MATH 303 — Measures and Integration
Lecture Notes, Fall 2024

Ethan Ackelsberg

Contents

Part 1

Motivation and Basics of Abstract Measure Theory

CHAPTER 1

Motivating Problems of Measure Theory

1. The Problem of Measurement

A basic (and very old) problem in mathematics is to compute the size (length, area, volume) of geometric objects. Areas of polygons and circles can be computed by elementary methods. More complicated regions bounded by continuous curves can be attacked with methods from calculus. But what about more general subsets of Euclidean space? Does it always make sense to talk about the (hyper-)volume of a subset of $\mathbb{R}^d$? What properties does volume have, and how do we compute it?

We will consider these general questions as the "problem of measurement" in Euclidean space and discuss some approaches to a solution.

2. Riemann Integration and Jordan Content

A good first attempt at solving the problem of measurement comes from the Riemann theory of integration. The basic strategy is to approximate general regions by finite collections of boxes (sets of the form $B = \prod_{i=1}^{d}[a_i, b_i]$). For such a box B, we declare the volume to be $\mathrm{Vol}(B) = \prod_{i=1}^{d}(b_i - a_i)$ and use this to define the volume of more general regions. We will now make this idea rigorous.

Definition 1.1: Darboux Integration

Let $B = \prod_{i=1}^{d}[a_i, b_i]$ be a box in $\mathbb{R}^d$, and let $f : B \to \mathbb{R}$ be a bounded function.

- A *Darboux partition* of B is a family of finite sequences $(x_{i,j})_{1 \le i \le d, 0 \le j \le n_i}$ such that $a_i = x_{i,0} < x_{i,1} < \cdots < x_{i,n_i} = b_i$ for each $i \in \{1, \ldots, d\}$.

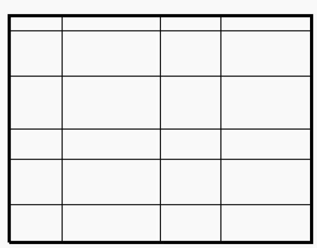


Figure 1.1. A Darboux partition in dimension $d = 2$ with $n_1 = 4$ and $n_2 = 6$.

- Given a Darboux partition $P = (x_{i,j})_{1 \le i \le d, 0 \le j \le n_i}$ of B, the *upper* and *lower Darboux sums of f over B* are given by

$$U_B(f, P) = \sum_{\mathbf{j} \in \prod_{i=1}^{d}\{1,\ldots,n_i\}} \sup_{\boldsymbol{x} \in B_{\mathbf{j}}} f(\boldsymbol{x}) \cdot \mathrm{Vol}(B_{\mathbf{j}})$$

and

$$L_B(f,P) = \sum_{\mathbf{j}\in\prod_{i=1}^d\{1,\ldots,n_i\}} \inf_{\boldsymbol{x}\in B_{\mathbf{j}}} f(\boldsymbol{x}) \cdot \mathrm{Vol}(B_{\mathbf{j}}),$$

where $B_{\mathbf{j}}$ is the box $\prod_{i=1}^d [x_{i,j_i-1}, x_{i,j_i}]$, and $\mathrm{Vol}(B_{\mathbf{j}}) = \prod_{i=1}^d (x_{i,j_i} - x_{i,j_i-1})$ is the volume of $B_{\mathbf{j}}$.

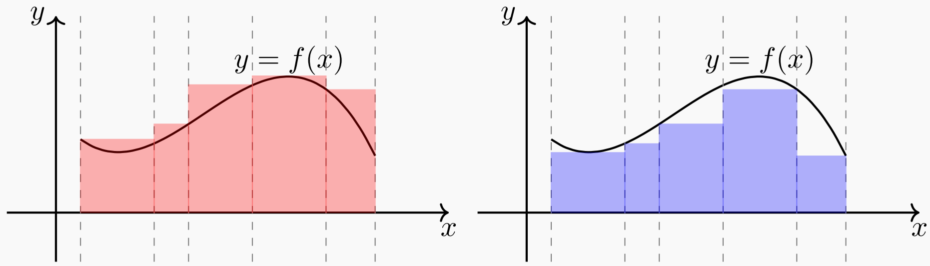


FIGURE 1.2. Upper (red) and lower (blue) Darboux sums of a function f over an interval ($d = 1$).

- The *upper* and *lower Darboux integral of f over B* are

$$U_B(f) = \inf\{U_B(f,P) : P \text{ is a Darboux partition of } B\}$$

and

$$L_B(f) = \sup\{L_B(f,P) : P \text{ is a Darboux partition of } B\}.$$

- The function f is *Darboux integrable over B* if $U_B(f) = L_B(f)$, and their common value is called the *Darboux integral of f over B* and is denoted by $\int_B f(\boldsymbol{x})\, d\boldsymbol{x}$.

PROPOSITION 1.2

A function f is Darboux integrable if and only if it is Riemann integrable. Moreover, the value of the Darboux integral and the Riemann integral (for a Riemann–Darboux integrable function) are the same.

DEFINITION 1.3

A bounded set $E \subseteq \mathbb{R}^d$ is a *Jordan measurable set* if $\mathbb{1}_E$ is Riemann–Darboux integrable over a box containing E. The *Jordan content* of a Jordan measurable set E is the value $J(E) = \int_B \mathbb{1}_E(\boldsymbol{x})\, d\boldsymbol{x}$, where B is any box containing E.

Jordan measurable sets include basic geometric objects such as polyhedra, conic sections, regions bounded by finitely many smooth curves/surfaces, etc.

DEFINITION 1.4

A set $S \subseteq \mathbb{R}^d$ is a *simple set* if it is a finite union of boxes $S = \bigcup_{j=1}^k B_j$.

If the boxes $B_1, \dots, B_k$ are disjoint, then the volume of the simple set $S = \bigcup_{j=1}^k B_j$ is $\mathrm{Vol}(S) = \sum_{j=1}^k \mathrm{Vol}(B_j)$. If some of the boxes intersect, then the volume of $S = \bigcup_{j=1}^k B_j$ can be computed using inclusion-exclusion:

$$\mathrm{Vol}(S) = \sum_{j=1}^k \mathrm{Vol}(B_j) - \sum_{1\le j_1<j_2\le k} \mathrm{Vol}(B_{j_1}\cap B_{j_2}) + \sum_{1\le j_1<j_2<j_3\le k} \mathrm{Vol}(B_{j_1}\cap B_{j_2}\cap B_{j_3}) - \dots$$

This expression is well-defined, since the intersection of two boxes is again a box. A Jordan measurable set is a set that is "well-approximated" by simple sets, as we will make precise now.

Definition 1.5

For a bounded set $E \subseteq \mathbb{R}^d$, define the *inner* and *outer Jordan content* by

$$J_*(E) = \sup\{\mathrm{Vol}(S) : S \subseteq E \text{ is a simple set}\}.$$

and

$$J^*(E) = \inf\{\mathrm{Vol}(S) : S \supseteq E \text{ is a simple set}\}.$$

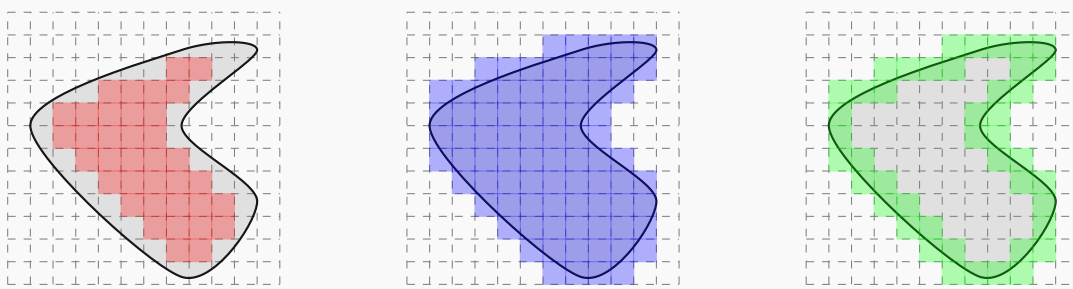


Figure 1.3. Simple sets approximating the inner (red) and outer Jordan content (blue) of a region in dimension $d = 2$. With the red boxes removed from the blue, we get a simple set covering the boundary (in green).

Theorem 1.6

Let $E \subseteq \mathbb{R}^d$ be a bounded set. The following are equivalent:

(i) E is Jordan measurable;
(ii) $J_*(E) = J^*(E)$ (in which case $J(E)$ is equal to this same value);
(iii) $J^*(\partial E) = 0$.

Proof. We will prove the $d = 1$ case. The multidimensional case is similar but more notationally cumbersome, so we omit it to avoid additional technical details that would largely obscure the main ideas.

(i) $\iff$ (ii). To establish this equivalence, it suffices to show

$$U_B(\mathbb{1}_E) = J^*(E) \qquad \text{and} \qquad L_B(\mathbb{1}_E) = J_*(E)$$

for any box (interval) $B \supseteq E$. Let us prove $U_B(\mathbb{1}_E) = J^*(E)$.

CLAIM 1. $U_B(\mathbb{1}_E) \leq J^*(E)$.

Let $\varepsilon > 0$. Then from the definition of the outer Jordan content, there exists a simple set $S \subseteq \mathbb{R}$ such that $E \subseteq S$ and $\text{Vol}(S) < J^*(E) + \varepsilon$. By assumption, B is an interval containing E, so $S \cap B$ is also a simple set containing E, and $\text{Vol}(S \cap B) \leq \text{Vol}(S) < J^*(E) + \varepsilon$. We may therefore assume without loss of generality that $S \subseteq B$. Write $B = [a, b]$ and $S = [a_1, b_1] \sqcup [a_2, b_2] \sqcup \cdots \sqcup [a_n, b_n]$ with $a \leq a_1 \leq b_1 < a_2 \leq b_2 < \cdots < a_n \leq b_n \leq b$. We define a Darboux partition[a] P of $[a, b]$ by $P = (x_i)_{i=0}^{2n+1}$ with $x_0 = a$, $x_1 = a_1$, $x_2 = b_1$, ..., $x_{2n-1} = a_n$, $x_{2n} = b_n$, $x_{2n+1} = b$. Then since $E \subseteq S$, we have

$$\begin{aligned} U_B(\mathbb{1}_E, P) &= \sum_{i=1}^{2n+1} \sup_{x_{i-1} \leq x \leq x_i} \mathbb{1}_E(x) \cdot (x_i - x_{i-1}) \\ &\leq 0 \cdot (a_1 - a) + 1 \cdot (b_1 - a_1) + 0 \cdot (a_2 - b_1) + \cdots + 1 \cdot (b_n - a_n) + 0 \cdot (b - b_n) \\ &= \text{Vol}(S). \end{aligned}$$

Hence, $U_B(\mathbb{1}_E) \leq U_B(\mathbb{1}_E, P) \leq \text{Vol}(S) < J^*(E) + \varepsilon$. This proves the claim.

[a] Strictly speaking, this may fail to be a Darboux partition, since some of the points are allowed to coincide. However, the value we compute for $U_B(\mathbb{1}_E, P)$ will be the correct value for the partition where we remove repetitions of the same point.

CLAIM 2. $J^*(E) \leq U_B(\mathbb{1}_E)$.

Let $\varepsilon > 0$. Write $B = [a, b]$. Then there exists a Darboux partition $a = x_0 < x_1 < \cdots < x_n = b$ such that $U_B(\mathbb{1}_E, P) < U_B(\mathbb{1}_E) + \varepsilon$. Let $M_i = \sup_{x_{i-1} \leq x \leq x_i} \mathbb{1}_E(x) \in \{0, 1\}$, and note that, by definition, $U_B(\mathbb{1}_E, P) = \sum_{i=1}^n M_i(x_i - x_{i-1})$. Let $I \subseteq \{1, \ldots, n\}$ be the set $I = \{1 \leq i \leq n : M_i = 1\}$, and let $S = \bigcup_{i \in I}[x_{i-1}, x_i]$. Then S is a simple set with length $\text{Vol}(S) = \sum_{i \in I}(x_i - x_{i-1}) = U_B(\mathbb{1}_E, P)$. Moreover, $E \subseteq S$, since S is the union of all intervals that have nonempty intersection with E. Thus, $J^*(E) \leq \text{Vol}(S) = U_B(\mathbb{1}_E, P) < U_B(\mathbb{1}_E) + \varepsilon$.

The identity $L_B(\mathbb{1}_E) = J_*(E)$ is proved similarly.

(ii) $\iff$ **(iii)**. It suffices to prove $J^*(\partial E) = J^*(E) - J_*(E)$. (See Figure 1.3.)

CLAIM 3. $J^*(\partial E) \leq J^*(E) - J_*(E)$.

Let $\varepsilon > 0$. Let S_1 be a simple set such that $E \subseteq S_1$ and $\text{Vol}(S_1) < J^*(E) + \frac{\varepsilon}{2}$. Since S_1 is closed, we have $\overline{E} \subseteq S_1$. Let S_2 be a simple set with $S_2 \subseteq E$ such that $\text{Vol}(S_2) > J_*(E) - \frac{\varepsilon}{2}$. Note that $\text{int}(S_2) \subseteq \text{int}(E)$. Therefore, $S = S_1 \setminus \text{int}(S_2)$ is a simple set and $\partial E = \overline{E} \setminus \text{int}(E) \subseteq S$, so $J^*(\partial E) \leq \text{Vol}(S) = \text{Vol}(S_2) - \text{Vol}(S_1) < J^*(E) - J_*(E) + \varepsilon$. But ε was arbitrary, so we conclude $J^*(\partial E) \leq J^*(E) - J_*(E)$.

CLAIM 4. $J^*(E) - J_*(E) \leq J^*(\partial E)$.

Let $\varepsilon > 0$, and let $S \supseteq \partial E$ be a simple set with $\mathrm{Vol}(S) < J^*(\partial E) + \frac{\varepsilon}{2}$. Write $S = \bigsqcup_{i=1}^{n} [a_i, b_i]$ with $a_1 \leq b_1 < a_2 \leq b_2 < \cdots < a_n \leq b_n$. Let $[a, b] \subseteq \mathbb{R}$ such that $E \subseteq [a, b]$ and $a < a_1$ and $b < b_n$. For notational convenience, let $b_0 = a$ and $a_{n+1} = b$. Let $I \subseteq \{0, \dots, n\}$ be the collection of indices i such that $(b_i, a_{i+1}) \cap E \neq \emptyset$. For each $i \in I$, we claim that $(b_i, a_{i+1}) \subseteq E$. If not, then (b_i, a_{i+1}) contains a boundary point of E, but $\partial E \subseteq S$, so this is a contradiction. Thus, $S' = \bigcup_{i \in I} [b_i, a_{i+1}]$ is a simple set with $\mathrm{int}(S') \subseteq E$. Shrinking slightly each interval in S', we obtain a simple set

$$S'' = \bigcup_{i \in I} \left[b_i + \frac{\varepsilon}{4(n+1)}, a_{i+1} - \frac{\varepsilon}{4(n+1)} \right]$$

such that $S'' \subseteq E$. Moreover, $\mathrm{Vol}(S'') \geq \mathrm{Vol}(S') - \frac{\varepsilon}{2(n+1)} |I| \geq \mathrm{Vol}(S') - \frac{\varepsilon}{2}$. Noting that $S \cup S'$ is a simple set containing E, we arrive at the inequality

$$J^*(E) - J_*(E) \leq \mathrm{Vol}(S \cup S') - \mathrm{Vol}(S'') = \mathrm{Vol}(S) + \mathrm{Vol}(S') - \mathrm{Vol}(S'') < J^*(\partial E) + \varepsilon.$$

This completes the proof of Theorem 1.6. □

Example 1.7

The sets $\mathbb{Q} \cap [0, 1]$ and $[0, 1] \setminus \mathbb{Q}$ are not Jordan measurable (see Exercise 1.1).

In addition to the above example, there are many other "nice" sets that are not Jordan measurable. There are, for instance, bounded open sets in $\mathbb{R}$ that are not Jordan measurable. We will work out one such example in detail.

Example 1.8

The complement U of the fat Cantor set (also known as the Smith–Volterra–Cantor set) $K \subseteq [0, 1]$ is Jordan non-measurable. We construct K iteratively, starting from $[0, 1]$, by removing intervals of length 4^{-n} at step n. In other words, at step n, we remove an interval of length 4^{-n} around each rational point with denominator 2^n.

Figure 1.4. Iterative construction of the fat Cantor set.

Let

$$U = \bigcup_{n=0}^{\infty} \bigcup_{j=1}^{2^n} \left(\frac{2j+1}{2^{n+1}} - \frac{1}{2 \cdot 4^{n+1}}, \frac{2j+1}{2^{n+1}} + \frac{1}{2 \cdot 4^{n+1}} \right).$$

Then $K = [0, 1] \setminus U$. The inner Jordan content of U is

$$J_*(U) = \sum_{n=0}^{\infty} \sum_{j=1}^{2^n} \mathrm{Len} \left(\frac{2j+1}{2^{n+1}} - \frac{1}{2 \cdot 4^{n+1}}, \frac{2j+1}{2^{n+1}} + \frac{1}{2 \cdot 4^{n+1}} \right) = \sum_{n=0}^{\infty} 2^n \cdot \frac{1}{4^{n+1}} = \frac{1}{4} \sum_{n=0}^{\infty} 2^{-n} = \frac{1}{2}.$$

However, $\overline{U} = [0, 1]$ (since U contains every rational number whose denominator is a power of 2), so the outer Jordan content of U is $J^*(U) = J^*([0, 1]) = 1$.

3. Limits of Integrable Functions

You may recall from the theory of Riemann integration that *uniform* limits of Riemann integrable functions are Riemann integrable, and one may in this case interchange the order of taking limits and computing the integral. More precisely:

Theorem 1.9

Let B be a box in $\mathbb{R}^d$. Let $(f_n)_{n\in\mathbb{N}}$ be a sequence of Riemann integrable functions on B, and suppose f_n converges uniformly to a function $f : B \to \mathbb{R}$. Then f is Riemann integrable, and

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

One of the deficiencies of the Riemann–Darboux–Jordan approach to integration and measurement is that pointwise (non-uniform) limits do not share this property.

Example 1.10

Enumerate the set $\mathbb{Q} \cap [0,1] = \{q_1, q_2, \dots\}$. Let $f_n : [0,1] \to [0,1]$ be the function

$$f_n(x) = \begin{cases} 1, & \text{if } x \in \{q_1, \dots, q_n\} \\ 0, & \text{otherwise.} \end{cases}$$

Then f_n is Riemann integrable and $f_n \to \mathbb{1}_{\mathbb{Q}\cap[0,1]}$ pointwise, but $\mathbb{1}_{\mathbb{Q}\cap[0,1]}$ is not Riemann integrable.

Since analysis so often deals with limits, it is desirable to develop a theory of integration that accommodates pointwise limits. The Lebesgue measure and Lebesgue integral resolve this shortcoming.

4. The Solution of Lebesgue

The Jordan non-measurable set in Example 1.8 appears to have a sensible notion of "length." Indeed, the complement U, being a disjoint union of intervals, could be reasonably assigned as a "length" the sum of the lengths of the (countably many) intervals of which it is made. This produces a value of $\frac{1}{2}$ for the length of U, and so we should take K to also have length $\frac{1}{2}$, since $K \sqcup U = [0,1]$ is an interval of length 1. The feature that U is a disjoint union of intervals turns out to not be any special feature of U at all but instead a general feature of open sets in $\mathbb{R}$.

Proposition 1.11

Let $U \subseteq \mathbb{R}$ be an open set. Then U can be expressed as a countable disjoint union of open intervals.

Proof. Exercise 1.2. □

By Proposition 1.11, it seems reasonable to define the length of an open set $U \subseteq \mathbb{R}$ as follows. Write $U = (a_1, b_1) \sqcup (a_2, b_2) \sqcup \dots$ as a disjoint union of open intervals, and define its length as $(b_1 - a_1) + (b_2 - a_2) + \dots$. Then open sets may play the role that simple sets played in the definition of the Jordan content, and this leads to the Lebesgue measure.

Remark. In higher dimensions, Proposition 1.11 needs to be modified, but one can still reasonably talk about the d-dimensional volume of open sets in $\mathbb{R}^d$. See Exercises 1.3 and 1.6.

Definition 1.12

Let $E \subseteq \mathbb{R}^d$.

- The *outer Lebesgue measure of E* is the quantity
$$\begin{aligned}\lambda^*(E) &= \inf\{\mathrm{Vol}(U) : U \supseteq E \text{ is open}\}\\ &= \inf\left\{\sum_{j=1}^{\infty}\mathrm{Vol}(B_j) : B_1, B_2, \dots \text{ are boxes, and } E \subseteq \bigcup_{j=1}^{\infty} B_j\right\}.\end{aligned}$$
- The set E is *Lebesgue measurable* (with *Lebesgue measure* $\lambda(E) = \lambda^*(E)$) if for every $\varepsilon > 0$, there exists an open set $U \subseteq \mathbb{R}^d$ such that $E \subseteq U$ and $\lambda^*(U \setminus E) < \varepsilon$.

