If A is infinite, then A is equipotent to a proper subset of itself. [4]

4. Give a bijection between $(0, 1)$ and $[0, 1]$, and prove that no continuous bijection exists between them.

5. Find the cardinality of $C(\mathbb{R}^1)$.

6. Find the cardinality of monotone functions on $\mathbb{R}^1$.[5]

7. Find the cardinality of all maps from $[0, 1]$ to $\mathbb{N}$.

8. P18 1,2,3

9. P24 5,6,7,9,10

10. P28 13,14,15,16,17

11. P64 5,7,9,10

12. P69 27

[3] This uses that the set of all rational sequences has cardinality $\aleph$; readers may prove this independently.
[4] This can also be used as a definition of infinite sets.
[5] 2011 midterm.

1.3 Point Sets in $\mathbb{R}^n$

1.3.1 Metrics in $\mathbb{R}^n$

The n-dimensional Euclidean space $\mathbb{R}^n$ naturally carries the Euclidean distance. For any

$$x = (x_1, x_2, \cdots, x_n) \in \mathbb{R}^n, y = (y_1, y_2, \cdots, y_n) \in \mathbb{R}^n,$$

define

$$\mathrm{d}(x, y) = \|x - y\|_{\mathbb{R}^n} = \Big(\sum_{i=1}^{n} |x_i - y_i|^2\Big)^{\frac{1}{2}}.$$

It is easy to prove that $\mathrm{d}(x, y)$ satisfies:

(1) [Positive definiteness] $\mathrm{d}(x, y) \geqslant 0$, with equality iff $x = y$;

(2) [Symmetry] $\mathrm{d}(x, y) = \mathrm{d}(y, x)$;

(3) [Triangle inequality] $\mathrm{d}(x, y) \leqslant \mathrm{d}(x, z) + \mathrm{d}(z, y)$.

Definition 1.3.1. *Let $\{x^k\}_{k=1}^{\infty} \subset \mathbb{R}^n$, $x \in \mathbb{R}^n$. We say x is the limit of $\{x^k\}$ as $k \to \infty$ if*

$$\lim_{k\to\infty} \mathrm{d}(x^k, x) = 0.$$

Theorem 1.3.1. *Let $\{x^k\}_{k=1}^{\infty} \subset \mathbb{R}^n$, $x \in \mathbb{R}^n$. Then*

$$x^k \to x \Leftrightarrow x_i^k \to x_i, i = 1, 2, \cdots .$$

Proof. Omitted. □

Definition 1.3.2. *Let $A \subset \mathbb{R}^n$. If there exists $M > 0$ such that for all $x \in A$,*

$$\|x\| \overset{\text{def.}}{=} d(x, 0) \leqslant M,$$

then A is called bounded.

Theorem 1.3.2 (Bolzano–Weierstrass). *If $\{x^k\} \subset \mathbb{R}^n$ is bounded, then it has a convergent subsequence in $\mathbb{R}^n$.*

Theorem 1.3.1 and Theorem 1.3.2 are standard results from analysis, so we omit proofs.

1.3.2 Neighborhoods, Interior Points, Open Sets, Closed Sets

Let $x_0 \in \mathbb{R}^n$. Define

$$B_\delta(x_0) = \{x \in \mathbb{R}^n | d(x, x_0) < \delta\}$$

as the δ-neighborhood of x_0, also denoted by $N_\delta(x_0)$ or $B(x_0, \delta)$.

Definition 1.3.3 (Open Set)**.** *Let $A \subset \mathbb{R}^n$, $x_0 \in A$. We call x_0 an interior point of A if there exists $\delta > 0$ such that*

$$N_\delta(x_0) \subset A.$$

The set of all interior points is called the interior of A, denoted by $\overset{\circ}{A}$. A is called open in $\mathbb{R}^n$ if $A = \overset{\circ}{A}$.

Remark 1.3.1. *The definition of openness depends on the chosen ambient space.*

Theorem 1.3.3. *Open sets in* $\mathbb{R}^n$ *satisfy:*

(1) $\varnothing$ *and* $\mathbb{R}^n$ *are open;*

(2) *if* $\{G_i\}_{i\in I}$ *is a family of open sets, then* $\bigcup\limits_{i\in I} G_i$ *is open;*

(3) *finite intersections of open sets are open.*

Proof. For (3), let G_i, $i = 1, 2, \cdots, n$, be open in $\mathbb{R}^n$. Take $x_0 \in \bigcap\limits_{i=1}^{n} G_i$. For each $1 \leqslant i \leqslant n$, since $x_0 \in G_i$, there exists $\delta_i > 0$ such that $B_{\delta_i}(x_0) \subset G_i$. Let $\delta = \min\{\delta_1, \delta_2, \cdots, \delta_n\}$, then

$$B_\delta(x_0) \subset G_i, i = 1, 2, \cdots, n,$$

so

$$B_\delta(x_0) \subset \Big(\bigcap_{i=1}^{n} G_i\Big).$$

Hence x_0 is an interior point of $\bigcap\limits_{i=1}^{n} G_i$. □

Definition 1.3.4 (Path Connectedness)**.** *Let* $A \subset \mathbb{R}^n$, $x, y \in A$. *We say* x, y *are path connected in* A *if there exists a continuous map* $\varphi : [0, 1] \to A$ *such that*

$$\varphi(0) = x, \varphi(1) = y.$$

A *is called path connected if every pair* $x, y \in A$ *is path connected. A subset* $B \subset A$ *is called a connected component of* A *if*

(1) B *is path connected;*

(2) B *is maximal: for every path-connected* $B_1 \subset A$, *one has* $B_1 \subset B$.

Theorem 1.3.4 (Structure Theorem of Open Sets in $\mathbb{R}^n$)**.** *Let* G *be open in* $\mathbb{R}^n$. *Then* G *can be written as the union of at most countably many path-connected components, each of which is open:*

$$G = \bigcup_{i\in I} G_i,$$

where I *is at most countable. Each component is called a component interval of* G.

Proof. See [3], Theorem 4.2. □

Remark 1.3.2. *For any* $x, y \in G$, *define* $x \sim y$ *if* x *and* y *are path connected. This is an equivalence relation. For each* $x \in G$, *let*

$$G_x = \{y \in G | y \sim x\},$$

then G_x *is a path-connected component of* G, *and is open.*

1.3.3 Limit Points and Closure

Definition 1.3.5 (Limit Point)**.** *Let* $A \subset \mathbb{R}^n$, $x_0 \in \mathbb{R}^n$. *We call* x_0 *a limit point (accumulation point) of* A *if for every* δ*-neighborhood* $N_\delta(x_0)$,

$$\Big(N_\delta(x_0)\backslash\{x_0\}\Big) \cap A \neq \varnothing.$$

The set of all limit points is called the derived set, denoted by A'.

x_0 is called an isolated point of A if there exists $\delta > 0$ such that

$$\Big(N_\delta(x_0)\backslash\{x_0\}\Big) \cap A = \varnothing.$$

x_0 is called a boundary point of A if for every δ-neighborhood $N_\delta(x_0)$,

$$\Big(N_\delta(x_0)\backslash\{x_0\}\Big) \cap A \neq \varnothing,$$

and

$$\Big(N_\delta(x_0)\backslash\{x_0\}\Big) \cap A^c \neq \varnothing.$$

The set of all boundary points is called the boundary of A, denoted by BdA or ∂A.

Definition 1.3.6 (Closed Set)**.** *$A \subset \mathbb{R}^n$ is called closed if $A' \subset A$. Its closure is $\overline{A} = A' \cup A$.*

Remark 1.3.3. *The definition of closedness depends on the ambient space.*

We have the following elementary fact.

Theorem 1.3.5. *$A \subset \mathbb{R}^n$ is closed in $\mathbb{R}^n$ iff A^c is open in $\mathbb{R}^n$.*

Proof. Omitted. □

From the definitions of closed set and derived set, one easily obtains:

Example 1.3.1. *For $E_1, E_2 \subset \mathbb{R}^n$,*

$$(E_1 \cup E_2)' = E_1' \cup E_2'.$$

Hence

$$\overline{E_1 \cup E_2} = \overline{E_1} \cup \overline{E_2}.$$

Corollary 1.3.1. *For $A \subset \mathbb{R}^n$, the closure $\overline{A}$ is closed in $\mathbb{R}^n$.*

Proof. This follows by contradiction. □

The next theorem corresponds exactly to Theorem 1.3.3.

Theorem 1.3.6. *Closed sets in $\mathbb{R}^n$ satisfy:*

(1) *$\varnothing$ and $\mathbb{R}^n$ are closed;*

(2) *if $\{G_i\}_{i\in I}$ is a family of closed sets, then $\bigcap_{i\in I} G_i$ is closed;*

(3) *finite unions of closed sets are closed.*

Proof. Consider complements, and apply De Morgan's law with Theorem 1.3.3. □

We now prove the famous Heine–Borel theorem (finite open cover principle). We first define refinement covers.

Definition 1.3.7 (Refinement Cover)**.** *Let $A \subset \mathbb{R}^n$ and let $\{G_i\}_{i\in I}$ be an open cover of A. We call $\{U_j\}_{j\in J}$ a refinement of $\{G_i\}_{i\in I}$ if $\{U_j\}_{j\in J}$ is also an open cover of A, and for each $j \in J$ there exists $i_j \in I$ such that*

$$U_j \subset G_{i_j}.$$

Density is another frequent concept.

Definition 1.3.8 (Dense)**.** *Let A be nonempty. A set E is dense in A if $A \subset \overline{E}$. If additionally $E \subset A$, then E is called a dense subset of A.*

With density, we define separability.

Definition 1.3.9 (Separable)**.** *A set A is called separable if it has a countable dense subset.*

Second countability is also important.

Definition 1.3.10 (Second Countable)**.** *A topological space (Ω, τ) is called second countable if τ has a countable base.*

Remark 1.3.4. *1. The proof of the next lemma uses that $\mathbb{R}^n$ is second countable.*[6]

2. For $\mathbb{R}^n$, separability, second countability, and the Lindelöff property below are equivalent. In general topological spaces, they need not be equivalent.

Lemma 1.3.1 (Lindelöff Property)**.** *Let $A \subset \mathbb{R}^n$ and $\{G_i\}_{i\in I}$ be any open cover of A. Then $\{G_i\}_{i\in I}$ has a countable subcover of A.*

Proof. For each $x \in A$, since $A \subset \bigcup\limits_{i\in I} G_i$, there exists $i_x \in I$ such that $x \in G_{i_x}$. Since G_{i_x} is open, there exists $\delta_x > 0$ with

$$N_{\delta_x}(x) \subset G_{i_x}.$$

Because $\mathbb{Q}^n$ is dense in $\mathbb{R}^n$, there exists $y_x \in \mathbb{Q}^n$ with $\mathrm{d}(y_x, x) < \delta_x/4$. Choose $\varepsilon_x \in \mathbb{Q}$ with $\delta_x/4 \leqslant \varepsilon_x \leqslant \delta_x/2$, then $x \in N_{\delta_x}(y_x) \subset G_{i_x}$. Let $X = \{N_{\delta_x}(y_x) | x \in A\}$. Then X is at most countable and is a refinement of $\{G_i\}$. Hence

$$A \subset \bigcup_{i=1}^{\infty} N_{\delta_{x_j}}(x_j) \subset \bigcup_{i=1}^{\infty} G_{i_j},$$

a countable cover of A. □

This lemma itself is also very useful.

Theorem 1.3.7 (Heine–Borel)**.** *$A \subset \mathbb{R}^n$ is bounded and closed iff every open cover of A has a finite subcover.*

Proof. Necessity. Let $\{G_i\}_{i\in I}$ be an open cover of A. By Lemma 1.3.1, assume it is at most countable. Suppose no finite subcover exists. Then for each k we can pick

$$x_{k+1} \in A \backslash \Big(\bigcup_{i=1}^{k} G_i \Big).$$

By Theorem 1.3.2, there is a subsequence (still denoted x_k) with $x_k \to x_0 \in A$. Then $x_0 \in G_{k_0}$ for some k_0. Since G_{k_0} is open, there exists $\delta > 0$ such that $N_\delta(x_0) \subset G_{k_0}$. From $x_k \to x_0$, for sufficiently large k we have $\mathrm{d}(x_k, x_0) < \delta$, i.e.

$$x_k \in N_\delta(x_0) \subset G_{k_0},$$

[6] The proof does not need explicit topological language; readers unfamiliar with point-set topology may ignore this remark.

contradicting the construction of x_k.

Sufficiency. First, boundedness is easy. Next prove closedness by contradiction. Assume A is not closed. Then $\exists x_0 \in A'$ with $x_0 \notin A$. Let

$$G_n = \{y \in \mathbb{R}^n | \mathrm{d}(y, x_0) > \frac{1}{n}\},$$

then

$$A = \bigcup_{n=1}^{\infty} G_n,$$

but this cover has no finite subcover, contradiction. □

Remark 1.3.5. *1. If every open cover of $A \subset \mathbb{R}^n$ has a finite subcover, then A is called compact in $\mathbb{R}^n$.*[7] *2. The necessity part fails in infinite-dimensional spaces. In fact, whether necessity holds can serve as a criterion for finite dimensionality. The sufficiency part always holds.*

For compactness, we give a sample problem.

Example 1.3.2. *Let $E \subset \mathbb{R}^n$ be compact and $x \in E^c$. Then x and E can be separated by open sets, i.e., there exist disjoint open sets U, V such that $x \in U$ and $E \subset V$.*

Proof. Sketch: first note any two distinct points in $\mathbb{R}^n$ can be separated by open sets; this yields an open cover of E. Then extract a finite subcover by compactness, and finally use closure of open sets under finite intersections. □

1.3.4 Extension Theorem for Continuous Functions on Closed Sets

In this section we discuss some properties of continuous functions and present an important extension theorem.

Definition 1.3.11 (Continuous Function)**.** *Let $E \subset \mathbb{R}^n$, $f : E \to \mathbb{R}$, $x_0 \in E$. If for every $\varepsilon > 0$ there exists $\delta > 0$ such that whenever $x \in E \cap B(x_0, \delta)$,*

$$|f(x) - f(x_0)| < \varepsilon,$$

then f is continuous at x_0. If f is continuous at every point of E, then f is continuous on E. The set of continuous functions on E is denoted by $C(E, \mathbb{R})$ or $C(E)$.

Continuity has the following equivalent formulation.

Theorem 1.3.8. *f is continuous at x iff for every neighborhood V_y of $y = f(x)$, there exists a neighborhood U_x of x such that $f(U_x) \subset V_y$.*

Continuous functions on bounded closed sets satisfy:

Proposition 1.3.1. *Let $F \subset \mathbb{R}^n$ be bounded closed and $f \in C(F, \mathbb{R})$. Then*

(1) *f is bounded on F;*

(2) *f attains its supremum and infimum on F;*

(3) *f is uniformly continuous on F;*

(4) *if F is path connected, then f has the intermediate value property.*

[7] More generally, for a topological space (Ω, τ), a subset $A \subset \Omega$ is compact if every open cover has a finite subcover.

Proof. Omitted. □

Our goal is to find a method to extend a continuous function on a closed set F to a continuous function on $\mathbb{R}^n$. First define continuous extension.

Definition 1.3.12 (Continuous Extension)**.** *Let $f \in C(E, \mathbb{R})$. A function $g \in C(\mathbb{R}^n, \mathbb{R})$ is called a continuous extension of f to $\mathbb{R}^n$ if*

$$g|_E \equiv f.$$

In general, continuous extension is difficult. To construct extensions on closed sets, we introduce distance to a set.

Definition 1.3.13. *Let $x \in \mathbb{R}^n$, $E \subset \mathbb{R}^n$, $E \neq \varnothing$. Define*

$$\operatorname{dist}(x, E) = \inf_{y \in E} \mathrm{d}(x, y)$$

as the distance from x to E. If $E_1, E_2 \subset \mathbb{R}^n$ are nonempty, define

$$\operatorname{dist}(E_1, E_2) = \inf_{x \in E_1} \operatorname{dist}(x, E_2)$$

as the distance between sets E_1 and E_2.

The next theorem says distance functions are continuous.

Theorem 1.3.9. *Let E be a nonempty subset of $\mathbb{R}^n$. Then*

$$f(x) = \operatorname{dist}(x, E)$$

is uniformly continuous on $\mathbb{R}^n$.

Proof. For every $z \in E$,

$$\mathrm{d}(x, z) \leqslant \mathrm{d}(x, y) + \mathrm{d}(y, z).$$

Taking infimum over z on both sides gives

$$\inf_{z \in E} \mathrm{d}(x, z) \leqslant \mathrm{d}(x, y) + \inf_{z \in E} \mathrm{d}(y, z),$$

i.e.

$$\operatorname{dist}(x, E) \leqslant \mathrm{d}(x, y) + \operatorname{dist}(y, E).$$

Similarly,

$$\operatorname{dist}(y, E) \leqslant \mathrm{d}(y, x) + \operatorname{dist}(x, E).$$

Since $d(x, y) = d(y, x)$,

$$\big|\operatorname{dist}(x, E) - \operatorname{dist}(y, E)\big| \leqslant \mathrm{d}(x, y) = |x - y|.$$

Done. □

Corollary 1.3.2. *Let F be a nonempty closed subset of $\mathbb{R}^n$, $x_0 \in \mathbb{R}^n$. Then there exists $y_0 \in F$ such that*

$$\operatorname{dist}(x_0, F) = \mathrm{d}(x_0, y_0).$$

Proof. Apply (2) of Proposition 1.3.1. □

Now we state and prove the extension theorem for continuous functions on closed sets.

Theorem 1.3.10 (Tietze Extension Theorem). *Let $F \subset \mathbb{R}^n$ be closed, $f \in C(F)$, and $|f(x)| < M$ for all $x \in F$. Then there exists $g \in C(\mathbb{R}^n)$ such that*

(1) $g|_F \equiv f$;

(2) $|g(x)| \leqslant M$ *for all* $x \in \mathbb{R}^n$.

Proof. Let

$$A = \{x \in F | \frac{M}{3} \leqslant f(x) \leqslant M\},$$

$$B = \{x \in F | -M \leqslant f(x) \leqslant \frac{M}{3}\},$$

$$C = \{x \in F | -\frac{M}{3} < f(x) < \frac{M}{3}\}.$$

Then A, B are closed and $A \cap B = \varnothing$.[8] Hence we can build a continuous function on $\mathbb{R}^n$ equal to 1 on A and -1 on B, e.g.

$$\varphi(x) = \frac{-\text{dist}(x, A) + \text{dist}(x, B)}{\text{dist}(x, A) + \text{dist}(x, B)}.$$

For all $x \in \mathbb{R}^n$, $|\varphi(x)| \leqslant 1$. Let

$$g_1(x) = \frac{M}{3}\varphi(x),$$

then $|g_1(x)| \leqslant M/3$ for all $x \in \mathbb{R}^n$. Also

$$|f(x) - g_1(x)| \leqslant \frac{2}{3}M, \forall x \in F = A \cup B \cup C.$$

Let $f_1(x) = f(x) - g_1(x)$. Applying the same argument to f_1, there exists a continuous function g_2 on $\mathbb{R}^n$ (constructed via distance functions) such that

$$|g_2(x)| \leqslant \frac{1}{3}(\frac{2}{3}M),$$

$$|f_1(x) - g_2(x)| = \left|f(x) - \sum_{i=1}^{2} g_i(x)\right| \leqslant (\frac{2}{3})^2 M, \forall x \in F.$$

Continuing inductively, there exist

$$G_n(x) = \sum_{i=1}^{n} g_i(x),$$

with

$$|g_n(x)| \leqslant \frac{1}{3}(\frac{2}{3})^{n-1} M, \forall x \in \mathbb{R}^n,$$

$$|f(x) - G_n(x)| \leqslant (\frac{2}{3})^n M, \forall x \in F.$$

[8] If both A, B are empty, replace M by $M/3$ and iterate. If the bound can be reduced arbitrarily, then $f \equiv 0$. Otherwise after finitely many steps at least one of A, B is nonempty. If one is empty, replace it with any disjoint closed set.

Now prove $\{G_n\}$ converges uniformly on $\mathbb{R}^n$:

$$\begin{aligned}|G_{n+p}(x)-G_n(x)| &= \Big|\sum_{i=n+1}^{n+p} g_i(x)\Big| \leqslant \sum_{i=n+1}^{n+p} |g_i(x)| \\ &\leqslant \sum_{i=n+1}^{\infty} \frac{1}{3}\big(\frac{2}{3}\big)^{i-1} M \\ &\leqslant \frac{1}{3}\big(\frac{2}{3}\big)^{n} M \sum_{k=0}^{\infty} \big(\frac{2}{3}\big)^{k}.\end{aligned}$$

Hence $\{G_n\}$ converges uniformly on $\mathbb{R}^n$. So

$$G(x) = \lim_{n\to\infty} G_n(x) = \sum_{n=1}^{\infty} g_n(x)$$

is continuous on $\mathbb{R}^n$, and

$$|G(x)| = \Big|\sum_{n=1}^{\infty} g_n(x)\Big| \leqslant M.$$

For $x \in F$, we have $f(x) \equiv G(x)$. □

Remark 1.3.6. *In the extension theorem above, we essentially used a special case of Urysohn's lemma in $\mathbb{R}^n$. Urysohn's lemma is one of the deeper theorems in topology, not easy to prove in full generality. But in $\mathbb{R}^n$ (more generally, metric spaces), the distance function makes the proof straightforward.*

Theorem 1.3.11 (Urysohn Lemma)**.** *Two closed sets in a normal space can be separated by a continuous function.*

Theorem 1.3.12 (Urysohn Lemma in $\mathbb{R}^n$)**.** *Let $E, F \subset \mathbb{R}^n$ be disjoint closed sets. Then there exists $f \in C(\mathbb{R}^n)$ such that $0 \leqslant f \leqslant 1$, with $E \subset f^{-1}(\{1\})$ and $F \subset f^{-1}(\{0\})$.*

Proof. Construct directly:

$$f(x) = \frac{\mathrm{d}(x,F)}{\mathrm{d}(x,E)+\mathrm{d}(x,F)}.$$

□

Another frequently used form is:

Theorem 1.3.13 (Urysohn')**.** *Let K be compact and V open with $K \subset V \subset \mathbb{R}^n$. Then there exists a continuous function f with compact support*[9] *such that $0 \leqslant f \leqslant 1$, $\mathrm{Supp}(f) \subset V$, and $f|_K = 1$. Equivalently, $\chi_K \leqslant f \leqslant \chi_V$.*

Think: (Partition of unity) Let $K \subset \mathbb{R}^n$ be compact and $\{V_i\}_{i=1}^k$ an open cover of K. Then there exists a family of continuous compactly supported functions $\{h_i\}_{i=1}^k$ such that $0 \leqslant h_i \leqslant 1$, $\mathrm{Supp}(h_i) \subset V_i$, and

$$\sum_{i=1}^{k} h_i(x) = 1, x \in K.$$

Finally, we briefly introduce semicontinuity.

[9] The support of f is

$$E = \overline{\{x | f(x) \neq 0\}},$$

denoted by $\mathrm{Supp}(f)$. By definition the support is closed; if also bounded, it is compact support.

Definition 1.3.14 (Semicontinuity)**.** *A function f is upper semicontinuous if for all $\varepsilon > 0$, there exists $\delta > 0$ such that for all $y \in \mathrm{B}(x, \delta)$,*

$$f(y) < f(x) + \varepsilon.$$

A function f is lower semicontinuous if for all $\varepsilon > 0$, there exists $\delta > 0$ such that for all $y \in \mathrm{B}(x, \delta)$,

$$f(x) - \varepsilon < f(y).$$

Remark 1.3.7. *f is upper semicontinuous iff $\{x|f(x) < \lambda\}$ is open for every $\lambda \in \mathbb{R}^1$. f is lower semicontinuous iff $\{x|f(x) > \lambda\}$ is open for every $\lambda \in \mathbb{R}^1$.*

Proposition 1.3.2. *Semicontinuous functions satisfy:*

(1) *If $\{f_\lambda\}_{\lambda\in\Lambda}$ is a family of lower semicontinuous functions and*

$$(\sup_{\lambda\in\Lambda} f_\lambda)(u) = \sup_{\lambda\in\Lambda} \{f_\lambda(u)\},$$

then $\sup\limits_{\lambda\in\Lambda} f_\lambda$ is lower semicontinuous;

(2) *finite sums of lower semicontinuous functions are lower semicontinuous;*

(3) *lower semicontinuous functions attain their minimum on compact sets;*

(4) *if F, G are open and closed subsets of $\mathbb{R}^1$, then χ_F is upper semicontinuous and χ_G is lower semicontinuous;*

(5) *if f_1 is upper semicontinuous, f_2 is lower semicontinuous, and $f_1 \leqslant f_2$, then there exists a continuous f such that $f_1 \leqslant f \leqslant f_2$.*

Proof. We only prove (3). Let K be compact and f lower semicontinuous. By definition, for fixed $\varepsilon > 0$ and each $x \in K$, there exists $\delta_x > 0$ such that for all $y \in \mathrm{B}(x, \delta_x)$,

$$f(y) < f(x) + \varepsilon.$$

Therefore $\{\mathrm{B}(x, \delta_x)\}_{x\in K}$ is an open cover of K. By Theorem 1.3.7, there is a finite subcover

$$\{\mathrm{B}(x_i, \delta_{x_i})\}, i = 1, 2, \cdots, n.$$

Hence f is bounded below and thus has an infimum. Let

$$\lambda = \inf_{x\in K} f(x).$$

By definition of infimum, there exists $\{x_n\}$ with $f(x_n) \to \lambda$. Since $K \subset \mathbb{R}^1$ is bounded and closed, by Bolzano–Weierstrass there is a subsequence $x_{k_i} \to x_0$. Then

$$\lambda = \inf_{i\to\infty} f(x_{k_i}) = \liminf_{i\to\infty} f(x_{k_i}) \geqslant f(x_0) \geqslant \lambda.$$

Hence $f(x_0) = \lambda$. □

Exercises 1.3

1. Let $X = \{G \subset \mathbb{R}^n | G \text{ open}\}$. Find $\overline{\overline{X}}$.

2. P36 1,2,3,4,5,7

3. P40 2,5

4. P62 1,2,3,4[10],5,6

5. P67 40,41

6. P69 28

1.3.5 Cantor Set

We now recursively define the middle-third Cantor set.