Proposition 1.13

If $E \subseteq \mathbb{R}^d$ is Jordan measurable, then E is Lebesgue measurable and $J(E) = \lambda(E)$.

The family of Lebesgue measurable sets is much larger than the family of Jordan measurable sets. Among the several nice properties of the Lebesgue measure (and abstract measures) that we will see later in the course are:

Proposition 1.14

(1) If $(E_n)_{n\in\mathbb{N}}$ are Lebesgue measurable sets, then $\bigcup_{n=1}^{\infty} E_n$ and $\bigcap_{n=1}^{\infty} E_n$ are Lebesgue measurable.
(2) If $(E_n)_{n\in\mathbb{N}}$ are pairwise disjoint and Lebesgue measurable, then $\lambda\left(\bigsqcup_{n=1}^{\infty} E_n\right) = \sum_{n=1}^{\infty}\lambda(E_n)$.
(3) If $E_1 \subseteq E_2 \subseteq \cdots \subseteq \mathbb{R}^d$ are Lebesgue measurable sets, then $\lambda\left(\bigcup_{n=1}^{\infty} E_n\right) = \lim_{n\to\infty}\lambda(E_n)$.
(4) If $E_1 \supseteq E_2 \supseteq \dots$ are Lebesgue measurable subsets of $\mathbb{R}^d$ and $\lambda(E_1) < \infty$, then $\lambda\left(\bigcap_{n=1}^{\infty} E_n\right) = \lim_{n\to\infty}\lambda(E_n)$.

5. Applications of Abstract Measure Theory

The mathematical language and tools encompassed in measure theory play a foundational role in many other areas of mathematics. A highly abbreviated sampling follows.

Probability Theory. Measure theory provides the axiomatic foundations of probability theory, providing rigorous notions of *random variables* and *probabilities* of *events*. Important limit laws (the law of large numbers and central limit theorem, for example) are phrased mathematically using measure-theoretic notions of convergence.

Fourier Analysis. Periodic (say, continuous or Riemann-integrable) functions on the real line have corresponding Fourier series representations $f(x) \sim \sum_{n\in\mathbb{Z}} \widehat{f}(n)e^{2\pi inx}$. The functions $e^{2\pi inx}$ are orthonormal, and Parseval's identity gives $\sum_{n\in\mathbb{Z}} |\widehat{f}(n)|^2 = \int_0^1 |f(x)|^2\,dx$. Given a sequence $(a_n)_{n\in\mathbb{N}}$, one may ask whether $\sum_{n\in\mathbb{Z}} a_n e^{2\pi inx}$ is the Fourier expansion of some function f, and if so, what properties does f have? Another natural question is whether the series $\sum_{n\in\mathbb{Z}} \widehat{f}(n)e^{2\pi inx}$ actually converges to the function f, and if so, in which sense? Both of these questions are properly answered in a measure-theoretic framework. If one is interested in decomposing functions defined on other groups (for instance, on compact abelian groups) into their Fourier series, then one also needs to develop a method of integrating functions on groups in order to compute Fourier coefficients and make sense of Parseval's identity.

FUNCTIONAL ANALYSIS AND OPERATOR THEORY. When one studies familiar concepts from linear algebra in infinite-dimensional spaces, measures become unavoidable for many tasks. For example, versions of the spectral theorem (generalizing the representation of suitable matrices in terms of their eigenvalues and eigenvectors) for operators on infinite-dimensional spaces require the abstract notion of a measure.

ERGODIC THEORY. Ergodic theory was developed to study the long-term statistical behavior of dynamical (time-dependent) systems, providing a framework to resolve important problems in physics related to the "ergodic hypothesis" in thermodynamics and the "stability" of the solar system. It turns out that the appropriate mathematical formalism for understanding these problems comes from abstract measure theory.

FRACTAL GEOMETRY. Self-similar geometric objects such as the Koch snowflake, Sierpiński carpet, and the middle-thirds Cantor set (see Figure 1.5) can be meaningfully assigned a notion of "dimension" that can take a non-integer value. How does one determine the dimension of a fractal object? There are several different approaches to dimension, but one of the most popular is the *Hausdorff dimension*, which relies on a family of measures that interpolate between the integer-dimensional Lebesgue measures.

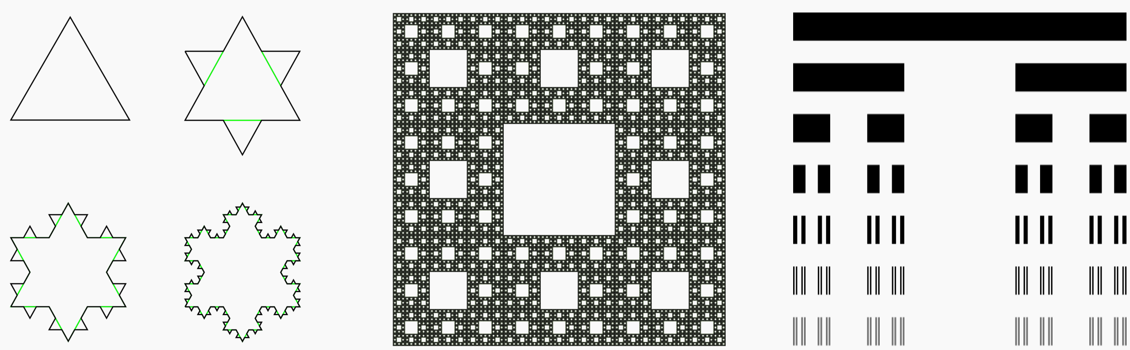


FIGURE 1.5. Fractal shapes: the Koch snowflake (left) of Hausdorff dimension $\frac{\log 4}{\log 3} \approx 1.26$, Sierpiński carpet (middle) of dimension $\frac{\log 8}{\log 3} \approx 1.89$, and middle-thirds Cantor set (right) of dimension $\frac{\log 2}{\log 3} \approx 0.63$.

Additional Reading

This introductory chapter is heavily influenced by the book of Tao [7] on measure theory. Many of the results in this chapter are discussed in greater detail in [7, Section 1.1].

Exercises

1.1 Show that $J^*(\mathbb{Q} \cap [0,1]) = J^*([0,1] \setminus \mathbb{Q}) = 1$, and $J_*(\mathbb{Q} \cap [0,1]) = J_*([0,1] \setminus \mathbb{Q}) = 0$.

1.2 Let $U \subseteq \mathbb{R}$ be an open set. Show that U can be written as a disjoint union of countably many open intervals.

1.3 Let $U = \{(x,y) : x^2 + y^2 < 1\} \subseteq \mathbb{R}^2$ be the open unit disk. Show that U cannot be expressed as a disjoint union of countably many open boxes.

1.4 Give an example to show that the statement

$$\lambda^*(E) = \sup_{U \subseteq E, U \text{ open}} \lambda^*(U).$$

is false.

1.5 (Area Interpretation of the Riemann Integral) Let $[a, b]$ be an interval and $f : [a, b] \to [0, \infty)$ a bounded function. Show that f is Riemann integrable if and only if the set

$$E_+ = \{(x, t) : a \leq x \leq b, 0 \leq t \leq f(x)\}$$

is a Jordan measurable set in $\mathbb{R}^2$, in which case

$$\int_a^b f(x)\ dx = J(E_+).$$

1.6 Let $U \subseteq \mathbb{R}^d$ be an open set. Show that U can be written as a disjoint union of countably many half-open boxes (i.e., sets of the form $B = \prod_{i=1}^d [a_i, b_i)$).

CHAPTER 2

Measure Spaces

1. σ-Algebras

Before defining measures, we must determine which subsets of a given set X we would like to be able to measure. The full set X should be measurable, and we should allow ourselves to perform the basic set-theoretic operations (complements, unions, and intersections). Allowing *finite* unions and intersections produces an *algebra* of sets. Algebras are a very useful notion, but (as with the Jordan content discussed in the previous chapter) they are insufficient for appropriately handling limits. We will therefore upgrade from algebras to *σ-algebras*:

Definition 2.1

Let X be a set. A *σ-algebra* on X is a family $\mathcal{B} \subseteq \mathscr{P}(X)$ of subsets of X with the following properties:

- $X \in \mathcal{B}$;
- If $B \in \mathcal{B}$, then $X \setminus B \in \mathcal{B}$;
- If $(B_n)_{n\in\mathbb{N}}$ is a countable family of elements of $\mathcal{B}$, then $\bigcup_{n\in\mathbb{N}} B_n \in \mathcal{B}$.

Remark. In the definition of a σ-algebra, we have made no explicit mention of intersections. However, by De Morgan's laws, we can also generate the countable intersection of sets: $\bigcap_{n\in\mathbb{N}} B_n = X \setminus \big(\bigcup_{n\in\mathbb{N}}(X \setminus B_n)\big)$.

Example 2.2

Some examples of σ-algebras include the following:

- For any set X, the power set $\mathscr{P}(X)$ is a σ-algebra, as is the pair $\{\emptyset, X\}$.
- The family $\mathcal{B} = \{B \subseteq \mathbb{R} : \text{either } B \text{ or } \mathbb{R} \setminus B \text{ is countable}\}$ of countable and co-countable subsets of $\mathbb{R}$ is a σ-algebra.
- Unions of unit-length intervals in $\mathbb{R}$ form a σ-algebra $\mathcal{B} = \big\{\bigcup_{n\in S}[n, n+1) : S \subseteq \mathbb{Z}\big\}$.

Proposition 2.3

Suppose $(\mathcal{B}_i)_{i\in I}$ is a family of σ-algebras on X. Then $\bigcap_{i\in I} \mathcal{B}_i$ is a σ-algebra.

Proof. Let $\mathcal{B} = \bigcap_{i\in I} \mathcal{B}_i$.

For every $i \in I$, we have $X \in \mathcal{B}_i$, so $X \in \mathcal{B}$.

Suppose $B \in \mathcal{B}$. Then $B \in \mathcal{B}_i$ for every $i \in I$, so $X \setminus B \in \mathcal{B}_i$ for every $i \in I$. Hence, $X \setminus B \in \mathcal{B}$.

Let $(B_n)_{n\in\mathbb{N}}$ be a countable family of sets in $\mathcal{B}$. For each $i \in I$, the sets $(B_n)_{n\in\mathbb{N}}$ belong to $\mathcal{B}_i$, so $\bigcup_{n\in\mathbb{N}} B_n \in \mathcal{B}_i$. Therefore, $\bigcup_{n\in\mathbb{N}} B_n \in \mathcal{B}$. $\square$

Definition 2.4

The *σ-algebra generated by a family* $\mathcal{S} \subseteq \mathscr{P}(X)$ is the smallest σ-algebra containing $\mathcal{S}$, denoted by $\sigma(\mathcal{S})$.

Remark. Note that $\sigma(\mathcal{S})$ is well-defined by Proposition 2.3:

$$\sigma(\mathcal{S}) = \bigcap \{\mathcal{B} : \mathcal{B} \text{ is a } \sigma\text{-algebra on } X, \mathcal{S} \subseteq \mathcal{B}\}.$$

In topological spaces (such as the real line), we will often consider the σ-algebra generated by the topology.

Definition 2.5

Let (X, τ) be a topological space. The *Borel σ-algebra* is the σ-algebra generated by the open subsets of X, i.e. $\mathrm{Borel}(X) = \sigma(\tau)$.

Borel sets can be placed in a hierarchy in terms of their level of complexity. At the simplest level are the open (G) and closed (F) sets. Next come countable intersections of open sets (G_δ sets) and countable unions of closed sets (F_σ sets) and so on.

$$\begin{array}{ccccccccccccc}
\Sigma^0_1 = G & \xrightarrow{\bigcap} & \Pi^0_2 = G_\delta & \xrightarrow{\bigcup} & \Sigma^0_3 = G_{\delta\sigma} & \xrightarrow{\bigcap} & \Pi^0_4 = G_{\delta\sigma\delta} & \xrightarrow{\bigcup} & \cdots & \xrightarrow{\bigcup} & \Sigma^0_\omega & \xrightarrow{\bigcap} & \cdots \\
\updownarrow c & & \updownarrow c & & \updownarrow c & & \updownarrow c & & & & \updownarrow c & & \\
\Pi^0_1 = F & \xrightarrow{\bigcup} & \Sigma^0_2 = F_\sigma & \xrightarrow{\bigcap} & \Pi^0_3 = F_{\sigma\delta} & \xrightarrow{\bigcup} & \Sigma^0_4 = F_{\sigma\delta\sigma} & \xrightarrow{\bigcap} & \cdots & \xrightarrow{\bigcap} & \Pi^0_\omega & \xrightarrow{\bigcup} & \cdots
\end{array}$$

Figure 2.1. The Borel hierarchy for subsets of a topological space.

The placement of a (Borel) set within the Borel hierarchy is a useful notion of "complexity" for sets. Intuitively speaking, if a set is lower down in the Borel hierarchy, then it is in some sense easier to define than a set higher up the hierarchy. Determining where sets occur in the Borel hierarchy (or if they are Borel at all) is a common theme in an area of mathematical logic known as *descriptive set theory*. We will largely not concern ourselves with such problems in this course, but some suggested additional reading appears at the end of this chapter for those who are interested.

In our development of the abstract theory of measures (where we may not even have a topology to work with), our object of study will be arbitrary sets X equipped with a σ-algebra.

Definition 2.6

A *measurable space* is a pair $(X, \mathcal{B})$, where X is a set and $\mathcal{B}$ is a σ-algebra on X. Elements of the σ-algebra $\mathcal{B}$ are called *measurable sets*.

2. Measurable Functions

Recall that a function $f : X \to Y$ from one topological space to another is continuous if the preimage of every open set in Y is open in X. Measurable functions are defined analogously, but with "open" replaced by "measurable."

Definition 2.7

Let $(X, \mathcal{B})$ and $(Y, \mathcal{C})$ be measurable spaces. A function $f : X \to Y$ is *measurable* if for every $C \in \mathcal{C}$, one has $f^{-1}(C) \in \mathcal{B}$.

Some basic properties of measurable functions that will be used frequently are as follows:

Proposition 2.8

(1) Let $(X, \mathcal{B})$, $(Y, \mathcal{C})$, and $(Z, \mathcal{D})$ be measurable spaces. Let $f : X \to Y$ and $g : Y \to Z$ be measurable functions. Then $g \circ f : X \to Z$ is measurable.
(2) Let $(X, \mathcal{B})$ and $(Y, \mathcal{C})$ be measurable spaces, and let $f : X \to Y$. Suppose $\mathcal{S} \subseteq \mathscr{P}(Y)$ is a family of sets such that $\sigma(\mathcal{S}) = \mathcal{C}$. If $f^{-1}(S) \in \mathcal{B}$ for every $S \in \mathcal{S}$, then f is a measurable function.
(3) Suppose X and Y are topological spaces and $\mathcal{B} = \text{Borel}(X)$ and $\mathcal{C} = \text{Borel}(Y)$ are the Borel σ-algebras on X and Y respectively. Then every continuous function $f : X \to Y$ is measurable.

Proof. **(1)** Let $D \in \mathcal{D}$. Since g is measurable, we have $C = g^{-1}(D) \in \mathcal{C}$. Then since f is measurable, $B = f^{-1}(C) \in \mathcal{B}$. But $B = f^{-1}(g^{-1}(D)) = (g \circ f)^{-1}(D)$, so $g \circ f$ is measurable.

(2) Let $\mathcal{F} = \{E \subseteq Y : f^{-1}(E) \in \mathcal{B}\}$. We claim that $\mathcal{F}$ is a σ-algebra. Then since $\mathcal{S} \subseteq \mathcal{F}$, we conclude that $\mathcal{C} = \sigma(\mathcal{S}) \subseteq \mathcal{F}$, so f is measurable. Let us now prove the claim:

- $f^{-1}(Y) = X \in \mathcal{B}$, so $Y \in \mathcal{F}$.
- Suppose $E \in \mathcal{F}$. Then $f^{-1}(Y \setminus E) = X \setminus \underbrace{f^{-1}(E)}_{\in \mathcal{B}} \in \mathcal{B}$, so $Y \setminus E \in \mathcal{F}$.
- Suppose $E_1, E_2, \dots \in \mathcal{F}$, and let $E = \bigcup_{n \in \mathbb{N}} E_n$. Then
$$f^{-1}(E) = \bigcup_{n \in \mathbb{N}} \underbrace{f^{-1}(E_n)}_{\in \mathcal{B}} \in \mathcal{B},$$
so $E \in \mathcal{F}$.

This proves that $\mathcal{F}$ is a σ-algebra on Y.

(3) This follows from (1) by taking $\mathcal{S}$ to be the collection of open sets in Y. $\square$

3. The Extended Real Numbers and Extended Real-Valued Functions

One obtains an important class of measurable functions when one considers functions defined on a measurable space taking real values. For many applications and in order to account more fully for limits of functions, it is often convenient to work with the slightly more general concept of *extended* real-valued functions.

Definition 2.9

The *extended real numbers* are the set $[-\infty, \infty] = \mathbb{R} \cup \{\infty, -\infty\}$ with the following topological and algebraic properties:

- The topology on $[-\infty, \infty]$ is generated by open intervals (a, b) with $a, b \in \mathbb{R}$ and sets of the form $(a, \infty] = (a, \infty) \cup \{\infty\}$ and $[-\infty, b) = (-\infty, b) \cup \{-\infty\}$ for $a, b \in \mathbb{R}$.

- Addition is extended as a commutative operation with $\infty + x = \infty$ and $-\infty + x = -\infty$ for real numbers $x \in \mathbb{R}$. For addition of two infinite quantities, we define $\infty + \infty = \infty$ and $-\infty + (-\infty) = -\infty$. However, $-\infty + \infty$ is undefined.
- Multiplication is also extended as a commutative operation with the properties

$$x \in (0, \infty) \implies \infty \cdot x = \infty \quad \text{and} \quad -\infty \cdot x = -\infty;$$
$$x \in (-\infty, 0) \implies \infty \cdot x = -\infty \quad \text{and} \quad -\infty \cdot x = \infty.$$

By convention, we define $\infty \cdot 0 = -\infty \cdot 0 = 0$. Multiplication of infinities is defined by $\infty \cdot \infty = (-\infty) \cdot (-\infty) = \infty$, and $-\infty \cdot \infty = -\infty$.

The topology we have defined on $[-\infty, \infty]$ is the *two-point compactification* of $\mathbb{R}$. You will check in the exercises (Exercise 2.1) that $[-\infty, \infty]$ is indeed a compact space (that is homeomorphic to a closed interval, say $[0, 1]$). The algebraic operations on $[-\infty, \infty]$ are all as one would expect, with one exception: $\infty \cdot 0$ is often considered as an "indeterminate form", but here we have given it a definite value of 0. The reason for this convention is the following proposition, which you will also prove in the exercises:

Proposition 2.10

Let $(x_n)_{n\in\mathbb{N}}$ be a sequence in $[-\infty, \infty]$, and let $c \in \mathbb{R}$. If $(x_n)_{n\in\mathbb{N}}$ converges to an extended real number, then the sequence $(cx_n)_{n\in\mathbb{N}}$ also converges, and

$$\lim_{n\to\infty} (cx_n) = c \cdot \lim_{n\to\infty} x_n. \tag{2.1}$$

Proof. Exercise 2.2. □

In order to have the desirable property (2.1), one has no choice but to define $\infty \cdot 0 = 0$: by taking the sequence $x_n = n$, we have

$$0 \cdot \infty = 0 \cdot \lim_{n\to\infty} n = \lim_{n\to\infty} (0 \cdot n) = 0.$$

Warning: Property (2.1) does not hold for $c \in \{\infty, -\infty\}$, as can be seen by taking a sequence $(x_n)_{n\in\mathbb{N}}$ that converges to 0.

We say that an extended real-valued function $f : X \to [-\infty, \infty]$ defined on a measurable space $(X, \mathcal{B})$ is *$\mathcal{B}$-measurable* (or simply *measurable*) if it is measurable as a function between the measurable spaces $(X, \mathcal{B})$ and $([-\infty, \infty], \mathrm{Borel}([-\infty, \infty]))$. Since we will always take the same σ-algebra on $[-\infty, \infty]$, we omit explicit reference to the Borel σ-algebra when discussing measurable extended real-valued functions.

Proposition 2.11

Let $(X, \mathcal{B})$ be a measurable space.

(1) Let $f : X \to [-\infty, \infty]$. The following are equivalent:
   (a) f is measurable;
   (b) for every $c \in \mathbb{R}$, $f^{-1}((c, \infty]) \in \mathcal{B}$;
   (c) for every $c \in \mathbb{R}$, $f^{-1}([c, \infty]) \in \mathcal{B}$;
   (d) for every $c \in \mathbb{R}$, $f^{-1}([-\infty, c)) \in \mathcal{B}$;
   (e) for every $c \in \mathbb{R}$, $f^{-1}([-\infty, c]) \in \mathcal{B}$.

(2) Suppose $(f_n)_{n\in\mathbb{N}}$ is a sequence of measurable functions from X to $[-\infty, \infty]$. The following functions are also measurable:

(a) $\sup_{n\in\mathbb{N}} f_n$;
(b) $\inf_{n\in\mathbb{N}} f_n$;
(c) $\limsup_{n\to\infty} f_n$;
(d) $\liminf_{n\to\infty} f_n$.

(3) Suppose $f, g : X \to \mathbb{R}$ are measurable functions. Then $f+g$ and $f\cdot g$ are measurable.

Notation. For convenience, we will often write sets of the form $f^{-1}((c,\infty])$ as $\{f > c\}$ and similarly for $\{f \geq c\}$, $\{f < c\}$, and $\{f \leq c\}$.

Proof of Proposition 2.11. **(1)** By Proposition 2.8(2), it suffices to check that each of the relevant collections of intervals generates the Borel σ-algebra on $[-\infty,\infty]$. Let us show that the collection of intervals $(c,\infty]$ for $c\in\mathbb{R}$ generates the Borel σ-algebra. All of the other proofs are similar, so we omit them.

Let $\mathcal{S} = \{(c,\infty] : c\in\mathbb{R}\}$. Note that every element of $\mathcal{S}$ is open in $[-\infty,\infty]$, so $\sigma(\mathcal{S}) \subseteq \mathrm{Borel}([-\infty,\infty])$. On the other hand, we can write $(a,b] = (a,\infty]\setminus(b,\infty]$ for $a,b\in\mathbb{R}, a<b$. Every open set in $\mathbb{R}$ is a countable (disjoint) union of such intervals, so every open subset of $\mathbb{R}$ is contained in $\sigma(\mathcal{S})$. We obtain the additional open sets in $[-\infty,\infty]$ from the rays $(c,\infty]\in\mathcal{S}$ and

$$[-\infty,c) = \bigcap_{n\in\mathbb{N}}\left[-\infty, c+\frac{1}{n}\right] = \bigcap_{n\in\mathbb{N}}\left([-\infty,\infty]\setminus\left(c+\frac{1}{n},\infty\right]\right)\in\sigma(\mathcal{S}).$$

Thus, $\mathrm{Borel}([-\infty,\infty])\subseteq\sigma(\mathcal{S})$.

(2) We will use (1).

(a) Let $f = \sup_{n\in\mathbb{N}} f_n$. Note that $\{f > c\} = \bigcup_{n\in\mathbb{N}}\{f_n > c\}$. Each of the sets $\{f_n > c\}$ belongs to $\mathcal{B}$, so $\{f > c\}\in\mathcal{B}$.

(b) Similarly to (a), letting $f = \inf_{n\in\mathbb{N}} f_n$, we may express $\{f < c\} = \bigcup_{n\in\mathbb{N}}\underbrace{\{f_n < c\}}_{\in\mathcal{B}}\in\mathcal{B}$.

(c) Recall that $\limsup_{n\to\infty} f_n = \inf_{k\in\mathbb{N}}\sup_{n\geq k} f_n$, so measurability of $\limsup_{n\to\infty} f_n$ follows from (a) and (b).

(d) Similar to (c): $\liminf_{n\to\infty} f_n = \sup_{k\in\mathbb{N}}\inf_{n\geq k} f_n$.

(3) Let $A:\mathbb{R}^2\to\mathbb{R}$ and $M:\mathbb{R}^2\to\mathbb{R}$ be the maps $A(x,y) = x+y$ and $M(x,y) = xy$. Both of the maps A and M are continuous and therefore (Borel) measurable. Moreover, $(f+g)(x) = A(f(x),g(x))$ and $(f\cdot g)(x) = M(f(x),g(x))$. Since the composition of measurable maps is measurable (see Proposition 2.8(1)), it suffices to prove $h : x\mapsto(f(x),g(x))$ is a measurable function from X to $\mathbb{R}^2$. By Proposition 2.8(2), we only need to check preimages of sets generating the Borel σ-algebra on $\mathbb{R}^2$. For convenience, we will take the boxes $[a,b)\times[c,d)$ (the first homework problem was to show that every open set in $\mathbb{R}^2$ is a countable (disjoint) union of such boxes, so they generate the Borel σ-algebra). Observe that

$$h^{-1}([a,b)\times[c,d)) = f^{-1}([a,b))\cap g^{-1}([c,d))\in\mathcal{B},$$

since f and g are measurable, so h is indeed a measurable function. □

Example 2.12

Let $(X,\mathcal{B})$ be a measurable space and $E\subseteq X$. The function $\mathbb{1}_E$ is measurable if and only if $E\in\mathcal{B}$.

4. Measures

We are now prepared to define measures on abstract measurable spaces.

Definition 2.13

Let $(X, \mathcal{B})$ be a measurable space. A *measure* on $(X, \mathcal{B})$ is a function $\mu : \mathcal{B} \to [0, \infty]$ such that

- $\mu(\emptyset) = 0$;
- COUNTABLE ADDITIVITY: for any sequence $(E_n)_{n\in\mathbb{N}}$ of pairwise disjoint elements of $\mathcal{B}$, one has $\mu\left(\bigsqcup_{n\in\mathbb{N}} E_n\right) = \sum_{n\in\mathbb{N}} \mu(E_n)$.

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

Nontrivial examples of measures take some effort to construct, and we will spend significant portions of the course discussing different methods for constructing interesting measures. However, there are a few immediate examples that do not require complicated constructions.

Example 2.14

Examples of measures include:

- For any set X, the *counting measure* is a measure defined on the σ-algebra $\mathscr{P}(X)$ by $\mu(E) = |E|$ if E is a finite set and $\mu(E) = \infty$ if E is an infinite set.
- Given a point $x \in X$, the *Dirac measure* defined on $\mathscr{P}(X)$ is the measure $\delta_x(E) = 1$ if $x \in E$ and $\delta_x(E) = 0$ if $x \notin E$.

We will use the following basic properties of measures frequently throughout this course:

Proposition 2.15

Let $(X, \mathcal{B}, \mu)$ be a measure space.

(1) MONOTONICITY: For any $A, B \in \mathcal{B}$, if $A \subseteq B$, then $\mu(A) \leq \mu(B)$.

(2) COUNTABLE SUB-ADDITIVITY: For any sequence $(E_n)_{n\in\mathbb{N}}$ in $\mathcal{B}$,

$$\mu\left(\bigcup_{n\in\mathbb{N}} E_n\right) \leq \sum_{n\in\mathbb{N}} \mu(E_n).$$

(3) CONTINUITY FROM BELOW: If $E_1 \subseteq E_2 \subseteq \cdots \in \mathcal{B}$, then

$$\mu\left(\bigcup_{n\in\mathbb{N}} E_n\right) = \lim_{n\to\infty} \mu(E_n).$$

(4) CONTINUITY FROM ABOVE: If $E_1 \supseteq E_2 \supseteq \cdots \in \mathcal{B}$ and $\mu(E_1) < \infty$, then

$$\mu\left(\bigcap_{n\in\mathbb{N}} E_n\right) = \lim_{n\to\infty} \mu(E_n).$$

PROOF. **(1)** Write $B = A \sqcup (B \setminus A)$. Then $\mu(B) = \mu(A) + \mu(B \setminus A) \geq \mu(A)$, since μ takes nonnegative values.

(2) Define a new sequence of sets E'_n by $E'_1 = E_1$ and $E'_n = E_n \setminus \bigcup_{j=1}^{n-1} E_j$ for $n \geq 2$. Then the sets $(E'_n)_{n\in\mathbb{N}}$ are pairwise disjoint and satisfy $E'_n \subseteq E_n$ and $\bigsqcup_{n\in\mathbb{N}} E'_n = \bigcup_{n\in\mathbb{N}} E_n$. Therefore,

$$\mu\left(\bigcup_{n\in\mathbb{N}} E_n\right) = \mu\left(\bigsqcup_{n\in\mathbb{N}} E'_n\right) = \sum_{n\in\mathbb{N}} \mu(E'_n) \leq \sum_{n\in\mathbb{N}} \mu(E_n),$$

where in the last step we have applied monotonicity of μ (property (1)).

(3) Let $E'_1 = E_1$ and $E'_n = E_n \setminus E_{n-1}$ for $n \geq 2$. For convenience, we will set $E_0 = \emptyset$ so that we also have $E'_1 = E_1 \setminus E_0$. Then

$$\mu\left(\bigcup_{n\in\mathbb{N}} E_n\right) = \mu\left(\bigsqcup_{n\in\mathbb{N}} E'_n\right) = \sum_{n\in\mathbb{N}} \mu(E'_n) \overset{(*)}{=} \sum_{n\in\mathbb{N}} (\mu(E_n) - \mu(E_{n-1})) \overset{(**)}{=} \lim_{n\to\infty} \mu(E_n).$$

The step $(*)$ uses additivity of μ, and $(**)$ comes from the telescoping of the sum.

(4) Define a new sequence $A_n = E_1 \setminus E_n$. Then $\emptyset = A_1 \subseteq A_2 \subseteq \ldots$, so

$$\mu\left(\bigcup_{n\in\mathbb{N}} A_n\right) = \lim_{n\to\infty} \mu(A_n)$$

by (3). But $\bigcup_{n\in\mathbb{N}} A_n = E_1 \setminus \bigcap_{n\in\mathbb{N}} E_n$, so

$$\mu(E_1) - \mu\left(\bigcap_{n\in\mathbb{N}} E_n\right) = \mu\left(\bigcup_{n\in\mathbb{N}} A_n\right) = \lim_{n\to\infty} \mu(A_n) = \mu(E_1) - \lim_{n\to\infty} \mu(E_n),$$

whence we deduce that (4) holds, since $\mu(E_1) < \infty$. □

EXAMPLE 2.16

Property (4) may fail if $\mu(E_1) = \infty$. Let $X = \mathbb{N}$, $\mathcal{B} = \mathscr{P}(\mathbb{N})$, and let μ be the counting measure. Let $E_n = \{m \in \mathbb{N} : m \geq n\}$. Then $\mu(E_n) = \infty$ for every $n \in \mathbb{N}$, but $\bigcap_{n\in\mathbb{N}} E_n = \emptyset$, so

$$\mu\left(\bigcap_{n\in\mathbb{N}} E_n\right) = 0 \neq \infty = \lim_{n\to\infty} \mu(E_n).$$

Additional Reading

The content of this chapter is common to every text on abstract measure theory, though the order of presentation differs. We have elected to follow more or less the order of presentation from Rudin's *Real and Complex Analysis* [**4**, Chapter 1]. Alternative presentations can be found in [**1**, Sections 1.2, 1.3, and 2.1], and [**7**, Section 1.4].

Introductory texts on measure theory tend not to give much treatment to the Borel hierarchy or other topics in descriptive set theory (and we will also not expand on such topics within these lecture notes). Those interested in learning more can take a look at the book of Kechris [**2**] and/or the lecture notes of Tserunyan [**8**], which draw quite heavily on [**2**].

Exercises

2.1 Prove that the extended real line $[-\infty, \infty]$ is homeomorphic to the closed unit interval $[0, 1]$.

2.2 Prove Proposition 2.10.

2.3 Let X, Y be sets and $f : X \to Y$ any function.

(a) Prove that if $\mathcal{C} \subseteq \mathscr{P}(Y)$ is a σ-algebra on Y, then $\mathcal{B} = \{f^{-1}(C) : C \in \mathcal{C}\}$ is a σ-algebra on X.
(b) Prove that for any family of sets $\mathcal{S} \subseteq \mathscr{P}(Y)$, we have $\sigma(f^{-1}(\mathcal{S})) = f^{-1}(\sigma(\mathcal{S}))$.

2.4 Let $(X, \mathcal{B}, \mu)$ be a finite measure space, and let $\mathcal{A} \subseteq \mathscr{P}(X)$ be an algebra on X such that $\sigma(\mathcal{A}) = \mathcal{B}$. Show that for every $B \in \mathcal{B}$ and every $\varepsilon > 0$, there exists $A \in \mathcal{A}$ such that $\mu(A \triangle B) < \varepsilon$.

2.5 Let X be a set. A family of subsets $\mathcal{S} \subseteq \mathscr{P}(X)$ is a *semi-algebra* if

- $\emptyset, X \in \mathcal{S}$;
- if $A, B \in \mathcal{S}$, then $A \cap B \in \mathcal{S}$;
- if $A, B \in \mathcal{S}$, then $A \setminus B = \bigsqcup_{i=1}^{n} C_i$ for some $C_1, \ldots, C_n \in \mathcal{S}$.

Show that if $\mathcal{S}$ is a semi-algebra, then the algebra generated by $\mathcal{S}$ is

$$\mathcal{A}(\mathcal{S}) = \left\{ \bigcup_{i=1}^{n} A_i : n \in \mathbb{N}, A_i \in \mathcal{S} \right\}.$$

Can $\bigcup$ be replaced by $\bigsqcup$?

2.6 Suppose $\mathcal{B}$ is an infinite σ-algebra (on an infinite set X).
(a) Show that $\mathcal{B}$ contains an infinite sequence $(E_n)_{n \in \mathbb{N}}$ of pairwise disjoint sets.
(b) Deduce that $\mathcal{B}$ has at least the cardinality of the continuum.

2.7 Prove that the following sets are Borel sets in $\mathbb{R}$:
(a) The set of points of continuity

$$C_f = \{x \in \mathbb{R} : f \text{ is continuous at } x\}$$

for an arbitrary function $f : \mathbb{R} \to \mathbb{R}$.
(b) The set of points of convergence

$$\mathrm{Conv} = \{x \in \mathbb{R} : \lim_{n \to \infty} f_n(x) \text{ exists}\}$$

for an arbitrary sequence of continuous functions $f_n : \mathbb{R} \to \mathbb{R}$.

2.8 Let $(X, \mathcal{B})$ be a measurable space, and let $\mu : \mathcal{B} \to [0, \infty]$. Prove that μ is a measure if and only if it satisfies the following three properties:

- $\mu(\emptyset) = 0$;
- FINITE ADDITIVITY: for any disjoint sets $A, B \in \mathcal{B}$,
$$\mu(A \sqcup B) = \mu(A) + \mu(B);$$
- CONTINUITY FROM BELOW: if $E_1 \subseteq E_2 \subseteq \cdots \in \mathcal{B}$, then
$$\mu\left(\bigcup_{n \in \mathbb{N}} E_n \right) = \lim_{n \to \infty} \mu(E_n).$$

CHAPTER 3

Integration Against a Measure

Our next task is to develop an integration theory for integrating measurable functions on abstract measure spaces. In the Riemann–Darboux approach to integration, we approximate a function $f : [a, b] \to [0, \infty)$ by step functions, for which we can easily define the integral. For the Lebesgue theory of integration, we will use a similar idea but with a more general class of functions: so-called simple functions.

1. Integration of Simple Functions

Definition 3.1

Let $(X, \mathcal{B})$ be a measurable space. A *simple function* is a measurable function $s : X \to \mathbb{C}$ taking only finitely many values.

Partitioning X into finitely many pieces corresponding to the values of a simple function s, we may write simple functions as linear combinations of indicator functions of measurable sets. That is, $s = \sum_{j=1}^n c_j \mathbb{1}_{E_j}$ for some numbers $c_j \in \mathbb{C}$ and measurable sets $E_j \in \mathcal{B}$. Given a measure μ on $(X, \mathcal{B})$, we define the integral of a simple function in the obvious way. To avoid issues with adding and subtracting infinities, we will deal for now only with nonnegative functions.

Definition 3.2

Let $(X, \mathcal{B}, \mu)$ be a measure space and $s : X \to [0, \infty)$ a simple function. Write $s = \sum_{j=1}^n c_j \mathbb{1}_{E_j}$ with $c_j \geq 0$ and $E_j \in \mathcal{B}$. The *integral of s with respect to μ* is given by

$$\int_X s \, d\mu = \sum_{j=1}^n c_j \mu(E_j).$$

Proposition 3.3

The integral of a nonnegative simple function is well-defined. That is, the value of the integral of a simple function s does not depend on the representation of s as a linear combination of indicator functions of measurable sets.

Proof. Suppose $s = \sum_{j=1}^n c_j \mathbb{1}_{E_j}$. Let $a_1, \ldots, a_m$ be the finite collection of values taken by s, and let $A_k = \{s = a_k\}$ for $k = 1, \ldots, m$. Then the sets $A_1, \ldots, A_k$ partition X, and $s = \sum_{k=1}^m a_k \mathbb{1}_{A_k}$. We will show $\sum_{j=1}^n c_j \mu(E_j) = \sum_{k=1}^m a_k \mu(A_k)$.

Define a new collection of sets $E'_J = \bigcap_{j \in J} E_j \setminus \bigcup_{i \notin J} E_j$ for $J \subseteq \{1, \ldots, n\}$. In other words, $x \in E'_J$ means that $x \in E_j$ if and only if $j \in J$. This defines a partition of X. Note that the value of s on the set E'_J is $c'_J = \sum_{j\in J} c_j$. We can therefore relate the sets E'_J to the sets A_k by

$$A_k = \bigsqcup_{J\subseteq\{1,\ldots,n\}, c'_J = a_k} E'_J.$$

Then on the one hand,

$$\sum_{k=1}^m a_k\mu(A_k) = \sum_{k=1}^m a_k \sum_{J\subseteq\{1,\ldots,n\}, c'_J = a_k} \mu(E'_J) = \sum_{J\subseteq\{1,\ldots,n\}} c'_J\mu(E'_J).$$

On the other hand,

$$\sum_{j=1}^n c_j\mu(E_j) = \sum_{j=1}^n c_j \sum_{\{j\}\subseteq J\subseteq\{1,\ldots,n\}} \mu(E'_J) = \sum_{J\subseteq\{1,\ldots,n\}} \sum_{j\in J} c_j\mu(E'_J) = \sum_{J\subseteq\{1,\ldots,n\}} c'_J\mu(E'_J).$$

This completes the proof. □

We used a particular representation of a simple function in the previous proof that will continue to be convenient to work with. Say that $\sum_{j=1}^n c_j \mathbb{1}_{E_j}$ is the *standard representation* of a simple function s if $s = \sum_{j=1}^n c_j \mathbb{1}_{E_j}$, and the sets $E_1, \ldots, E_n$ partition X (that is, they are pairwise disjoint and their union is X).

Proposition 3.4

Let $(X, \mathcal{B}, \mu)$ be a measure space, let $s, t : X \to [0, \infty)$ be simple functions, and let $c \in \mathbb{R}$, $c \geq 0$. Then
(1) $\int_X cs\ d\mu = c \cdot \int_X s\ d\mu$;
(2) $\int_X (s + t)\ d\mu = \int_X s\ d\mu + \int_X t\ d\mu$;
(3) if $s \leq t$, then $\int_X s\ d\mu \leq \int_X t\ d\mu$.

Proof. **(1)** Let $s = \sum_{j=1}^n c_j \mathbb{1}_{E_j}$. Then $cs = \sum_{j=1}^n (cc_j)\mathbb{1}_{E_j}$, so

$$\int_X cs\ d\mu = \sum_{j=1}^n (cc_j)\mu(E_j) = c \cdot \sum_{j=1}^n c_j\mu(E_j) = c \cdot \int_X s\ d\mu.$$

For (2) and (3), it will be helpful to work with the standard representation, so let $s = \sum_{j=1}^n c_j \mathbb{1}_{E_j}$ and $t = \sum_{k=1}^m d_k \mathbb{1}_{F_k}$ be the standard representations. Define sets $A_{j,k} = E_j \cap F_k$ for $j \in \{1, \ldots, n\}$ and $k \in \{1, \ldots, m\}$. Then $E_j = \bigsqcup_{k=1}^m A_{j,k}$ and $F_k = \bigsqcup_{j=1}^n A_{j,k}$.

(2) The function $s + t$ takes the value $c_j + d_k$ on $A_{j,k}$, so

$$\int_X (s+t)\ d\mu = \sum_{j,k} (c_j + d_k)\mu(A_{j,k}) = \sum_{j=1}^n c_j \underbrace{\sum_{k=1}^m \mu(A_{j,k})}_{\mu(E_j)} + \sum_{k=1}^m d_k \underbrace{\sum_{j=1}^n \mu(A_{j,k})}_{\mu(F_k)} = \int_X s\ d\mu + \int_X t\ d\mu.$$

(3) By assumption, if $A_{j,k} \neq \emptyset$, then $c_j \leq d_k$. Thus,

$$\int_X s\ d\mu = \sum_{j=1}^n c_j\mu(E_j) = \sum_{j,k} c_j\mu(A_{j,k}) \leq \sum_{j,k} d_k\mu(A_{j,k}) = \sum_{k=1}^m d_k\mu(F_k) = \int_X t\ d\mu.$$

□

Definition 3.5

Let $(X, \mathcal{B}, \mu)$ be a measure space, $s : X \to [0, \infty)$ a simple function, and $E \in \mathcal{B}$ a measurable set. The *integral of s with respect to μ over E* is given by

$$\int_E s \, d\mu = \int_X s \cdot \mathbb{1}_E \, d\mu.$$

Note that if s is simple, then $s \cdot \mathbb{1}_E$ is also simple, so the above definition makes sense.