At step 1, set $I_0 = (0,1)$, remove the middle closed interval of length 1/3, namely $I^{1,1} = [\frac{1}{3}, \frac{2}{3}]$, and obtain $I_1 = I_0 \backslash I^{1,1}$, which has two path-connected components $I_{1,1} = (a_{1,1}, b_{1,1})$ and $I_{1,2} = (a_{1,2}, b_{1,2})$. Write $I^1 = I^{1,1}$.

At step 2, for each $I_{1,j}$, $j \in \{1,2\}$, remove the middle closed third,

$$I^{2,j} = [a_{1,j} + \frac{1}{3}(b_{1,j} - a_{1,j}), a_{1,j} + \frac{2}{3}(b_{1,j} - a_{1,j})],$$

and write $I^2 = \bigcup_{j=1}^{2} I^{2,j}$. The remaining set is $I_2 = I_1 \backslash I^2$, with four path-connected components $I_{2,j} = (a_{2,j}, b_{2,j})$, $j \in \{1,2,3,4\}$.

Continue inductively. Suppose step n has been completed, yielding I_n with 2^n path-connected components $I_{n,j} = (a_{n,j}, b_{n,j})$, $j \in \{1, 2, \cdots, 2^n\}$. At step $n+1$, for each $I_{n,j}$ remove the middle closed third,

$$I^{n+1,j} = [a_{n,j} + \frac{1}{3}(b_{n,j} - a_{n,j}), a_{n,j} + \frac{2}{3}(b_{n,j} - a_{n,j})],$$

set $I^{n+1} = \bigcup_{j=1}^{2^n} I^{n+1,j}$, and let $I_{n+1} = I_n \backslash I^{n+1}$. Then I_{n+1} has 2^{n+1} path-connected components $I_{n+1,j} = (a_{n+1,j}, b_{n+1,j})$, $j \in \{1, 2, \cdots, 2^{n+1}\}$.

Thus I_k and I^k are defined recursively, and the Cantor middle-third set is

$$\mathcal{C} = \Big(\bigcap_{k=1}^{\infty} I_k\Big) \cup \{0,1\} = [0,1] \backslash \Big(\bigcup_{k=1}^{\infty} I^k\Big).$$

The Cantor set $\mathcal{C}$ has the following properties.

Proposition 1.3.3. *Let $\mathcal{C}$ be the Cantor middle-third set. Then*

(1) *$\mathcal{C}$ is closed;*

(2) $\sum_{k=1}^{\infty} |I^k| = \sum_{k=1}^{\infty} \frac{2^{k+1}}{3^k} = 1$*;*

(3) $\overline{\overline{\mathcal{C}}} = \aleph$*;*

(4) $\overset{\circ}{\mathcal{C}} = \varnothing$*;*

(5) $\mathcal{C} \subset \mathcal{C}' (\Rightarrow \mathcal{C} = \mathcal{C}')$.

The general notion is the Cantor set, namely a nowhere dense perfect set. We first define perfect sets.

[10] This problem seems misstated; F should likely be replaced by a bounded closed subset of G.

Definition 1.3.15 (Perfect Set)**.** *A set E is called perfect if*

$$E = E'.$$

To aid understanding, we give a characterization in one dimension.

Proposition 1.3.4. *$E \subset \mathbb{R}^1$ is perfect iff E^c is open and*

$$E^c = \bigcup_{n=1}^{N} (a_n, b_n),$$

where N may be ∞, and for $n \neq m$, (a_m, b_m) and (a_n, b_n) share no endpoints.

Proof. On one hand, since E is perfect, i.e. $E = E'$, we have $\overline{E} = E \cup E' = E$, so E is closed and E^c is open. By Theorem 1.3.4, E^c is a countable union of disjoint connected open sets, and in $\mathbb{R}^1$ connected open sets are open intervals. Also, perfectness implies no isolated points in E, equivalently component intervals of E^c cannot share endpoints.

On the other hand,

$$E^c \text{ open} \Rightarrow E \text{ closed} \Rightarrow E' \subset E,$$

$$E \text{ has no isolated points} \Rightarrow E \subset E',$$

thus the claim holds. □

Next define nowhere dense sets.

Definition 1.3.16 (Nowhere Dense Set)**.** *Let $A, B \subset \mathbb{R}^n$. If $\overline{A} \supseteq B$, then A is dense in B. If $\overset{\circ}{\overline{A}} = \varnothing$, equivalently for any $x_0 \in \mathbb{R}^n$ and any $\varepsilon > 0$, there exist $x \in B_\varepsilon(x_0)$ and $\delta > 0$ such that $B_\delta(x) \subseteq B_\varepsilon(x_0)$ and*

$$B_\delta(x) \cap A = \varnothing,$$

then A is called nowhere dense (also sparse).

Further, a countable union of nowhere dense sets is called a first category set; sets that are not first category are second category. These notions depend on the ambient space.

Example 1.3.3. *The x-axis is of second category in $\mathbb{R}^1$, but of first category in $\mathbb{R}^2$.*

Now we characterize one-dimensional nowhere dense perfect sets.

Let $E \in \mathbb{R}^n$ be a bounded nowhere dense perfect set, and define

$$m = \inf\{x | x \in E\}, M = \sup\{x | x \in E\}.$$

Since E is closed, $m \in E$ and $M \in E$. By Proposition 1.3.4,

$$E^c = (-\infty, m) \cup (M, \infty) \bigcup_{n=1}^{\infty} (a_n, b_n],$$

and for $n \neq m$, intervals (a_m, b_m) and (a_n, b_n) share no endpoints. Assume $b_n < a_m$. Since E is nowhere dense, there are points of E^c in b_n, a_m. Hence there exists an open interval $(c, d) \subset E^c$, which is another component interval lying between (a_m, b_m) and (a_n, b_n).

This shows that between any two component intervals of E^c, there is another one, and the three have no common endpoints.

The following proposition reveals the structure of complements of one-dimensional nowhere dense perfect sets and also gives the famous Cantor function.

Proposition 1.3.5. *Let E be as above. Then there exists $\varphi : [m, M] \to [m, M]$ with:*

(1) *φ is monotone, continuous, and surjective;*

(2) *$\varphi|_{(a_n, b_n)} \equiv const.$*

Moreover, one can reorder the component intervals of E^c so that their order resembles that of the complement of the Cantor set.

Proof. Step 1. Reorder component intervals of E^c analogously to the Cantor complement order. By Proposition 1.3.4,

$$E^c = (-\infty, m) \cup (M, \infty) \bigcup_{n=1}^{\infty} (a_n, b_n],$$

set $(c_{1,1}, d_{1,1}) = (a_1, b_1)$. Choose the first interval in original order between $(-\infty, m)$ and (a_1, b_1), denote it by $(c_{2,1}, d_{2,1})$, i.e.

$$n_{2,1} = \inf\{n | a_n > m, b_n < a_1\},$$

and set $(c_{2,1}, d_{2,1}) = (a_{n_{2,1}}, b_{n_{2,1}})$. Similarly choose the first interval between (a_1, b_1) and $(M, +\infty)$, denoted $(c_{2,2}, d_{2,2})$, i.e.

$$n_{2,2} = \inf\{n | a_n > b_1, b_n < M\},$$

and set $(c_{2,2}, d_{2,2}) = (a_{n_{2,2}}, b_{n_{2,2}})$. Suppose after step k we have chosen

$$O_1 = (c_{1,1}, d_{1,1}),$$

$$O_2 = (c_{2,1}, d_{2,1}) \cup (c_{2,2}, d_{2,2}),$$

$$\cdots$$

$$O_k = \bigcup_{j=1}^{2^{k-1}} (c_{k,j}, d_{k,j}).$$

Now reorder component intervals in $\bigcup_{i=1}^{k} O_i$ by position; between each adjacent pair choose one component interval of E^c with minimal original index. This gives 2^k intervals

$$(c_{k+1,j}, d_{k+1,j}), j = 1, 2, \cdots, 2^k,$$

and define

$$O_{k+1} = \bigcup_{j=1}^{2^k} (c_{k+1,j}, d_{k+1,j}), k = 0, 1, 2, \cdots.$$

We claim $E^c = (-\infty, m) \cup (M, +\infty) \bigcup_{i=1}^{\infty} O_i$. Clearly $\bigcup_{i=1}^{k} O_i \subset E^c$. Conversely, every (a_n, b_n) is eventually chosen: if $(a_1, b_1), \cdots, (a_n, b_n)$ are chosen by step k but (a_{n+1}, b_{n+1}) is not, then it lies between two component intervals of $\bigcup_{i=1}^{k} O_i$, so it must be chosen at step $k+1$ by construction.

Step 2. Construct $\varphi : [m, M] \to [m, M]$. Let $O_0 = \{m, M\}$ and define on O_0:

$$\varphi_0(x) = \begin{cases} m & x = m \\ M & x = M \end{cases},$$

so jump size between adjacent components is $\frac{M-m}{2^0}$. On $O_0 \cup O_1$, define

$$\varphi_1(x) = \begin{cases} \frac{m+M}{2}, & x \in (c_{1,1}, d_{1,1}) \\ \varphi_0(x) & x \in O_0 \end{cases},$$

then jump size is $\frac{M-m}{2^1}$. On $O_0 \cup O_1 \cup O_2$, define

$$\varphi_2(x) = \begin{cases} \frac{m+\varphi_1|_{(c_{1,1},d_{1,1})}}{2} & x \in (c_{2,1}, d_{2,1}) \\ \frac{M+\frac{m+M}{2}}{2} & x \in (c_{2,2}, d_{2,2}) \\ \varphi_1(x) & x \in O_1 \end{cases},$$

then jump size is $\frac{M-m}{2^2}$. Continuing this process, define φ_{k+1} from φ_k. On each $(c_{k+1,j}, d_{k+1,j})$, set the value to the average of neighboring component values. Then the jump size between adjacent components is

$$\frac{M-m}{2^k}.$$

Proceeding inductively yields a monotone increasing function ψ on $\{m, M\} \cup \left(\bigcup_{i=1}^{\infty} O_i\right)$, constant on each $(c_{k,j}, d_{k,j})$, and with jump size $\frac{M-m}{2^k}$ between adjacent components of $\bigcup_{i=1}^{k} O_i$.

Now define $\varphi|_{O_n} = \psi|_{O_n}$ and

$$\varphi(x) = \sup\Big\{\psi(y) | y \leqslant x, y \in \{m, M\} \cup \Big(\bigcup_{i=1}^{\infty} O_i\Big)\Big\},$$

so φ is monotone increasing.

Step 3. Prove continuity of φ. For any $\varepsilon > 0$ and fixed $x \in [m, M]$, choose k large so that $\frac{M-m}{2^k} < \varepsilon$. If $x \in \bigcup_{i=1}^{k} O_i$, continuity at x is clear. If $x \in [m, M] \backslash \bigcup_{i=1}^{k} O_i$, then x lies between two neighboring components of $\bigcup_{i=1}^{k} O_i$, say left $(c_{i,j}, d_{i,j})$ and right $(c'_{i,j}, d'_{i,j})$. Let $\delta = \max\{x - c_{i,j}, d'_{i,j} - x\}$. For $y \in (x - \delta, x + \delta)$,

$$|\varphi(y) - \varphi(x)| \leqslant \psi\Big(\frac{c'_{i,j} + d'_{i,j}}{2}\Big) - \psi\Big(\frac{c_{i,j} + d_{i,j}}{2}\Big) < \frac{M-m}{2^k} < \varepsilon.$$

Hence φ is continuous. □

Remark 1.3.8. *This proposition shows that, up to ordering of component intervals, complements of one-dimensional nowhere dense perfect sets have the same structure as the complement of the Cantor set.*

1.4 Borel Sets and the Category Theorem

We now introduce Borel sets, which are fundamental in real analysis and probability theory. First we define the σ-ring mentioned in the overview.

Definition 1.4.1 (σ-Ring)**.** *Let X be nonempty and $\Gamma \subset \mathcal{P}(X)$. We call Γ a σ-ring on X if*

(1) *for all $A, B \in \Gamma$, $A \backslash B \in \Gamma$;*

(2) *for $\{A_k\}_{k=1}^{\infty} \subset \Gamma$, $\bigcup\limits_{k=1}^{\infty} A_k \in \Gamma$.*

Remark 1.4.1. *1. From the definition,*

$$A \cap B = (A \cup B) \backslash ((A \backslash B) \cup (B \backslash A)) \in \Gamma.$$

Also from $A_1 \backslash \Big(\bigcap\limits_{k=1}^{\infty} A_k\Big) = \bigcup\limits_{k=1}^{\infty} (A_1 \backslash A_k) \in \Gamma$, we get

$$\bigcap_{k=1}^{\infty} A_k = A_1 \backslash \Big(A_1 \backslash \big(\bigcap_{k=1}^{\infty} A_k\big)\Big) \in \Gamma.$$

2. The intersection of any family of σ-rings is still a σ-ring.

Definition 1.4.2 (Generated σ-Ring)**.** *Let X be nonempty and $\Sigma \subset \mathcal{P}(X)$. Define*

$$U = \{\Gamma' \in \mathcal{P}(X) | \Sigma \subset \Gamma', \textit{and } \Gamma' \textit{ is a } \sigma\textit{-ring}\}.$$

Then

$$\Gamma(\Sigma) = \bigcap_{\Gamma' \in U} \Gamma'$$

is called the σ-ring generated by Σ.

Remark 1.4.2. *In fact, $\Gamma(\Sigma)$ is the smallest σ-ring containing Σ.*

Now we introduce an important generated σ-ring: the Borel ring.

Definition 1.4.3 (Borel Ring)**.** *Let $\mathcal{B}(\mathbb{R}^n)$ be the σ-ring generated by all open and closed subsets of $\mathbb{R}^n$. This is called the Borel ring on $\mathbb{R}^n$.*

Remark 1.4.3. *1. Clearly $\mathcal{B}(\mathbb{R}^n) \subset \mathcal{P}(\mathbb{R}^n)$, so $\overline{\overline{\mathcal{B}(\mathbb{R}^n)}} \leqslant 2^{\aleph}$. Is $\overline{\overline{\mathcal{B}(\mathbb{R}^n)}} = \aleph$? Intuitively yes, but the proof is nontrivial and requires transfinite induction, which we omit.*

2. From the definition, $(a, b] = \bigcap_{n=1}^{\infty}(a, b + 1/n)$ and $\mathbb{Q} = \bigcup_{n=1}^{\infty} r_n$ are both in $\mathcal{B}(\mathbb{R}^n)$.

We can further classify elements of $\mathcal{B}(\mathbb{R}^n)$; the simplest are G_δ and F_σ sets.

Definition 1.4.4 (G_δ, F_σ Sets)**.** *$A \subset \mathbb{R}^n$ is called a G_δ set if there exists a sequence of open sets $\{G_k\}_{k=1}^{\infty}$ such that*

$$A = \bigcap_{k=1}^{\infty} G_k;$$

$A \subset \mathbb{R}^n$ is called an F_σ set if there exists a sequence of closed sets $\{F_k\}_{k=1}^{\infty}$ such that

$$A = \bigcup_{k=1}^{\infty} F_k.$$

Remark 1.4.4. *G_δ sets are necessarily uncountable, but there exist G_δ sets of very small "length".*

Let us look at examples.

Example 1.4.1. *Any closed set $F \subset \mathbb{R}^n$ is a G_δ set. Indeed, define*

$$G_n = \{x \in \mathbb{R}^n | \mathrm{dist}(x, F) < \frac{1}{n}\},$$

then

$$F = \bigcap_{n=1}^{\infty} G_n.$$

Example 1.4.2. *$\mathbb{R} \backslash \mathbb{Q}$ is not an F_σ set, and $\mathbb{Q}$ is not a G_δ set.*

To prove Example 1.4.2, we need the famous Baire category theorem.

Example 1.4.3 (Baire Category Theorem)**.** *$\mathbb{R}^n$ cannot be expressed as a countable union of nowhere dense sets*[11]*.*

[11] Equivalently, $\mathbb{R}^n$ is of second category. More generally: a nonempty complete metric space cannot be written as a countable union of nowhere dense sets.

Proof. Assume $\mathbb{R}^n = \bigcup_{n=1}^{\infty} A_n$, where each A_n is nowhere dense. Choose $x_0 \in \mathbb{R}^n$, let $B_1 = \overline{B(x_0, 1)}$. Since A_1 is nowhere dense, it is not dense in $B(x_0, 1)$, so there exist $x_1 \in B(x_0, 1) \backslash \overline{A_1}$ and $\delta_1 (< 1/2)$ with

$$\overline{B(x_1, \delta_1)} \cap \overline{A_1} = \varnothing, \overline{B(x_1, \delta_1)} \subset B_1.$$

Let $B_2 = \overline{B(x_1, \delta_1)}$. Since A_2 is nowhere dense, there exist $x_2 \in B(x_1, \delta_1) \backslash \overline{A_2}$ and $\delta_2 (< 1/3)$ such that

$$\overline{B(x_2, \delta_2)} \cap \overline{A_2} = \varnothing, \overline{B(x_2, \delta_2)} \subset B_2.$$

Continue inductively to obtain $\{x_n\} \subset \mathbb{R}^n$ and $\delta_n (< 1/n)$ such that

$$\overline{B(x_{k+1}, \delta_{k+1})} \cap \overline{\bigcup_{i=1}^{k} A_i} = \varnothing,$$

$$\overline{B(x_{k+1}, \delta_{k+1})} \subset \overline{B(x_k, \delta_k)} \subset \cdots \subset \overline{B(x_0, 1)}.$$

Since these are closed balls and $\delta_k \to 0$, $\{x_n\}$ is Cauchy. Let $x^* \in \mathbb{R}^n$ be its limit. Then

$$x^* \in \bigcap_{k=1}^{\infty} \overline{B(x_k, \delta_k)}. \tag{1.2}$$

By assumption $\mathbb{R}^n = \bigcup_{n=1}^{\infty} A_n$, so $x^* \in A_{k_0}$ for some k_0. From (1.2),

$$x^* \in \overline{B(x_{k_0+1}, \delta_{k_0+1})},$$

contradicting $\overline{B(x_{k_0+1}, \delta_{k_0+1})} \cap A_{k_0} = \varnothing$. □

Now reconsider Example 1.4.2. Suppose $\mathbb{R} \backslash \mathbb{Q} = \bigcup_{n=1}^{\infty} F_n$ with each F_n closed. Then

$$\begin{aligned} \mathbb{Q} \cap F_n = \varnothing &\Rightarrow F_n \text{ is nowhere dense} \\ &\Rightarrow \mathbb{R} = \Big(\bigcup_{n=1}^{\infty} F_n \Big) \cup \Big(\bigcup_{r \in \mathbb{Q}} r \Big) \\ &\Rightarrow \mathbb{R} \text{ can be written as a union of countably many nowhere dense closed sets.} \end{aligned}$$

This contradicts Theorem 1.4.3.

Exercises 1.4

1. P59 4,5,6,7

Chapter 2

Lebesgue Measurable Sets

2.1 Measures on Rings

2.1.1 The Ring $\mathcal{R}_0^n$

Definition 2.1.1. *Let $\mathcal{R}^n$ be a family of subsets of $\mathbb{R}^n$. If whenever $A, B \in \mathcal{R}^n$ we have $A \cup B, A \backslash B \in \mathcal{R}^n$, then $\mathcal{R}^n$ is called a ring on $\mathbb{R}^n$. In particular, if $\mathbb{R}^n \in \mathcal{R}^n$, then $\mathcal{R}^n$ is called an algebra on $\mathbb{R}^n$.*

Remark 2.1.1. *Note the difference from the σ-ring defined in Section 1.4.*

Now let us examine a concrete example. The ring in this example and its higher-dimensional generalization are crucial in what follows.

Example 2.1.1. *Let $\mathcal{R}_0^1$ be the family of all finite unions of bounded left-open right-closed intervals in $\mathbb{R}^1$:*

$$\mathcal{R}_0^1 = \Big\{ \bigcup_{k=1}^{m} (a_k, b_k] \mid a_k, b_k \in \mathbb{R}^1,\ k = 1, 2, \dots, m \Big\}.$$

Then $\mathcal{R}_0^1$ is a ring on $\mathbb{R}^1$.

Proof. Let

$$A = \bigcup_{k=1}^{m} (a_k, b_k], \quad B = \bigcup_{j=1}^{l} (c_j, d_j] \in \mathcal{R}_0^1.$$

Then

$$A \cup B = \bigcup_{k=1}^{m} \bigcup_{j=1}^{l} \Big((a_k, b_k] \cup (c_j, d_j] \Big) \in \mathcal{R}_0^1,$$

and

$$\begin{aligned} A \backslash B &= \bigcup_{k=1}^{m} \Big((a_k, b_k] \backslash \bigcup_{j=1}^{l} (c_j, d_j] \Big) \\ &= \bigcup_{k=1}^{m} \bigcap_{j=1}^{l} \Big((a_k, b_k] \backslash (c_j, d_j] \Big). \end{aligned}$$

So it suffices to verify $(a_k, b_k] \backslash (c_j, d_j] \in \mathcal{R}_0^1$ and closure under finite intersections. □

We now give a canonical decomposition for elements in $\mathcal{R}_0^1$.

Lemma 2.1.1 (Canonical Decomposition)**.** *If $A \in \mathcal{R}_0^1$, then A can be written as a finite union of pairwise disjoint left-open right-closed intervals.*

Proof. Write

$$A = \bigcup_{k=1}^{m} (a_k, b_k].$$

Use induction on m. The case $m = 1$ is clear. Assume true for $m = l$. For $m = l + 1$,

$$A = \bigcup_{k=1}^{l} (a_k, b_k] \cup (a_{l+1}, b_{l+1}].$$

Sort left endpoints and assume $a_{l+1} \geq a_k$ for $k = 1, \dots, l$.

(1) If $a_{l+1} > \max_{1 \leq k \leq l} b_k$, then by induction

$$\widehat{A} = \bigcup_{k=1}^{l} (a_k, b_k] = \bigcup_{k=1}^{\widehat{l}} (\widehat{a}_k, \widehat{b}_k],$$

so

$$A = \bigcup_{k=1}^{\widehat{l}} (\widehat{a}_k, \widehat{b}_k] \cup (a_{l+1}, b_{l+1}]$$

is the required form.

(2) If there exists k_0 with $1 \leq k_0 \leq l$ such that $b_{k_0} \geq a_{l+1} \geq a_{k_0}$, let

$$b'_{k_0} = \max\{b_{k_0}, b_{l+1}\}.$$

Then

$$(a_{k_0}, b_{k_0}] \cup (a_{l+1}, b_{l+1}] = (a_{k_0}, b'_{k_0}],$$

and the conclusion follows from induction. □

Next we generalize Example 2.1.1 to higher dimensions. We first need a lemma: the Cartesian product of rings is still a ring.

Lemma 2.1.2. *Let $\mathcal{R}^n$ and $\mathcal{R}^m$ be rings of subsets of $\mathbb{R}^n$ and $\mathbb{R}^m$, respectively. Define*

$$\mathcal{R}^n \times \mathcal{R}^m = \Big\{ \bigcup_{k=1}^{r} A_k \times B_k \mid A_k \in \mathcal{R}^n,\ B_k \in \mathcal{R}^m,\ k = 1, \dots, r \Big\}.$$

Then $\mathcal{R}^n \times \mathcal{R}^m$ is a ring on $\mathbb{R}^{n+m}$.