Proposition 3.6

Let $(X, \mathcal{B}, \mu)$ be a measure space, and let $s : X \to [0, \infty)$ be a simple function. Then

$$\nu(E) = \int_E s \, d\mu$$

defines a measure on $(X, \mathcal{B})$.

Proof. Note that $s \cdot \mathbb{1}_\emptyset = 0$, so $\nu(\emptyset) = 0$. Suppose $(E_n)_{n\in\mathbb{N}}$ is a pairwise disjoint family of measurable sets, and let $E = \bigsqcup_{n\in\mathbb{N}} E_n$. Write $s = \sum_{j=1}^m a_j \mathbb{1}_{A_j}$. Then $s \cdot \mathbb{1}_E = \sum_{j=1}^m a_j \mathbb{1}_{A_j \cap E}$, so

$$\nu(E) = \sum_{j=1}^m a_j \mu(A_j \cap E) = \sum_{j,n} a_j \mu(A_j \cap E_n) = \sum_{n\in\mathbb{N}} \int_X s \cdot \mathbb{1}_{E_n} \, d\mu = \sum_{n\in\mathbb{N}} \nu(E_n).$$

Note that the sum over n is an infinite sum so reordering requires some justification. Fortunately, all of the values $a_j \mu(A_j \cap E_n)$ are nonnegative, so the sum can be computed in any order without changing the value. □

2. Integration of Nonnegative Measurable Functions

We now want to extend the definition of the integral against a measure to all nonnegative measurable functions. The next proposition shows that simple functions are a sufficiently general class to approximate arbitrary measurable functions.

Proposition 3.7

Let $(X, \mathcal{B})$ be a measurable space, and let $f : X \to [0, \infty]$ be measurable. Then there exists a sequence $(s_n)_{n\in\mathbb{N}}$ of simple functions such that $0 \le s_1 \le s_2 \le \cdots \le f$, and $s_n \to f$ pointwise.

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

$$s_n(x) = \begin{cases} \frac{a}{2^n}, & \text{if } \frac{a}{2^n} \le f(x) < \frac{a+1}{2^n} \text{ and } a < n \cdot 2^n. \\ n, & \text{if } f(x) \ge n. \end{cases}$$

□

It is therefore reasonable to define the integral of an arbitrary nonnegative measurable function as follows.

Definition 3.8

Let $(X, \mathcal{B}, \mu)$ be a measure space, and let $f : X \to [0, \infty]$ be measurable. We define the *integral of f with respect to μ* as

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

Given a measurable set $E \in \mathcal{B}$, the *integral of f with respect to μ over E* is defined by

$$\int_E f \, d\mu = \int_X f \cdot \mathbb{1}_E \, d\mu.$$

One may object at this point and suggest an alternative definition. Since $f : X \to [0, \infty]$ can be obtained as $f = \lim_{n\to\infty} s_n$ for an increasing sequence of simple functions $0 \le s_1 \le s_2 \le \dots$, why not define $\int_X f \, d\mu = \lim_{n\to\infty} \int_X s_n \, d\mu$? As we will see shortly, this is in fact an equivalent definition that is extremely useful for many applications. However, *as a definition*, it has two serious defects: why should the limit exist? and why should the value be the same for all possible approximations by simple functions? This is why we prefer Definition 3.8 above (and why this is the standard definition across measure theory textbooks).

Proposition 3.9

Let $(X, \mathcal{B}, \mu)$ be a measure space, and let $f, g : X \to [0, \infty]$ be measurable. If $f \le g$, then

$$\int_X f \, d\mu \le \int_X g \, d\mu.$$

Proof. It suffices to observe $\{s \text{ simple function} : 0 \le s \le f\} \subseteq \{s \text{ simple function} : 0 \le s \le g\}$. □

Theorem 3.10: Monotone Convergence Theorem

Let $(f_n)_{n\in\mathbb{N}}$ be a sequence of measurable functions $0 \le f_1 \le f_2 \le \dots$, and let $f = \lim_{n\to\infty} f_n$. Then

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

Remark. Note that a consequence of the monotone convergence theorem is that $\int_X f \, d\mu$ can be computed by taking a sequence of simple functions $0 \le s_1 \le s_2 \le \dots \to f$ and computing $\lim_{n\to\infty} \int_X s_n \, d\mu$.

Proof of Monotone Convergence Theorem. First, f is a measurable function by Proposition 2.11. By monotonicity of the integral (Proposition 3.9), the sequence $\int_X f_n \, d\mu$ is increasing, so $\lim_{n\to\infty} \int_X f_n \, d\mu = \sup_{n\in\mathbb{N}} \int_X f_n \, d\mu \in [0, \infty]$ exists as an extended real number. Moreover,

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

since the inequality holds for each $n \in \mathbb{N}$. Therefore, it suffices to show

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

If $\lim_{n\to\infty} \int_X f_n \, d\mu = \infty$, there is nothing to prove, so assume $\lim_{n\to\infty} \int_X f_n \, d\mu < \infty$.

Let $c < 1$. Let $s : X \to [0, \infty)$ be a simple function, $0 \le s \le f$. For $n \in \mathbb{N}$, let $E_n = \{f_n \ge cs\}$. Then $E_1 \subseteq E_2 \subseteq \dots$ and $X = \bigcup_{n\in\mathbb{N}} E_n$. By Proposition 3.6, let $\nu : \mathcal{B} \to [0, \infty]$ be the measure $\nu(E) = \int_E s \, d\mu$. We have

$$\begin{aligned} c \cdot \int_X s \, d\mu &= c \cdot \nu(X) \\ &= c \cdot \lim_{n\to\infty} \nu(E_n) && \text{(continuity from below)} \\ &= \lim_{n\to\infty} c \cdot \nu(E_n) && \text{(Proposition 2.10)} \\ &= \lim_{n\to\infty} \int_{E_n} cs \, d\mu && \text{(Proposition 3.4)} \\ &\le \lim_{n\to\infty} \int_X f_n \, d\mu && \text{(monotonicity).} \end{aligned}$$

Taking a supremum over all such simple functions, we conclude

$$c \cdot \int_X f \, d\mu \le \lim_{n\to\infty} \int_X f_n \, d\mu.$$

Letting $c \to 1$ yields the desired result. □

Proposition 3.11

Let $(X, \mathcal{B}, \mu)$ be a measure space, and let $f, g : X \to [0, \infty]$ be measurable functions. Let $c \in [0, \infty)$.

(1) $\int_X cf \, d\mu = c \cdot \int_X f \, d\mu$.

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

Proof. **(1)** This follows quickly from the definition of the integral and Proposition 3.4.

(2) We use the monotone convergence theorem. Let $0 \le s_1 \le s_n \le \dots \le f$ with $s_n \to f$ and $0 \le t_1 \le t_2 \le \dots \le g$ with $t_n \to g$. Then $0 \le s_1 + t_1 \le s_2 + t_2 \le \dots \le f + g$ and $s_n + t_n \to f + g$. Thus,

$$\begin{aligned} \int_X (f + g) \, d\mu &= \lim_{n\to\infty} \int_X (s_n + t_n) \, d\mu && \text{(MCT)} \\ &= \lim_{n\to\infty} \int_X s_n \, d\mu + \lim_{n\to\infty} \int_X t_n \, d\mu && \text{(Proposition 3.4)} \\ &= \int_X f \, d\mu + \int_X g \, d\mu && \text{(MCT).} \end{aligned}$$

□

Theorem 3.12

Let $(X, \mathcal{B}, \mu)$ be a measure space, and let $(f_n)_{n\in\mathbb{N}}$ be a sequence of nonnegative measurable functions, $f_n : X \to [0, \infty]$. Then

$$\int_X \left(\sum_{n=1}^{\infty} f_n \right) d\mu = \sum_{n=1}^{\infty} \int_X f_n \, d\mu.$$

Proof. We have

$$\begin{aligned}
\int_X \left(\sum_{n=1}^{\infty} f_n\right) d\mu &= \int_X \lim_{N\to\infty} \left(\sum_{n=1}^{N} f_n\right) d\mu \\
&= \lim_{N\to\infty} \int_X \left(\sum_{n=1}^{N} f_n\right) d\mu && \text{(MCT)} \\
&= \lim_{N\to\infty} \sum_{n=1}^{N} \int_X f_n \, d\mu && \text{(additivity of the integral)} \\
&\sum_{n=1}^{\infty} \int_X f_n \, d\mu.
\end{aligned}$$

□

Theorem 3.13: Fatou's Lemma

Let $(X, \mathcal{B}, \mu)$ be a measure space. Let $(f_n)_{n\in\mathbb{N}}$ be a sequence of measurable functions, $f_n : X \to [0, \infty]$. Then

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

Proof. Let $f = \liminf_{n\to\infty} f_n$. Define $F_N = \inf_{n\geq N} f_n$. Then $0 \leq F_1 \leq F_2 \leq \dots$ and $F_N \to f$. Therefore,

$$\begin{aligned}
\int_X f \, d\mu &= \lim_{N\to\infty} \int_X F_N \, d\mu && \text{(MCT)} \\
&\leq \lim_{N\to\infty} \inf_{n\geq N} \int_X f_n \, d\mu && \text{(monotonicity of the integral)} \\
&= \liminf_{N\to\infty} \int_X f_n \, d\mu.
\end{aligned}$$

□

3. Integration of Real and Complex-Valued Functions

The method for integrating real and complex-valued functions involves decomposing these functions as linear combinations of nonnegative functions. An important observation is that such a decomposition can be done in a measurable way.

Definition 3.14

Let X be a set and $f : X \to [-\infty, \infty]$. The *positive part* f^+ and *negative part* f^- of f are defined by

$$f^+ = \max\{f, 0\} \qquad \text{and} \qquad f^- = \max\{-f, 0\}.$$

Note that $f = f^+ - f^-$ and $|f| = f^+ + f^-$. Moreover, if $(X, \mathcal{B})$ is a measurable space and $f : X \to [-\infty, \infty]$ is measurable, then f^+ and f^- are measurable by Proposition 2.11.

Definition 3.15

Let $(X, \mathcal{B}, \mu)$ be a measure space.

- An extended real-valued measurable function $f : X \to [-\infty, \infty]$ is *integrable* if
$$\int_X |f| \, d\mu < \infty.$$
In this case, the *integral of f* is defined by
$$\int_X f \, d\mu = \int_X f^+ \, d\mu - \int_X f^- \, d\mu.$$
- A complex-valued measurable function $f : X \to \mathbb{C}$ is *integrable* if
$$\int_X |f| \, d\mu < \infty,$$
and the *integral of f* is defined by
$$\int_X f \, d\mu = \int_X \operatorname{Re}(f) \, d\mu + i \int_X \operatorname{Im}(f) \, d\mu.$$
- Given a measurable set $E \in \mathcal{B}$, a measurable function f taking extended real or complex values is *integrable over E* if $f \cdot \mathbb{1}_E$ is integrable, and the *integral of f over E* is
$$\int_E f \, d\mu = \int_X f \cdot \mathbb{1}_E \, d\mu.$$

Remark. By monotonicity of the integral (Proposition 3.9), if a function is integrable, then it is also integrable over every measurable subset of X.

4. Integral Identities and Inequalities

Proposition 3.16: Triangle Inequality for the Integral

Suppose $(X, \mathcal{B}, \mu)$ is a measure space and $f : X \to \mathbb{C}$ is an integrable function. Then
$$\left| \int_X f \, d\mu \right| \le \int_X |f| \, d\mu.$$

Proof. First, suppose f is real-valued. Then by the triangle inequality and linearity,
$$\left| \int_X f \, d\mu \right| = \left| \int_X f^+ \, d\mu - \int_X f^- \, d\mu \right| \le \int_X f^+ \, d\mu + \int_X f^- \, d\mu = \int_X |f| \, d\mu.$$

Now suppose f is complex-valued. Let $\lambda \in \mathbb{C}$ with $|\lambda| = 1$ such that $\left| \int_X f \, d\mu \right| = \lambda \int_X f \, d\mu$. Then
$$\left| \int_X f \, d\mu \right| = \operatorname{Re}\left(\int_X \lambda f \, d\mu \right) = \int_X \operatorname{Re}(\lambda f) \, d\mu \le \int_X |\operatorname{Re}(\lambda f)| \, d\mu \le \int_X |f| \, d\mu.$$
□

Proposition 3.17: Linearity of the Integral

Let $(X, \mathcal{B}, \mu)$ be a measure space. Let $f, g : X \to \mathbb{C}$ be integrable functions, and let $c \in \mathbb{C}$. Then

(1) $f + g$ is integrable, and $\int_X (f + g) \, d\mu = \int_X f \, d\mu + \int_X g \, d\mu$.

(2) cf is integrable, and $\int_X cf\ d\mu = c\int_X f\ d\mu$.

Proof. **(1)** First, by the triangle inequality, we have $|f+g| \le |f| + |g|$. Therefore,

$$\int_X |f+g|\ d\mu \overset{(*)}{\le} \int_X (|f|+|g|)\ d\mu \overset{(**)}{=} \int_X |f|\ d\mu + \int_X |g|\ d\mu < \infty.$$

In step $(*)$, we have used monotonicity of the integral (Proposition 3.9), and in $(**)$, we have used additivity (Proposition 3.11).

Decomposing f and g into their real and imaginary parts, it suffices to prove the identity $\int_X (f+g)\ d\mu = \int_X f\ d\mu + \int_X g\ d\mu$ for real-valued functions f and g. Let $h = f+g$. Then $h = h^+ - h^- = f^+ - f^- + g^+ - g^-$. This can be rearranged to the identity $h^+ + f^- + g^- = h^- + f^+ + g^+$. Then using additivity of the integral for nonnegative functions (Proposition 3.11), we have

$$\begin{aligned}\int_X h^+\ d\mu + \int_X f^-\ d\mu + \int_X g^-\ d\mu &= \int_X (h^+ + f^- + g^-)\ d\mu \\ &= \int_X (h^- + f^+ + g^+)\ d\mu \qquad (3.1)\\ &= \int_X h^-\ d\mu + \int_X f^+\ d\mu + \int_X g^+\ d\mu.\end{aligned}$$

Rearranging again,

$$\begin{aligned}\int_X (f+g)\ d\mu &= \int_X h^+\ d\mu - \int_X h^-\ d\mu && \text{(Definition 3.15)}\\ &= \int_X f^+\ d\mu - \int_X f^-\ d\mu + \int_X g^+\ d\mu - \int_X g^-\ d\mu && \text{(by (3.1))}\\ &= \int_X f\ d\mu + \int_X g\ d\mu && \text{(Definition 3.15)}\end{aligned}$$

(2) Note that $|cf| = |c||f|$, so

$$\int_X |cf|\ d\mu = \int_X |c||f|\ d\mu \overset{(*)}{=} |c|\int_X |f|\ d\mu < \infty,$$

where $(*)$ follows from Proposition 3.11. Hence, cf is integrable.

For computing the integral of cf, we consider several different cases.

Case 1. $c \ge 0$

When f is nonnegative, we have

$$\int_X cf\ d\mu = c\int_X f\ d\mu$$

by Proposition 3.11. The identity follows for a general complex-valued function f by decomposing $f = \left(\mathrm{Re}\,(f)^+ - \mathrm{Re}\,(f)^-\right) + i\left(\mathrm{Im}\,(f)^+ - \mathrm{Im}\,(f)^-\right)$.

Case 2. $c = -1$

For real-valued $f : X \to \mathbb{R}$, we use the identities $(-f)^+ = f^-$ and $(-f)^- = f^+$ to obtain

$$\int_X (-f)\ d\mu = \int_X f^-\ d\mu - \int_X f^+\ d\mu = -\int_X f\ d\mu.$$

Complex-valued functions can be handled by decomposing into real and imaginary parts.

CASE 3. $c = i$

Noting that $\mathrm{Re}\,(if) = -\mathrm{Im}\,(f)$ and $\mathrm{Im}\,(if) = \mathrm{Re}\,(f)$, we have

$$\begin{aligned}
\int_X if\ d\mu &= \int_X (-\mathrm{Im}\,(f))\ d\mu + i\int_X \mathrm{Re}\,(f)\ d\mu && \text{(Definition 3.15)}\\
&= -\int_X \mathrm{Im}\,(f)\ d\mu + i\int_X \mathrm{Re}\,(f)\ d\mu && \text{(Case 2)}\\
&= i\left(\int_X \mathrm{Re}\,(f)\ d\mu + i\int_X \mathrm{Im}\,(f)\ d\mu\right)\\
&= i\int_X f\ d\mu && \text{(Definition 3.15)}
\end{aligned}$$

CASE 4. $c \in \mathbb{R}$

Combine Case 1 and Case 2.

CASE 5. $c \in \mathbb{C}$

Write $c = a + ib$ with $a, b \in \mathbb{R}$. Then

$$\begin{aligned}
\int_X cf\ d\mu &= \int_X (af + ibf)\ d\mu\\
&= \int_X af\ d\mu + \int_X ibf\ d\mu && \text{(by (1))}\\
&= \int_X af\ d\mu + i\int_X bf\ d\mu && \text{(Case 3)}\\
&= a\int_X f\ d\mu + ib\int_X f\ d\mu && \text{(Case 4)}\\
&= c\int_X f\ d\mu.
\end{aligned}$$

□

Let $(X, \mathcal{B}, \mu)$ be a measure space, and denote by $L^1(\mu)$ the set of integrable functions. Proposition 3.17 shows that $L^1(\mu)$ is a (complex) vector space. Moreover, in the course of the proof, we showed

$$\int_X |cf|\ d\mu = |c|\int_X |f|\ d\mu \qquad \text{and} \qquad \int_X |f+g|\ d\mu \le \int_X |f|\ d\mu + \int_X |g|\ d\mu.$$

In other words, if we let

$$\|f\|_1 = \int_X |f|\ d\mu,$$

then $\|\cdot\|_1$ defines a *seminorm* on the vector space of integrable functions on $(X, \mathcal{B}, \mu)$.

Definition 3.18

Let V be a real or complex vector space. A function $\|\cdot\| : V \to [0, \infty)$ is a *seminorm* if it satisfies:

- TRIANGLE INEQUALITY: $\|u + v\| \le \|u\| + \|v\|$ for all $u, v \in V$, and
- ABSOLUTE HOMOGENEITY: $\|cv\| = |c| \|v\|$ for all $v \in V$ and all scalars c.

A seminorm is a *norm* if it satisfies the additional property

- POSITIVE DEFINITE: if $v \in V$ and $\|v\| = 0$, then $v = 0$.

The seminorm $\|\cdot\|_1$ on the space of integrable functions may not be a norm in general, but a small modification will turn it into a norm. This will be discussed in greater detail later in the course, in the context of so-called L^p spaces. One of the important ingredients is a deeper understanding of *null sets*, which we will discuss now.

5. Sets of Measure Zero

Definition 3.19

Let $(X, \mathcal{B}, \mu)$ be a measure space.

- A measurable set $N \in \mathcal{B}$ is a *null set* if $\mu(N) = 0$.
- We say that a property holds *almost everywhere* if there exists a null set $N \in \mathcal{B}$ such that the property holds for every point $x \in X \setminus N$.

REMARK. An easy consequence of countable additivity and monotonicity of measures is that the family $\mathcal{N}$ of null sets forms a *σ-ideal* of $\mathcal{B}$:

- $\emptyset \in \mathcal{N}$;
- if $A \in \mathcal{N}$ and $B \in \mathcal{B}$ with $B \subseteq A$, then $B \in \mathcal{N}$; and
- if $(N_n)_{n\in\mathbb{N}}$ is a countable family of null sets, then $\bigcup_{n\in\mathbb{N}} N_n \in \mathcal{N}$.

NOTATION. The phrases "almost everywhere" or "almost every" are often abbreviated by a.e. or μ-a.e. if the measure needs to be specified. In a statement of the form "Property P holds a.e.," we interpret a.e. as "almost everywhere." For a statement of the form "Property P holds for a.e. $x \in X$," we read a.e. as "almost every," and the meaning is the same as in the previous example statement.

Null sets naturally arise and play an important role in integration theory. Some examples are provided by the next three propositions.

Proposition 3.20

Let $(X, \mathcal{B}, \mu)$ be a measure space. Suppose $f : X \to [-\infty, \infty]$ is an integrable function. Then $f(x) \in \mathbb{R}$ for μ-a.e. $x \in X$.