Proof. Let

$$A = \bigcup_{k=1}^{r} (A_k \times B_k), \quad B = \bigcup_{j=1}^{s} (C_j \times D_j) \in \mathcal{R}^n \times \mathcal{R}^m,$$

with $A_k, C_j \in \mathcal{R}^n$ and $B_k, D_j \in \mathcal{R}^m$. Then

$$A \cup B = \bigcup_{k=1}^{r} \bigcup_{j=1}^{s} \Big(A_k \times B_k \cup C_j \times D_j \Big) \in \mathcal{R}^n \times \mathcal{R}^m.$$

Also,

$$A \cap B = \bigcup_{k=1}^{r} \bigcup_{j=1}^{s} \Big((A_k \cap C_j) \times (B_k \cap D_j) \Big) \in \mathcal{R}^n \times \mathcal{R}^m.$$

And

$$\begin{aligned} A\backslash B &= \bigcup_{k=1}^{r} \bigcap_{j=1}^{s} \Big((A_k \times B_k) \backslash (C_j \times D_j) \Big) \\ &= \bigcup_{k=1}^{r} \bigcap_{j=1}^{s} \Big(((A_k \backslash C_j) \times B_k) \cup (A_k \times (B_k \backslash D_j)) \Big) \in \mathcal{R}^n \times \mathcal{R}^m. \end{aligned}$$

So the family is closed under union and difference, hence it is a ring. □

Now define the higher-dimensional analog.

Definition 2.1.2 (The Ring $\mathcal{R}_0^n$)**.** *Let $\mathcal{R}_0^n$ be the family of all finite unions of bounded left-open right-closed boxes in $\mathbb{R}^n$:*

$$\mathcal{R}_0^n = \Big\{ \bigcup_{k=1}^{l} I_k \mid I_k \textit{ is a bounded left-open right-closed box in } \mathbb{R}^n \Big\},$$

where

$$I_k = \{(x_1, \dots, x_n) \in \mathbb{R}^n \mid a_i < x_i \le b_i,\ i = 1, \dots, n\}.$$

We will see that $\mathcal{R}_0^n$ is also a ring on $\mathbb{R}^n$.

Theorem 2.1.1. *With the definition above, $\mathcal{R}_0^n$ is a ring on $\mathbb{R}^n$, and*

$$\mathcal{R}_0^n = \mathcal{R}_0^1 \times \mathcal{R}_0^1 \times \cdots \times \mathcal{R}_0^1.$$

Proof. By Lemma 2.1.2,

$$\mathcal{R}_0^1 \times \mathcal{R}_0^1 \times \cdots \times \mathcal{R}_0^1$$

is a ring on $\mathbb{R}^n$, and it contains all bounded left-open right-closed boxes. Hence

$$\mathcal{R}_0^n \subset \mathcal{R}_0^1 \times \cdots \times \mathcal{R}_0^1.$$

It remains to prove the reverse inclusion.

It is enough to show every element of $\mathcal{R}_0^1 \times \cdots \times \mathcal{R}_0^1$ can be written as a finite union of left-open right-closed boxes. Use induction on dimension. For $n = 1$, trivial. Assume true in dimension $n-1$. Let

$$A \in \mathcal{R}_0^1 \times \cdots \times \mathcal{R}_0^1 \subset \mathbb{R}^n, \quad A = B \times C,$$

where

$$B \in \mathcal{R}_0^1 \times \cdots \times \mathcal{R}_0^1 \subset \mathbb{R}^{n-1}, \quad C \in \mathcal{R}_0^1.$$

By induction,

$$B = \bigcup_{k=1}^{r} I_k^{n-1}, \quad C = \bigcup_{j=1}^{s} (c_j, d_j].$$

Hence

$$B \times C = \bigcup_{k=1}^{r} \bigcup_{j=1}^{s} (I_k^{n-1} \times (c_j, d_j]),$$

and each factor is a box in $\mathbb{R}^n$. So

$$\mathcal{R}_0^n \supset \mathcal{R}_0^1 \times \cdots \times \mathcal{R}_0^1.$$

□

The next lemma is the higher-dimensional version of Lemma 2.1.1.

Lemma 2.1.3 (Canonical Decomposition)**.** *If $A \in \mathcal{R}_0^n$, then A can be written as a finite union of pairwise disjoint left-open right-closed boxes.*

Proof. Induct on n. For $n = 1$, this is Lemma 2.1.1. Assume true for $n = k$. For $n = k+1$, let

$$A = \bigcup_{r=1}^{m} I_r^{k+1} \in \mathcal{R}_0^{k+1},$$

with

$$I_r^{k+1} = I_r^k \times (a_r, b_r],$$

where I_r^k is a box in $\mathbb{R}^k$. Reorder

$$a_1, b_1, a_2, b_2, \dots, a_m, b_m$$

as

$$a^0 < a^1 < \cdots < a^l, \quad 2 \le l \le 2m.$$

For each I_r^{k+1}, there are boxes $I_{r,i}^{k+1}$ (possibly empty) such that

$$I_r^{k+1} = \bigcup_{i=1}^{l} (I_{r,i}^{k+1} \times (a^{i-1}, a^i]).$$

Apply the induction hypothesis in $\mathbb{R}^k$. □

2.1.2 Measure on $\mathcal{R}_0^n$

Now we construct a measure on the ring $\mathcal{R}_0^n$.

Definition 2.1.3. *Let $\mathcal{R}_0^n$ be the ring of finite unions of bounded left-open right-closed boxes in $\mathbb{R}^n$. A real-valued function μ on $\mathcal{R}_0^n$ is called a measure on $\mathcal{R}_0^n$ if:*

(1) $\mu(\varnothing) = 0$*;*

(2) *(nonnegativity) $\mu(A) \ge 0$ for all $A \in \mathcal{R}_0^n$;*

(3) *(countable additivity) for any pairwise disjoint sequence $\{A_i\}_{i=1}^{\infty} \subset \mathcal{R}_0^n$, if*

$$\bigcup_{i=1}^{\infty} A_i \in \mathcal{R}_0^n, \tag{2.1}$$

then

$$\mu\Big(\bigcup_{i=1}^{\infty} A_i\Big) = \sum_{i=1}^{\infty} \mu(A_i).$$

Remark 2.1.2. *In* (2.1)*, we require the countable union to lie in $\mathcal{R}_0^n$ because $\mathcal{R}_0^n$ is only closed under finite unions.*

Definition 2.1.4. *For $A \in \mathcal{R}_0^n$, suppose a canonical decomposition is*

$$A = \bigcup_{i=1}^{k} I_i.$$

Define

$$m_0(A) = \sum_{i=1}^{k} |I_i|,$$

where

$$I_i = \{(x_1^i, \dots, x_n^i) \mid a_j^i < x_j^i \le b_j^i,\ j = 1, \dots, n\},$$

$$|I_i| = \prod_{j=1}^{n} (b_j^i - a_j^i) = \int_{I_i} \chi_{I_i}(x)\, dx.$$

Clearly, $m_0(A) \ge 0$ *and* $m_0(\varnothing) = 0$.

Since a given $A \in \mathcal{R}_0^n$ can have multiple canonical decompositions, we must show $m_0(A)$ is independent of the decomposition.

Lemma 2.1.4. *The definition of* $m_0(A)$ *does not depend on the chosen canonical decomposition of* A.

Proof. Step 1. If I is a single box, then $m_0(I) = |I|$, independent of decomposition.

Suppose

$$I = \bigcup_{i=1}^{k} I_i$$

is a canonical decomposition of I. Then

$$\begin{aligned} |I| &= \int_I \chi_I(x)\, dx = \int_{\bigcup_{i=1}^{k} I_i} \chi_{\cup I_i}(x)\, dx \\ &= \sum_{i=1}^{k} \int_{I_i} \chi_{I_i}(x)\, dx = \sum_{i=1}^{k} |I_i|. \end{aligned}$$

Step 2. Suppose $A \in \mathcal{R}_0^n$ has two canonical decompositions

$$A = \bigcup_{i=1}^{k} I_i = \bigcup_{j=1}^{l} J_j.$$

Then

$$A = \bigcup_{i=1}^{k} \bigcup_{j=1}^{l} (I_i \cap J_j),$$

$$J_j = \bigcup_{i=1}^{k} (I_i \cap J_j), \quad I_i = \bigcup_{j=1}^{l} (I_i \cap J_j).$$

By Step 1,

$$\begin{aligned} m_0(A) &= \sum_{i=1}^{k} |I_i| = \sum_{i=1}^{k} \sum_{j=1}^{l} |I_i \cap J_j| \\ &= \sum_{j=1}^{l} \sum_{i=1}^{k} |I_i \cap J_j| = \sum_{j=1}^{l} |J_j|. \end{aligned}$$

So $m_0(A)$ is well-defined. □

Lemma 2.1.5. *The set function m_0 on $\mathcal{R}_0^n$ has:*

(1) *(finite additivity) if $A_1, \ldots, A_k \in \mathcal{R}_0^n$ are pairwise disjoint, then*

$$m_0\Big(\bigcup_{i=1}^k A_i\Big) = \sum_{i=1}^k m_0(A_i);$$

(2) *(monotonicity) if $A_1, \ldots, A_k \in \mathcal{R}_0^n$ are pairwise disjoint and $\bigcup_{i=1}^k A_i \subset A \in \mathcal{R}_0^n$, then*

$$\sum_{i=1}^k m_0(A_i) \leq m_0(A);$$

(3) *(finite subadditivity under covers) if $A, A_1, \ldots, A_k \in \mathcal{R}_0^n$ and $A \subset \bigcup_{i=1}^k A_i$, then*

$$m_0(A) \leq \sum_{i=1}^k m_0(A_i).$$

Proof. (1) Let

$$A_i = \bigcup_{j=1}^{m_i} I_j^i, \quad i = 1, \ldots, k$$

be canonical decompositions. Then

$$\bigcup_{i=1}^k A_i = \bigcup_{i=1}^k \bigcup_{j=1}^{m_i} I_j^i$$

is a canonical decomposition, so

$$m_0\Big(\bigcup_{i=1}^k A_i\Big) = \sum_{i=1}^k \sum_{j=1}^{m_i} |I_j^i| = \sum_{i=1}^k m_0(A_i).$$

(2) Let

$$A_{k+1} = A \backslash \bigcup_{i=1}^k A_i.$$

Then $A_1, \ldots, A_{k+1}$ are pairwise disjoint and

$$m_0(A) = \sum_{i=1}^{k+1} m_0(A_i) \geq \sum_{i=1}^k m_0(A_i).$$

(3) By (1) and (2), m_0 is monotone on $\mathcal{R}_0^n$: if $A \subset B$, then $m_0(A) \leq m_0(B)$. For $A \subset \bigcup_{i=1}^k A_i$, define

$$B_1 = A_1, \quad B_2 = A_2 \backslash A_1, \quad \ldots, \quad B_k = A_k \backslash \Big(\bigcup_{i=1}^{k-1} A_i\Big).$$

Then $B_1, \ldots, B_k$ are pairwise disjoint and $\bigcup_{i=1}^k A_i = \bigcup_{i=1}^k B_i$. Hence

$$m_0(A) \leq m_0\Big(\bigcup_{i=1}^k A_i\Big) = m_0\Big(\bigcup_{i=1}^k B_i\Big) = \sum_{i=1}^k m_0(B_i) \leq \sum_{i=1}^k m_0(A_i).$$

□

Theorem 2.1.2. *m_0 is a measure on $\mathcal{R}_0^n$.*

Proof. By Lemma 2.1.5, it remains to prove countable additivity. Let $\{A_k\}$ be pairwise disjoint in $\mathcal{R}_0^n$ and

$$A = \bigcup_{k=1}^{\infty} A_k \in \mathcal{R}_0^n.$$

By monotonicity,

$$\sum_{i=1}^{k} m_0(A_i) \le m_0(A), \quad \forall k \ge 1.$$

Let $k \to \infty$:

$$\sum_{i=1}^{\infty} m_0(A_i) \le m_0(A).$$

For the reverse inequality, let

$$A = \bigcup_{i=1}^{l} I_i$$

be a canonical decomposition of A, and

$$A_i = \bigcup_{j=1}^{m_i} I_j^i$$

one for A_i. Write

$$I_i = \{(x_1, \dots, x_n) \mid a_r < x_r \le b_r,\ r = 1, \dots, n\}.$$

Given $\varepsilon > 0$, choose inner closed boxes

$$I_i^{\delta_i} = \{(y_1, \dots, y_n) \mid a_r + \delta_i \le y_r \le b_r,\ r = 1, \dots, n\}$$

so that

$$|I_i| \ge |I_i^{\delta_i}| \ge |I_i| - \frac{\varepsilon}{2l}.$$

Let

$$A^\delta = \bigcup_{i=1}^{l} I_i^{\delta_i},$$

so A^δ is bounded closed.

For each I_j^i, enlarge slightly to open boxes I_j^{i,δ_j^i} with

$$|I_j^{i,\delta_j^i}| - \frac{\varepsilon}{m_i 2^{i+1}} \le |I_j^i| \le |I_j^{i,\delta_j^i}|.$$

Then

$$\{I_j^{i,\delta_j^i} \mid i = 1, 2, \dots;\ j = 1, 2, \dots, m_i\}$$

is an open cover of A^δ. By compactness, there exists k_0 such that

$$A^\delta \subset \bigcup_{i=1}^{k_0} \bigcup_{j=1}^{m_i} I_j^{i,\delta_j^i}.$$

Hence

$$\begin{aligned} m_0(A) - \varepsilon \le \sum_{i=1}^{l} |I_i^{\delta_i}| &\le \sum_{i=1}^{k_0} \sum_{j=1}^{m_i} |I_j^{i,\delta_j^i}| \\ &\le \sum_{i=1}^{k_0} \sum_{j=1}^{m_i} \left(|I_j^i| + \frac{\varepsilon}{m_i 2^{i+1}} \right) \\ &\le \sum_{i=1}^{k_0} \left(m_0(A_i) + \frac{\varepsilon}{2^{i+1}} \right) \le \sum_{i=1}^{\infty} m_0(A_i) + \varepsilon. \end{aligned}$$

By arbitrariness of ε,

$$m_0(A) \leq \sum_{i=1}^{\infty} m_0(A_i).$$

$\square$

Corollary 2.1.1. *If* $A, A_i \in \mathcal{R}_0^n$ *and*

$$A \subset \bigcup_{i=1}^{\infty} A_i,$$

then

$$m_0(A) \leq \sum_{i=1}^{\infty} m_0(A_i).$$

Proof. Since

$$A = \bigcup_{i=1}^{\infty} (A_i \cap A),$$

let $B_i = A_i \cap A$ and

$$C_i = B_i \backslash \Big(\bigcup_{j=1}^{i-1} B_j \Big) \in \mathcal{R}_0^n.$$

Then $C_i \cap C_j = \varnothing$ for $i \neq j$, and

$$A = \bigcup_{i=1}^{\infty} C_i.$$

Thus

$$m_0(A) = \sum_{i=1}^{\infty} m_0(C_i) \leq \sum_{i=1}^{\infty} m_0(A_i).$$

$\square$

Exercises 2.1

1. Construct a measure μ on a ring $\mathcal{R}_0^n$ such that

   (1) for every $r \in \mathbb{Q}$, $\mu^*(\{r\}) = 0$;

   (2) for every closed set $F \subset \mathbb{R} \backslash \mathbb{Q}$, $\mu^*(F) = 0$.

2. Let g be increasing and right-continuous on $\mathbb{R}$. For each box in $\mathbb{R}^n$

   $$I = \{(x_1, \ldots, x_n) \mid a_i \leq x_i \leq b_i,\ i = 1, \ldots, n\},$$

   define

   $$\mu(I) = \prod_{i=1}^{n} (g(b_i) - g(a_i)).$$

   Prove that μ is a measure on $\mathcal{R}_0^n$.

3. Let μ, ν be measures on $\mathcal{R}_0^n$ and $\mathcal{R}_0^m$. For $A \in \mathcal{R}_0^n$, $B \in \mathcal{R}_0^m$ with $\mu(A), \nu(B) < \infty$, define

   $$(\mu * \nu)(A \times B) = \mu(A)\nu(B).$$

   Is $\mu * \nu$ a measure on $\mathcal{R}_0^{n+m}$?

4. Let $\mathcal{S}, \mathcal{R}$ be σ-rings on X, Y. Define

$$\mathcal{T} = \Big\{ \bigcup_{i=1}^{\infty} A_i \times B_i \mid A_i \in \mathcal{S},\ B_i \in \mathcal{R} \Big\}.$$

Is $\mathcal{T}$ a σ-ring on $X \times Y$?

2.2 Outer Measure

Definition 2.2.1. *For $A \subset \mathbb{R}^n$, define*

$$m_0^*(A) = \inf \Big\{ \sum_{i=1}^{\infty} m_0(A_i) \mid A_i \in \mathcal{R}_0^n,\ A \subset \bigcup_{i=1}^{\infty} A_i \Big\}.$$

This is the outer measure induced by m_0.

Remark 2.2.1. *In abstract form, for a measure space $(X, \mathcal{R}, \mu)$ with $\mathcal{R}$ a ring and μ a measure on $\mathcal{R}$, define*

$$\mu^*(A) = \inf \Big\{ \sum_{i=1}^{\infty} \mu(A_i) \mid A_i \in \mathcal{R},\ A \subset \bigcup_{i=1}^{\infty} A_i \Big\},$$

called the outer measure induced by μ (see Definition 2.2.2).

Theorem 2.2.1. *The outer measure m_0^* satisfies:*

(1) $m_0^*(\varnothing) = 0$;

(2) *if $A \subset B \subset \mathbb{R}^n$, then $m_0^*(A) \le m_0^*(B)$;*

(3) *if $A \in \mathcal{R}_0^n$, then $m_0^*(A) = m_0(A)$;*

(4)

$$m_0^*\Big(\bigcup_{i=1}^{\infty} A_i \Big) \le \sum_{i=1}^{\infty} m_0^*(A_i).$$

Proof. Proofs of (1) and (2) are omitted.

(3) Set $A_1 = A$, $A_i = \varnothing$ for $i \ge 2$. Then

$$A \subset \bigcup_{i=1}^{\infty} A_i,$$

so

$$m_0^*(A) \le \sum_{i=1}^{\infty} m_0(A_i) = m_0(A).$$

Conversely, for any cover $A \subset \bigcup_{i=1}^{\infty} A_i$ with $A_i \in \mathcal{R}_0^n$,

$$m_0(A) \le \sum_{i=1}^{\infty} m_0(A_i).$$

Taking infimum gives $m_0(A) \le m_0^*(A)$.

(4) Assume without loss of generality that

$$\sum_{i=1}^{\infty} m_0^*(A_i) < \infty.$$

For any $\varepsilon > 0$, by definition of $m_0^*(A_i)$ there exist $A_i^j \in \mathcal{R}_0^n$ such that

$$A_i \subset \bigcup_{j=1}^{\infty} A_i^j, \quad \sum_{j=1}^{\infty} m_0(A_i^j) \le m_0^*(A_i) + \frac{\varepsilon}{2^i}.$$

Hence

$$A := \bigcup_{i=1}^{\infty} A_i \subset \bigcup_{i=1}^{\infty}\bigcup_{j=1}^{\infty} A_i^j,$$

and

$$\begin{aligned} m_0^*(A) &\le \sum_{i=1}^{\infty}\sum_{j=1}^{\infty} m_0(A_i^j) \le \sum_{i=1}^{\infty}\Big(m_0^*(A_i) + \frac{\varepsilon}{2^i}\Big) \\ &= \sum_{i=1}^{\infty} m_0^*(A_i) + \varepsilon. \end{aligned}$$

Let $\varepsilon \downarrow 0$. □

Example 2.2.1. *We have:*

(1) *If $x_0 \in \mathbb{R}^n$, then $m_0^*(\{x_0\}) = 0$;*

(2) $m_0^*(\mathbb{Q}^n) = 0$*;*

(3) *If $\mathcal{C}$ is the middle-third Cantor set, then*

$$m_0^*(\mathcal{C}) \le m_0^*\Big(\bigcup_{i=1}^{2^n} I_i\Big) \le \Big(\frac{2}{3}\Big)^n \to 0 \quad (n \to \infty);$$

(4) *If I is a bounded box, then*

$$m_0^*(\mathring{I}) = m_0^*(\overline{I}).$$

The measure m_0 on $\mathcal{R}_0^n$ is countably additive, but the outer measure m_0^* is not.

Example 2.2.2 (Outer Measure Is Not Additive)**.** *First, if m_0^* were finitely additive, then it would be countably additive. Indeed, if $\{E_n\}_{n=1}^N$ are disjoint and*

$$m_0^*\Big(\bigcup_{n=1}^{N} E_n\Big) = \sum_{n=1}^{N} m_0^*(E_n),$$

then for any disjoint $\{E_n\}_{n=1}^\infty$,

$$m_0^*\Big(\bigcup_{n=1}^{\infty} E_n\Big) \ge m_0^*\Big(\bigcup_{n=1}^{N} E_n\Big) = \sum_{n=1}^{N} m_0^*(E_n) \to \sum_{n=1}^{\infty} m_0^*(E_n),$$

while subadditivity gives the reverse inequality.

Now give a counterexample to countable additivity. For each $x \in (0,1)$, let

$$L_x = \{\xi \in (0,1) \mid \xi - x \in \mathbb{Q}\}.$$

Then $(0,1)$ is a disjoint union of such classes. Choose one representative from each class to form S (using the axiom of choice). Let $S_k = S + r_k$ with $r_k \in (-1,1)\cap\mathbb{Q}$. Then S_k are disjoint and

$$(0,1) \subset \bigcup_{k=1}^{\infty} S_k \subset (-1,2).$$

So

$$1 = m_0^*((0,1)) \le m_0^*\Big(\bigcup_{k=1}^{\infty} S_k\Big) = \sum_{k=1}^{\infty} m_0^*(S_k) \le m_0^*((-1,2)) = 3.$$

Since $m_0^*(S_k) = m_0^*(S)$ and the series converges,

$$m_0^*(S) = 0 \Rightarrow \sum_{k=1}^{\infty} m_0^*(S_k) = 0,$$

which yields $1 \le 0 \le 3$, contradiction.

Theorem 2.2.2. *For any $A \subset \mathbb{R}^n$, define*

$$m_{0,\delta}^*(A) = \inf\Big\{\sum_{i=1}^{\infty} m_0(A_i) \mid A_i \in \mathcal{R}_0^n,\ \operatorname{diam}(A_i) < \delta,\ A \subset \bigcup_{i=1}^{\infty} A_i\Big\}.$$

Then

$$m_{0,\delta}^*(A) = m_0^*(A).$$

Proof. Clearly, $m_0^*(A) \le m_{0,\delta}^*(A)$.

For the opposite inequality, assume $m_0^*(A) < \infty$. Given $\varepsilon > 0$, choose $A_i \in \mathcal{R}_0^n$ such that

$$A \subset \bigcup_{i=1}^{\infty} A_i, \quad \sum_{i=1}^{\infty} m_0(A_i) < m_0^*(A) + \varepsilon.$$

Subdivide each box in each A_i into finitely many boxes of diameter $< \delta$:

$$A_i = \bigcup_{j=1}^{m_i} I_j^i = \bigcup_{j=1}^{m_i} \bigcup_{k=1}^{l_j^i} I_j^{i,k}, \quad \operatorname{diam}(I_j^{i,k}) < \delta.$$

Then

$$m_{0,\delta}^*(A) \le \sum_{i=1}^{\infty} \sum_{j=1}^{m_i} \sum_{k=1}^{l_j^i} |I_j^{i,k}| < m_0^*(A) + \varepsilon.$$

Let $\varepsilon \downarrow 0$. ◻

Definition 2.2.2 (Abstract Outer Measure)**.** *Let X be a set, $\mathcal{R} \subset \mathcal{P}(X)$ a ring, and μ a measure on $\mathcal{R}$. Define*

$$\Sigma(\mathcal{R}) = \Big\{A \subset X \mid \exists A_i \in \mathcal{R},\ A \subset \bigcup_{i=1}^{\infty} A_i\Big\}.$$

For $A \in \Sigma(\mathcal{R})$, define

$$\mu^*(A) = \inf\Big\{\sum_{i=1}^{\infty} \mu(A_i) \mid A \subset \bigcup_{i=1}^{\infty} A_i,\ A_i \in \mathcal{R}\Big\}.$$

This is called the outer measure induced by μ.