PROOF. Let $N = \{x \in \mathbb{R} : |f(x)| = \infty\}$. We want to show that N is a null set. By monotonicity of the integral (Proposition 3.9),

$$\int_X |f| \, d\mu \ge \int_N |f| \, d\mu = \infty \cdot \mu(N).$$

On the other hand, by integrability of f,

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

Thus, $\infty \cdot \mu(N) < \infty$, so $\mu(N) = 0$. □

Corollary 3.21: Borel–Cantelli Lemma

Let $(X, \mathcal{B}, \mu)$ be a measure space. Suppose $(E_n)_{n\in\mathbb{N}}$ is a sequence of measurable sets and $\sum_{n=1}^{\infty} \mu(E_n) < \infty$. Then

$$\mu\left(\{x \in X : x \in E_n \text{ for infinitely many } n \in \mathbb{N}\}\right) = 0.$$

Proof. One possible proof uses continuity from above and was given in the exercises (see Exercise 3.1). We will now give a different proof using integration.

Let $f = \sum_{n=1}^{\infty} \mathbb{1}_{E_n}$. Note that $f(x) = \infty$ if and only if $x \in E_n$ for infinitely many $n \in \mathbb{N}$. By Theorem 3.12,

$$\int_X f \; d\mu = \sum_{n=1}^{\infty} \underbrace{\int_X \mathbb{1}_{E_n}}_{\mu(E_n)} < \infty.$$

So by Proposition 3.20, $f < \infty$ a.e. That is,

$$\mu\left(\{x \in X : x \in E_n \text{ for infinitely many } n \in \mathbb{N}\}\right) = \mu(\{f = \infty\}) = 0.$$

□

Proposition 3.22

Let $(X, \mathcal{B}, \mu)$ be a measure space, and let $f, g : X \to \mathbb{C}$ be measurable functions. Suppose $f = g$ a.e. Then f is integrable if and only if g is integrable. Moreover, if f and g are integrable, then

$$\int_X f \; d\mu = \int_X g \; d\mu.$$

Proof. Let $N = \{x \in X : f(x) \neq g(x)\}$. By assumption, N is a null set.

Step 1. Integrability

Suppose f is integrable. Then

$$\begin{aligned}
\int_X |g| \; d\mu &= \int_{X\setminus N} |f| \; d\mu + \int_N |g| \; d\mu && \text{(linearity of the integral)} \\
&\leq \int_X |f| \; d\mu + \underbrace{\infty \cdot \mu(N)}_{0} && \text{(monotonicity of the integral)} \\
&= \int_X |f| \; d\mu < \infty,
\end{aligned}$$

so g is integrable. Reversing the roles of f and g proves the converse.

Step 2. Integral

Assume f and g are integrable. Then

$$\begin{aligned}
\left|\int_X g\, d\mu - \int_X f\, d\mu\right| &= \left|\int_X (g-f)\, d\mu\right| && \text{(linearity of the integral)} \\
&\le \int_X |g-f|\, d\mu && \text{(triangle inequality for the integral)} \\
&= \int_{X\setminus N} 0\, d\mu + \int_N |g-f|\, d\mu && \text{(linearity of the integral)} \\
&\le 0\cdot\mu(X\setminus N) + \infty\cdot\mu(N) = 0.
\end{aligned}$$

□

Proposition 3.23

Let $(X, \mathcal{B}, \mu)$ be a measure space, and let $f : X \to [0,\infty]$ be a measurable function. Then $\int_X f\, d\mu = 0$ if and only if $f = 0$ a.e.

Proof. If $f = 0$ a.e., then by Proposition 3.22, f is integrable and

$$\int_X f\, d\mu = \int_X 0\, d\mu = 0\cdot\mu(X) = 0.$$

Conversely, suppose $\int_X f\, d\mu = 0$. Then by Markov's inequality (Exercise 3.2),

$$\mu\left(\{f > c\}\right) \le \frac{1}{c}\int_X f\, d\mu = 0$$

for every $c > 0$. Therefore, by continuity of μ from below,

$$\mu\left(\{f \ne 0\}\right) = \mu\left(\bigcup_{n\in\mathbb{N}}\left\{f > \frac{1}{n}\right\}\right) = \lim_{n\to\infty}\mu\left(\left\{f > \frac{1}{n}\right\}\right) = 0.$$

That is, $f = 0$ a.e. □

The examples above (especially Proposition 3.22) show that null sets are negligible from the point of view of integration, and we can very often ignore modifications that happen on null sets. There is one subtle issue that requires care, however: in general, a subset of a null set may not be measurable and non-measurable modifications on null sets may create issues. For this reason, it is often convenient to work with *complete* measure spaces, as defined below.

Definition 3.24

A measure space $(X, \mathcal{B}, \mu)$ is *complete* if every subset of every null set is measurable. That is, if $E \subseteq X$ and there exists $N \in \mathcal{B}$ with $E \subseteq N$ and $\mu(N) = 0$, then $E \in \mathcal{B}$.

The following proposition is a useful tool for passing to complete measure spaces.

Proposition 3.25

Let $(X, \mathcal{B}, \mu)$ be a measure space. Let $\mathcal{N} = \{N \in \mathcal{B} : \mu(N) = 0\}$ be the σ-ideal of μ-null sets. Then the family $\overline{\mathcal{B}} = \{E \cup F : E \in \mathcal{B}, F \subseteq N \in \mathcal{N}\}$ is a σ-algebra, and there is a unique extension $\overline{\mu}$ of μ to $\overline{\mathcal{B}}$.

Proof. Exercise 3.9. □

Definition 3.26

The *completion* of a measure space $(X, \mathcal{B}, \mu)$ is the space $(X, \overline{\mathcal{B}}, \overline{\mu})$, where $\overline{\mathcal{B}}$ and $\overline{\mu}$ are as defined in Proposition 3.25.

6. The Dominated Convergence Theorem

We have already seen two fundamental convergence theorems for integration against a measure: the monotone convergence theorem and Fatou's lemma. We are nearly ready to state another fundamental result about integration: the dominated convergence theorem. First, we need to introduce the two notions of convergence that will be related by the dominated convergence theorem.

Definition 3.27

Let $(X, \mathcal{B}, \mu)$ be a measure space.

- We say that a sequence $(f_n)_{n\in\mathbb{N}}$ of functions on X *converges almost everywhere* to a function f if $\lim_{n\to\infty} f_n(x) = f(x)$ for almost every $x \in X$.
- A sequence $(f_n)_{n\in\mathbb{N}}$ of integrable functions *converges in L^1* to $f \in L^1(\mu)$ if

$$\|f_n - f\|_1 = \int_X |f_n - f| \; d\mu \to 0$$

in $\mathbb{R}$ as $n \to \infty$.

The dominated convergence theorem says that any sequence that converges almost everywhere and is "L^1-dominated" will converge in L^1. The precise mathematical formulation is as follows:

Theorem 3.28: Dominated Convergence Theorem

Let $(X, \mathcal{B}, \mu)$ be a measure space. Let $(f_n)_{n\in\mathbb{N}}$ be a sequence of integrable functions, $f_n : X \to \mathbb{C}$, and let $f : X \to \mathbb{C}$ be measurable. Suppose

- $f_n \to f$ a.e., and
- there is an integrable function $g : X \to [0, \infty)$ such that $\sup_{n\in\mathbb{N}} |f_n| \leq g$ a.e.

Then f is integrable and $f_n \to f$ in $L^1(\mu)$. In particular,

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

Proof. First, $|f| \leq |g|$ a.e., so f is integrable.

Observe:

$$\begin{aligned}
\int_X 2g \; d\mu - \limsup_{n\to\infty} \int_X |f - f_n| \; d\mu &= \liminf_{n\to\infty} \int_X (2g - |f - f_n|) \; d\mu \\
&\geq \int_X \liminf_{n\to\infty} (2g - |f - f_n|) \; d\mu && \text{(Fatou's lemma)} \\
&= \int_X 2g \; d\mu && (f_n \to f)
\end{aligned}$$

Rearranging, we conclude

$$\limsup_{n\to\infty} \int_X |f - f_n| \; d\mu \leq 0.$$

Using the triangle inequality for the integral,

$$\left|\int_X f\, d\mu - \int_X f_n\, d\mu\right| \leq \int_X |f - f_n|\, d\mu \to 0,$$

so

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

□

The assumption that the sequence $(f_n)_{n\in\mathbb{N}}$ is "dominated" by an integrable function g is a necessary assumption to avoid "escape of mass to infinity," as the following example demonstrates.

EXAMPLE 3.29

Let $X = \mathbb{Z}$, $\mathcal{B} = \mathscr{P}(\mathbb{Z})$, and let μ be the counting measure. Let $f_n = \mathbb{1}_{\{n\}}$. Then $f_n(x) \to 0$ for every $x \in X$. However,

$$\int_X f_n\, d\mu = 1$$

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

$$\int_X \lim_{n\to\infty} f_n\, d\mu = \int_X 0\, d\mu = 0 \neq 1.$$

Additional Reading

For other presentations of integration on abstract measures spaces, see [**1**, Section 2.1–2.3], [**4**, Chapter 1], [**6**, Sections 2.1 and 6.2], and/or [**7**, Section 1.3 and Subsection 1.4.4]. The development of integration in the books of Folland [**1**] and Rudin [**4**] is very similar to the presentation in these notes. By contrast, Stein and Shakarchi [**6**] and Tao [**7**] first develop integration in the special case of the Lebesgue measure before moving to abstract spaces. The book of Stein and Shakarchi [**6**] also proves the fundamental convergence theorems in a different order, starting with a special case of the dominated convergence theorem known as the *bounded convergence theorem*, and then deducing Fatou's lemma, the monotone convergence theorem, and the general case of the dominated convergence theorem.

There is a very nice book of Oxtoby [**3**] that develops useful analogies between measure spaces and topological spaces and includes a discussion of null sets in relation to a σ-ideal of "topologically negligible" sets called *meager* sets or sets of *first category*.

Exercises

3.1 Prove the Borel–Cantelli lemma: if $(A_n)_{n\in\mathbb{N}}$ is a family of measurable subsets of a probability space $(X, \mathcal{B}, \mu)$ and $\sum_{n\in\mathbb{N}} \mu(A_n) < \infty$, then

$$\mu\left(\{x \in X : x \in A_n \text{ for infinitely many } n \in \mathbb{N}\}\right) = 0.$$

3.2 Let $(X, \mathcal{B}, \mu)$ be a measure space and f a measurable function. Prove Markov's inequality: for any $c > 0$,

$$\mu\left(\{|f| \geq c\}\right) \leq \frac{1}{c} \int_{\{|f|>c\}} |f|\, d\mu \leq \frac{1}{c} \int_X |f|\, d\mu.$$

3.3 Let $(X, \mathcal{B}, \mu)$ and $(Y, \mathcal{C}, \nu)$ be measure space, and let $T : X \to Y$ be a measurable function. Define $T\mu : \mathcal{C} \to [0, \infty]$ by $(T\mu)(A) = \mu(T^{-1}(A))$. Prove $T\mu = \nu$ if and only if for every integrable function $f : Y \to \mathbb{C}$,

$$\int_Y f\, d\nu = \int_X f \circ T\, d\mu.$$

3.4 Let $(X, \mathcal{B}, \mu)$ be a probability space. Let $(A_n)_{n\in\mathbb{N}}$ be a family of measurable sets with $a = \inf_{n\in\mathbb{N}} \mu(A_n) > 0$. Show that there is a set $E \subseteq \mathbb{N}$ such that $\overline{d}(E) := \limsup_{N\to\infty} \frac{|E\cap\{1,\ldots,N\}|}{N} \geq a$, and for any finite set $F \subseteq E$, $F \neq \emptyset$, one has $\mu\left(\bigcap_{n\in F} A_n\right) > 0$ by proving the following intermediate steps:

(a) Justify that we can assume without loss of generality that $\bigcap_{n\in F} A_n \neq \emptyset$ if and only if $\mu\left(\bigcap_{n\in F} A_n\right) > 0$ for every finite set $F \subseteq \mathbb{N}$. It may help to define the countable set

$$\mathcal{F} = \left\{ F \subseteq \mathbb{N} : |F| < \infty, \bigcap_{n\in F} A_n \neq \emptyset, \text{and } \mu\left(\bigcap_{n\in F} A_n\right) = 0 \right\}.$$

(b) Prove

$$\int_X \limsup_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} \mathbb{1}_{A_n} \, d\mu \geq a.$$

(c) Define $E = \{n \in \mathbb{N} : x \in A_n\}$ for a suitable choice of $x \in X$.

3.5 Let $(X, \mathcal{B}, \mu)$ be a measure space. Suppose $f : X \to [0,\infty]$ is a measurable function. Define $\nu : \mathcal{B} \to [0,\infty]$ by

$$\nu(E) = \int_E f \, d\mu.$$

Prove that ν is a measure.

3.6 Let $(X, \mathcal{B}, \mu)$ be a measure space, and let $f : X \to \mathbb{C}$ be an integrable function. Prove that for any $\varepsilon > 0$, there exists $\delta > 0$ with the following property: if $E \in \mathcal{B}$ and $\mu(E) < \delta$, then $\left|\int_E f \, d\mu\right| < \varepsilon$.

3.7 Let $(X, \mathcal{B}, \mu)$ be a measure space, and let $f, g : X \to \mathbb{C}$ be integrable functions. Show that $f = 0$ a.e. if and only if $\int_E f \, d\mu$ for every $E \in \mathcal{B}$.

3.8 Show that a measure space $(X, \mathcal{B}, \mu)$ is complete if and only if it satisfies the following property: for functions $f, g : X \to \mathbb{C}$, if f is measurable and $f = g$ a.e., then g is measurable.

3.9 Prove Proposition 3.25.

3.10 Show that simple functions are dense in L^1. That is, if $(X, \mathcal{B}, \mu)$ is a measure space and $f \in L^1(\mu)$, then for every $\varepsilon > 0$, there exists a simple function $s : X \to \mathbb{C}$ such that $\|f - s\|_1 < \varepsilon$.

3.11 Let $(X, \mathcal{B}, \mu)$ be a measure space and $E \in \mathcal{B}$. If $(E_n)_{n\in\mathbb{N}}$ is a sequence of measurable sets and $E = \bigcup_{n\in\mathbb{N}} E_n$, prove that for every integrable function $f \in L^1(\mu)$,

$$\lim_{n\to\infty} \int_{E_n} f \, d\mu = \int_E f \, d\mu.$$

State and prove an analogous result for decreasing sequences.

3.12 Prove

$$\lim_{n\to\infty} \left(1 + \frac{x}{n}\right)^n = \sum_{k\geq 0} \frac{x^k}{k!}.$$

3.13 Let $(X, \mathcal{B}, \mu)$ be a measure space, and let $(f_n)_{n\in\mathbb{N}}$ be a sequence of measurable functions $f_n : X \to \mathbb{C}$. Suppose

$$\sum_{n\in\mathbb{N}} \int_X |f_n| \, d\mu < \infty.$$

Prove that $\sum_{n\in\mathbb{N}} f_n$ converges a.e. to an integrable function $f \in L^1(\mu)$, and

$$\int_X f \, d\mu = \sum_{n\in\mathbb{N}} \int_X f_n \, d\mu.$$

3.14 In this exercise, we will use measure-theoretic tools in order to carry out computations with Riemann integrals. Assume for the purposes of this exercise that there is a measure λ on the Borel subsets of $\mathbb{R}$ with the property: if $f : [a, b] \to \mathbb{R}$ is a Riemann integrable function, then

$$\int_{[a,b]} f \; d\lambda = \int_a^b f(x) \; dx,$$

where the integral on the left is the measure-theoretic integral and the integral on the right is the Riemann integral. (We will discuss multiple methods of constructing such a measure λ (the *Lebesgue measure*) in future lectures.)

(a) Compute

$$\lim_{n\to\infty} \int_0^\infty \frac{n \sin\left(\frac{x}{n}\right)}{x(1+x^2)} \; dx.$$

(b) Show that for $a > -1$,

$$\int_0^1 \frac{x^a \log x}{1-x} \; dx = -\sum_{k=1}^\infty \frac{1}{(a+k)^2}.$$

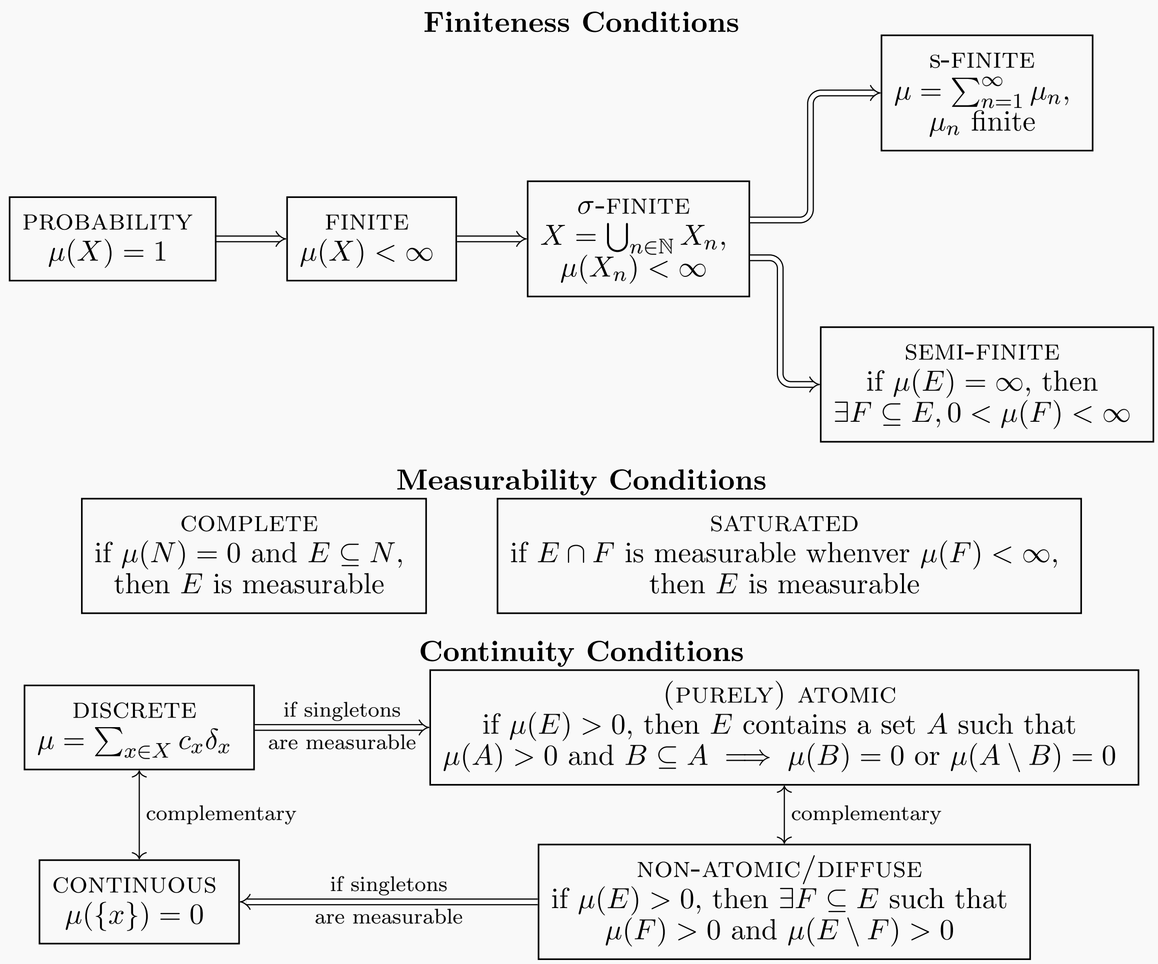
Part 2

Constructions of Measures

CHAPTER 4

Taxonomy of Measures

In this short chapter, we give a taxonomy of measures based on various properties that may be desirable or undesirable in certain circumstances. Some of these properties have been seen in previous chapters, while others are introduced for the first time here. The taxonomy is summarized in Figure 4.1, and precise definitions are given below.



Regularity Conditions (for Borel Measures)

LOCALLY FINITE	$\mu(K) < \infty$ for compact $K \subseteq X$
OUTER REGULAR	$\mu(E) = \inf\{\mu(U) : U \supseteq E \text{ open}\}$
INNER REGULAR	$\mu(E) = \sup\{\mu(K) : K \subseteq E \text{ compact}\}$

FIGURE 4.1. Taxonomy of Measures

1. Properties of Measures

Definition 4.1

Let $(X, \mathcal{B}, \mu)$ be a measure space. A set E is

- *locally measurable* if $E \cap K \in \mathcal{B}$ for every set $K \in \mathcal{B}$ with $\mu(K) < \infty$;
- an *atom* if $\mu(E) > 0$ and every measurable subset $F \subseteq E$, $F \in \mathcal{B}$ satisfies either $\mu(F) = 0$ or $\mu(E \setminus F) = 0$.