Remark 2.2.2. *We call $(X, \mathcal{R}, \mu)$ a measure space. If additionally $\mu(X) = 1$, it is a probability space.*

Theorem 2.2.3. *Let $(X, \mathcal{R}, \mu)$ be a measure space and μ^* the induced outer measure. Then:*

(1) *if $A \subset B$, then $\mu^*(A) \le \mu^*(B)$;*
(2) *if $A \in \mathcal{R}$, then $\mu^*(A) = \mu(A)$;*
(3) *(subadditivity) for $A_i \in \Sigma(\mathcal{R})$,*

$$\mu^*\Big(\bigcup_{i=1}^{\infty} A_i\Big) \le \sum_{i=1}^{\infty} \mu^*(A_i).$$

Proof. We only prove (2). The inequality $\mu^*(A) \le \mu(A)$ is immediate. For the reverse, assume $\mu^*(A) < \infty$. By definition, for any $\varepsilon > 0$ there exist $A_i \in \mathcal{R}$ such that

$$A \subset \bigcup_{i=1}^{\infty} A_i, \quad \sum_{i=1}^{\infty} \mu(A_i) \le \mu^*(A) + \varepsilon.$$

We may assume $A_i \cap A_j = \varnothing$ for $i \neq j$ (by standard disjointization). Then

$$A = A \cap \Big(\bigcup_{i=1}^{\infty} A_i\Big) = \bigcup_{i=1}^{\infty} (A \cap A_i) \in \mathcal{R},$$

and

$$\mu(A) = \sum_{i=1}^{\infty} \mu(A \cap A_i) \le \sum_{i=1}^{\infty} \mu(A_i) \le \mu^*(A) + \varepsilon.$$

Let $\varepsilon \downarrow 0$. □

Remark 2.2.3. *This proof pattern has already appeared several times and is worth noting.*

2.3 Lebesgue Measure

In this section we define Lebesgue measure. We begin with Lebesgue measurable sets.

Definition 2.3.1 (Lebesgue Measurable Set)**.** *Let $\mathcal{R}_0^n, m_0, m_0^*$ be as above. A set $A \subset \mathbb{R}^n$ is called Lebesgue measurable if for every $E \subset \mathbb{R}^n$,*

$$m_0^*(E) = m_0^*(E \cap A) + m_0^*(E \cap A^c). \tag{2.2}$$

Remark 2.3.1. *1. The family of all n-dimensional Lebesgue measurable sets is denoted by $\mathcal{L}^n$. For $A \in \mathcal{L}^n$, set $m(A) = m_0^*(A)$ and call it the Lebesgue measure of A.*

2. Equation (2.2) is called the Carathéodory condition.
3. If $E \in \mathcal{L}^n$, then

$$\begin{aligned} m(E) &= m(E \cap (A \cup A^c)) \\ &= m((E \cap A) \cup (E \cap A^c)) \\ &= m(E \cap A) + m(E \cap A^c). \end{aligned}$$

So the Carathéodory identity always holds inside measurable sets; this is the original motivation for introducing it.

Now show every set in $\mathcal{R}_0^n$ is Lebesgue measurable.

Theorem 2.3.1. *If $A \in \mathcal{R}_0^n$, then $A \in \mathcal{L}^n$.*

Proof. One side is immediate from subadditivity:

$$m_0^*(E) \le m_0^*(E \cap A) + m_0^*(E \cap A^c).$$

For the reverse, assume $m_0^*(E) < \infty$. For any $\varepsilon > 0$, choose $A_i \in \mathcal{R}_0^n$ with

$$E \subset \bigcup_{i=1}^{\infty} A_i, \quad \sum_{i=1}^{\infty} m_0(A_i) \le m_0^*(E) + \varepsilon.$$

Then

$$E \cap A \subset \bigcup_{i=1}^{\infty} (A_i \cap A), \quad E \cap A^c \subset \bigcup_{i=1}^{\infty} (A_i \cap A^c),$$

and each intersection lies in $\mathcal{R}_0^n$. Hence

$$\begin{aligned} m_0^*(E \cap A) + m_0^*(E \cap A^c) &\le \sum_{i=1}^{\infty} m_0(A_i \cap A) + \sum_{i=1}^{\infty} m_0(A_i \cap A^c) \\ &= \sum_{i=1}^{\infty} m_0(A_i) \le m_0^*(E) + \varepsilon. \end{aligned}$$

Let $\varepsilon \downarrow 0$. □

Remark 2.3.2. *In later proofs of measurability, by subadditivity it is enough to prove*

$$m_0^*(E) \ge m_0^*(E \cap A) + m_0^*(E \cap A^c).$$

The following theorem implies that $\mathcal{L}^n$ is a ring.

Theorem 2.3.2. *We have:*

(1) *if $m_0^*(A) = 0$, then $A \in \mathcal{L}^n$;*
(2) *the Cantor middle-third set $\mathcal{C}$ is Lebesgue measurable;*
(3) *if $A \in \mathcal{L}^n$, then $A^c \in \mathcal{L}^n$;*
(4) *if $A, B \in \mathcal{L}^n$, then $A \cup B, A \cap B \in \mathcal{L}^n$, and when $A \cap B = \varnothing$,*

$$m(A \cup B) = m(A) + m(B).$$

Proof. We prove only (4). For any $E \subset \mathbb{R}^n$,

$$\begin{aligned} m_0^*(E) &= m_0^*(E \cap A) + m_0^*(E \cap A^c) \\ &= m_0^*(E \cap A \cap B) + m_0^*(E \cap A \cap B^c) + m_0^*(E \cap A^c). \end{aligned}$$

Since $(A \cap B^c) \cup A^c = (A \cap B)^c$, we get

$$m_0^*(E) \ge m_0^*(E \cap (A \cap B)) + m_0^*(E \cap (A \cap B)^c).$$

The opposite inequality is obvious, so $A \cap B \in \mathcal{L}^n$. Then

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

and by (3), $A \cup B \in \mathcal{L}^n$. If $A \cap B = \varnothing$,

$$\begin{aligned} m_0^*(A \cup B) &= m_0^*((A \cup B) \cap A) + m_0^*((A \cup B) \cap A^c) \\ &= m_0^*(A) + m_0^*(B). \end{aligned}$$

Hence $\mathcal{L}^n$ is a ring. □

In fact, $\mathcal{L}^n$ is a σ-ring.

Theorem 2.3.3. *Let $\mathcal{L}^n, m_0^*, m$ be as above. If $A_i \in \mathcal{L}^n$, $i = 1, 2, \dots,$ then*

$$\bigcup_{i=1}^{\infty} A_i \in \mathcal{L}^n.$$

Moreover, if $A_i \cap A_j = \varnothing$ for $i \neq j$, then

$$m\Big(\bigcup_{i=1}^{\infty} A_i\Big) = \sum_{i=1}^{\infty} m(A_i).$$

Proof. By disjointization, let

$$B_1 = A_1, \quad B_2 = A_2 \backslash A_1, \quad \dots, \quad B_k = A_k \backslash \Big(\bigcup_{i=1}^{k-1} A_i\Big), \ \dots$$

Then each $B_k \in \mathcal{L}^n$ (Theorem 2.3.2),

$$\bigcup_{i=1}^{k} B_i = \bigcup_{i=1}^{k} A_i \quad (k = 1, 2, \dots, \infty),$$

and $B_i \cap B_j = \varnothing$ for $i \neq j$. So it suffices to treat disjoint $\{B_i\}$.

For any $E \subset \mathbb{R}^n$ and any $k \in \mathbb{N}$,

$$\begin{aligned} m_0^*(E) &= m_0^*\Big(E \cap (\bigcup_{i=1}^{k} B_i)\Big) + m_0^*\Big(E \cap (\bigcup_{i=1}^{k} B_i)^c\Big) \\ &= \sum_{i=1}^{k} m_0^*(E \cap B_i) + m_0^*\Big(E \cap (\bigcup_{i=1}^{k} B_i)^c\Big) \\ &\geq \sum_{i=1}^{k} m_0^*(E \cap B_i) + m_0^*\Big(E \cap (\bigcup_{i=1}^{\infty} B_i)^c\Big). \end{aligned}$$

Let $k \to \infty$:

$$\begin{aligned} m_0^*(E) &\geq \sum_{i=1}^{\infty} m_0^*(E \cap B_i) + m_0^*\Big(E \cap (\bigcup_{i=1}^{\infty} B_i)^c\Big) \\ &\geq m_0^*\Big(E \cap (\bigcup_{i=1}^{\infty} B_i)\Big) + m_0^*\Big(E \cap (\bigcup_{i=1}^{\infty} B_i)^c\Big). \end{aligned}$$

Hence $\bigcup_{i=1}^{\infty} B_i \in \mathcal{L}^n$. Taking $E = \bigcup_{i=1}^{\infty} B_i$ yields

$$m\Big(\bigcup_{i=1}^{\infty} B_i\Big) = \sum_{i=1}^{\infty} m(B_i).$$

□

The next two theorems are used frequently.

Theorem 2.3.4 (Measure Limit for Increasing Measurable Sets)**.** *Let $\{E_i\}_{i=1}^{\infty}$ be measurable with*

$$E_1 \subset E_2 \subset \cdots \subset E_n \subset \cdots.$$

Then

$$m\Big(\lim_{k\to\infty} E_k\Big) = \lim_{k\to\infty} m(E_k).$$

Proof. Assume $m(E_k) < \infty$ for all k; otherwise trivial. Define

$$B_1 = E_1, \quad B_2 = E_2 \backslash E_1, \quad \ldots, \quad B_k = E_k \backslash E_{k-1}, \ldots$$

Then $B_k \in \mathcal{L}^n$, pairwise disjoint, and

$$\bigcup_{k=1}^{\infty} B_k = \lim_{k\to\infty} E_k.$$

Hence

$$\begin{aligned} m\Big(\lim_{k\to\infty} E_k\Big) &= m\Big(\bigcup_{k=1}^{\infty} B_k\Big) = \sum_{k=1}^{\infty}(m(E_k) - m(E_{k-1})) \\ &= \lim_{k\to\infty} \sum_{i=1}^{k}(m(E_i) - m(E_{i-1})) = \lim_{k\to\infty} m(E_k), \end{aligned}$$

with $E_0 = \varnothing$. □

With a finiteness condition, we get the decreasing version.

Theorem 2.3.5 (Measure Limit for Decreasing Measurable Sets)**.** *Let* $\{E_i\}_{i=1}^{\infty}$ *be measurable with*

$$E_1 \supset E_2 \supset \cdots \supset E_n \supset \cdots,$$

and

$$m(E_1) < \infty.$$

Then

$$m\Big(\lim_{k\to\infty} E_k\Big) = \lim_{k\to\infty} m(E_k).$$

Proof. Since

$$E_1 \backslash E_k \subset E_1 \backslash E_{k+1}, \quad k = 1, 2, \ldots,$$

$\{E_1 \backslash E_k\}$ is increasing. By Theorem 2.3.4,

$$m\Big(E_1 \backslash \lim_{k\to\infty} E_k\Big) = m\Big(\lim_{k\to\infty} (E_1 \backslash E_k)\Big) = \lim_{k\to\infty} m(E_1 \backslash E_k).$$

Using $m(E_1) < \infty$,

$$m(E_1) - m\Big(\lim_{k\to\infty} E_k\Big) = m(E_1) - \lim_{k\to\infty} m(E_k),$$

so

$$m\Big(\lim_{k\to\infty} E_k\Big) = \lim_{k\to\infty} m(E_k).$$

□

Exercises 2.3

1. P92 1,2,3,4,5,6,7

2.4 Relation Between Lebesgue Measurable Sets and Borel Sets

(Real variable theory) does not require as broad a background as is sometimes imagined. Roughly speaking, three principles suffice: every measurable set is close to a finite union of intervals; every measurable function is close to a continuous function; every convergent function sequence is close to a uniformly convergent one. Most results are rather direct applications of these ideas. When students face a real-variable problem, these are the three principles they most need to master. If one principle appears to solve a genuine problem, one naturally asks whether this "closeness" is sufficient; in practice, the answer is generally yes.

This section concerns Littlewood's first principle.

Theorem 2.4.1. *Every closed set $F \subset \mathbb{R}^n$ is Lebesgue measurable, i.e. $F \in \mathcal{L}^n$.*

Proof. Step 1. If F is bounded and closed, then $F \in \mathcal{L}^n$. For any open box $\overset{\circ}{I}$, it is measurable because it is I minus ∂I, and ∂I has measure zero. For each $x \in F$, define

$$I_k(x) = \{y = (y_1, \ldots, y_n) \mid |y_i - x_i| < 1/k\}.$$

Then $\{I_k(x) \mid x \in F\}$ is an open cover of F, so there exists a finite subcover

$$I_k(x^1), \ldots, I_k(x^{l_k}).$$

Let

$$A_k = \bigcup_{j=1}^{l_k} I_k(x^j).$$

Since $\mathcal{L}^n$ is a σ-ring, $A_k \in \mathcal{L}^n$. We claim

$$F = \bigcap_{k=1}^{\infty} A_k.$$

Clearly $F \subset A_k$ for every k, so $F \subset \bigcap_k A_k$. Conversely, if $y \in \bigcap_k A_k$, then for each k there exists $x^k \in F$ with

$$|y_i - x_i^k| < 1/k, \quad i = 1, \ldots, n,$$

thus

$$\|y - x^k\| < \frac{\sqrt{n}}{k}.$$

Hence $x^k \to y$, and since F is closed, $y \in F$.

Step 2. If F is an arbitrary closed subset of $\mathbb{R}^n$, define

$$F_k = F \cap \overline{B(0,k)}.$$

By Step 1, $F_k \in \mathcal{L}^n$. Also

$$F = \bigcup_{k=1}^{\infty} F_k.$$

By σ-ring closure, $F \in \mathcal{L}^n$. □

By definitions of open and Borel sets, we immediately get:

Corollary 2.4.1. *Every open set and every Borel set in $\mathbb{R}^n$ is Lebesgue measurable.*

Theorem 2.4.2. *If $A \in \mathcal{L}^n$, then for every $\varepsilon > 0$ there exist open G and closed F in $\mathbb{R}^n$ such that*

$$F \subset A \subset G, \quad m(G \backslash A) < \varepsilon, \quad m(A \backslash F) < \varepsilon.$$

Proof. Step 1. If A is bounded, there exists open $G \supset A$ with $m(G\backslash A) < \varepsilon$. By $m(A) = m^*(A)$, for any $\varepsilon > 0$ there exist $A_k \in \mathcal{R}_0^n$ such that

$$A \subset \bigcup_{k=1}^{\infty} A_k, \quad m(A) + \frac{\varepsilon}{2} \geq \sum_{k=1}^{\infty} m(A_k).$$

Write canonical decompositions $A_k = \bigcup_{j=1}^{l_k} I_{k,j}$. Enlarge each box to an open box $I_{k,j}^{\varepsilon}$ with

$$I_{k,j} \subset I_{k,j}^{\varepsilon}, \quad m(I_{k,j}^{\varepsilon}) - m(I_{k,j}) < \frac{\varepsilon}{2^k l_k}.$$

Then

$$\sum_{k=1}^{\infty} \sum_{j=1}^{l_k} m(I_{k,j}^{\varepsilon}) < m(A) + \varepsilon.$$

Let

$$G = \bigcup_{k=1}^{\infty} \bigcup_{j=1}^{l_k} I_{k,j}^{\varepsilon}.$$

Then G is open, $A \subset G$, and

$$m(G\backslash A) = m(G) - m(A) < \varepsilon.$$

Step 2. For general measurable A, there exists open $G \supset A$ with $m(G\backslash A) < \varepsilon$. Define

$$A_k = \{x \in A \mid k < \|x\| \leq k+1\}.$$

By Step 1, for each k there is open $G_k \supset A_k$ such that

$$m(G_k \backslash A_k) < \frac{\varepsilon}{2^k}.$$

Let

$$G = \bigcup_{k=1}^{\infty} G_k.$$

Then G is open, $A \subset G$, and

$$\begin{aligned} m(G\backslash A) &\leq m\Big(\bigcup_{k=1}^{\infty} (G_k \backslash A)\Big) \leq \sum_{k=1}^{\infty} m(G_k \backslash A) \\ &< \varepsilon. \end{aligned}$$

Step 3. Existence of closed $F \subset A$ with $m(A\backslash F) < \varepsilon$. Apply Step 2 to A^c: there is open $G \supset A^c$ with

$$m(G \backslash A^c) < \varepsilon.$$

Let $F = G^c$. Then F is closed, $F \subset A$, and

$$m(A\backslash F) = m(G\backslash A^c) < \varepsilon.$$

□

The next theorem is a direct consequence of Theorem 2.4.2.

Theorem 2.4.3. *If $A \in \mathcal{L}^n$, then*

(1) *there exists a G_δ set G^* with $A \subset G^*$ and $m(G^* \backslash A) = 0$;*

(2) *there exists an F_σ set F^* with $F^* \subset A$ and $m(A \backslash F^*) = 0$.*

Remark 2.4.1. *Since both G_δ and F_σ sets are Borel, Lebesgue measurable sets differ from Borel sets only by null sets from both outside and inside.*

Theorem 2.4.4. *Let A be a Borel set in $\mathbb{R}^n$, and let $B \subset \mathbb{R}^n$ with $m^*(B) = 0$. Then $A \cup B$ and $A \backslash B$ are in $\mathcal{L}^n$, and*

$$m(A \cup B) = m(A), \quad m(A \backslash B) = m(A).$$

Proof. Since $B \in \mathcal{L}^n$, both $A \cup B$ and $A \backslash B$ are measurable. Also,

$$m(A) \leq m(A \cup B) \leq m(A) + m(B) = m(A),$$

$$m(A \backslash B) \leq m(A) \leq m(A \backslash B) + m(B) = m(A \backslash B).$$

So equalities hold. □

Theorem 2.4.5 (Translation Invariance)**.** *If $A \in \mathcal{L}^n$, then $A + \{x_0\} \in \mathcal{L}^n$ and*

$$m(A + \{x_0\}) = m(A).$$

Proof. Because $A \in \mathcal{L}^n$, there exists G_δ set $G^* \supset A$ with

$$m(G^* \backslash A) = 0.$$

Then

$$A + \{x_0\} \subset G^* + \{x_0\},$$

and

$$m^*\Big((G^* + \{x_0\}) \backslash (A + \{x_0\})\Big) = m^*(G^* \backslash A) = 0.$$

So $(G^* + \{x_0\}) \backslash (A + \{x_0\})$ is measurable. Since $G^* + \{x_0\}$ is Borel (hence measurable), $A + \{x_0\}$ is measurable. The measure identity follows from standard translation invariance on boxes and approximation. □

Remark 2.4.2. *Let μ be a measure on $\mathcal{B}^n$, finite on compact sets. If μ is translation invariant, then there exists constant λ such that*

$$\mu(B) = \lambda m(B), \quad \forall B \in \mathcal{B}^n.$$

We omit the proof.

Exercises 2.4

1. P98 1,2,3

2. P102 1,2,3 (replace the conclusion by $m(E) \geq \delta$)

2.5 Completeness of Lebesgue Measure

2.5.1 Existence of Non-Lebesgue-Measurable Sets

Example 2.5.1 (Nonmeasurable Set). *Let $I = [0,1]^n$ be the unit cube in $\mathbb{R}^n$, and let $\mathbb{Q}^n$ be the set of rational points. We aim to construct sets $\{A_k\}$ such that:*

(1) $A_k \cap A_j = \varnothing$ *for* $k \neq j$*;*

(2) $\{A_k\}$ *is uniformly bounded;*

(3) $I \subset \bigcup_{k=1}^{\infty} A_k$*;*

(4) $m^*(A_k) = m^*(A_j)$*;*

(5) $A_k \in \mathcal{L}^n \Leftrightarrow A_j \in \mathcal{L}^n$*.*

These imply A_1 is not Lebesgue measurable.

Construction: for each $x \in (0,1)$, define

$$A_x = \{\xi \in (0,1) \mid \xi - x \in \mathbb{Q}\}.$$

Then $(0,1)$ is a disjoint union of the classes A_x. Choose one representative from each class to form A (axiom of choice). Let

$$A_k = A + r_k, \quad r_k \in (-1,1) \cap \mathbb{Q}.$$

Then $\{A_k\}$ has the required properties.

Remark 2.5.1. *1. For any translation-invariant measure on $\mathbb{R}^n$ assigning positive measure to cubes, the set A_1 above is nonmeasurable.*

2. This example also shows m^* is neither countably additive nor finitely additive.

2.5.2 Extension of Measures

For any $A \subset \mathbb{R}^n$, define

$$m^{**}(A) = \inf\Big\{\sum_{i=1}^{\infty} m(A_i) \mid A \subset \bigcup_{i=1}^{\infty} A_i,\ A_i \in \mathcal{L}^n\Big\}.$$

If for all $E \subset \mathbb{R}^n$,

$$m^{**}(E) = m^{**}(E \cap A) + m^{**}(E \cap A^c),$$

then A is called second-order Lebesgue measurable.

Clearly every Lebesgue measurable set is second-order measurable. It might appear this extends the measurable class, but the next theorem shows otherwise.

Theorem 2.5.1. *For all $A \subset \mathbb{R}^n$,*

$$m^{**}(A) = m^*(A).$$

Proof. Since $\mathcal{R}_0^n \subset \mathcal{L}^n$,

$$m^{**}(A) \leq m^*(A).$$

For the reverse inequality, assume $m^{**}(A) < \infty$. By definition, for any $\varepsilon > 0$ there exist $A_k \in \mathcal{L}^n$ such that

$$A \subset \bigcup_{k=1}^{\infty} A_k, \quad \sum_{k=1}^{\infty} m(A_k) \leq m^{**}(A) + \varepsilon.$$

For each k, by definition of $m^*(A_k)$ there exist $B_j^k \in \mathcal{R}_0^n$ with

$$A_k \subset \bigcup_{j=1}^{\infty} B_j^k, \quad \sum_{j=1}^{\infty} m_0(B_j^k) \leq m^*(A_k) + \frac{\varepsilon}{2^k}.$$

Hence

$$A \subset \bigcup_{k=1}^{\infty} \bigcup_{j=1}^{\infty} B_j^k,$$

so

$$\begin{aligned} m^*(A) &\leq \sum_{k=1}^{\infty} \sum_{j=1}^{\infty} m_0(B_j^k) \leq \sum_{k=1}^{\infty} \left(m^*(A_k) + \frac{\varepsilon}{2^k} \right) \\ &\leq m^{**}(A) + 2\varepsilon. \end{aligned}$$

Let $\varepsilon \downarrow 0$ to get $m^*(A) \leq m^{**}(A)$. □

Theorem 2.5.2 (Uniqueness of Measure Extension)**.** *Let $\mathcal{R}$ be a σ-ring of subsets of $\mathbb{R}^n$ with $\mathcal{R}_0^n \subset \mathcal{R}$. Let μ be a measure on $\mathcal{R}$ satisfying*

$$\mu|_{\mathcal{R}_0^n} = m|_{\mathcal{R}_0^n},$$

i.e. $\mu(A) = m(A)$ for all $A \in \mathcal{R}_0^n$. Then

$$\mu|_{\mathcal{B}^n} = m|_{\mathcal{B}^n},$$

where $\mathcal{B}^n$ is the smallest σ-ring containing $\mathcal{R}_0^n$.

Proof. Step 1. For every open set $G \subset \mathbb{R}^n$, $\mu(G) = m(G)$.

For each $x \in G$, there is a left-open right-closed box $I(y_x, r_x)$ with rational center y_x and rational side length $r_x > 0$ such that

$$I(y_x, r_x) \subset G.$$

Hence

$$G = \bigcup_{x \in G} I(y_x, r_x).$$

This family is at most countable, so

$$G = \bigcup_{k=1}^{\infty} I(y_{x_k}, r_{x_k}).$$

Disjointize into boxes $\{B_k\}$. By countable additivity of μ and m,

$$\mu(G) = \sum_{k=1}^{\infty} \mu(B_k) = \sum_{k=1}^{\infty} m(B_k) = m(G).$$

Step 2. For every bounded closed set $F \subset \mathbb{R}^n$, $\mu(F) = m(F)$.

Take open $G \supset F$. Then

$$G = (G \backslash F) \cup F,$$

so

$$\mu(G) = m(G) = m(G \backslash F) + m(F) = \mu(G \backslash F) + m(F),$$

hence

$$\mu(F) = \mu(G) - \mu(G \backslash F) = m(F).$$

Step 3. For every closed set $F \subset \mathbb{R}^n$, $\mu(F) = m(F)$.

Let

$$F_k = F \cap \overline{B(0,k)}.$$

By Step 2, $\mu(F_k) = m(F_k)$. Since $F_k \uparrow F$, by continuity from below,

$$\mu(F) = \lim_{k\to\infty} \mu(F_k) = \lim_{k\to\infty} m(F_k) = m(F).$$

Step 4. For every Borel set $E \in \mathcal{B}^n$, $\mu(E) = m(E)$.