Definition 4.2

Let $(X, \mathcal{B})$ be a measurable space. A measure μ on $(X, \mathcal{B})$ is

- a *probability measure* if $\mu(X) = 1$;
- *finite* if $\mu(X) < \infty$;
- *σ-finite* if there is a countable sequence of measurable sets $(X_n)_{n\in\mathbb{N}}$ in $\mathcal{B}$ such that $X = \bigcup_{n\in\mathbb{N}} X_n$ and $\mu(X_n) < \infty$ for each $n \in \mathbb{N}$;
- *s-finite* if μ is a countable sum $\mu = \sum_{n\in\mathbb{N}} \mu_n$ of finite measures $\mu_n : \mathcal{B} \to [0, \infty)$;
- *semi-finite* if every set of infinite measure contains a subset of positive finite measure, i.e. if $E \in \mathcal{B}$ and $\mu(E) = \infty$, then there exists $F \in \mathcal{B}$ with $F \subseteq E$ and $\mu(F) < \infty$;
- *complete* if every subset of every null set is measurable, i.e. if $E \subseteq X$ and there exists $N \in \mathcal{B}$ with $E \subseteq N$ and $\mu(N) = 0$, then $E \in \mathcal{B}$;
- *saturated* if every locally measurable set is measurable;
- *discrete* if μ is a combination of Dirac measures, $\mu = \sum_{x\in X} c_x \delta_x$ for some coefficients $c_x \in [0, \infty]$;
- *continuous* if μ has no point masses, i.e. $\mu(\{x\}) = 0$ for every $x \in X$;
- *(purely) atomic* if every set of positive measure contains an atom;
- *non-atomic* or *diffuse* if there are no atoms.

If X is a topological space and $\mathcal{B} = \mathrm{Borel}(X)$, then μ is

- *locally finite* if every compact set has finite measure;
- *outer regular* if $\mu(E) = \inf\{\mu(U) : U \supseteq E \text{ open}\}$ for every $E \in \mathcal{B}$;
- *inner regular* if $\mu(E) = \sup\{\mu(K) : K \subseteq E \text{ compact}\}$ for every $E \in \mathcal{B}$.

The relationships between the various properties in Definition 4.2 are displayed in Figure 4.1.

2. Finiteness Properties

As shown in Figure 4.1, every probability measure is finite, every finite measure is σ-finite, and every σ-finite measure is both s-finite and semi-finite. The next example shows that s-finiteness and semi-finiteness are rather different notions from one another, neither one implying the other.

Example 4.3

An s-finite measure that is not semi-finite: Let X be a non-empty set, and let $x \in X$. Define $\mu(E) = \infty$ if $x \in E$ and $\mu(E) = 0$ if $x \notin E$. Then $\mu = \sum_{n=1}^{\infty} \delta_x$, so μ is s-finite. However, the set $\{x\}$ has infinite measure and no subsets of non-zero finite measure, so μ is not semi-finite.

A semi-finite measure that is not s-finite: Let X be an uncountable set, and let $\mu : \mathscr{P}(X) \to [0, \infty]$ be the counting measure on X. If $E \subseteq X$ and $\mu(E) = \infty$, then taking any point $x \in E$, we have $\mu(\{x\}) = 1 < \infty$, so μ is a semi-finite measure. However, since

X is uncountable, μ cannot be expressed as a countable sum of finite measures, so μ is not s-finite.

Most texts on measure theory focus on σ-finite measures and omit mention of the more general concept of s-finite measures. However, as we will see, s-finite measures are the natural class of measures for many important results in measure theory. One reason to appreciate the generality of s-finite measures is provided by the next theorem. We will encounter other advantages of working with s-finite measures later on in the course.

Theorem 4.4

(1) Let $(X, \mathcal{B}, \mu)$ be an s-finite measure space, and let $(Y, \mathcal{C})$ be a measurable space. Suppose $\pi : X \to Y$ is a measurable map. Then the measure $\pi_*\mu : \mathcal{C} \to [0, \infty]$ defined by $\pi_*\mu(C) = \mu(\pi^{-1}(C))$ is s-finite.
(2) There exists a σ-finite measure space $(X, \mathcal{B}, \mu)$, a measurable space $(Y, \mathcal{C})$, and a measurable map $\pi : X \to Y$ such that $\pi_*\mu$ is not σ-finite.

Proof. **(1)** First note that the projection of a finite measure is finite. Indeed, $\pi_*\mu(Y) = \mu(\pi^{-1}(Y)) = \mu(X)$. Noting that $\pi_* \left(\sum_{n=1}^{\infty} \mu_n\right) = \sum_{n=1}^{\infty} \pi_*\mu_n$ then completes the proof.

(2) Let $X = \mathbb{Z}^2$, $\mathcal{B} = \mathscr{P}(\mathbb{Z}^2)$, and let $\mu : \mathcal{B} \to [0, \infty]$ be the counting measure. Let $Y = \mathbb{Z}$ and $\mathcal{C} = \mathscr{P}(\mathbb{Z})$, and let $\pi : \mathbb{Z}^2 \to \mathbb{Z}$ be the projection onto the first coordinate, i.e. $\pi(n, m) = n$ for $(n, m) \in \mathbb{Z}^2$. Then $\pi_*\mu(E)$ counts the number of points in $\mathbb{Z}^2$ whose first coordinate belongs to E. Hence, $\pi_*\mu(E) = \infty$ whenever $E \neq \emptyset$. Therefore, $\pi_*\mu$ is not σ-finite (nor even semi-finite). □

3. Decompositions of Measures

When we say that two notions are "complementary," we mean that they are mutually exclusive and every (σ-finite) measure can be decomposed into pieces satisfying one or the other property. Namely, for the complementary notions shown in Figure 4.1, we have the following decomposition result:

Proposition 4.5

(1) Let $(X, \mathcal{B}, \mu)$ be a σ-finite measure space. Then there is a unique decomposition $\mu = \mu_a + \mu_{na}$ as a sum of a purely atomic measure μ_a and a non-atomic measure μ_{na}.
(2) Let $(X, \mathcal{B}, \mu)$ be an s-finite measure space, and suppose $\{x\} \in \mathcal{B}$ for every $x \in X$. Then there is a unique decomposition $\mu = \mu_d + \mu_c$ as a sum of a discrete measure μ_d and a continuous measure μ_c.

In general, *atomic* and *discrete* are different notions.

Example 4.6

Let X be an uncountable set, and let $\mathcal{B} = \{E \subseteq X : E \text{ is countable or } X \setminus E \text{ is countable}\}$. Define a probability measure $\mu : \mathcal{B} \to [0, 1]$ by $\mu(E) = 0$ if E is countable and $\mu(E) = 1$ is $X \setminus E$ is uncountable. Then E is atomic (each co-countable set is an atom) but also continuous.

However, in many frequently-encountered situations, atomic and discrete measures coincide.

Theorem 4.7

Let X be a separable metric space. Suppose μ is a locally finite Borel measure on X. If $A \in \mathrm{Borel}(X)$ is an atom of μ, then there is a point $x \in A$ such that $\mu(\{x\}) = \mu(A) > 0$. Hence, every atomic locally finite Borel measure on X is discrete.

Proving the decomposition of a σ-finite measure into atomic and non-atomic components is a bit lengthy, so we will prove only part (2) of Proposition 4.5. Because of Theorem 4.7, the decomposition into discrete and continuous parts is sufficient for most purposes.

Proof of Proposition 4.5(2).

Step 1. Existence.

Since μ is s-finite, we may write $\mu = \sum_{n=1}^{\infty} \mu_n$ for some finite measures $\mu_n : \mathcal{B} \to [0, \infty)$. For each $k \in \mathbb{N}$, let $X_{n,k} = \left\{x \in X : \mu_n(\{x\}) \geq \frac{1}{k}\right\}$. Note that $X_{n,k}$ has at most $k\mu_n(X)$ elements for each $n, k \in \mathbb{N}$. Therefore, $X_0 = \{x \in X : \mu(\{x\}) > 0\} = \bigcup_{n\in\mathbb{N}} \bigcup_{k\in\mathbb{N}} X_{n,k}$ is a countable set.

For $x \in X_0$, let $c_x = \mu(\{x\})$. Define $\mu_d = \sum_{x\in X_0} c_x \delta_x$, and let $\mu_c : \mathcal{B} \to [0, \infty)$ be the measure $\mu_c(E) = \mu(E \setminus X_0)$ for $E \in \mathcal{B}$. Then μ_d is manifestly a discrete measure. Moreover, for any $x \in X$,

$$\mu_c(\{x\}) = \mu(\{x\} \setminus X_0) = \begin{cases} \mu(\{x\}), & \text{if } x \notin X_0; \\ 0, & \text{if } x \in X_0. \end{cases}$$

Since X_0 is the set of all point masses for μ, it follows that $\mu_c(\{x\}) = 0$ for every $x \in X$; that is, μ_c is continuous. Finally, for any $E \in \mathcal{B}$,

$$\mu(E) = \mu(E \cap X_0) + \mu_c(E)$$

and

$$\mu(E \cap X_0) = \sum_{x\in E\cap X_0} \mu(\{x\}) = \sum_{x\in X_0} c_x \delta_x(E) = \mu_d(E).$$

Step 2. Uniqueness.

Let $\mu = \mu_d + \mu_c$ be the decomposition obtained in Step 1. Suppose $\mu = \mu'_d + \mu'_c$ is another decomposition into a discrete measure μ'_d and a continuous measure μ'_c. We want to show $\mu'_d = \mu_d$ and $\mu'_c = \mu_c$.

Let $x \in X_0$. Since μ'_c is continuous, we have $\mu'_c(\{x\}) = 0$, so $\mu'_d(\{x\}) = \mu(\{x\}) = c_x$. On the other hand, if $x \in X$ is any point and $\mu'_d(\{x\}) > 0$, then $\mu(\{x\}) \geq \mu'_d(\{x\}) > 0$, so $x \in X_0$. Therefore, the point masses of μ'_d are exactly the elements of X_0, and $\mu'_d(\{x\}) = c_x$ for $x \in X_0$. Since μ'_d is discrete, it can thus be represented as $\mu'_d = \sum_{x\in X_0} c_x \delta_x$. That is, $\mu'_d = \mu_d$, and it follows that we also have $\mu'_c = \mu_c$.

□

The condition of semi-finiteness also leads to a decomposition result.

Proposition 4.8

Let $(X, \mathcal{B}, \mu)$ be a measure space. There exists a decomposition $\mu = \mu_{\text{sf}} + \mu_{\text{inf}}$ such that μ_{sf} is semi-finite and μ_{inf} takes only the values 0 and ∞.

Unlike the decompositions in Proposition 4.5, the decomposition in Proposition 4.8 is not unique in general. One way of obtaining the decomposition is to define

$$\mu_{\text{sf}}(E) = \sup \left\{ \mu(F) : F \in \mathcal{B}, F \subseteq E, \text{and } \mu(F) < \infty \right\},$$

and

$$\mu_{\text{inf}}(E) = \begin{cases} 0, & \text{if } E \text{ is semi-finite;} \\ \infty, & \text{if } E \text{ is not semi-finite.} \end{cases}$$

Here, we say that a measurable set E is semi-finite if the measure $\mu_E : \mathcal{B} \to [0, \infty]$ defined by $\mu_E(A) = \mu(A \cap E)$ is a semi-finite measure. In other words, $E \in \mathcal{B}$ is semi-finite if every subset of E of infinite measure has a further subset of positive finite measure.

Exercises

4.1 Show that every σ-finite measure is s-finite and semi-finite.

4.2 In this exercise, we will construct a *saturation* of a measure space similarly to how we constructed the completion of a measure space.

Let $(X, \mathcal{B}, \mu)$ be a measure space.

(a) Show that the family

$$\widetilde{\mathcal{B}} = \{ E \subseteq X : E \cap F \in \mathcal{B} \text{ for every } F \in \mathcal{B} \text{ with } \mu(F) < \infty \}$$

of locally measurable sets is a σ-algebra on X.

(b) Define a function $\widetilde{\mu} : \widetilde{\mathcal{B}} \to [0, \infty]$ by

$$\widetilde{\mu}(E) = \begin{cases} \mu(E), & \text{if } E \in \mathcal{B}; \\ \infty, & \text{if } E \notin \mathcal{B}. \end{cases}$$

Show that $\widetilde{\mu}$ is a measure and $(X, \widetilde{B}, \widetilde{\mu})$ is saturated.

4.3 Suppose $(X, \mathcal{B}, \mu)$ is a non-atomic probability space. The goal of this problem is to prove Sierpiński's theorem: if $t \in [0, 1]$, then there exists a set $E \in \mathcal{B}$ with $\mu(E) = t$.

(a) Show for any $s \in (0, 1)$, there exists a set $E \in \mathcal{B}$ such that $0 < \mu(E) < s$.

(b) Let $t \in (0, 1)$. Construct a sequence of disjoint measurable sets $(E_n)_{n \in \mathbb{N}}$ such that

  (i) for each $n \in \mathbb{N}$, $\mu(E_1 \cup \cdots \cup E_n) < t$, and

  (ii) if possible, E_n is chosen so that $\mu(E_n) > \frac{1}{n}$.

Show that $\mu\left(\bigcup_{n \in \mathbb{N}} E_n\right) = t$.

CHAPTER 5

Lebesgue–Stieltjes Measures

Let us rephrase (an instance of) the problem of measurement using the language of abstract measure theory developed in Part 1.

Problem 5.1: Problem of Measurement in One Dimension

Construct a measure $\lambda : \mathrm{Borel}(\mathbb{R}) \to [0, \infty]$ such that $\lambda(I) = \mathrm{length}(I)$ for every interval $I \subseteq \mathbb{R}$. Is there a unique such measure? Can the measure be defined on all subsets of $\mathbb{R}$?

We will address this problem in a more general framework, where we allow for different assignments of measure to intervals.

Definition 5.2

A Borel measure $\mu : \mathrm{Borel}(\mathbb{R}) \to [0, \infty]$ is *locally finite* if $\mu(K) < \infty$ for every compact set $K \subseteq \mathbb{R}$. The *distribution function* of a locally finite Borel measure μ is the function $F_\mu : \mathbb{R} \to \mathbb{R}$ defined by

$$F_\mu(x) = \begin{cases} \mu((0, x]), & \text{if } x > 0; \\ 0, & \text{if } x = 0; \\ -\mu((x, 0]), & \text{if } x < 0. \end{cases}$$

By monotonicity of the measure μ, its distribution function F_μ is necessarily increasing. Moreover, by continuity from above and below, F_μ is a right-continuous function. The goal of this chapter is to prove the following theorem:

Theorem 5.3: Existence and Uniqueness of Lebesgue–Stieltjes Measures

Let $F : \mathbb{R} \to \mathbb{R}$ be an increasing, right-continuous function with $F(0) = 0$. There exists a σ-algebra $\mathcal{M}_F$ containing the Borel subsets of $\mathbb{R}$ and a complete measure $\mu_F : \mathcal{M}_F \to [0, \infty]$ such that $F = F_{\mu_F}$. Moreover, if $\nu : \mathrm{Borel}(\mathbb{R}) \to [0, \infty]$ is a Borel measure satisfying $F_\nu = F$, then $\nu = \mu_F|_{\mathrm{Borel}(\mathbb{R})}$, and μ_F is the completion of ν.

Definition 5.4

Let $F : \mathbb{R} \to \mathbb{R}$ be an increasing, right-continuous function. The unique complete measure μ_F given by Theorem 5.3 is called the *Lebesgue–Stieltjes measure* associated to F.

Proposition 5.5

Let F be an increasing, right-continuous function, and let μ_F be the Lebesgue–Stieltjes measure associated to F. Then

$$\lim_{x\to\infty} F(x) = \sup_{x\in\mathbb{R}} F(x) = \mu_F((0,\infty)) \qquad \text{and} \qquad \lim_{x\to-\infty} F(x) = \inf_{x\in\mathbb{R}} F(x) = -\mu_F((-\infty,0]).$$

Proof. This is an application of continuity from below of the measure μ_F. □

Notation. Given an increasing, right-continuous function, we will write $F(\infty)$ for the value $\lim_{x\to\infty} F(x)$ and $F(-\infty) = \lim_{x\to-\infty} F(x)$. In general, $F(\pm\infty)$ is an extended real number.

1. The π-λ Theorem and Uniqueness of Lebesgue–Stieltjes Measures

Before constructing Lebesgue–Stieltjes measures, let us prove that every locally finite Borel measure is uniquely determined by its distribution function. The key tool will be the π-λ theorem, for which we need a new definition.

Definition 5.6

Let X be a set.

- A family $\mathcal{P} \subseteq \mathscr{P}(X)$ of subsets of X is a *π-system* if $\mathcal{P}$ is closed under finite intersections.
- A family $\mathcal{L} \subseteq \mathscr{P}(X)$ is a *λ-system* if $\emptyset \in \mathcal{L}$ and $\mathcal{L}$ is closed under complements and countable disjoint unions.

Example 5.7

The following are examples of π systems:

- the collection $\mathcal{P} = \{(a,b] : a,b \in \mathbb{R}\}$ of half-open intervals in $\mathbb{R}$;
- the family of open sets of any topological space;
- given a measure space $(X,\mathcal{B},\mu)$, the family $\mathcal{P} = \{E \in \mathcal{B} : \mu(X \setminus E) = 0\}$ of co-null sets;
- given two measurable spaces $(X,\mathcal{B})$ and $(Y,\mathcal{C})$, the family $\mathcal{P} = \{B \times C : B \in \mathcal{B}, C \in \mathcal{C}\}$ of "rectangles" in $X \times Y$.

Examples of λ-systems include:

- for two probability measures μ,ν on a measurable space $(X,\mathcal{B})$, the family $\mathcal{L} = \{E \in \mathcal{B} : \mu(E) = \nu(E)\}$.

Another characterization of λ-systems is given by the following proposition:

Proposition 5.8

Let X be a set. A family $\mathcal{L} \subseteq \mathscr{P}(X)$ is a λ-system if and only if it satisfies the following three properties:

(1) $X \in \mathcal{L}$;
(2) if $A,B \in \mathcal{L}$ and $A \subseteq B$, then $B \setminus A \in \mathcal{L}$;
(3) if $A_1 \subseteq A_2 \subseteq \dots$ is an increasing sequence in $\mathcal{L}$, then $\bigcup_{n\in\mathbb{N}} A_n \in \mathcal{L}$.

Proof. Suppose $\mathcal{L}$ is a λ-system. We check that $\mathcal{L}$ satisfies properties (1)–(3).

(1) Since $\emptyset \in \mathcal{L}$ and $\mathcal{L}$ is closed under complements, we have $X \in \mathcal{L}$.

(2) Let $A, B \in \mathcal{L}$ with $A \subseteq B$. Then $B \setminus A = B \cap A^c = (B^c \cup A)^c$. The assumption $A \subseteq B$ means $B^c \cap A = \emptyset$, so we have represented $B \setminus A$ in terms of A and B using complementation and disjoint union. Hence, $B \setminus A \in \mathcal{L}$.

(3) Let $A_1 \subseteq A_2 \subseteq \dots$ be an increasing sequence in $\mathcal{L}$. Let $B_1 = A_1$ and $B_n = A_n \setminus A_{n-1}$ for $n \geq 2$. By property (ii), $B_n \in \mathcal{L}$ for every $n \in \mathbb{N}$. Therefore, $\bigcup_{n\in\mathbb{N}} A_n = \bigsqcup_{n\in\mathbb{N}} B_n \in \mathcal{L}$.

Conversely, suppose $\mathcal{L} \subseteq \mathscr{P}(X)$ is a family of sets satisfying (1), (2), and (3).

Applying property (2) with $A = B = X$, we have $\emptyset = X \setminus X \in \mathcal{L}$.

Let $A \in \mathcal{L}$. Combining (1) and (2), $A^c = X \setminus A \in \mathcal{L}$.

Finally, let $(A_n)_{n\in\mathbb{N}}$ be a pairwise disjoint sequence of sets in $\mathcal{L}$. Then $B_n = A_1 \sqcup \dots \sqcup A_n$ forms an increasing sequence, so by property (3), it suffices to prove that $B_n \in \mathcal{L}$. By induction, this reduces to showing that the disjoint union of two sets in $\mathcal{L}$ is an element of $\mathcal{L}$. Let $C, D \in \mathcal{L}$ with $C \cap D = \emptyset$. Then $C \sqcup D = (C^c \cap D^c)^c = (C^c \setminus D)^c$. We have already checked that $\mathcal{L}$ is closed under complementation. The disjointness of C and D implies $D \subseteq C^c$, so $C^c \setminus D \in \mathcal{L}$ by property (2). Thus, $C \sqcup D \in \mathcal{L}$. □

Theorem 5.9: π-λ Theorem (Sierpiński–Dynkin)

Let X be a set, and suppose $\mathcal{P} \subseteq \mathscr{P}(X)$ is a π-system. If $\mathcal{L} \subseteq \mathscr{P}(X)$ is a λ-system and $\mathcal{P} \subseteq \mathcal{L}$, then $\sigma(\mathcal{P}) \subseteq \mathcal{L}$.

We will prove the π-λ theorem with the help of several lemmas.

Lemma 5.10

Let X be a set. A family $\mathcal{B} \subseteq \mathscr{P}(X)$ of subsets of X is a σ-algebra if and only if $\mathcal{B}$ is both a π-system and a λ-system.

Proof. The definition of a λ-system is the same as the definition of a σ-algebra, except that one is only allowed to take unions of *disjoint* sets in the definition of a λ-system. It therefore suffices to check that being a π-system as well allows for taking countable unions of not necessarily disjoint sets.

Suppose $E_1, E_2, \dots \in \mathcal{B}$. Define $E_1' = E_1$, $E_2' = E_2 \setminus E_1$, $\dots$, $E_n' = E_n \setminus \bigcup_{i=1}^{n-1} E_i$. Then $E_1', E_2', \dots$ are pairwise disjoint and satisfy $\bigsqcup_{n\in\mathbb{N}} E_n' = \bigcup_{n\in\mathbb{N}} E_n$, so it suffices to check that $E_n' \in \mathcal{B}$ for each $n \in \mathbb{N}$. But this is clear upon rewriting $E_n' = E_n \cap \bigcap_{i=1}^{n-1} E_i^c$, since $E_i^c = X \setminus E_i \in \mathcal{B}$ (by the axioms of a λ-system) and a finite intersection of sets from $\mathcal{B}$ belongs to $\mathcal{B}$ (by the axioms of a π-system). □

Lemma 5.11

Let X be a set, and suppose $(\mathcal{L}_i)_{i\in I}$ is a collection of λ-systems $\mathcal{L}_i \subseteq \mathscr{P}(X)$. Then $\bigcap_{i\in I} \mathcal{L}_i$ is a λ-system.

Proof. The proof is the same as the proof of Proposition 2.3, except we only allow disjoint unions. □

Definition 5.12

Let X be a set and $\mathcal{S} \subseteq \mathscr{P}(X)$ a family of subsets of X. The *λ-system generated by $\mathcal{S}$* is the smallest λ-system containing $\mathcal{S}$:

$$\lambda(\mathcal{S}) = \bigcap\{\mathcal{L} \subseteq \mathscr{P}(X) : \mathcal{L} \text{ is a } \lambda\text{-system}, \mathcal{S} \subseteq \mathcal{L}\}.$$

Lemma 5.13

Let X be a set, and let $\mathcal{P} \subseteq \mathscr{P}(X)$ be a π-system. The λ-system $\lambda(\mathcal{P})$ generated by $\mathcal{P}$ is a σ-algebra.

Proof. By Lemma 5.10, it suffices to show that $\lambda(\mathcal{P})$ is a π-system.

Claim 1. For any set $A \in \lambda(\mathcal{P})$, the family $\mathcal{L}_A := \{B \subseteq X : A \cap B \in \lambda(\mathcal{P})\}$ is a λ-system.

Since $A \in \lambda(\mathcal{P})$, we see that $X \in \mathcal{L}_A$.

Suppose $B_1, B_2 \in \mathcal{L}_A$ and $B_1 \subseteq B_2$. Then

$$A \cap (B_2 \setminus B_1) = \underbrace{(A \cap B_2)}_{\in\lambda(\mathcal{P})} \setminus \underbrace{(A \cap B_1)}_{\in\lambda(\mathcal{P})} \in \lambda(\mathcal{P}),$$

so $B_2 \setminus B_1 \in \mathcal{L}_A$.

Finally, suppose $B_1 \subseteq B_2 \subseteq \cdots \in \mathcal{L}_A$. Then $A \cap \bigcup_{n\in\mathbb{N}} B_n = \bigcup_{n\in\mathbb{N}}(A \cap B_n) \in \lambda(\mathcal{P})$, so $\bigcup_{n\in\mathbb{N}} B_n \in \mathcal{L}_A$.

This proves the claim.

Claim 2. For any $A \in \lambda(\mathcal{P})$ and any $B \in \mathcal{P}$, we have $A \cap B \in \lambda(\mathcal{P})$.

This follows from Claim 1: the family $\mathcal{L}_B$ is a λ-system, and $\mathcal{P} \subseteq \mathcal{L}_B$ by the definition of a π-system, so $\mathcal{L}_B \supseteq \lambda(\mathcal{P}) \ni A$.

Let $A, B \in \lambda(\mathcal{P})$. The family $\mathcal{L}_A$ is a λ-system (by Claim 1) containing $\mathcal{P}$ (by Claim 2), so $\mathcal{L}_A \supseteq \lambda(\mathcal{P}) \ni B$. Hence, $A \cap B \in \lambda(\mathcal{P})$. □

Now we can complete the proof of the π-λ theorem.

Proof of π-λ Theorem (Theorem 5.9). Let $\mathcal{P}$ be a π-system, $\mathcal{L}$ a λ-system, and suppose $\mathcal{P} \subseteq \mathcal{L}$. On the one hand, by Lemma 5.13, the λ-system $\lambda(\mathcal{P})$ generated by $\mathcal{P}$ is a σ-algebra, so $\sigma(\mathcal{P}) \subseteq \lambda(\mathcal{P})$. On the other hand, $\mathcal{L}$ is a λ-system containing $\mathcal{P}$, so $\lambda(\mathcal{P}) \subseteq \mathcal{L}$. Combining these two observations completes the proof. □

Corollary 5.14: Uniqueness of Lebesgue–Stieltjes Measures

Suppose μ and ν are locally finite Borel measures on $\mathbb{R}$ with the same distribution function $F_\mu = F_\nu = F$. Then $\mu = \nu$.

Proof. Let $\mathcal{P}$ be the π-system $\mathcal{P} = \{(a, b] : a, b \in \mathbb{R}\}$ of half-open intervals. Define

$$\mathcal{L} = \{E \in \mathrm{Borel}(\mathbb{R}) : \mu(E \cap (-N, N]) = \nu(E \cap (-N, N]) \text{ for every } N \in \mathbb{N}\}.$$

Claim 1. $\mathcal{L}$ is a λ-system

For every $N \in \mathbb{N}$,
$$\mu((-N, N]) = (\mu((-N, 0]) + \mu((0, N])) = F(N) - F(-N).$$
The same holds for ν, so $\mathbb{R} \in \mathcal{L}$.

Suppose $E \in \mathcal{L}$, and let $N \in \mathbb{N}$. By additivity of μ and ν, we have
$$\begin{aligned}\mu(E^c \cap (-N, N]) &= \mu((-N, N]) - \mu(E \cap (-N, N]) \\ &= \nu((-N, N]) - \nu(E \cap (-N, N]) \\ &= \nu(E^c \cap (-N, N]),\end{aligned}$$
so $E^c \in \mathcal{L}$.

Finally, $\mathcal{L}$ is closed under countable disjoint unions as a consequence of countable additivity of the measures μ and ν.

Claim 2. $\mathcal{P} \subseteq \mathcal{L}$

The sets $(-N, N]$ belong to $\mathcal{P}$, which is a π-system, so it suffices to prove $\mu(P) = \nu(P)$ for every $P \in \mathcal{P}$. Let $P = (a, b] \in \mathcal{P}$. If $b \leq a$, then $P = \emptyset$, so $\mu(P) = \nu(P) = 0$. Suppose $a < b$. If $a \leq b$, then $\mu(P) = F(b) - F(a) = \nu(P)$.

By the π-λ theorem, $\sigma(\mathcal{P}) \subseteq \mathcal{L}$. But $\mathcal{P}$ generates the Borel σ-algebra (we essentially showed this in the proof of Proposition 2.11), so $\mathcal{L} = \text{Borel}(\mathbb{R})$. Hence, applying continuity from below, we have
$$\mu(E) = \lim_{N\to\infty} \mu\left(E \cap (-N, N]\right) = \lim_{N\to\infty} \nu\left(E \cap (-N, N]\right) = \nu(E)$$
for every $E \in \text{Borel}(\mathbb{R})$. □

Example 5.15

The locally finite condition cannot be dropped from Corollary 5.14. As an example, define a measure $\mu : \text{Borel}(\mathbb{R}) \to [0, \infty]$ by
$$\mu(B) = \#(B \cap \mathbb{Q}),$$
and let $\nu = \infty \cdot \lambda$, where λ is the Lebesgue measure. (We will construct λ later in this chapter, but for now, take it as a given that the Lebesgue measure exists.) Every non-empty interval in $\mathbb{R}$ contains infinitely many rational points, so
$$F_\mu(x) = \begin{cases} \infty, & \text{if } x > 0; \\ 0, & \text{if } x = 0; \\ -\infty, & \text{if } x < 0. \end{cases}$$
Similarly, every non-empty interval in $\mathbb{R}$ has positive Lebesgue measure, so multiplying by ∞, the measure ν has the same distribution function $F_\nu = F_\mu$. However, μ and ν are not the same measure, since, for instance, $\mu(\mathbb{R} \setminus \mathbb{Q}) = 0$, while $\nu(\mathbb{R} \setminus \mathbb{Q}) = \infty$, and $\mu(\{0\}) = 1$, while $\nu(\{0\}) = 0$.

2. Half-Open Intervals

Now we begin the construction of Lebesgue–Stieltjes measures. Let us define some basic objects that we will work with for the construction.

Definition 5.16

A *left-open, right-closed interval* is a set of the form

- $\mathbb{R}$,
- $\emptyset$,
- $(a, b]$ with $a, b \in \mathbb{R}$, $a < b$,
- $(-\infty, b]$ with $b \in \mathbb{R}$, or
- (a, ∞) with $a \in \mathbb{R}$.

The intersection of two left-open, right-closed intervals is a left-open, right-closed intervals, and the complement of a left-open, right-closed interval is either a left-open, right-closed interval or a disjoint union of two left-open, right-closed intervals. Therefore, the family of left-open, right-closed intervals forms a *semi-algebra* on $\mathbb{R}$. We recall the definition below.

Definition 5.17

Let X be a set. A family $\mathcal{S} \subseteq \mathscr{P}(X)$ of subsets of X is a *semi-algebra* if

- $\emptyset, X \in \mathcal{S}$;
- if $A, B \in \mathcal{S}$, then $A \cap B \in \mathcal{S}$;
- if $A \in \mathcal{S}$, then $X \setminus A = \bigsqcup_{i=1}^{n} C_i$ for some $C_1, \ldots, C_n \in \mathcal{S}$.

In Exercise 2.5, you showed the following fact:

Proposition 5.18

Let $\mathcal{S}$ be a semi-algebra on a set X. Then

$$\mathcal{A} = \left\{ \bigsqcup_{i=1}^{n} S_i : n \in \mathbb{N}, S_1, \ldots, S_n \in \mathcal{S} \right\}$$

is an algebra.

Notation. We will denote the algebra generated by the semi-algebra of left-open, right-closed intervals by

$$\mathcal{A}_{int} = \left\{ \bigsqcup_{i=1}^{n} I_i : n \in \mathbb{N}, I_i \text{ is a left-open, right-closed interval} \right\}.$$

Note that the σ-algebra generated by $\mathcal{A}_{int}$ is the Borel σ-algebra.

3. Premeasures and Outer Measures

We will begin the construction of the Lebesgue–Stieltjes measure associated to a distribution function F by assigning a measure to each element of $\mathcal{A}_{int}$.

Definition 5.19

Let X be a set and $\mathcal{A} \subseteq \mathscr{P}(X)$ an algebra. A *premeasure* is a function $\mu_0 : \mathcal{A} \to [0, \infty]$ such that

- $\mu_0(\emptyset) = 0$;
- if $(A_n)_{n\in\mathbb{N}}$ is a sequence of pairwise disjoint elements of $\mathcal{A}$ and $A = \bigsqcup_{n\in\mathbb{N}} A_n \in \mathcal{A}$, then $\mu_0(A) = \sum_{n=1}^{\infty} \mu_0(A_n)$.

Note that if $\mathcal{A}$ is a σ-algebra, then a premeasure on $\mathcal{A}$ is the same thing as a measure. More generally, if $\mu : \mathcal{B} \to [0, \infty]$ is a measure on a measurable space $(X, \mathcal{B})$ and $\mathcal{A} \subseteq \mathcal{B}$ is an algebra, then $\mu_0 = \mu|_{\mathcal{A}}$ defines a premeasure on $\mathcal{A}$.

The next proposition shows that we can associate to an increasing right-continuous function F a premeasure on the algebra $\mathcal{A}_{int}$ generated by left-open, right-closed intervals. We will see afterwards how to extend this premeasure to a measure on the Borel subsets of $\mathbb{R}$.

Proposition 5.20

Let $F : \mathbb{R} \to \mathbb{R}$ be increasing and right-continuous. Define a function $\mu_{F,0} : \mathcal{A}_{int} \to [0, \infty]$ by

$$\mu_{F,0}\left(\bigsqcup_{i=1}^{n}(a_i, b_i]\right) = \sum_{i=1}^{n} (F(b_i) - F(a_i)) .$$

Then $\mu_{F,0}$ is a premeasure on $\mathcal{A}_{int}$.

Proof. We will first show that $\mu_{F,0}$ is a well-defined function on $\mathcal{A}_{int}$ and then prove that it is a premeasure.

Step 1. $\mu_{F,0}$ is well-defined

Every element of $\mathcal{A}_{int}$ can always be written uniquely in the form

$$\bigsqcup_{i=1}^{n}(a_i, b_i]$$

with $-\infty \leq a_1 < b_1 < a_2 < b_2 < \cdots < a_n < b_n \leq \infty$. Indeed, after writing the intervals in increasing order, if $a_{i+1} = b_i$ for some i, then the intervals $(a_i, b_i]$ and $(a_{i+1}, b_{i+1}]$ can be merged into the single interval $(a_i, b_{i+1}]$. This process of merging leaves the expression for $\mu_{F,0}$ unchanged, since if $b_i = a_{i+1}$, we have a telescoping phenomenon

$$(F(b_i) - F(a_i)) + (F(b_{i+1}) - F(a_{i+1})) = F(b_{i+1}) - F(a_i).$$

Thus, the formula for $\mu_{F,0}$ gives the same value for every possible expression of $A \in \mathcal{A}_{int}$ as a disjoint union of left-open, right-closed intervals.

Step 2. If $(a, b] = \bigsqcup_{i=1}^{\infty}(a_i, b_i]$, then $\mu_{F,0}((a, b]) \leq \sum_{i=1}^{\infty} \mu_{F,0}((a_i, b_i])$.

Let $\delta > 0$ and let $\varepsilon_i > 0$ for $i \in \mathbb{N}$. Then $[a+\delta, b]$ is a closed interval covered by the union of open intervals $\bigcup_{i=1}^{\infty}(a_i, b_i+\varepsilon_i)$. By the Heine–Borel theorem (compactness of closed intervals in $\mathbb{R}$), there is a finite subcover $i_1, \ldots, i_n$ such that $[a+\delta, b] \subseteq \bigcup_{j=1}^{n}(a_{i_j}, b_{i_j}+\varepsilon_{i_j})$. Therefore, $(a+\delta, b] \subseteq \bigcup_{j=1}^{n}(a_{i_j}, b_{i_j}+\varepsilon_{i_j}]$, so by Step 1,

$$\mu_{F,0}\left((a+\delta, b]\right) \leq \sum_{j=1}^{n} \mu_{F,0}\left((a_{i_j}, b_{i_j}+\varepsilon_{i_j}]\right) \leq \sum_{i=1}^{\infty} \mu_{F,0}((a_i, b_i+\varepsilon_i]).$$

Letting $\delta \to 0$,

$$\lim_{\delta\to 0^+} \mu_{F,0}\left((a+\delta, b]\right) = F(b) - \lim_{\delta\to 0^+} F(a+\delta) = F(b) - F(a) = \mu_{F,0}((a,b]),$$

since F is right-continuous. Similarly, given $\varepsilon > 0$, we can take ε_i sufficiently small so that $\sum_{i=1}^{\infty}\mu_{F,0}((a_i, b_i+\varepsilon_i]) \leq \sum_{i=1}^{\infty}\mu_{F,0}((a_i, b_i]) + \varepsilon$. Then letting $\varepsilon \to 0$ proves the desired inequality.

Step 3. $\mu_{F,0}$ is countably additive.

Suppose $(A_n)_{n\in\mathbb{N}}$ is a sequence of pairwise disjoint elements of $\mathcal{A}_{int}$, and $A = \bigsqcup_{n\in\mathbb{N}} A_n \in \mathcal{A}_{int}$. Each of the sets A_n, belonging to the algebra $\mathcal{A}_{int}$, can be written in the form $A_n = \bigsqcup_{m=1}^{M_n} S_{n,m}$, where $S_{n,m}$ is a left-open, right-closed interval. We know $\mu_{F,0}(A_n) = \sum_{m=1}^{M_n}\mu_{F,0}(S_{n,m})$ by definition. Replacing $(A_n)_{n\in\mathbb{N}}$ by $(S_{n,m})_{n\in\mathbb{N}, 1\leq m\leq M_n}$, we may assume from the start that A_n is a left-open, right-closed interval for each $n \in \mathbb{N}$.

Since $A \in \mathcal{A}$, we may also write $A = \bigsqcup_{m=1}^{M} S_m$ for some left-open, right-closed intervals S_m. Then

$$\begin{aligned}
\mu_{F,0}(A) &= \sum_{m=1}^{M}\mu_{F,0}(S_m) && \text{(definition of } \mu_{F,0}) \\
&= \sum_{m=1}^{M}\mu_{F,0}\left(\bigsqcup_{n\in\mathbb{N}}(S_m \cap A_n)\right) && (A = \sqcup_{n\in\mathbb{N}}A_n) \\
&\leq \sum_{n,m}\mu_{F,0}(S_m \cap A_n) && \text{(Step 2)} \\
&= \sum_{n=1}^{\infty}\mu_{F,0}\left(\bigsqcup_{m=1}^{M}(S_m \cap A_n)\right) && \text{(definition of } \mu_{F,0}) \\
&= \sum_{n=1}^{\infty}\mu_{F,0}(A_n) && \text{(definition of } \mu_{F,0})
\end{aligned}$$

On the other hand, for $N \in \mathbb{N}$,

$$\sum_{n=1}^{N}\mu_{F,0}(A_n) = \mu_{F,0}\left(\bigsqcup_{n=1}^{N} A_n\right) \leq \mu_{F,0}(A),$$

so taking a limit as $N \to \infty$, $\sum_{n=1}^{\infty}\mu_{F,0}(A_n) \leq \mu_{F,0}(A)$.

□

The next stage in the construction is to extend the premeasure $\mu_{F,0}$ to an *outer measure* defined on all subsets of $\mathbb{R}$.

Definition 5.21

Let X be a set. A function $\mu^* : \mathscr{P}(X) \to [0, \infty]$ is an *outer measure* if

- $\mu^*(\emptyset) = 0$;
- MONOTONE: if $A \subseteq B$, then $\mu^*(A) \leq \mu^*(B)$; and
- COUNTABLY SUBADDITIVE: for any sequence of sets $(A_n)_{n\in\mathbb{N}}$, one has $\mu^*\left(\bigcup_{n\in\mathbb{N}} A_n\right) \leq \sum_{n=1}^{\infty} \mu^*(A_n)$.

A premeasure can always be extended to an outer measure, as shown by the following proposition.

Proposition 5.22

Let $\mathcal{A}$ be an algebra on a set X, and suppose $\mu_0 : \mathcal{A} \to [0, \infty]$ is a premeasure. Then

$$\mu^*(E) = \inf\left\{\sum_{n=1}^{\infty} \mu_0(A_n) : E \subseteq \bigcup_{n\in\mathbb{N}} A_n, A_n \in \mathcal{A}\right\}$$

defines an outer measure on X with $\mu^*|_{\mathcal{A}} = \mu_0$.

PROOF. Let us check the properties one at a time.

First, $\emptyset \in \mathcal{A}$, so $\mu^*(\emptyset) \leq \mu_0(\emptyset) = 0$.

Next, suppose $A \subseteq B$. Then any set containing B also contains A, so the expression defining $\mu^*(A)$ involves an infimum over a larger collection than the expression defining $\mu^*(B)$. Hence, $\mu^*(A) \leq \mu^*(B)$.