Given $\varepsilon > 0$, choose open G_ε and closed F_ε with

$$F_\varepsilon \subset E \subset G_\varepsilon,$$

$$m(E \backslash F_\varepsilon) < \varepsilon, \quad m(G_\varepsilon \backslash E) < \varepsilon.$$

Then

$$m(G_\varepsilon) - m(F_\varepsilon) < 2\varepsilon.$$

Also

$$m(F_\varepsilon) \le m(E) \le m(G_\varepsilon),$$

$$\mu(F_\varepsilon) \le \mu(E) \le \mu(G_\varepsilon),$$

and by previous steps $\mu(F_\varepsilon) = m(F_\varepsilon)$, $\mu(G_\varepsilon) = m(G_\varepsilon)$. Let $\varepsilon \downarrow 0$ to conclude $\mu(E) = m(E)$. □

Remark 2.5.2. *For measures on $\mathbb{R}^1$ induced by monotone right-continuous functions, all sets may be measurable; this does not contradict the theorem above.*

Finally, we give an example of a Lebesgue measurable set that is not Borel.

Example 2.5.2 (Lebesgue Measurable but Non-Borel Set)**.** *Let $\mathcal{C}$ be the middle-third Cantor set, and let $\{(a_k, b_k)\}$ be component intervals of its complement. Let $\varphi : [0,1] \to [0,1]$ be the Cantor function constructed as in Proposition 1.3.5, so*

$$\varphi|_{(a_k,b_k)} \equiv c_k.$$

Define

$$f(x) = \frac{1}{2}(x + \varphi(x)), \quad x \in [0,1].$$

Then $f(0) = 0$, $f(1) = 1$, and f is strictly increasing; hence bijective. For every k,

$$m(f((a_k, b_k))) = \frac{1}{2}(b_k - a_k).$$

So

$$m\Big(f\Big(\bigcup_{k=1}^{\infty}(a_k,b_k)\Big)\Big)=\frac{1}{2}.$$

Since

$$m\Big(f(\mathcal{C})\cup f\Big(\bigcup_{k=1}^{\infty}(a_k,b_k)\Big)\Big)=1,$$

we obtain

$$m(f(\mathcal{C}))=\frac{1}{2}.$$

By the construction in Example 2.5.1, choose a Lebesgue nonmeasurable set $A\subset f(\mathcal{C})$. Let

$$B=f^{-1}(A)\subset\mathcal{C}.$$

Then

$$0\le m^*(B)\le m(\mathcal{C})=0,$$

so $B\in\mathcal{L}^1$. But B is not Borel; otherwise $f(B)=A$ would be Borel (see exercise), contradicting nonmeasurability of A.

Exercises 2.5

1. If $f(x)$ is strictly increasing, show that both $f^{-1}:\mathcal{B}^n\to\mathcal{B}^n$ and $f:\mathcal{B}^n\to\mathcal{B}^n$ are bijections.

Chapter 3

Lebesgue Measurable Functions

3.1 Definition and Basic Properties

First, for convenience in later arguments, we adopt the convention

$$0 \cdot \infty = 0.$$

Definition 3.1.1 (Almost Everywhere)**.** *Let $E \subset \mathbb{R}^n$ be Lebesgue measurable, and let $P(x)$ be a statement about $x \in E$. If there exists a null set $E_0 \subset E$ such that $P(x)$ holds for all $x \in E \backslash E_0$, then we say $P(x)$ holds almost everywhere on E, written as*

$$P(x) \text{ holds for a.e. } x \in E.$$

Example 3.1.1. *By Definition 3.1.1, we can define the notions "finite almost everywhere" and "bounded almost everywhere." These are different concepts.*

Definition 3.1.2 (Lebesgue Measurable Function)**.** *Let $E \subset \mathbb{R}^n$ be Lebesgue measurable, and let*

$$f : E \to \mathbb{R} \cup \{\pm\infty\}.$$

We call f a Lebesgue measurable function if for every $c \in \mathbb{R} \cup \{\pm\infty\}$,

$$E(x|f(x) > c) = \{x \in E | f(x) > c\}$$

is Lebesgue measurable.[1]

Remark 3.1.1. *1. Sometimes one does not assume E is measurable, but then one assumes f is finite-valued*[2] *and*

$$E = \bigcup_{n=1}^{\infty} E(x|f(x) > -n)$$

is measurable.

2. For extended real-valued f, one may equivalently require $E(x|f(x) \geq c)$ to be measurable for all $c \in \mathbb{R} \cup \{\pm\infty\}$. In particular, taking $c = -\infty$ implies E is measurable.

[1]Unless otherwise stated, we abbreviate $\{x \in E|f(x) > c\}$ by $E(x|f(x) > c)$.

[2]Finite-valued means the range of f does not contain $\pm\infty$.

3. (Borel measurable function) Let $E \in \mathcal{B}^n$ and $f : E \to \mathbb{R} \cup \{\pm\infty\}$. If for every $c \in \mathbb{R} \cup \{\pm\infty\}$,

$$E(x|f(x) > c) \in \mathcal{B}^n,$$

then f is called Borel measurable. Every Borel measurable function is Lebesgue measurable.

Definition 3.1.3 (Measurable Function on an Abstract Measure Space)**.** *Let $(X, \mathcal{R}, \mu)$ be a measure space, $E \in \mathcal{R}$, and $f : E \to \mathbb{R} \cup \{\pm\infty\}$. If for every $c \in \mathbb{R} \cup \{\pm\infty\}$,*

$$E(x|f(x) > c) \in \mathcal{R},$$

then f is called μ-measurable.

Theorem 3.1.1. *Let $E \in \mathcal{L}^n$, $f : E \to \mathbb{R} \cup \{\pm\infty\}$, and let $D \subset \mathbb{R}$ be a dense subset. If for every $r \in D$,*

$$E(x|f(x) > r) \in \mathcal{L}^n,$$

then f is Lebesgue measurable on E.

Proof. Left as an exercise. □

Theorem 3.1.2. *Let $E \in \mathcal{L}^n$ and $f : E \to \mathbb{R} \cup \{\pm\infty\}$. The following are equivalent:*

(1) *f is Lebesgue measurable on E;*

(2) *for every $c \in \mathbb{R} \cup \{\pm\infty\}$,*

$$E(x|f(x) \le c) \in \mathcal{L}^n;$$

(3) *for every $c \in \mathbb{R} \cup \{\pm\infty\}$,*

$$E(x|f(x) < c) \in \mathcal{L}^n;$$

(4) *for every $c \in \mathbb{R} \cup \{\pm\infty\}$,*

$$E(x|f(x) \ge c) \in \mathcal{L}^n;$$

(5) *for every $c, d \in \mathbb{R} \cup \{\pm\infty\}$,*

$$E(x|c < f(x) \le d) \in \mathcal{L}^n.$$

Proof. Left as an exercise. □

Corollary 3.1.1. *Let $E \in \mathcal{L}^n$ and $f : E \to \mathbb{R} \cup \{\pm\infty\}$. If f is Lebesgue measurable, then for every $c \in \mathbb{R} \cup \{\pm\infty\}$,*

$$E(x|f(x) = c) \in \mathcal{L}^n.$$

Proof. This follows directly from Theorem 3.1.2. □

Example 3.1.2. *Let $E \in \mathcal{L}^n$. The indicator function χ_E is Lebesgue measurable. In particular, the Dirichlet function is both Lebesgue measurable and Borel measurable.*

Theorem 3.1.3. *Let $E \in \mathcal{L}^n$, and let f, g be Lebesgue measurable on E. Then:*

(1) *for every $\alpha \in \mathbb{R}$, αf is Lebesgue measurable;*

(2) *$f + g$ is Lebesgue measurable;*

(3) *fg is Lebesgue measurable;*

(4) *if $g(x) \neq 0$ for $x \in E$, then f/g is Lebesgue measurable;*

(5) *$\max\{f, g\}$ and $\min\{f, g\}$ are Lebesgue measurable;*

(6) *$\sup\{f_k\}$, $\inf\{f_k\}$, $\overline{\lim\limits_{k\to\infty}} f_k$, and $\underline{\lim\limits_{k\to\infty}} f_k$ are Lebesgue measurable.*

Proof. For (4), it is enough to show $1/g$ is measurable.

For (5),

$$E(x|\max\{f(x),g(x)\}\geq c)=E(x|f(x)\geq c)\cup E(x|g(x)\geq c).$$

For (6),

$$\overline{\lim_{k\to\infty}} f_k(x)=\inf_m\Big\{\sup_{k\geq m} f_k(x)\Big\}.$$

The remaining parts are standard. □

Example 3.1.3. *Let $E_i\in\mathcal{L}^n$, $i=1,2,\ldots,k$. Then*

$$\chi(x)=\sum_{i=1}^k \alpha_i\chi_{E_i}(x) \tag{3.1}$$

is Lebesgue measurable.

Remark 3.1.2. *A function of the form* (3.1) *is called a simple function (or step function).*

Example 3.1.4. *If f is Lebesgue measurable on E, and $A\subset E$ with $A\in\mathcal{L}^n$, then the restriction $f|_A$ is Lebesgue measurable.*[3]

Example 3.1.5. *Let $E\in\mathcal{L}^n$. If f is continuous on E, then f is Lebesgue measurable.*

Exercises 3.1

1. P127 1,2,3,4,5,6,7
2. P129 8,9

3.2 Structure of Lebesgue Measurable Functions

Theorem 3.2.1. *Let $E\subset\mathbb{R}^n$ be Lebesgue measurable.*

(1) *If f is a nonnegative Lebesgue measurable function on E and finite a.e., then there exists an increasing*[4] *sequence of simple functions $\{\varphi_k\}_{k=1}^\infty$ such that*

$$\varphi_k\to f\quad \text{a.e. on } E.$$

If in addition f is bounded, then

$$\varphi_k\rightrightarrows f.$$

(2) *If f is Lebesgue measurable on E and finite a.e., then there exists a sequence of simple functions $\{\varphi_k\}_{k=1}^\infty$ such that*

$$|\varphi_k|\leq|f|\quad \text{a.e. on } E,$$

and

$$\varphi_k\to f\quad \text{a.e. on } E.$$

If f is bounded, then

$$\varphi_k\rightrightarrows f.$$

[3] That is, $f|_A$ is the restriction of f to A.

[4] Namely, $\varphi_k\leq\varphi_{k+1}$.

Proof. Step 1. For each natural number k, define

$$E_{k,i} = E\Big(x\Big|\frac{i-1}{2^k} \le f(x) \le \frac{i}{2^k}\Big), \quad i = 1, 2, \dots, k2^k.$$

Set

$$\varphi_k(x) = \begin{cases} \frac{i-1}{2^k}, & x \in E_{k,i}, \\ k, & x \in E \setminus \bigcup_{i=1}^{k2^k} E_{k,i}. \end{cases}$$

Then $\varphi_k \le \varphi_{k+1}$, and if $x \in E_{k,i}$,

$$0 \le f(x) - \varphi_k(x) \le \frac{1}{2^k}.$$

Since f is finite a.e. on E, the set $E_\infty = \{x \in E | f(x) = +\infty\}$ has measure zero. For $x \in E\setminus E_\infty$ and any $\varepsilon > 0$, choose k_0 so that $f(x) < k_0$ and $2^{-k_0} < \varepsilon$. Then $x \in E_{k_0,i_0}$ for some i_0, hence

$$0 \le f(x) - \varphi_{k_0}(x) \le \frac{1}{2^{k_0}} < \varepsilon.$$

Thus $\varphi_k \to f$ a.e. If $f(x) \le M$ on E, then for $k > M$ the same estimate holds for all $x \in E$, so $\varphi_k \rightrightarrows f$.

Step 2. Let

$$f^+ = \max\{f, 0\}, \quad f^- = \max\{-f, 0\}, \quad f = f^+ - f^-.$$

Apply Step 1 to f^+ and f^- to obtain simple functions $\varphi_k \nearrow f^+$ and $\psi_k \nearrow f^-$ a.e., with $\varphi_k \le f^+$ and $\psi_k \le f^-$. Then

$$\varphi_k - \psi_k \to f^+ - f^- = f,$$

and

$$|\varphi_k - \psi_k| \le \varphi_k + \psi_k \le f^+ + f^- = |f|.$$

So part (2) follows. □

Following this theorem, a standard strategy for statements about measurable functions is:

1. prove it for indicator functions;
2. prove it for simple functions;
3. prove it for nonnegative measurable functions;
4. decompose into positive and negative parts, then prove it for general measurable functions.

Lemma 3.2.1. *Let $F_1, F_2, \dots, F_k \subset \mathbb{R}^n$ be pairwise disjoint closed sets, and define*

$$\varphi(x) = \sum_{i=1}^{k} \alpha_i \chi_{F_i}(x).$$

Then $\varphi|_{\cup_{i=1}^k F_i}$ is continuous.

Proof. Fix $x_0 \in \bigcup_{i=1}^k F_i$. Then $x_0 \in F_{i_0}$ for some i_0. Since the F_i are pairwise disjoint and closed, for each $i \neq i_0$,

$$\delta_i = \operatorname{dist}(x_0, F_i) > 0.$$

Set $\delta_0 = \min_{i \neq i_0} \delta_i$. If

$$x \in \Big(\bigcup_{i=1}^k F_i\Big) \cap B(x_0, \delta_0),$$

then $x \in F_{i_0}$, hence

$$|\varphi(x) - \varphi(x_0)| = 0 < \varepsilon \quad \forall \varepsilon > 0.$$

So $\varphi|_{\cup_{i=1}^k F_i}$ is continuous. □

The next theorem is Lusin's theorem, corresponding to Littlewood's second principle: every measurable function is close to a continuous one.

Theorem 3.2.2 (Lusin's Theorem)**.** *Let $E \subset \mathbb{R}^n$ be Lebesgue measurable, and let f be Lebesgue measurable on E and finite a.e. Then for every $\delta > 0$, there exists a closed set $F \subset E$ such that $f|_F$ is continuous and*

$$m(E \backslash F) < \delta.$$

Proof. Step 1: Assume $m(E) < \infty$. For each $k \in \mathbb{N}$, define

$$E_{k,i} = E\Big(x \Big| \frac{i-1}{k} \leq f(x) < \frac{i}{k}\Big), \quad i = 0, \pm 1, \pm 2, \dots$$

Then $E_{k,i} \cap E_{k,j} = \varnothing$ for $i \neq j$, and

$$E = \Big(\bigcup_{i=-\infty}^{\infty} E_{k,i}\Big) \cup E_{-\infty} \cup E_{+\infty},$$

where $E_{\pm\infty} = E(x | f(x) = \pm\infty)$ and $m(E_{\pm\infty}) = 0$. Since

$$m(E) = \sum_{i=-\infty}^{\infty} m(E_{k,i}) + m(E_{\pm\infty}),$$

for given $\delta > 0$ we can choose i_k such that

$$\Big(\sum_{i=-\infty}^{-i_k-1} + \sum_{i=i_k+1}^{\infty}\Big) m(E_{k,i}) < \frac{\delta}{2^{k+1}}.$$

Define

$$\varphi_k(x) = \sum_{i=-i_k}^{i_k} \frac{i-1}{k} \chi_{E_{k,i}}(x).$$

For each i, choose closed $F_{k,i} \subset E_{k,i}$ with

$$m(E_{k,i} \backslash F_{k,i}) < \frac{\delta}{2^{k+2} i_k}.$$

Let

$$F_k = \bigcup_{i=-i_k}^{i_k} F_{k,i}, \quad \psi_k = \varphi_k|_{F_k}.$$

By Lemma 3.2.1, ψ_k is continuous on F_k, and for $x \in F_k$,

$$|f(x) - \psi_k(x)| = |f(x) - \varphi_k(x)| \le \frac{1}{k}.$$

Also,

$$m(E \backslash F_k) \le \frac{\delta}{2^k}.$$

Set

$$F = \bigcap_{k=1}^{\infty} F_k.$$

Then F is closed and

$$m(E \backslash F) = m\Big(\bigcup_{k=1}^{\infty} (E \backslash F_k)\Big) \le \sum_{k=1}^{\infty} m(E \backslash F_k) \le \delta.$$

Moreover, each $\psi_k|_F$ is continuous and

$$|\psi_k(x) - f(x)| \le \frac{1}{k} \quad \forall x \in F,$$

so $\psi_k \rightrightarrows f$ on F. Hence $f|_F$ is continuous.

Step 2: If $m(E) = \infty$, apply Step 1 on truncated pieces and combine them to obtain the same conclusion. □

Corollary 3.2.1. *Let $E \subset \mathbb{R}^n$ be Lebesgue measurable, and let f be Lebesgue measurable on E and finite a.e. For every $\delta > 0$, there exists a continuous function φ on $\mathbb{R}^n$ such that*

$$m\big(E(x|f(x) \ne \varphi(x))\big) < \delta.$$

Proof. This follows immediately from Theorem 3.2.2 and Theorem 1.3.10. □

Corollary 3.2.2. *Let $E \subset \mathbb{R}^n$ be Lebesgue measurable, and let f be Lebesgue measurable on E and finite a.e. Then there exists a sequence of continuous functions $\{\varphi_k\}$ such that*

$$\varphi_k \to f \quad a.e. \ on \ E.$$

Proof. Apply Theorem 3.2.2 recursively. Choose closed sets $F_k \subset E$ such that

$$m\Big(E \backslash \bigcup_{i=1}^{k} F_i\Big) < \frac{1}{k},$$

and f is continuous on $\bigcup_{i=1}^{k} F_i$. Let φ_k be a continuous extension to $\mathbb{R}^n$ of $f|_{\cup_{i=1}^{k} F_i}$. Then $\varphi_k = f$ on $\cup_{i=1}^{k} F_i$, and therefore

$$m\Big(E \backslash \bigcup_{k=1}^{\infty} F_k\Big) = 0.$$

Hence $\varphi_k \to f$ a.e. on E. □

Theorem 3.2.3. *Let $f : [0,1] \to [0,1]$ be strictly increasing and continuous, with $f([0,1]) = [0,1]$. Then*

$$\mathcal{B} = f^{-1}(\mathcal{B}).$$

Proof. Let

$$\mathcal{A} = \{f^{-1}(B) \mid B \in \mathcal{B},\ B \subset [0,1]\}.$$

We prove in two steps.

Step 1: Prove $\mathcal{B} \subset \mathcal{A}$.

(1) If $A_1, A_2 \in \mathcal{A}$, then $A_1 = f^{-1}(B_1)$ and $A_2 = f^{-1}(B_2)$ for some $B_1, B_2 \in \mathcal{B}$. Thus

$$A_1 \backslash A_2 = f^{-1}(B_1 \backslash B_2) \in \mathcal{A}.$$

(2) If $A_i \in \mathcal{A}$ with $A_i = f^{-1}(B_i)$, then

$$\bigcup_{i=1}^{\infty} A_i = f^{-1}\Big(\bigcup_{i=1}^{\infty} B_i\Big) \in \mathcal{A}.$$

So $\mathcal{A}$ is a σ-ring.

(3) Since f is continuous, preimages of open or closed subsets of $[0,1]$ are open or closed in $[0,1]$, hence belong to $\mathcal{A}$. Therefore $\mathcal{B} \subset \mathcal{A}$.

Step 2: Let $g = f^{-1}$. Then g is also strictly increasing and continuous. By Step 1, $\mathcal{B} \subset g^{-1}(\mathcal{B}) = f(\mathcal{B})$, so

$$f^{-1}(\mathcal{B}) \subset f^{-1}(f(\mathcal{B})) = \mathcal{B}.$$

Combining with Step 1 gives $\mathcal{B} = f^{-1}(\mathcal{B})$. □

Remark 3.2.1. *This proof method is typical and is often called the "good-set principle": first collect all sets with the desired property, then show that this family forms an appropriate class.*

3.3 Almost Everywhere Convergence and Convergence in Measure

In this section we introduce two important notions of convergence and their relationship.

3.3.1 Almost Everywhere Convergence

Definition 3.3.1 (Almost Everywhere Convergence)**.** *Let $E \subset \mathbb{R}^n$ be Lebesgue measurable, and*

$$f, f_k : E \to \mathbb{R} \cup \{\pm\infty\}, \quad k = 1, 2, \ldots,$$

be Lebesgue measurable. If

$$f_k(x) \to f(x)$$

for a.e. $x \in E$, then we say $\{f_k\}$ converges to f almost everywhere on E, written

$$f_k \to f \quad a.e.\ on\ E,$$

or

$$f_k \overset{a.e.}{\to} f \quad x \in E.$$

In Theorem 3.4.1 we will see that measurable functions generalize continuous functions. In real analysis, even pointwise limits of continuous functions need not be continuous. In measure theory, however, a.e. limits preserve Lebesgue measurability.

Theorem 3.3.1. *Let $E \subset \mathbb{R}^n$ be Lebesgue measurable, and let $f_k : E \to \mathbb{R} \cup \{\pm\infty\}$ be Lebesgue measurable for $k = 1, 2, \ldots$. If*

$$f_k \overset{a.e.}{\to} f \quad x \in E,$$

then f is measurable on E.

Proof. Left as an exercise. □

Next we present the classical Egorov theorem. We first need a lemma.

Lemma 3.3.1. *Let $E \subset \mathbb{R}^n$ be Lebesgue measurable with*

$$m(E) < \infty.$$

Let f, f_k be measurable extended real-valued functions on E, finite a.e., and suppose

$$f_k \overset{a.e.}{\to} f \quad x \in E.$$

For every $\varepsilon > 0$, define

$$E_k(\varepsilon) = E(x||f_k(x) - f(x)| > \varepsilon).$$

Then

$$\lim_{j\to\infty} m\Big(\bigcup_{k=j}^{\infty} E_k(\varepsilon)\Big) = 0.$$

Proof. Let

$$E(\varepsilon) = \overline{\lim_{k\to\infty}} E_k(\varepsilon) = \bigcap_{j=1}^{\infty}\bigcup_{k=j}^{\infty} E_k(\varepsilon).$$

The set $E(\varepsilon)$ contains points where the deviation $|f_k - f| > \varepsilon$ happens infinitely often. Since $f_k \to f$ a.e., $m(E(\varepsilon)) = 0$. Then the conclusion follows from continuity of measure from above (using $m(E) < \infty$). □

The Egorov theorem below corresponds to Littlewood's third principle: a.e. convergence is nearly uniform convergence on large subsets.

Theorem 3.3.2 (Egorov)**.** *Let $E \subset \mathbb{R}^n$ be Lebesgue measurable with*

$$m(E) < \infty.$$

Let f, f_k be measurable on E, finite a.e., and assume

$$f_k \overset{a.e.}{\to} f \quad x \in E.$$

Then for every $\delta > 0$, there exists a measurable subset $E_\delta \subset E$ such that $m(E_\delta) < \delta$ and

$$f_k \rightrightarrows f \quad \text{on } E \backslash E_\delta.$$

Proof. Let

$$E_0 = E(x|\exists k: \ |f_k(x)| = \infty \text{ or } |f(x)| = \infty),$$

so $m(E_0) = 0$.

In Lemma 3.3.1, take $\varepsilon = 1/m$ for $m = 1, 2, \ldots$. Then for each m, there exists j_m such that

$$m\Big(\bigcup_{k=j_m}^{\infty} E_k\Big(\frac{1}{m}\Big)\Big) < \frac{\delta}{2^m}.$$

Define

$$E_\delta = E_0 \cup \Big(\bigcup_{m=1}^{\infty}\bigcup_{k=j_m}^{\infty} E_k\Big(\frac{1}{m}\Big)\Big).$$

Then

$$m(E_\delta) \le \sum_{m=1}^{\infty} \frac{\delta}{2^m} = \delta.$$

Now fix $\varepsilon > 0$, choose m_0 with $1/m_0 < \varepsilon$. If $x \in E \backslash E_\delta$, then for all $k \ge j_{m_0}$,

$$|f_k(x) - f(x)| \le \frac{1}{m_0} < \varepsilon.$$

Hence $f_k \rightrightarrows f$ on $E \backslash E_\delta$. □

Remark 3.3.1. *The finite measure assumption in Egorov's theorem cannot be removed.*

Example 3.3.1. *Let $f_k(x) = \chi_{(0,k)}(x)$. Then*

$$f_k \overset{a.e.}{\to} \chi_{(0,\infty)} \quad x \in \mathbb{R}.$$

But for any set E_δ with $m(E_\delta) < \infty$, uniform convergence on $\mathbb{R}\backslash E_\delta$ fails.

3.3.2 Convergence in Measure

Definition 3.3.2 (Convergence in Measure)**.** *Let $E \subset \mathbb{R}^n$ be Lebesgue measurable with $m(E) < \infty$, and let f, f_k be measurable and finite a.e. on E. If for every $\varepsilon > 0$,*

$$\lim_{k\to\infty} m(E(x||f_k(x) - f(x)| > \varepsilon)) = 0,$$

then $\{f_k\}$ is said to converge to f in measure on E, written

$$f_k \overset{m}{\to} f \quad x \in E,$$

or

$$f_k \Rightarrow f \quad x \in E.$$

Convergence in measure is fundamental in probability theory. It is weaker than uniform convergence and typically weaker than almost everywhere convergence.

Definition 3.3.3 (True in Measure)**.** *Let $P_n(x)$ be a sequence of statements about $x \in E$. If*

$$\lim_{n\to\infty} m(E(x|P_n(x) \text{ is false})) = 0,$$

then we say this sequence of statements is true in measure.

The next theorem gives a basic implication between the two convergences.

Theorem 3.3.3. *Let $E \subset \mathbb{R}^n$ be Lebesgue measurable with $m(E) < \infty$. Suppose f, f_k are measurable and finite a.e. on E, and*

$$f_k \overset{a.e.}{\to} f \quad x \in E.$$

Then

$$f_k \overset{m}{\to} f \quad x \in E.$$

Proof. Fix $\varepsilon > 0$ and $\delta > 0$. By Theorem 3.3.2, there exists $E_\delta \subset E$ with $m(E_\delta) < \delta$ such that $f_k \rightrightarrows f$ on $E \backslash E_\delta$. Hence there is k_0 such that for $k > k_0$,

$$|f_k(x) - f(x)| < \varepsilon \quad \forall x \in E \backslash E_\delta.$$

Therefore

$$E(x||f_k(x) - f(x)| \geq \varepsilon) \subset E_\delta,$$

so

$$m(E(x||f_k(x) - f(x)| \geq \varepsilon)) < \delta \quad (k > k_0).$$

This is exactly $f_k \overset{m}{\to} f$. □

Remark 3.3.2. *The condition $m(E) < \infty$ in Theorem 3.3.3 is essential. For example,*

$$f_k(x) = \begin{cases} 1, & |x| \geq k, \\ 0, & |x| < k, \end{cases}$$

satisfies $f_k \overset{a.e.}{\to} 0$ on $\mathbb{R}$, but $f_k \not\Rightarrow 0$ on $\mathbb{R}$.

Theorem 3.3.3 shows that a.e. convergence implies convergence in measure on finite-measure sets. The converse generally fails.

Example 3.3.2. *Construct a sequence on $[0, 1]$ by blocks:*

$$f_{1,1}(x) = 1, \quad x \in [0, 1],$$

$$f_{2,1}(x) = \begin{cases} 1, & x \in [0, \frac{1}{2}), \\ 0, & x \in [\frac{1}{2}, 1], \end{cases} \qquad f_{2,2}(x) = \begin{cases} 0, & x \in [0, \frac{1}{2}), \\ 1, & x \in [\frac{1}{2}, 1], \end{cases} ,$$

and in general,

$$f_{k,i}(x) = \begin{cases} 1, & x \in [\frac{i}{2^{k-1}}, \frac{i+1}{2^{k-1}}), \\ 0, & \textit{otherwise}, \end{cases} \qquad k = 1, 2, \ldots, \ i = 1, 2, \ldots, 2^{k-1}.$$

Then this sequence converges to 0 in measure but not almost everywhere.

Hence convergence in measure is strictly weaker than almost everywhere convergence.

Theorem 3.3.4 (Riesz)**.** *Let $E \subset \mathbb{R}^n$ be Lebesgue measurable. Let f, f_k be measurable and finite a.e. on E. If*

$$f_k \overset{m}{\to} f \quad x \in E,$$

then there exists a subsequence $\{f_{k_j}\}$ such that

$$f_{k_j} \overset{a.e.}{\to} f \quad x \in E.$$

Before proving Theorem 3.3.4, we record a corollary.

Corollary 3.3.1. *Let $E \subset \mathbb{R}^n$ be Lebesgue measurable with $m(E) < \infty$, and let f, f_k be measurable and finite a.e. on E. Then $f_k \overset{m}{\to} f$ on E if and only if every subsequence $\{f_{k_i}\}$ has a further subsequence $\{f_{k_{i_j}}\}$ such that*

$$f_{k_{i_j}} \overset{a.e.}{\to} f \quad x \in E.$$

Proof. Necessity follows from Theorem 3.3.4.

For sufficiency, argue by contradiction. If $f_k \not\Rightarrow f$, then there exist $\varepsilon_0, \delta_0 > 0$ and a subsequence $\{f_{k_i}\}$ such that

$$m(E(x||f_{k_i}(x) - f(x)| > \varepsilon_0)) \geq \delta_0 \quad \forall i.$$

By the assumption on subsequences, $\{f_{k_i}\}$ has a further subsequence $\{f_{k_{i_j}}\}$ with $f_{k_{i_j}} \overset{a.e.}{\to} f$. By Theorem 3.3.3, this implies $f_{k_{i_j}} \overset{m}{\to} f$, contradiction. □

Proof of Theorem 3.3.4. Since $f_k \overset{m}{\to} f$, for each $j \in \mathbb{N}$ we can choose k_j (strictly increasing) such that

$$m\Big(E\Big(x\Big||f_{k_j}(x) - f(x)| \geq \frac{1}{2^j}\Big)\Big) < \frac{1}{2^j}.$$

Define

$$E_j = E\Big(x\Big||f_{k_j}(x) - f(x)| \geq \frac{1}{2^j}\Big).$$

Then

$$\sum_{j=1}^{\infty} m(E_j) < \infty,$$

so

$$m\Big(\overline{\lim_{j\to\infty}} E_j\Big) = 0.$$

For $x \notin \overline{\lim} E_j$, there exists $j_0(x)$ such that for all $j \geq j_0(x)$,

$$|f_{k_j}(x) - f(x)| < \frac{1}{2^j},$$

which implies $f_{k_j}(x) \to f(x)$. Hence $f_{k_j} \overset{a.e.}{\to} f$ on E. □

When introducing a new convergence notion, one naturally asks about uniqueness of limits. To address this, we introduce Cauchy sequences in measure.

Definition 3.3.4 (Cauchy in Measure)**.** *Let $E \subset \mathbb{R}^n$ be Lebesgue measurable with $m(E) < \infty$. A sequence of measurable functions $\{f_k\}$ on E is called Cauchy in measure if for every $\varepsilon > 0$ and $\delta > 0$, there exists K such that whenever $i, j > K$,*

$$m(E(x||f_i(x) - f_j(x)| > \varepsilon)) < \delta.$$

Remark 3.3.3. *A Cauchy sequence in measure is sometimes also called a measure-Cauchy sequence.*

Theorem 3.3.5. *Let $E \subset \mathbb{R}^n$ be Lebesgue measurable with $m(E) < \infty$. Then $\{f_k\}$ is Cauchy in measure if and only if there exists a Lebesgue measurable function f on E such that*

$$f_k \overset{m}{\to} f \quad x \in E.$$

Proof. Sufficiency is immediate from the definition of convergence in measure.

For necessity, choose inductively a subsequence $\{f_{k_j}\}$ such that

$$m\Big(E\Big(x\Big||f_{k_{j+1}}(x) - f_{k_j}(x)| \geq \frac{1}{2^j}\Big)\Big) < \frac{1}{2^j}.$$

Set

$$A_j = E\Big(x\Big||f_{k_{j+1}}(x) - f_{k_j}(x)| \geq \frac{1}{2^j}\Big).$$

Then $\sum_j m(A_j) < \infty$, so $m(\overline{\lim}\, A_j) = 0$. For $x \notin \overline{\lim}\, A_j$, the series of successive differences is summable, hence $\{f_{k_j}(x)\}$ is Cauchy in $\mathbb{R}$ and converges pointwise. Define

$$f(x) = \lim_{j\to\infty} f_{k_j}(x)$$

on this full-measure set and extend on the null set arbitrarily; then f is measurable and

$$f_{k_j} \overset{a.e.}{\to} f \quad \text{on } E.$$

Therefore $f_{k_j} \overset{m}{\to} f$ by Theorem 3.3.3.

Finally, use the Cauchy-in-measure property of the full sequence: for given $\varepsilon, \delta > 0$, choose J so that for all $k > k_J$,

$$m(E(x||f_k - f_{k_J}| > \varepsilon/2)) < \delta/2,$$

and choose $j \geq J$ large enough so that

$$m(E(x||f_{k_j} - f| > \varepsilon/2)) < \delta/2.$$

Then for all $k > k_J$,

$$m(E(x||f_k - f| > \varepsilon)) \leq m(E(x||f_k - f_{k_j}| > \varepsilon/2)) + m(E(x||f_{k_j} - f| > \varepsilon/2)) < \delta.$$

Hence $f_k \overset{m}{\to} f$. □

Exercises 3.3

1. P140 1,3,4,5,6
2. P144 1,2
3. P149 10,11,13,15

3.4 Characterizations of Measurability and Composition Measurability

We now present a necessary and sufficient condition for function measurability.

Theorem 3.4.1 (Characterization of Measurable Functions)**.** *Let $E \subset \mathbb{R}^n$ be Lebesgue measurable. Then $f : E \to \mathbb{R}$ is Lebesgue measurable if and only if for every Borel set $B \subset \mathbb{R}$,*

$$E(x|f(x) \in B)$$

is Lebesgue measurable.

Proof. Sufficiency is immediate: for each $t \in \mathbb{R}$, $(t, \infty) \in \mathcal{B}$, so $E(x|f(x) > t)$ is measurable.

For necessity, use the good-set principle again. Define

$$\mathcal{S} = \{B \subset \mathbb{R} \mid f^{-1}(B) \text{ is Lebesgue measurable in } E\}.$$

(1) If $B_k \in \mathcal{S}$, then $f^{-1}(B_k)$ is measurable, so

$$f^{-1}\Big(\bigcup_{k=1}^{\infty} B_k\Big) = \bigcup_{k=1}^{\infty} f^{-1}(B_k)$$

is measurable. Hence $\mathcal{S}$ is closed under countable unions.

(2) If $B_1, B_2 \in \mathcal{S}$, then

$$f^{-1}(B_1 \cap B_2) = f^{-1}(B_1) \cap f^{-1}(B_2)$$

is measurable. So $\mathcal{S}$ is closed under finite intersections.

(3) If $B \in \mathcal{S}$, then

$$E = f^{-1}(\mathbb{R}) = f^{-1}(B \cup B^c) = f^{-1}(B) \cup f^{-1}(B^c),$$

so $f^{-1}(B^c)$ is measurable. Thus $\mathcal{S}$ is closed under complements in $\mathbb{R}$, hence under differences.

So $\mathcal{S}$ is a σ-ring. Also, for every interval $(c, d) \subset \mathbb{R}$,

$$f^{-1}((c, d)) = E(x|f(x) > c) \cap E(x|f(x) < d) \in \mathcal{S}.$$

Therefore all Borel sets belong to $\mathcal{S}$, and necessity follows. ☐

Remark 3.4.1. *Recall: a real-valued function on $\mathbb{R}$ is continuous iff the preimage of every open set is open. In this sense, measurable functions extend continuity.*

Now consider

$$E \xrightarrow{f} \mathbb{R} \xrightarrow{g} \mathbb{R},$$

where both f and g are Lebesgue measurable. Is $g \circ f : E \to \mathbb{R}$ always measurable? We have

$$\begin{aligned} E(x|g \circ f(x) > t) &= E(x|g \circ f(x) \in (t, \infty)) \\ &= E(x|f(x) \in g^{-1}((t, \infty))). \end{aligned}$$

The set $g^{-1}((t, \infty))$ is Lebesgue measurable, but not necessarily Borel. So Theorem 3.4.1 does not directly imply measurability of $g \circ f$. In fact, $g \circ f$ need not be measurable.

Exercises 3.4

1. Give a direct example showing that the composition of two Lebesgue measurable functions need not be measurable.

Chapter 4

Lebesgue Integral

In this chapter we build Lebesgue integration step by step. We begin with nonnegative measurable functions on finite-measure sets, then extend to general measurable functions on general measurable sets.

4.1 Integrals of Nonnegative Measurable Functions on Finite-Measure Sets

With the foundations of previous chapters, we now formalize the Lebesgue integral beyond the introductory definition.

Definition 4.1.1 (Lebesgue Integral)**.** *Let $E \subset \mathbb{R}^n$ be Lebesgue measurable with $m(E) < \infty$. Let $f : E \to \mathbb{R}$ be nonnegative, bounded, and measurable, so there exist $0 \le m_0 < M_0 < \infty$ such that*

$$m_0 \le f(x) < M_0, \quad x \in E.$$

Take a partition D of $[m_0, M_0)$:

$$m_0 = y_0 < y_1 < \cdots < y_k = M_0.$$

Set

$$\delta(D) = \max_{0 \le i \le k-1} (y_{i+1} - y_i),$$

and for arbitrary $\xi_i \in [y_i, y_{i+1}]$ define

$$S(f, D) = \sum_{i=0}^{k-1} \xi_i \, m(E_i),$$

where

$$E_i = \{x \in E \mid y_i \le f(x) < y_{i+1}\}.$$

If there exists a constant S such that for every $\varepsilon > 0$ there is $\delta > 0$ with

$$\delta(D) < \delta \implies |S(f, D) - S| < \varepsilon,$$

then f is called Lebesgue integrable on E, and S is called the Lebesgue integral of f on E, denoted by

$$(L) \int_E f(x) \, dx.$$

When no confusion occurs, (L) is omitted.

Example 4.1.1. *If $E \subset \mathbb{R}^n$ is Lebesgue measurable and $m(E) = 0$, then every bounded measurable function on E is Lebesgue integrable and*

$$\int_E f(x)\,dx = 0.$$

Example 4.1.2 (Dirichlet Function)**.** *The Dirichlet function (see 5) is Lebesgue integrable on $[0,1]$, and*

$$\int_{[0,1]} D(x)\,dx = 0.$$

Theorem 4.1.1. *Let $E \subset \mathbb{R}^n$ be Lebesgue measurable with $m(E) < \infty$. If f is bounded and Lebesgue measurable on E, then f is Lebesgue integrable on E.*

Proof. Define

$$\underline{S}(f,D) = \sum_{i=0}^{k-1} y_i\, m(E_i), \quad \overline{S}(f,D) = \sum_{i=0}^{k-1} y_{i+1}\, m(E_i).$$

Standard estimates give

$$\underline{S}(f,D) \le \overline{S}(f,D), \quad \overline{S}(f,D) - \underline{S}(f,D) \le \delta(D)m(E).$$

For nested refinements $D_1 \subset D_2 \subset \cdots$ with $\delta(D_j) \to 0$, the lower sums are increasing and upper sums are decreasing, with the same limit. This common limit is the desired integral value, and every sufficiently fine partition gives a Riemann-type sum close to it. □

Theorem 4.1.2. *Let $\{E_i\}_{i=1}^m \subset \mathcal{P}(\mathbb{R}^n)$ be Lebesgue measurable, pairwise disjoint, and let $E = \bigcup_{i=1}^m E_i$. If $m(E) < \infty$, then*

$$\int_E f(x)\,dx = \sum_{i=1}^m \int_{E_i} f(x)\,dx.$$

Proof. Write the defining sums for E and split each level set by E_i. Then pass to the limit in the partition mesh. □

Theorem 4.1.3 (Linearity)**.** *Let $E \subset \mathbb{R}^n$ be Lebesgue measurable with $m(E) < \infty$. If f, g are nonnegative bounded measurable functions on E, and $\alpha, \beta \ge 0$, then $\alpha f + \beta g$ is integrable and*

$$\int_E (\alpha f(x) + \beta g(x))\,dx = \alpha \int_E f(x)\,dx + \beta \int_E g(x)\,dx.$$

Proof. Scalar multiplication follows directly from the definition. For addition, partition both ranges and use a common refinement in the (f,g)-plane. Upper and lower estimates squeeze the integral of $f+g$ between sums converging to $\int_E f + \int_E g$. □

Theorem 4.1.4 (Monotonicity)**.** *Let $E \subset \mathbb{R}^n$ be Lebesgue measurable with $m(E) < \infty$. If f, g are integrable on E and $f \le g$ a.e. on E, then*

$$\int_E f(x)\,dx \le \int_E g(x)\,dx.$$

Proof. Set $h = g - f \ge 0$ a.e. Then

$$\int_E g - \int_E f = \int_E h \ge 0.$$

□

Theorem 4.1.5. *Let $E \subset \mathbb{R}^n$ with $m(E) < \infty$, and let f be nonnegative, bounded, and measurable on E. If*

$$\int_E f(x)\,dx = 0,$$

then $f = 0$ a.e. on E.

Proof. Left as an exercise. □

Theorem 4.1.6. *If f is nonnegative and Riemann integrable on $[a,b]$, then f is Lebesgue integrable on $[a,b]$ and*

$$(L)\int_{[a,b]} f(x)\,dx = (R)\int_a^b f(x)\,dx.$$

Proof. Since f is Riemann integrable, it is bounded. Let $\{D_k\}$ be nested partitions with mesh going to 0, and let φ_k, ψ_k be the step functions built from infima and suprema on partition subintervals. Then

$$\varphi_k \le f \le \psi_k, \quad \int \varphi_k \to \int_a^b f, \quad \int \psi_k \to \int_a^b f.$$

Passing to limits gives equality of the two integrals. □

Exercises 4.1

1. Let f be bounded and measurable on $\mathbb{R}^n$. Prove that

   $$I(x) = \int_{B(0,\|x\|)} f(y)\,dy$$

   is continuous.

2. Prove that f is Riemann integrable on $[a,b]$ if and only if f is continuous a.e. on $[a,b]$.

3. P220 1,2,3,4,5,6

4.2 Integrals of General Measurable Functions on General Measurable Sets

We now remove boundedness of f and finite-measure assumption on E, while first keeping $f \ge 0$.

Definition 4.2.1. *Let $E \subset \mathbb{R}^n$ be Lebesgue measurable and $f \ge 0$ measurable on E. Let $\{E_k\}$ be an increasing measurable exhaustion of E with*

$$E_k \subset E_{k+1}, \quad m(E_k) < \infty, \quad E = \bigcup_{k=1}^{\infty} E_k.$$

Define truncations

$$[f]_m(x) = \min\{f(x), m\}.$$

If

$$\lim_{k\to\infty} \int_{E_k} [f]_k(x)\,dx < \infty,$$

then f is Lebesgue integrable on E, and we define

$$\int_E f(x)\,dx = \lim_{k\to\infty} \int_{E_k} [f]_k(x)\,dx.$$

The definition must be independent of the chosen exhaustion and truncation indices.

Lemma 4.2.1. *Let $E \subset \mathbb{R}^n$ be Lebesgue measurable and $f \geq 0$ measurable on E. Let $\{E_k^{(j)}\}$, $j = 1, 2$, be two increasing finite-measure exhaustions of E, and let $\{m_k^{(i)}\}$, $i = 1, 2$, be increasing sequences with $m_k^{(i)} \to \infty$. If*

$$\lim_{k\to\infty} \int_{E_k^{(1)}} [f]_{m_k^{(1)}} \, dx < \infty,$$

then

$$\lim_{k\to\infty} \int_{E_k^{(2)}} [f]_{m_k^{(2)}} \, dx = \lim_{k\to\infty} \int_{E_k^{(1)}} [f]_{m_k^{(1)}} \, dx.$$

Proof. Let

$$S^{(1)} = \lim_{k\to\infty} \int_{E_k^{(1)}} [f]_{m_k^{(1)}} \, dx.$$

It suffices to prove for every finite-measure measurable $A \subset E$ and $M > 0$,

$$\int_A [f]_M \, dx \leq S^{(1)}. \tag{4.1}$$

Choose k large enough that $M \leq m_k^{(1)}$. Then

$$\int_A [f]_M \, dx \leq \int_{E_k^{(1)}} [f]_{m_k^{(1)}} \, dx + M \, m(A \backslash E_k^{(1)}).$$

Let $k \to \infty$. Since $A \backslash E_k^{(1)} \downarrow \varnothing$, the second term goes to 0, proving (4.1). Taking $A = E_k^{(2)}$ and $M = m_k^{(2)}$ yields

$$\lim_{k\to\infty} \int_{E_k^{(2)}} [f]_{m_k^{(2)}} \, dx \leq S^{(1)}.$$

By symmetry we get the reverse inequality. □

Definition 4.2.2. *For measurable $f : E \to \mathbb{R} \cup \{\pm\infty\}$, define*

$$f^+ = \max\{f, 0\}, \quad f^- = \max\{-f, 0\}.$$

These are called the positive and negative parts of f.

By definition,

$$|f| = f^+ + f^-.$$

Definition 4.2.3. *Let $E \subset \mathbb{R}^n$ be Lebesgue measurable and $f : E \to \mathbb{R} \cup \{\pm\infty\}$ measurable. If both f^+ and f^- are Lebesgue integrable, then f is called Lebesgue integrable on E, and*

$$\int_E f(x) \, dx = \int_E f^+(x) \, dx - \int_E f^-(x) \, dx.$$

Theorem 4.2.1. *Let $E \subset \mathbb{R}^n$ be Lebesgue measurable and f measurable on E. Define*

$$E^\infty = E(x | f(x) = \pm\infty).$$

If f is Lebesgue integrable, then

$$m(E^\infty) = 0.$$

Proof. Since $|f|$ is integrable, for any k,

$$k\, m(E_k \cap E^\infty) \leq \int_{E_k} [|f|]_k \, dx \leq \int_E |f| \, dx.$$

Let $k \to \infty$. □

Theorem 4.2.2 (Finite Additivity in the Region)**.** *Let $E_1, E_2 \subset \mathbb{R}^n$ be disjoint measurable sets and $E = E_1 \cup E_2$. If f is Lebesgue integrable on E, then f is integrable on E_1 and E_2, and*

$$\int_E f(x)\, dx = \int_{E_1} f(x)\, dx + \int_{E_2} f(x)\, dx.$$

Proof. Apply the definition to f^+ and f^- with the induced exhaustions $E_i \cap F_k$ from an exhaustion $\{F_k\}$ of E, then subtract. □

Theorem 4.2.3. *Let $E \subset \mathbb{R}^n$ be measurable. If $|f| \leq F$ a.e. on E and F is Lebesgue integrable on E, then f is Lebesgue integrable on E.*

Proof. From

$$0 \leq f^+ \leq F, \quad 0 \leq f^- \leq F \quad \text{a.e.},$$

we get

$$\int_E f^+ \leq \int_E F < \infty, \quad \int_E f^- \leq \int_E F < \infty.$$

Hence f is integrable. □

Theorem 4.2.4 (Linearity)**.** *Let $E \subset \mathbb{R}^n$ be measurable, $f, g \in L^1(E)$, and $\alpha, \beta \in \mathbb{R}$. Then $\alpha f + \beta g \in L^1(E)$ and*

$$\int_E (\alpha f + \beta g)\, dx = \alpha \int_E f\, dx + \beta \int_E g\, dx.$$

Proof. The scaling rule follows from truncations and sign decomposition for $\alpha > 0$, then for $\alpha < 0$ by $\alpha f = -(|\alpha| f)$. For addition, first prove for nonnegative functions using

$$[f+g]_M \leq [f]_M + [g]_M \leq [f+g]_{2M},$$

then pass to general functions via positive and negative parts. □

Theorem 4.2.5 (Monotonicity)**.** *Let $E \subset \mathbb{R}^n$ be measurable and $f, g \in L^1(E)$. If $f \leq g$ a.e. on E, then*

$$\int_E f\, dx \leq \int_E g\, dx.$$

Proof. Apply linearity to $h = g - f \geq 0$ a.e. □

Corollary 4.2.1. *If $f \in L^1(E)$, then*

$$\left| \int_E f(x)\, dx \right| \leq \int_E |f(x)|\, dx.$$

Theorem 4.2.6. *Let $E \subset \mathbb{R}^n$ be measurable and $f \in L^1(E)$ with $f \geq 0$ a.e. If*

$$\int_E f(x)\, dx = 0,$$

then $f = 0$ a.e. on E.

Proof. It is enough to prove $m(E(x|f(x) \geq \alpha)) = 0$ for every $\alpha > 0$. □

Example 4.2.1. *The improper Riemann integral may exist while the Lebesgue integral fails due to lack of absolute integrability. For example,* $\sin x/x$ *on* $[1, \infty)$ *is conditionally convergent in the improper Riemann sense but not Lebesgue integrable on* $[1, \infty)$.

Theorem 4.2.7 (Absolute Continuity of the Integral)**.** *Let* $E \subset \mathbb{R}^n$ *be measurable and* $f \in L^1(E)$. *For every* $\varepsilon > 0$, *there exists* $\delta > 0$ *such that for every measurable* $e \subset E$,

$$m(e) < \delta \implies \left| \int_e f(x)\, dx \right| < \varepsilon.$$

Proof. Choose k_0 so that

$$\int_E |f|\, dx - \int_E [|f|]_{k_0}\, dx < \frac{\varepsilon}{2},$$

and set $\delta = \varepsilon/(2k_0)$. Then for $m(e) < \delta$,

$$\int_e |f| \leq \int_E (|f| - [|f|]_{k_0})\, dx + k_0 m(e) < \varepsilon.$$

Hence $|\int_e f| \leq \int_e |f| < \varepsilon$. □

Theorem 4.2.8 (Countable Additivity in the Region)**.** *Let* $E \subset \mathbb{R}^n$ *be measurable, and let* $\{E_k\}$ *be pairwise disjoint measurable subsets with*

$$E = \bigcup_{k=1}^{\infty} E_k.$$

Then $f \in L^1(E)$ *if and only if:*

(1) $f \in L^1(E_k)$ *for all* k;

(2)

$$\sum_{k=1}^{\infty} \int_{E_k} |f(x)|\, dx < \infty.$$

Moreover, when $f \in L^1(E)$,

$$\int_E f(x)\, dx = \sum_{k=1}^{\infty} \int_{E_k} f(x)\, dx.$$

Proof. Necessity: from finite additivity on partial unions and monotonicity for $|f|$.

Sufficiency: choose an exhaustion F_m of E and define

$$G_m = F_m \cap \Big(\bigcup_{k=1}^{m} E_k \Big).$$

Then

$$\int_{G_m} |f| \leq \sum_{k=1}^{m} \int_{E_k} |f| < \infty,$$

so $f \in L^1(E)$. The series formula follows by controlling the tail with $\sum_{k>m} \int_{E_k} |f|$. □

Corollary 4.2.2. *Let* $f \geq 0$ *on* $\mathbb{R}^n$ *with* $f \in L^1(\mathbb{R}^n)$. *Define*

$$\varphi(E) = \int_E f(x)\, dx$$

for measurable E. *Then* φ *is a measure on* $(\mathbb{R}^n, \mathcal{L}^n)$.

Proof. This follows from $\varphi(\varnothing) = 0$, monotonicity, and countable additivity in the region. $\square$

Theorem 4.2.9. *If $f \in L^1(E)$ on measurable $E \subset \mathbb{R}^n$, then*

$$\lim_{k\to\infty} k\, m(E(x||f(x)| \geq k)) = 0.$$

Proof. For each k,

$$k\, m(E(x||f| \geq k)) \leq \int_{E(x||f|\geq k)} |f|\, dx.$$

Since $m(E(x||f| \geq k)) \to 0$ and the integral is absolutely continuous in the region (Theorem 4.2.7), the right side tends to 0. $\square$

For $E_k = E(x|k \leq f(x) < k+1)$, one gets

$$\sum_{k=1}^{\infty} k\, m(E_k) < \infty$$

for $f \in L^1(E)$; conversely, on finite-measure E, this condition implies integrability.

Theorem 4.2.10. *Let $E \subset \mathbb{R}^n$ be measurable and $f \geq 0$ measurable on E. Then $f \in L^1(E)$ if and only if*

$$\underline{S}(f) \stackrel{\text{def}}{=} \sup\left\{ \int_E h(x)\, dx \mid h \leq f \text{ a.e., and } h \text{ is simple}\right\} < \infty.$$

Moreover, if $f \in L^1(E)$, then

$$\int_E f(x)\, dx = \underline{S}(f).$$

Proof. If $f \in L^1(E)$, monotonicity gives $\underline{S}(f) \leq \int_E f$. For the reverse inequality, approximate f by level simple functions on E_k and use truncations. Taking limits yields $\int_E f \leq \underline{S}(f)$. $\square$

Exercises 4.2

1. P221 7,8,9

4.3 Limit Theorems for the Lebesgue Integral

This section presents core limit-exchange results in Lebesgue integration.

Theorem 4.3.1 (Lebesgue Dominated Convergence Theorem)**.** *Let $E \subset \mathbb{R}^n$ be measurable. Let $f_n, f \in L^1(E)$ with either*

$$f_n \stackrel{a.e.}{\to} f \quad or \quad f_n \stackrel{m}{\to} f.$$

Assume there exists $F \in L^1(E)$ such that

$$|f_n| \leq F \quad \text{a.e. on } E \text{ for all } n.$$

Then

$$\lim_{n\to\infty} \int_E f_n(x)\, dx = \int_E f(x)\, dx.$$

Proof. First $|f| \le F$ a.e., so $f \in L^1(E)$. Fix $\varepsilon > 0$. Choose measurable $E_m \subset E$ with $m(E_m) < \infty$ and

$$\int_{E \setminus E_m} F < \frac{\varepsilon}{6}.$$

On E_m, convergence in measure and absolute continuity of the integral imply

$$\int_{E_m} |f_n - f| < \frac{2\varepsilon}{3}$$

for n large. On $E \setminus E_m$,

$$\int_{E \setminus E_m} |f_n - f| \le 2 \int_{E \setminus E_m} F < \frac{\varepsilon}{3}.$$

Hence $\int_E |f_n - f| < \varepsilon$, giving the conclusion. □

Example 4.3.1. *Without integrable domination, the conclusion can fail even under uniform convergence. Let*

$$f_k(x) = \begin{cases} 1/k, & x \in [0,k), \\ 0, & x \in [k,\infty), \end{cases} \qquad x \in [0,\infty).$$

Then $f_k \rightrightarrows 0$, *but*

$$\int_{[0,\infty)} f_k(x)\,dx = 1 \not\to 0.$$

Example 4.3.2. *Pointwise a.e. convergence alone is not enough. Let*

$$f_k(x) = \begin{cases} k, & x \in [0,1/k), \\ 0, & x \in [1/k,1], \end{cases}$$

on $[0,1]$. *Then* $f_k \xrightarrow{a.e.} 0$ *but*

$$\int_{[0,1]} f_k(x)\,dx = 1 \not\to 0.$$

Theorem 4.3.2 (Levi's Lemma)**.** *Let $E \subset \mathbb{R}^n$ be measurable, and let $\{f_n\}$ be a monotone increasing sequence of integrable functions on E (or monotone decreasing, with the analogous sign condition). If*

$$\lim_{n\to\infty} \int_E f_n(x)\,dx < \infty,$$

then there exists $f \in L^1(E)$ such that

$$f_n \xrightarrow{a.e.} f,$$

and

$$\lim_{n\to\infty} \int_E f_n(x)\,dx = \int_E f(x)\,dx.$$

Proof. Assume $f_n \uparrow$. Define $f(x) = \lim_n f_n(x)$, measurable. Apply dominated convergence to truncations on finite-measure exhaustions to show $f \in L^1(E)$ and identify limits of integrals. □

Theorem 4.3.3 (Fatou's Lemma)**.** *Let $E \subset \mathbb{R}^n$ be measurable and $f_n \in L^1(E)$.*

If there exists $h \in L^1(E)$ with $h \le f_n$ a.e. for all n, then $\underline{\lim}_{n\to\infty} f_n$ is integrable and

$$\int_E \underline{\lim}_{n\to\infty} f_n(x)\,dx \le \underline{\lim}_{n\to\infty} \int_E f_n(x)\,dx.$$

If there exists $H \in L^1(E)$ with $f_n \le H$ a.e. for all n, then $\overline{\lim}_{n\to\infty} f_n$ is integrable and

$$\overline{\lim}_{n\to\infty} \int_E f_n(x)\,dx \le \int_E \overline{\lim}_{n\to\infty} f_n(x)\,dx.$$

Proof. For the first statement, define

$$\varphi_k(x) = \inf\{f_k(x), f_{k+1}(x), \dots\}.$$

Then $\varphi_k \uparrow \underline{\lim} f_n$ and each $\varphi_k \in L^1(E)$, with

$$\int_E \varphi_k \leq \varliminf_{n\to\infty} \int_E f_n.$$

Apply Levi (Theorem 4.3.2) and pass to the limit. The second statement follows by applying the first to $-f_n$. □

Remark 4.3.1. *The dominated convergence theorem, Levi's lemma, and Fatou's lemma are equivalent in strength.*

Example 4.3.3. *Without one-sided integrable control, Fatou-type conclusions can fail. A moving spike sequence on $[0,1]$ with value $-k$ on one $1/k$-subinterval and 0 elsewhere has constant integral -1, but liminf can be $-\infty$.*

Theorem 4.3.4 (Vitali-Type Limit Theorem)**.** *Let $E \subset \mathbb{R}^n$ be measurable, and let $\{f_k\} \subset L^1(E)$ satisfy:*

(1) *(uniform absolute continuity of integrals) for every $\varepsilon > 0$, there exists $\delta > 0$ independent of k such that*

$$m(e) < \delta \implies \int_e |f_k| < \varepsilon;$$

(2) *(uniform decay on the unbounded part) for every $\varepsilon > 0$, there exists measurable $F \subset E$ independent of k such that*

$$\int_{E \setminus F} |f_k| < \varepsilon;$$

(3) *$f_k \overset{a.e.}{\to} f$ or $f_k \overset{m}{\to} f$ on E.*

Then

$$\lim_{k\to\infty} \int_E f_k(x)\,dx = \int_E f(x)\,dx.$$

Proof. Choose bounded measurable $F \subset E$ so the tails on $E\setminus F$ are uniformly small. On F, use convergence in measure plus uniform absolute continuity to bound

$$\int_F |f_k - f|.$$

Together with the tail bound and Fatou for f, this gives

$$\int_E |f_k - f| \to 0,$$

hence convergence of integrals. □

Theorem 4.3.5 (Term-by-Term Integration)**.** *Let $E \subset \mathbb{R}^n$ be measurable and $f_n \in L^1(E)$ with*

$$\sum_{n=1}^{\infty} \int_E |f_n(x)|\,dx < \infty.$$

Then $\sum_{n=1}^{\infty} f_n(x)$ converges a.e. on E. If

$$f(x) = \sum_{n=1}^{\infty} f_n(x),$$

then $f \in L^1(E)$ and

$$\int_E f(x)\,dx = \sum_{n=1}^{\infty} \int_E f_n(x)\,dx.$$

Proof. Set

$$F_m = \sum_{n=1}^{m} |f_n|.$$

Then $F_m \uparrow F$ and by Levi,

$$\int_E F = \sum_{n=1}^{\infty} \int_E |f_n| < \infty.$$

So the series converges absolutely a.e. Let $S_m = \sum_{n=1}^{m} f_n$. Then $S_m \overset{a.e.}{\to} f$ and $|S_m| \le F$. Apply dominated convergence. □

Theorem 4.3.6 (Continuity of Parameter-Dependent Integrals)**.** *Let $E \subset \mathbb{R}^n$ be measurable, $B(y_0, \delta) \subset \mathbb{R}^m$, and*

$$f : E \times B(y_0, \delta) \to \mathbb{R}$$

with:

(1) *for each y, $f(\cdot, y)$ is measurable on E;*

(2) *for a.e. x, $f(x, \cdot)$ is continuous on $B(y_0, \delta)$;*

(3) *there exists $F \in L^1(E)$ such that $|f(x,y)| \le F(x)$ a.e. for all y.*

Then

$$I(y) = \int_E f(x,y)\,dx$$

is continuous on $B(y_0, \delta)$.

Proof. Take $y_k \to y$, set $g_k(x) = f(x, y_k)$ and $g(x) = f(x,y)$. Then $g_k \overset{a.e.}{\to} g$ and $|g_k| \le F$. Apply dominated convergence:

$$I(y_k) = \int_E g_k \to \int_E g = I(y).$$

□

Theorem 4.3.7 (Differentiability of Parameter-Dependent Integrals)**.** *Let $E \subset \mathbb{R}^n$ be measurable and*

$$f : E \times (a,b) \to \mathbb{R}$$

with:

(1) *for each $y \in (a,b)$, $f(\cdot, y)$ is measurable;*

(2) *for a.e. x, $f(x, \cdot)$ is differentiable on (a,b);*

(3) *there exists $F \in L^1(E)$ such that*

$$\left| \frac{\partial}{\partial y} f(x,y) \right| \le F(x)$$

a.e. for all $y \in (a,b)$.

Then

$$I(y) = \int_E f(x,y)\,dx$$

is differentiable on (a,b) and

$$I'(y) = \int_E \frac{\partial}{\partial y} f(x,y)\,dx.$$

Proof. Use difference quotients

$$g_h(x) = \frac{f(x, y+h) - f(x, y)}{h}.$$

Then $g_h \overset{a.e.}{\to} \partial_y f(\cdot, y)$ and $|g_h| \leq F$ a.e. Dominated convergence yields differentiation under the integral sign. □

Exercises 4.3

1. P185 1,5,6,7

2. P191 11

3. P222 12,18,19,20,21,23

4.4 Relation Between Lebesgue Integrable Functions and Continuous Functions

Lebesgue integrable functions are closely connected to continuous functions.

Theorem 4.4.1. *If $f \in L^1(\mathbb{R}^n)$, then for every $\varepsilon > 0$ there exists a continuous compactly supported function g on $\mathbb{R}^n$ such that*

$$\int_{\mathbb{R}^n} |f(x) - g(x)|\, dx < \varepsilon.$$

Proof. Choose k large and M large so that the tail and truncation errors satisfy

$$\int_{\mathbb{R}^n \setminus B(0,k)} |f| < \frac{\varepsilon}{3}, \quad \int_{B(0,k)} (|f| - [|f|]_M) < \frac{\varepsilon}{3}.$$

By Lusin's theorem, there exists closed $F \subset B(0, k)$ with small complement where $[f]_M$ is continuous. Extend continuously to $\mathbb{R}^n$ by a compactly supported g with support in $B(0, k)$. Then split the integral over $\mathbb{R}^n \setminus B(0, k)$, $B(0, k) \setminus F$, and F. □

Remark 4.4.1. *This means every L^1 function can be decomposed into a compactly supported continuous part plus an arbitrarily small L^1 remainder.*

Corollary 4.4.1. *If $f \in L^1(\mathbb{R}^n)$, then*

$$\lim_{h \to 0} \int_{\mathbb{R}^n} |f(x+h) - f(x)|\, dx = 0.$$

Proof. Left as an exercise. □

4.5 Fubini Theorem

In this section we study the relation between multiple and iterated integrals, and when integration order can be exchanged.

Theorem 4.5.1 (Tonelli Theorem)**.** *Let $f(x,y) \geq 0$ be Lebesgue measurable on $\mathbb{R}^n = \mathbb{R}^p \times \mathbb{R}^q$. Then:*

(A) *for a.e. $x \in \mathbb{R}^p$, $f(x,\cdot)$ is measurable on $\mathbb{R}^q$;*

(B)

$$F(x) = \int_{\mathbb{R}^q} f(x,y)\,dy$$

is measurable on $\mathbb{R}^p$;

(C)

$$\int_{\mathbb{R}^n} f(x,y)\,dx\,dy = \int_{\mathbb{R}^p} F(x)\,dx.$$

Define

$$\mathcal{F} = \{f \geq 0 \text{ measurable on } \mathbb{R}^n \mid f \text{ satisfies (A), (B), (C)}\}.$$

Lemma 4.5.1. *For $\mathcal{F}$ above:*

(1) *if $f \in \mathcal{F}$ and $\alpha \geq 0$, then $\alpha f \in \mathcal{F}$;*

(2) *if $f, g \in \mathcal{F}$, then $f + g \in \mathcal{F}$;*

(3) *if $f, g \in \mathcal{F}$, $g \in L^1(\mathbb{R}^n)$, and $f - g \geq 0$ a.e., then $f - g \in \mathcal{F}$;*

(4) *if $f_k \in \mathcal{F}$ and $f_k \uparrow f$ a.e., then $f \in \mathcal{F}$.*

Proof. (1) and (2) are immediate. For (3), use finiteness of the section integral of g a.e. and linearity of section integration. For (4), apply Levi in $\mathbb{R}^n$ and then in $\mathbb{R}^p$ to pass limits through both integrals. □

Proof of Theorem 4.5.1. Start with indicator functions of rectangles $I_1 \times I_2$, for which all statements are explicit. Then extend to open sets, closed sets, null sets, and general measurable sets by the good-set principle and monotone approximation. Next extend from indicators to simple functions, and then from simple functions to arbitrary nonnegative measurable functions using Lemma 4.5.1. □

Remark 4.5.1. *1. The roles of p and q are symmetric, so integration order can be exchanged under Tonelli assumptions.*

2. The theorem also holds on any measurable subset $E \subset \mathbb{R}^n$ by replacing f with $f\chi_E$.

Theorem 4.5.2 (Fubini)**.** *Let $f \in L^1(\mathbb{R}^n)$ on $\mathbb{R}^n = \mathbb{R}^p \times \mathbb{R}^q$. Then:*

(A) *for a.e. $x \in \mathbb{R}^p$, $f(x,\cdot)$ is measurable on $\mathbb{R}^q$;*

(B)

$$F(x) = \int_{\mathbb{R}^q} f(x,y)\,dy$$

is measurable on $\mathbb{R}^p$;

(C)

$$\int_{\mathbb{R}^n} f(x,y)\,dx\,dy = \int_{\mathbb{R}^q}\int_{\mathbb{R}^p} f(x,y)\,dx\,dy = \int_{\mathbb{R}^p}\int_{\mathbb{R}^q} f(x,y)\,dy\,dx.$$

Proof. Apply Tonelli theorem to f^+ and f^- separately and subtract. □

Definition 4.5.1 (Section). *Let $E \subset \mathbb{R}^n = \mathbb{R}^p \times \mathbb{R}^q$. For each $x \in \mathbb{R}^p$, define*

$$E(x) = \{y \in \mathbb{R}^q \mid (x, y) \in E\}.$$

This is the section of E at x. If E is Lebesgue measurable, Tonelli implies: for a.e. x, $E(x)$ is measurable, $m(E(x))$ is measurable in x (defined a.e.), and

$$m(E) = \int_{\mathbb{R}^p} m(E(x))\, dx.$$

Remark 4.5.2. *Similarly one defines $E(y)$; these are sometimes denoted by E_x and E^y.*

Corollary 4.5.1. *Let $f \geq 0$ be measurable on measurable $E \subset \mathbb{R}^n$, and define*

$$E^y = E(x|f(x) > y).$$

Then

$$\int_E f(x)\, dx = \int_0^\infty m(E^y)\, dy.$$

Proof. Consider

$$D = \{(x, y) \mid x \in E,\ 0 \leq y \leq f(x)\}.$$

By Fubini/Tonelli, compute $m(D)$ first by integrating in y, then in x, to obtain the layer-cake formula. □

Corollary 4.5.2. *If $A \subset \mathbb{R}^p$ and $B \subset \mathbb{R}^q$ are Lebesgue measurable, then $A \times B \subset \mathbb{R}^n$ is measurable and*

$$m(A \times B) = m(A)\, m(B).$$

Proof. Apply Tonelli to $\chi_{A\times B}(x, y) = \chi_A(x)\chi_B(y)$. □

Chapter 5

Measure Derivatives and the Newton-Leibniz Formula

At the beginning of this chapter we introduce an abstract measure class and then develop differentiation of measures and applications.

Definition 5.0.1 (Radon Measure)**.** *A measure μ on a topological space containing all Borel sets is called a Radon measure if it is inner regular and locally finite. Inner regular means: for every measurable set E,*

$$\mu(E) = \sup\{\mu(F) \mid F \subset E,\ F \text{ closed}\}.$$

Locally finite means: for every x, there is an open neighborhood U of x with $\mu(U) < \infty$.

Our discussion in this chapter is centered around Radon measures.

5.1 Vitali Covering Theorem

Definition 5.1.1 (Vitali Covering)**.** *Let $E \subset \mathbb{R}^n$, and let $\mathcal{F}$ be a family of closed balls in $\mathbb{R}^n$. We call $\mathcal{F}$ a Vitali (fine) cover of E if for every $x \in E$ and every $\varepsilon > 0$, there exists $B_x^\varepsilon \in \mathcal{F}$ such that*

$$x \in B_x^\varepsilon, \quad \text{diam}(B_x^\varepsilon) < \varepsilon.$$

If $\mathcal{F}$ is a fine cover of E, then

$$\mathcal{F}_\delta = \{B \in \mathcal{F} \mid \text{diam}(B) < \delta\}$$

is also a fine cover of E.

If additionally $\mu^*(E) < \infty$ and there exists open $G \supset E$ with $\mu(G) < \infty$, then

$$\mathcal{F}_\delta^G = \{B \in \mathcal{F} \mid B \subset G,\ \text{diam}(B) < \delta\}$$

is still a fine cover of E.

Theorem 5.1.1 (Vitali)**.** *Let $\mathcal{F}$ be a family of nondegenerate closed balls in $\mathbb{R}^n$ with*

$$d^* = \sup\{\text{diam}(B) \mid B \in \mathcal{F}\} < \infty.$$

Then there exists an at most countable disjoint subfamily $\mathcal{G} \subset \mathcal{F}$ such that

$$\bigcup_{B \in \mathcal{F}} B \subset \bigcup_{B \in \mathcal{G}} \widehat{B},$$

where $\widehat{B}$ is the concentric ball with radius 5 times that of B.

Proof. Partition $\mathcal{F}$ by dyadic diameter levels and choose maximal disjoint subfamilies inductively. Any remaining ball at level $k+1$ must intersect a previously chosen ball of level at most $k+1$. The diameter comparison yields containment inside a 5-dilate of the chosen intersecting ball. □

Corollary 5.1.1. *Let $E \subset \mathbb{R}^n$ and let $\mathcal{F}$ be a fine cover of E by closed balls, with bounded diameters. Then there exists a countable disjoint subfamily $\mathcal{G} \subset \mathcal{F}$ such that for every finite subset*

$$\{B_1, \ldots, B_k\} \subset \mathcal{G},$$

one has

$$E \backslash \bigcup_{j=1}^{k} B_j \subset \bigcup_{B \in \mathcal{G} \backslash \{B_1, \ldots, B_k\}} \widehat{B}.$$

Proof. Apply Theorem 5.1.1 to the remaining part after removing finitely many selected balls. □

Corollary 5.1.2. *Let μ be a nonnegative Radon measure on $\mathbb{R}^n$, let $E \subset \mathbb{R}^n$ with $\mu^*(E) < \infty$, and let $\mathcal{F}$ be a fine cover of E. Then for every $\varepsilon > 0$, there are finitely many pairwise disjoint balls $B_1, \ldots, B_k \in \mathcal{F}$ such that*

$$\mu^*\Big(E \backslash \bigcup_{j=1}^{k} B_j\Big) < \varepsilon.$$

Proof. Restrict to balls contained in an open $G \supset E$ with finite measure, use Corollary 5.1.1, and then use countable additivity and tail control of measure. □

Corollary 5.1.3. *Let μ be a nonnegative Radon measure on $\mathbb{R}^n$, let $E \subset \mathbb{R}^n$, and let $\mathcal{F}$ be a fine cover of E. Then there is a countable disjoint family $\{B_j\} \subset \mathcal{F}$ such that*

$$\mu^*\Big(E \backslash \bigcup_{j=1}^{\infty} B_j\Big) = 0.$$

Proof. Localize on bounded annuli, apply Corollary 5.1.2 on each localization with geometric errors, then add a final fine-cover selection on the remainder. □

Remark 5.1.1. *1. These results play the same role as finite covering arguments in classical analysis: local estimates on balls are upgraded to global statements by almost full disjoint-ball coverings.*

2. Compare the strength of the four statements carefully. Theorem 5.1.1 assumes only a bounded-diameter ball family and gives coverage after dilation. Corollary 5.1.3 requires fine covering and yields measure-theoretic almost-full coverage by disjoint original balls.

Exercises 5.1

1. P248 1,2

5.2 Hahn Decomposition Theorem

Definition 5.2.1 (Positive Set, Negative Set)**.** *Let μ be a Radon measure on $\mathbb{R}^n$. A set $A \subset \mathbb{R}^n$ is called positive for μ if for every μ-measurable E,*

$$\mu(A \cap E) \geq 0.$$

A set $B \subset \mathbb{R}^n$ is called negative for μ if for every μ-measurable E,

$$\mu(B \cap E) \leq 0.$$

Remark 5.2.1. *If A is positive, then A is measurable and $\mu(A) \geq 0$. But $\mu(A) \geq 0$ alone does not imply A is positive.*

The goal is to decompose $\mathbb{R}^n$ into one positive and one negative part.

Lemma 5.2.1. *Let μ be a Radon measure on $\mathbb{R}^n$.*

(1) *If A_1, A_2 are positive sets, then $A_1 \cup A_2$, $A_1 \cap A_2$, and $A_1 \backslash A_2$ are positive.*

(2) *If B_1, B_2 are negative sets, then $B_1 \cup B_2$, $B_1 \cap B_2$, and $B_1 \backslash B_2$ are negative.*

(3) *If A is positive and B is negative, then for every measurable E,*

$$\mu(E \cap A \cap B) = 0.$$

Proof. For (1), use set decompositions of intersections and sign constraints term by term. Part (2) follows from (1) by replacing μ with $-\mu$. Part (3) is immediate from both inequalities. □

Theorem 5.2.1 (Hahn Decomposition Theorem)**.** *For any Radon measure μ on $\mathbb{R}^n$, there exist a positive set A and a negative set B such that*

$$A \cap B = \varnothing, \quad \mathbb{R}^n = A \cup B.$$

Proof. Work first on each annulus

$$M_k = \{x \in \mathbb{R}^n \mid k \leq \|x\| < k+1\}.$$

Take the infimum of μ over negative subsets of M_k, build an increasing sequence of negative sets approaching this infimum, and pass to the union. The complement is then shown to be positive by contradiction using maximality of the infimum and Lemma 5.2.1. Combine annuli. □

Remark 5.2.2. *Define*

$$\mu^+(E) = \mu(E \cap A), \quad \mu^-(E) = -\mu(E \cap B).$$

Then

$$\mu = \mu^+ - \mu^-.$$

This is the Jordan decomposition of a signed Radon measure.

5.3 Derivatives of Radon Measures

We now extend the derivative concept from analysis to Radon measures.

Definition 5.3.1 (Derivative)**.** *Let ν, μ be nonnegative Radon measures on $\mathbb{R}^n$. For each $x \in \mathbb{R}^n$, define*

$$\overline{D}_\mu \nu(x) = \begin{cases} \varlimsup_{r\to 0^+} \dfrac{\nu(B(x,r))}{\mu(B(x,r))}, & \mu(B(x,r)) > 0 \ \forall r > 0, \\ \infty, & \exists r > 0 \textit{ with } \mu(B(x,r)) = 0, \end{cases}$$

$$\underline{D}_\mu \nu(x) = \begin{cases} \varliminf_{r\to 0^+} \dfrac{\nu(B(x,r))}{\mu(B(x,r))}, & \mu(B(x,r)) > 0 \ \forall r > 0, \\ \infty, & \exists r > 0 \textit{ with } \mu(B(x,r)) = 0. \end{cases}$$

If upper and lower derivatives coincide, we denote the common value by

$$D_\mu \nu(x).$$

Example 5.3.1 (Average Value for Continuous Density)**.** *Let $f \geq 0$ be continuous and locally integrable on $\mathbb{R}^n$. Let $\mu = m$ (Lebesgue measure) and*

$$\nu(E) = \int_E f(x)\, dx.$$

Then

$$D_\mu \nu(x) = \lim_{r\to 0} \frac{\int_{B(x,r)} f(y)\, dy}{m(B(x,r))} = f(x).$$

Lemma 5.3.1. *Let ν, μ be nonnegative Radon measures on $\mathbb{R}^n$.*

(1) *If*

$$E \subset \{x \in \mathbb{R}^n \mid \underline{D}_\mu \nu(x) \leq a\},$$

then

$$\nu^*(E) \leq a\, \mu^*(E).$$

(2) *If*

$$E \subset \{x \in \mathbb{R}^n \mid \overline{D}_\mu \nu(x) \geq b\},$$

then

$$\nu^*(E) \geq b\, \mu^*(E).$$

Proof. Use Vitali covering selections built from balls where the local ratio is near the liminf or limsup threshold, then apply Corollary 5.1.3 and countable additivity. □

Theorem 5.3.1. *Let ν, μ be nonnegative Radon measures on $\mathbb{R}^n$.*

(1) *$D_\mu \nu$ exists and is finite μ-a.e. on $\mathbb{R}^n$.*
(2) *$D_\mu \nu$ is μ-measurable.*

Proof. For existence and finiteness, apply Lemma 5.3.1 to sets where $\overline{D}_\mu \nu = \infty$ and where $\underline{D}_\mu \nu < \overline{D}_\mu \nu$. Both have zero μ-measure. For measurability, approximate derivatives through measurable ball-measure maps

$$x \mapsto \nu(B(x,r)), \quad x \mapsto \mu(B(x,r)).$$

Then pass to limsup/liminf operations. □

For a measurable set E with $\mu(E) > 0$, write

$$\fint_E f\,d\mu = \frac{1}{\mu(E)} \int_E f\,d\mu.$$

Theorem 5.3.2 (Average Value Theorem)**.** *If $f \in L^1_{\mathrm{loc}}(\mathbb{R}^n)$, then*

$$\lim_{r\to 0} \fint_{B(x,r)} f(y)\,dy = f(x)$$

for a.e. $x \in \mathbb{R}^n$.

Proof. This is the Lebesgue differentiation theorem. Approximate f in L^1_{loc} by continuous functions and apply Lemma 5.3.1 to the error measure. □

Remark 5.3.1. *A stronger form is*

$$\lim_{r\to 0} \fint_{B(x,r)} |f(y) - f(x)|\,dy = 0$$

a.e. Such points are called Lebesgue points.

Corollary 5.3.1. *Let $E \subset \mathbb{R}^n$ be Lebesgue measurable. Then for a.e. $x \in E$,*

$$\lim_{r\to 0} \frac{m(E \cap B(x,r))}{m(B(x,r))} = 1.$$

Proof. Apply Theorem 5.3.2 to $f = \chi_E$. □

Remark 5.3.2. *For positive-measure nowhere-dense perfect sets, the pointwise ratio is always < 1 for each fixed $r > 0$, yet the limit above is still 1 a.e.*

Exercises 5.3

1. Let $\{x_k\}, \{y_k\} \subset \mathbb{R}^n$ be discrete sequences and define

$$\mu(E) = \sum_{k=1}^{\infty} \frac{1}{2^k}\delta_{x_k}(E) + m(E), \quad \nu(E) = \sum_{k=1}^{\infty} \frac{1}{2^k}\delta_{y_k}(E).$$

   Find $D_\mu \nu(x)$.

5.4 Radon-Nikodym Theorem

For Radon measures we have the analog of the Newton-Leibniz principle.

Definition 5.4.1 (Absolute Continuity)**.** *Let ν, μ be Radon measures on $\mathbb{R}^n$. We say ν is absolutely continuous with respect to μ (written $\nu \ll \mu$) if for every measurable E,*

$$|\mu|(E) = 0 \implies |\nu|(E) = 0,$$

where $|\mu| = \mu^+ + \mu^-$. When $\mu \geq 0$, this is equivalent to $\mu(E) = 0 \Rightarrow \nu(E) = 0$.

Theorem 5.4.1 (Radon-Nikodym)**.** *Let μ be a nonnegative Radon measure on $\mathbb{R}^n$, and let ν be a Radon measure with $\nu \ll \mu$. Then there exists a locally μ-integrable function f such that for every measurable $E \subset \mathbb{R}^n$,*

$$\nu(E) = \int_E f\, d\mu = \int_E D_\mu \nu\, d\mu. \tag{5.1}$$

Proof. Step 1: every μ-measurable set is ν-measurable under $\nu \ll \mu$.

Step 2: for

$$E^\infty = \{D_\mu \nu = \infty\}, \quad E_0 = \{D_\mu \nu = 0\},$$

prove

$$\nu(E^\infty) = \int_{E^\infty} D_\mu \nu\, d\mu, \quad \nu(E_0) = \int_{E_0} D_\mu \nu\, d\mu.$$

Step 3: decompose any measurable E into derivative level sets

$$E_k = \{x \in E \mid D_\mu \nu(x) \in [t^k, t^{k+1})\}, \quad t > 1,$$

and sandwich $\nu(E)$ between $t \int_E D_\mu \nu\, d\mu$ and $t^{-1} \int_E D_\mu \nu\, d\mu$. Let $t \to 1$. ◻

Corollary 5.4.1. *Let μ be a nonnegative Radon measure on $\mathbb{R}^n$ and let $f \in L^1_{\mathrm{loc}}(\mu)$. Then*

$$\lim_{r \to 0} \fint_{B(x,r)} f(y)\, d\mu = f(x)$$

for μ-a.e. x.

Proof. Apply Theorem 5.4.1 to the signed measure

$$\nu(E) = \int_E f\, d\mu.$$

Then $f = D_\mu \nu$ a.e. ◻

Remark 5.4.1. *Formally,*

$$D_\mu \left(\int f\, d\mu \right) = f \quad \textit{a.e.}$$

so measure differentiation is the inverse of integration.

Definition 5.4.2 (Mutual Singularity)**.** *Let ν, μ be nonnegative Radon measures on $\mathbb{R}^n$. We write $\nu \perp \mu$ if there exists measurable $B \subset \mathbb{R}^n$ such that*

$$\mu(\mathbb{R}^n \backslash B) = 0, \quad \nu(B) = 0.$$

Remark 5.4.2. *In general, singularity does not imply absolute continuity.*

Theorem 5.4.2 (Lebesgue Decomposition Theorem)**.** *Let $\mu \geq 0$ be a Radon measure on $\mathbb{R}^n$, and let ν be a Radon measure. Then*

$$\nu = \nu_{\mathrm{AC}} + \nu_s,$$

where $\nu_{\mathrm{AC}} \ll \mu$ and $\nu_s \perp \mu$.

Proof. Define

$$\mathcal{E} = \{A \subset \mathbb{R}^n \mid \mu(\mathbb{R}^n \backslash A) = 0\},$$

choose $B \in \mathcal{E}$ minimizing $\nu(A)$ over $\mathcal{E}$ (via infimizing sequence and intersection), and set

$$\nu_{\mathrm{AC}} = \nu|_B, \quad \nu_s = \nu|_{\mathbb{R}^n \backslash B}.$$

Minimality gives $\nu_{\mathrm{AC}} \ll \mu$, while $\nu_s \perp \mu$ follows from definition of $\mathcal{E}$. □

Remark 5.4.3. *For nonnegative μ,*

$$D_\mu \nu = D_\mu \nu_{\mathrm{AC}} \quad \mu\text{-a.e.}, \qquad D_\mu \nu_s = 0 \quad \mu\text{-a.e.}$$

under the theorem assumptions.

Example 5.4.1. *With $\mu = m$ on $\mathbb{R}$,*

$$\nu(E) = \int_E f(x)\, d\mu + \sum_{i=1}^{\infty} \alpha_i \delta_{x_i}(E)$$

is the sum of an absolutely continuous part and a singular atomic part.

Exercises 5.4

1. Does there exist a non-atomic Radon measure singular to Lebesgue measure?

5.5 Monotone Functions and Functions of Bounded Variation

5.5.1 Monotone Functions

This section parallels the previous one because a monotone function induces a measure.

Theorem 5.5.1. *Let f be monotone increasing on $[a, b]$. Then f is differentiable a.e. on $[a, b]$.*

Example 5.5.1 (Monotone Function Discontinuous on a Dense Set)**.** *Let $\sum(\alpha_n + \beta_n) < \infty$ with $\alpha_n, \beta_n > 0$, let $\{x_n\} = \mathbb{Q}$, and define*

$$f(x) = \sum_{n=1}^{\infty} \alpha_n H_0(x - x_n) + \sum_{n=1}^{\infty} \beta_n H_1(x - x_n),$$

where

$$H_0(x) = \begin{cases} 1, & x > 0, \\ 0, & x \le 0, \end{cases} \qquad H_1(x) = \begin{cases} 1, & x \ge 0, \\ 0, & x < 0. \end{cases}$$

Then f has dense discontinuities but is still differentiable a.e. by Theorem 5.5.1.

Remark 5.5.1. *For background on the Heaviside function, see [4].*

Lemma 5.5.1. *Let f be monotone increasing on $[a, b]$. If*

$$E \subset \{x \in [a, b] \mid \underline{D}f(x) \le \alpha\},$$

then

$$f^*(E) \le \alpha\, m^*(E).$$

If

$$E \subset \{x \in [a, b] \mid \overline{D}f(x) \ge \beta\},$$

then

$$f^*(E) \ge \beta\, m^*(E),$$

where

$$\underline{D}f(x) = \underline{\lim}_{h \to 0} \frac{f(x+h) - f(x)}{h}, \quad \overline{D}f(x) = \overline{\lim}_{h \to 0} \frac{f(x+h) - f(x)}{h},$$

and $f^*(E) = m^*(f(E))$.

Proof. Construct fine covers by intervals on which difference quotients satisfy the desired bound, apply Vitali covering, and estimate image measure by interval lengths. □

Proof of Theorem 5.5.1. Define

$$E^\infty = \{\overline{D}f = \infty\}.$$

By Lemma 5.5.1, $m(E^\infty) = 0$. For

$$E_\alpha^\beta = \{\underline{D}f \le \alpha < \beta \le \overline{D}f\},$$

Lemma 5.5.1 gives

$$\beta m^*(E_\alpha^\beta) \le f^*(E_\alpha^\beta) \le \alpha m^*(E_\alpha^\beta),$$

so $m(E_\alpha^\beta) = 0$. Take countable rational pairs $\alpha < \beta$. □

Theorem 5.5.2. *If f is monotone increasing on $[a, b]$, then*

$$\int_a^b f'(x)\, dx \le f(b) - f(a).$$

Proof. Use difference quotients

$$g_k(x) = \frac{f(x + 1/k) - f(x)}{1/k} \ge 0,$$

then apply Fatou/Levi type passage and monotonicity bounds. □

Theorem 5.5.3 (Fubini Term-by-Term Differentiation)**.** *Let f_k be monotone increasing on $[a, b]$, and suppose $\sum_{k=1}^\infty f_k(x)$ converges a.e. on $[a, b]$. Then*

$$\frac{d}{dx}\Big(\sum_{k=1}^\infty f_k(x)\Big) = \sum_{k=1}^\infty f_k'(x)$$

a.e. on $[a, b]$.

Proof. Set $S = \sum f_k$, so S is monotone and differentiable a.e. Then write

$$S'(x) = \sum_{k=1}^n f_k'(x) + r_n'(x),$$

with $r_n = \sum_{k=n+1}^\infty f_k$. Show $r_n' \downarrow 0$ a.e. via Theorem 5.5.2 and convergence of the tail at endpoints. Let $n \to \infty$. □

5.5.2 Functions of Bounded Variation

Definition 5.5.1 (Function of Bounded Variation)**.** *Let f be finite-valued on $[a, b]$. For a partition*

$$a = x_0 \le x_1 \le \cdots \le x_n = b,$$

define

$$\bigvee_f(\Delta) = \sum_{i=1}^{n} |f(x_i) - f(x_{i-1})|.$$

If

$$\bigvee_a^b(f) = \sup\{\bigvee_f(\Delta) \mid \Delta \text{ partition of } [a, b]\} < \infty,$$

then $f \in BV[a, b]$ and $\bigvee_a^b(f)$ is called total variation.

Example 5.5.2. *Every bounded monotone function on $[a, b]$ is in $BV[a, b]$, and*

$$\bigvee_a^b(f) = |f(b) - f(a)|.$$

Example 5.5.3. *If f is Lipschitz on $[a, b]$ with constant L, then $f \in BV[a, b]$ and*

$$\bigvee_a^b(f) \le L|b - a|.$$

Example 5.5.4 (Continuous but Not of Bounded Variation)**.**

$$f(x) = \begin{cases} x \sin\left(\dfrac{1}{x^2}\right), & x \in (0, 1], \\ 0, & x = 0. \end{cases}$$

Theorem 5.5.4. *Functions in $BV[a, b]$ satisfy:*

(1) *boundedness;*

(2) *linear closure and*

$$\bigvee_a^b(\alpha f + \beta g) \le |\alpha| \bigvee_a^b(f) + |\beta| \bigvee_a^b(g);$$

(3) *product closure with standard estimate;*

(4) $\bigvee_a^b(f) = 0 \iff f$ *constant;*

(5) *restriction monotonicity on subintervals;*

(6) *additivity of total variation at a split point c:*

$$\bigvee_a^b(f) = \bigvee_a^c(f) + \bigvee_c^b(f);$$

(7) *if $f_n \in BV[a, b]$, $\sup_n \bigvee_a^b(f_n) \le M$, and $f_n \to f$ pointwise, then $f \in BV[a, b]$ and*

$$\bigvee_a^b(f) \le M.$$

Proof. Only (6) and (7) are nontrivial here. For (6), refine partitions by inserting c and pass to suprema in both directions. For (7), for each fixed partition pass limit inside finite sums:

$$\bigvee_f(\Delta) = \lim_{n\to\infty} \bigvee_{f_n}(\Delta) \leq M,$$

then take supremum over Δ. ☐

Remark 5.5.2. *Property* (6) *implies*

$$x \mapsto \bigvee_a^x (f)$$

is monotone increasing.

Theorem 5.5.5 (Jordan Decomposition Theorem)**.** *If $f \in BV[a,b]$ and*

$$g(x) = \frac{1}{2}\Big(\bigvee_a^x(f) + f(x)\Big), \quad h(x) = \frac{1}{2}\Big(\bigvee_a^x(f) - f(x)\Big),$$

then g, h are increasing on $[a,b]$ and

$$f = g - h.$$

Proof. For $x \leq y$,

$$|f(y) - f(x)| \leq \bigvee_x^y(f) = \bigvee_a^y(f) - \bigvee_a^x(f),$$

which gives monotonicity of both $\bigvee_a^x(f) \pm f(x)$. ☐

Corollary 5.5.1. *If $f \in BV[a,b]$, then its discontinuity set is at most countable, f is differentiable a.e., and $f' \in L^1[a,b]$.*

Theorem 5.5.6. *If $f \in BV[a,b]$, then f and $\bigvee_a^x(f)$ have the same continuity/discontinuity points, and at each discontinuity x_0 the one-sided oscillations satisfy*

$$\omega_f^+(x_0) = \lim_{x\to x_0^+} |f(x) - f(x_0)| = \lim_{x\to x_0^+} \bigvee_{x_0}^x(f),$$

$$\omega_f^-(x_0) = \lim_{x\to x_0^-} |f(x) - f(x_0)| = \lim_{x\to x_0^-} \bigvee_{x_0}^x(f).$$

Proof. One inequality is immediate from

$$|f(x) - f(x_0)| \leq \bigvee_{x_0}^x(f).$$

The reverse inequality follows by choosing near-optimal partitions and isolating the first subinterval. ☐

Theorem 5.5.7. *If $f \in BV[a,b]$, then*

$$\frac{d}{dx}\Big(\bigvee_a^x(f)\Big) = |f'(x)|$$

a.e. on $[a,b]$.

Proof. Approximate variation by piecewise monotone envelopes and use the term-by-term differentiation theorem (Theorem 5.5.3) on a suitable positive series of monotone error terms. □

Remark 5.5.3. *Consequently,*

$$\int_a^b |f'(x)|\,dx \le \bigvee_a^b(f).$$

Theorem 5.5.8 (Hally)**.** *Let $\mathcal{F} \subset BV[a,b]$ with*

$$\sup_{f\in\mathcal{F}} \Big(|f(a)| + \bigvee_a^b(f)\Big) < M.$$

Then every sequence in $\mathcal{F}$ has a subsequence converging pointwise to a function in $BV[a,b]$.

Proof. First handle monotone families via diagonal extraction on rational points. Extend limits by monotonicity and adjust values at countably many discontinuities. Then reduce general BV case through Jordan decomposition. □

Exercises 5.5

1. P255 10,11,12
2. P281 11
3. P283 7,8

5.6 Absolutely Continuous Functions

This section gives necessary and sufficient conditions for the Newton-Leibniz formula.

Theorem 5.6.1. *If $f \in L^1[a,b]$ and*

$$F(x) = \int_a^x f(t)\,dt,$$

then F is differentiable a.e. and

$$F'(x) = f(x) \quad a.e.$$

Proof. Write

$$F(x) = \int_a^x f^+(t)\,dt - \int_a^x f^-(t)\,dt.$$

Then $F \in BV[a,b]$, hence differentiable a.e. Use approximation of f by continuous functions in L^1 together with variation estimates to conclude $F' = f$ a.e. □

Definition 5.6.1 (Absolute Continuity)**.** *A real-valued function f on $[a,b]$ is absolutely continuous if for every $\varepsilon > 0$ there exists $\delta > 0$ such that for every finite family of pairwise disjoint intervals*

$$[x_1, y_1], \ldots, [x_n, y_n] \subset [a,b],$$

if

$$\sum_{i=1}^{n}(y_i - x_i) < \delta,$$

then

$$\sum_{i=1}^{n}|f(y_i) - f(x_i)| < \varepsilon.$$

The class is denoted by $AC[a,b]$.

Lemma 5.6.1 (Singular Function)**.** *Suppose* f *is nonconstant on* $[a,b]$*, differentiable a.e., and*

$$f'(x) = 0 \quad a.e.$$

Then $f \notin AC[a,b]$.

Proof. If f were absolutely continuous, then by definition and Vitali-type interval selection one obtains arbitrarily short disjoint interval families with arbitrarily small total increment, contradicting nonconstancy under $f' = 0$ a.e. □

Theorem 5.6.2. (1) *If* $f \in AC[a,b]$*, then* $f \in C[a,b]$.

(2) *If* $f, g \in AC[a,b]$ *and* $\alpha, \beta \in \mathbb{R}$*, then* $\alpha f + \beta g \in AC[a,b]$.

(3) *If* $f \in AC[a,b]$*, then* $f \in BV[a,b]$.

(4) *If* $f \in L^1[a,b]$ *and*

$$F(x) = \int_a^x f(t)\,dt,$$

then $F \in AC[a,b]$.

Proof. Standard and omitted. □

Now let $f \in AC[a,b]$ and define

$$g(x) = \int_a^x f'(t)\,dt.$$

By part (4) of Theorem 5.6.2, $g \in AC[a,b]$. Hence $h = g - f \in AC[a,b]$. By Theorem 5.6.1,

$$h'(x) = g'(x) - f'(x) = 0 \quad \text{a.e.}$$

and by Lemma 5.6.1, h must be constant. Therefore

$$f(x) = \int_a^x f'(t)\,dt + \text{const}.$$

Evaluating at $x = b$ yields:

Theorem 5.6.3 (Newton-Leibniz)**.** *If* $f \in AC[a,b]$*, then*

$$f(b) - f(a) = \int_a^b f'(t)\,dt.$$

For a.e. differentiable f with $f' \in L^1[a,b]$, define

$$f_s(x) = f(x) - \int_a^x f'(t)\,dt.$$

Then $f_s'(x) = 0$ a.e., so

$$f(x) = \int_a^x f'(t)\,dt + f_s(x) = f_{\text{AC}}(x) + f_s(x),$$

which is the decomposition into absolutely continuous and singular parts.

This parallels the Radon-Nikodym section. Compare the two frameworks.

Appendix

A Another Construction of Lebesgue Measure

Below we reconstruct Lebesgue measure from the viewpoint of abstract measure theory.

Measurable Sets and Measures

Definition A.1 (Measurable Space)**.** *Let X be a nonempty set, and let $\mathcal{M}$ be a σ-algebra on X. Then the pair $(X, \mathcal{M})$ is called a measurable space.*

Definition A.2 (Measure)**.** *Let $(X, \mathcal{M})$ be a measurable space. A function $\mu : \mathcal{M} \to [0, +\infty]$ is called a measure on this measurable space if it is countably additive.*

Remark A.1. *Naturally, we assume that μ is not identically $+\infty$.*

Proposition A.1. *Let $(X, \mathcal{M})$ be a measurable space, and let μ be a measure on it. Then μ has the following properties:*

(1) $\mu(\varnothing) = 0$*;*

(2) *finite additivity;*

(3) *monotonicity;*

(4) *continuity from below: if $A_1 \subset A_2 \subset \cdots$, then*

$$\lim_{n\to\infty} \mu(A_n) = \mu\Big(\bigcup_{n=1}^{\infty} A_n\Big);$$

(5) *continuity from above: if $A_1 \supset A_2 \supset \cdots$ and $\mu(A_1) < +\infty$, then*

$$\lim_{n\to\infty} \mu(A_n) = \mu\Big(\bigcap_{n=1}^{\infty} A_n\Big).$$

Now let us look at some examples of abstract measures.

Example A.1 (Counting Measure)**.** *Take the measurable space $(X, \mathcal{P}(X))$. Define $\mu : \mathcal{P}(X) \to \mathbb{N} \cup \{\infty\}$ by mapping each set A to its cardinality.*

Remark A.2. *In fact, all series theory can be viewed as integration theory for counting measure.*

Example A.2 (Dirac Measure)**.** *If $x \in E$, then $\delta_x(E) = 1$. If $x \notin E$, then $\delta_x(E) = 0$.*

Bibliography

[1] Minqiang Zhou, *Theory of Real Functions*, Peking University Press, Beijing, 2001.

[2] Daoxing Xia et al., *Real Functions and Functional Analysis* (2nd ed.), Volume I, Higher Education Press, Beijing, 2010.

[3] Shengwang Wang and Weixing Zheng, *Outline of Real Functions and Functional Analysis* (4th ed.), Volume I, Higher Education Press, Beijing, 2011.

[4] Minyou Qi, *Revisiting Calculus*, Higher Education Press, Beijing, 2004.

[5] Elias M. Stein, *Real Analysis*, Princeton University Press, 2005.

[6] P. R. Halmos, *Measure Theory*, Springer-Verlag, New York, 1974.

Index