Now let us prove countable subadditivity. Let $(A_n)_{n\in\mathbb{N}}$ be a sequence of subsets of X. If

$$\sum_{n=1}^{\infty} \mu^*(A_n) = \infty,$$

there is nothing to show, so assume

$$\sum_{n=1}^{\infty} \mu^*(A_n) < \infty.$$

Let $\varepsilon > 0$. For each n, let $(A_{n,k})_{k\in\mathbb{N}}$ be a sequence of elements of $\mathcal{A}$ such that $A_n \subseteq \bigcup_{k\in\mathbb{N}} A_{n,k}$ and

$$\sum_{k=1}^{\infty} \mu_0(A_{n,k}) < \mu^*(A_n) + \frac{\varepsilon}{2^n}.$$

Then $(A_{n,k})_{n,k\in\mathbb{N}}$ is a countable family of elements of the algebra $\mathcal{A}$, and $\bigcup_{n\in\mathbb{N}} A_n \subseteq \bigcup_{n,k\in\mathbb{N}} A_{n,k}$, so

$$\mu^*\left(\bigcup_{n\in\mathbb{N}} A_n\right) \leq \sum_{n,k} \mu_0(A_{n,k}) \leq \sum_{n=1}^{\infty}\left(\mu^*(A_n) + \frac{\varepsilon}{2^n}\right) = \sum_{n=1}^{\infty} \mu^*(A_n) + \varepsilon.$$

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

Finally, let us show $\mu^*|_{\mathcal{A}} = \mu_0$. Let $A \in \mathcal{A}$. Then by definition $\mu^*(A) \leq \mu_0(A)$. It remains to show $\mu^*(A) \geq \mu_0(A)$. Let $(A_n)_{n\in\mathbb{N}}$ be a sequence in $\mathcal{A}$ such that $A \subseteq \bigcup_{n\in\mathbb{N}} A_n$. Define a new sequence $(B_n)_{n\in\mathbb{N}}$ by $B_1 = A \cap A_1$ and $B_n = A \cap A_n \setminus (A_1 \cup \cdots \cup A_{n-1})$. Since $\mathcal{A}$ is an algebra, the sets B_n belong to $\mathcal{A}$. Moreover, $(B_n)_{n\in\mathbb{N}}$ is a sequence of pairwise disjoint sets whose union is $A \in \mathcal{A}$, so

$$\mu_0(A) = \sum_{n=1}^{\infty} \mu_0(B_n) = \sum_{n=1}^{\infty} (\mu_0(B_n) + \mu_0(A_n \setminus B_n)) \leq \sum_{n=1}^{\infty} \mu_0(A_n).$$

Taking an infimum over all such collections $(A_n)_{n\in\mathbb{N}}$ gives the desired inequality $\mu_0(A) \leq \mu^*(A)$. □

The outer measure μ_F^* obtained from the premeasure $\mu_{F,0}$ is not in general a measure on $\mathscr{P}(X)$. The problem is that, while μ_F^* is *subadditive*, it may fail to be additive. In order to obtain a measure, we restrict to the sets with better additive behavior.

Definition 5.23

Let μ^* be an outer measure on a set X. A set $E \subseteq X$ is *μ^*-measurable* if for every $A \subseteq X$,

$$\mu^*(A) = \mu^*(A \cap E) + \mu^*(A \setminus E). \tag{5.1}$$

Remark. Outer measures are subadditive, so (5.1) is equivalent to the *a priori* weaker inequality

$$\mu^*(A) \geq \mu^*(A \cap E) + \mu^*(A \setminus E).$$

Theorem 5.24: Carathéodory's Theorem

Let μ^* be an outer measure on a set X. Let $\mathcal{M} \subseteq \mathscr{P}(X)$ be the family of μ^*-measurable sets. Then $\mathcal{M}$ is a σ-algebra, and $\mu^*|_{\mathcal{M}}$ is a complete measure.

Proof. We break the proof into several steps.

Claim 1. $X \in \mathcal{M}$.

Given $A \subseteq X$, we have $\mu^*(A \cap X) + \mu^*(A \setminus X) = \mu^*(A) + \mu^*(\emptyset) = \mu^*(A)$.

Claim 2. $\mathcal{M}$ is closed under complementation.

Rewriting $A \setminus E = A \cap E^c$, the measurability condition (5.1) is symmetric in E and E^c.

Claim 3. $\mathcal{M}$ is closed under finite unions.

Suppose $E, F \in \mathcal{M}$, and let $A \subseteq X$. We want to show

$$\mu^*(A) \geq \mu^*(A \cap (E \cup F)) + \mu^*(A \setminus (E \cup F)).$$

Writing $A \cap (E \cup F) = (A \cap E) \cup (A \cap F \cap E^c)$ and $A \setminus (E \cup F) = A \cap F^c \cap E^c$ and applying subadditivity of μ^*, we have

$$\mu^*(A \cap (E \cup F)) + \mu^*(A \setminus (E \cup F)) \leq \mu^*(A \cap E) + \underbrace{\mu^*(A \cap F \cap E^c) + \mu^*(A \cap F^c \cap E^c)}_{\mu^*(A \cap E^c)} = \mu^*(A).$$

Claims 1–3 show that $\mathcal{M}$ is an algebra. The next claim upgrades $\mathcal{M}$ to a σ-algebra and proves that $\mu^*|_{\mathcal{M}}$ is a measure.

Claim 4. $\mathcal{M}$ is closed under countable disjoint unions, and $\mu^*|_{\mathcal{M}}$ is countably additive.

Suppose $(E_n)_{n\in\mathbb{N}}$ be a sequence of pairwise disjoint sets in $\mathcal{M}$, and let $E = \bigsqcup_{n\in\mathbb{N}} E_n$. Let $A \subseteq X$. As in Claim 3, we want to show

$$\mu^*(A) \geq \mu^*(A \cap E) + \mu^*(A \setminus E). \tag{5.2}$$

If $\mu^*(A) = \infty$, there is nothing to check, so assume $\mu^*(A) < \infty$. Let $F_N = \bigsqcup_{n=1}^{N} E_n$. By Claim 3, we have

$$\mu^*(A \cap F_N) = \sum_{n=1}^{N} \mu^*(A \cap E_n).$$

Hence, by countable subadditivity of μ^*,

$$\mu^*(A \cap E) \leq \sum_{n=1}^{\infty} \mu^*(A \cap E_n) = \lim_{N\to\infty} \mu^*(A \cap F_N).$$

For fixed $N \in \mathbb{N}$, $F_N \in \mathcal{M}$ by Claim 3, so

$$\mu^*(A) = \mu^*(A \cap F_N) + \mu^*(A \setminus F_N) \geq \mu^*(A \cap F_N) + \mu^*(A \setminus E).$$

Taking a limit as $N \to \infty$ gives (5.2).

Note that we actually proved the stronger inequality

$$\mu^*(A) \geq \sum_{n=1}^{\infty} \mu^*(A \cap E_n) + \mu^*(A \setminus E).$$

Taking $A = E$ establishes countable additivity of μ^*.

Finally, we check that $(X, \mathcal{M}, \mu^*|_{\mathcal{M}})$ is complete.

Claim 5. If $N \subseteq X$ and $\mu^*(N) = 0$, then $N \in \mathcal{M}$.

Let $A \subseteq X$. Then by monotonicity,

$$\mu^*(A \cap N) + \mu^*(A \setminus N) \leq \mu^*(N) + \mu^*(A) = \mu^*(A).$$

□

The last remaining piece to tie everything together is relating the σ-algebra $\mathcal{M}$ in Carathéodory's theorem to the algebra on which the premeasure μ_0 was defined.

Lemma 5.25

Let $\mathcal{A}$ be an algebra on a set X. Let $\mu_0 : \mathcal{A} \to [0, \infty]$ be a premeasure, and let μ^* be the outer measure extending μ_0 as in Proposition 5.22. Then every element of $\mathcal{A}$ is μ^*-measurable.

Proof. Let $A \in \mathcal{A}$, and let $B \subseteq X$ be an arbitrary set. We want to show

$$\mu^*(B) \geq \mu^*(B \cap A) + \mu^*(B \setminus A). \tag{5.3}$$

Let $(A_n)_{n\in\mathbb{N}}$ be family of elements of $\mathcal{A}$ such that $B \subseteq \bigcup_{n\in\mathbb{N}} A_n$. Then

$$\sum_{n=1}^{\infty} \mu_0(A_n) = \sum_{n=1}^{\infty} (\mu_0(A_n \cap A) + \mu_0(A_n \setminus A)) = \sum_{n=1}^{\infty} \mu_0(A_n \cap A) + \sum_{n=1}^{\infty} \mu_0(A_n \setminus A).$$

The union $\bigcup_{n\in\mathbb{N}}(A_n \cap A)$ contains $B \cap A$, and similarly, $\bigcup_{n\in\mathbb{N}}(A_n \setminus A)$ contains $B \setminus A$, so by the definition of the outer measure μ^*,

$$\sum_{n=1}^{\infty} \mu_0(A_n) \geq \mu^*(B \cap A) + \mu^*(B \setminus A).$$

Taking an infimum over all such collections $(A_n)_{n\in\mathbb{N}}$ gives (5.3). □

Putting everything together, we get the following theorem.

Theorem 5.26: Hahn–Kolmogorov Extension Theorem

Let $\mathcal{A} \subseteq \mathscr{P}(X)$ be an algebra on a set X, and let $\mu_0 : \mathcal{A} \to [0, \infty]$ be a premeasure. Then μ_0 extends to a complete measure $\mu : \mathcal{M} \to [0, \infty]$ defined on a σ-algebra $\mathcal{M} \supseteq \sigma(\mathcal{A})$. Moreover, if μ_0 is σ-finite, then the extension of μ_0 to $\sigma(\mathcal{A})$ is unique, and μ is the completion of this unique extension.

Proof. By Proposition 5.22, μ_0 extends to an outer measure $\mu^* : \mathscr{P}(X) \to [0, \infty]$. Let $\mathcal{M}$ be the σ-algebra of μ^*-measurable sets, and let $\mu = \mu^*|_{\mathcal{M}}$. Then μ is a complete measure by Carathéodory's theorem (Theorem 5.24), and by Lemma 5.25, $\mathcal{A} \subseteq \mathcal{M}$.

Uniqueness in the σ-finite case is a consequence of the π-λ theorem and follows on exactly the same lines as the proof of Corollary 5.14. □

Remark. When μ_0 is not σ-finite, it may have several different extensions to $\sigma(\mathcal{A})$. The outer measure construction is the maximal such extension in the sense that given any other extension $\nu : \sigma(\mathcal{A}) \to [0, \infty]$ of μ_0, one has $\nu(E) \leq \mu(E)$ for every $E \in \sigma(\mathcal{A})$. We will return to this subject in the context of product measures later in the course, where it will sometimes be useful to work with a different extension of a premeasure than the one obtained by Carathéodory's theorem.

Corollary 5.27: Existence of Lebesgue–Stieltjes Measures

Let $F : \mathbb{R} \to \mathbb{R}$ be increasing and right-continuous with $F(0) = 0$. There exists a σ-algebra $\mathcal{M}_F$ containing the Borel subsets of $\mathbb{R}$ and a complete measure $\mu_F : \mathcal{M}_F \to [0, \infty]$ with distribution function $F_{\mu_F} = F$.

Proof. Let $\mu_{F,0}$ be the premeasure on $\mathcal{A}_{int}$ given by Proposition 5.20. Then the extension μ_F given by the Hahn–Kolmogorov extension theorem (Theorem 5.26) is a complete measure with distribution function F. □

4. Lebesgue Measure

Definition 5.28

The *Lebesgue measure* on $\mathbb{R}$ is the Lebesgue–Stieltjes measure associated to the distribution function $F(x) = x$.

Proposition 5.29

Let λ be the Lebesgue measure on $\mathbb{R}$, and let $\mathcal{M}$ be the σ-algebra of Lebesgue measurable sets.

(1) translation-invariance: $\lambda(E + t) = \lambda(E)$ for every $E \in \mathcal{M}$ and $t \in \mathbb{R}$;

(2) REFLECTION-INVARIANCE: $\lambda(-E) = \lambda(E)$ for every $E \in \mathcal{M}$;
(3) DILATION PROPERTY: $\lambda(tE) = |t|\lambda(E)$ for every $E \in \mathcal{M}$ and $t \in \mathbb{R}$;

PROOF. Compute the distribution function of the transformed measure and apply uniqueness of Lebesgue–Stieltjes measures (Corollary 5.14). We leave the details as an exercise. □

Using the translation-invariance property of the Lebesgue measure, we can prove the existence of a non-measurable set.

THEOREM 5.30

There exists a Lebesgue non-measurable subset of $\mathbb{R}$.

PROOF. Define an equivalence relation on $[0, 1)$ by $x \sim y$ if $y - x \in \mathbb{Q}$. By the axiom of choice, let $E \subseteq [0, 1)$ be a set containing exactly one representative of each equivalence class. For each $t \in \mathbb{Q} \cap [0, 1)$, let $E_t = \{x + t \bmod 1 : x \in E\} \subseteq [0, 1)$.

CLAIM 1. The sets $(E_t)_{t\in\mathbb{Q}\cap[0,1)}$ are pairwise disjoint.

For $t, s \in \mathbb{Q} \cap [0, 1)$ and $x, y \in E$, if $x + t \equiv y + s \pmod 1$, then
$$y - x \equiv t - s \pmod 1,$$
so $x \sim y$. But E contains only one element from each equivalence class, so $x = y$ and $t = s$.

CLAIM 2. $\bigsqcup_{t\in\mathbb{Q}\cap[0,1)} E_t = [0, 1)$

Let $x \in [0, 1)$. Then there exists $y \in E$ with $y \sim x$, since E has a representative of each equivalence class. Let $t = x - y \bmod 1 \in \mathbb{Q} \cap [0, 1)$. Then
$$y + t \equiv x \pmod 1).$$
so $x \in E_t$.

Assume for contradiction that E is Lebesgue measurable.

CLAIM 3. For every $t \in \mathbb{Q} \cap [0, 1)$, E_t is Lebesgue measurable and $\lambda(E_t) = \lambda(E)$.

We can write
$$E_t = ((E + t) \cap [0, 1)) \sqcup ((E + t) \cap [1, 2) - 1).$$
Therefore, by translation invariance of the Lebesgue measure,
$$\lambda(E_t) = \lambda(E + t) = \lambda(E).$$

Combining Claims 1–3 and using countable additivity of the Lebesgue measure,
$$1 = \lambda([0, 1)) = \sum_{t\in\mathbb{Q}\cap[0,1)} \lambda(E_t) = \sum_{t\in\mathbb{Q}\cap[0,1)} \lambda(E) = \infty \cdot \lambda(E).$$
There is no value of $\lambda(E)$ that can satisfy this equation. We have thus reached a contradiction, so E is non-measurable. □

Remark. The axiom of choice plays a crucial role in Theorem 5.30. Using the set-theoretic notion of an *inaccessible cardinal*, Robert Solovay constructed a model of set theory under the ZF axioms without choice in which every subset of $\mathbb{R}$ is Lebesgue measurable [**5**].

5. Cantor Measure

Another interesting example of a Borel measure on the real line is the "uniform" measure on the middle-thirds Cantor set, which we will construct now. Recall that the Cantor set $C \subseteq [0,1]$ is obtained by starting with the full interval $[0,1]$ and iteratively removing the middle third of each remaining interval at each step. We can therefore write $[0,1] \setminus C = \bigcup_{n=0}^{\infty} \bigcup_{k=0}^{2^n-1} I_{n,k}$, where $I_{n,0}, \ldots, I_{n,2^{n-1}}$ is an enumeration of the removed intervals of length $3^{-(n+1)}$ in increasing order. How should we define the distribution function for a uniform measure on C? Well, after step $n = 0$, where we remove the interval $(\frac{1}{3}, \frac{2}{3})$, we have half of of the Cantor set to the left of this interval and the other half to the right, so the distribution function should take the value $\frac{1}{2}$ on the entirety of this interval. Arguing similarly, the distribution function should take the value $\frac{2k+1}{2^{n+1}}$ on the interval $I_{n,k}$ for each $n \geq 0$ and $k \in \{0, 1, \ldots, 2^n - 1\}$. Since the union of the intervals $\bigcup_{n=0}^{\infty} \bigcup_{k=0}^{2^n-1} I_{n,k}$ are dense in $[0,1]$, there is a unique way of interpolating between the values on the intervals $I_{n,k}$ in order to obtain a continuous function. We call this continuous function the *Cantor function* and its associated Lebesgue–Stieltjes measure the *Cantor measure*.

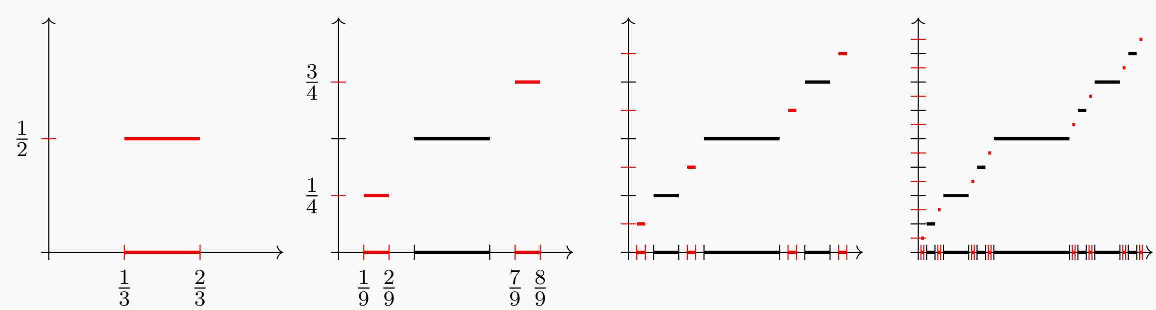


Figure 5.1. Iterative construction of the Cantor function

A more explicit description of the Cantor set is the collection of numbers in the interval $[0,1]$ whose binary expansion consists entirely of the digits 0 and 2. From this description of the Cantor set, we can also obtain a formula for the Cantor function, namely

$$c(x) = \begin{cases} \sum_{j=1}^{\infty} \frac{a_j/2}{2^j}, & \text{if } x = \sum_{j=1}^{\infty} \frac{a_j}{3^j} \in C; \\ \sup_{y \leq x, y \in C} c(y), & \text{if } x \notin C. \end{cases}$$

The Cantor measure has a surprising combination of properties. Suppose you have a perfectly fair coin to flip, and you record the sequence of heads and tails as you flip the coin repeatedly. Recording heads as the digit 2 and tails as the digit 0, this sequence of coin flips produces a random element of the Cantor set in terms of its base 3 expansion. The distribution of this random element of the Cantor set is described by the Cantor measure.

The Cantor function is continuous, and so by Exercise 5.6, the Cantor measure is a continuous probability measure. This is despite the fact that all of the mass of the Cantor measure is concentrated on the Cantor set, which is a set of Lebesgue measure zero! This makes the Cantor measure an example of what is called a *singular measure*, and as a result, even though we have given a reasonable probabilistic method for constructing Cantor-distributed random variables, the Cantor measure does not have a probability density function. That is, there is no function $f : [0,1] \to [0, \infty)$

for which we can express $\mu_C((a,b]) = \int_a^b f(x)\,dx$. We will reencounter singular measures and deal with them systematically later in the course.

6. Regularity of Lebesgue–Stieltjes Measures

Built into the definition of Lebesgue–Stieltjes measures is the fact that they are locally finite, but it is not at all obvious that Lebesgue–Stieltjes measures should have additional regularity properties. However, from the outer measure construction, we can quickly deduce several useful and nontrivial properties about Lebesgue–Stieltjes measures.

Proposition 5.31

Let μ be a Lebesgue–Stieltjes measure on $\mathbb{R}$, and let $\mathcal{M}_\mu$ be the σ-algebra of μ-measurable sets.

(1) Let $E \subseteq \mathbb{R}$. The following are equivalent:
   (i) $E \in \mathcal{M}_\mu$;
   (ii) for any $\varepsilon > 0$, there exists a closed set F and an open set G such that $F \subseteq E \subseteq G$ and $\mu(G \setminus F) < \varepsilon$;
   (iii) there exists an F_σ set A and a G_δ set B such that $A \subseteq E \subseteq B$ and $\mu(B \setminus A) = 0$.

(2) OUTER REGULARITY: If $E \in \mathcal{M}_\mu$, then
$$\mu(E) = \inf\{\mu(U) : U \supseteq E \text{ open}\}.$$

(3) INNER REGULARITY: If $E \in \mathcal{M}_\mu$, then
$$\mu(E) = \sup\{\mu(K) : K \subseteq E \text{ compact}\}.$$

Proof. We will first prove outer regularity (2), then prove the measurability conditions (1), and end with inner regularity (3).

(2) Let $E \in \mathcal{M}_\mu$. By monotonicity of μ, it suffices to show $\mu(E) \geq \inf\{\mu(U) : U \supseteq E \text{ open}\}$. As usual, if $\mu(E) = \infty$, there is nothing to check, so assume $\mu(E) < \infty$, and let $\varepsilon > 0$. From the outer measure construction, we have
$$\mu(E) = \inf\left\{\sum_{n=1}^{\infty} \mu((a_n, b_n]) : E \subseteq \bigcup_{n\in\mathbb{N}} (a_n, b_n]\right\}.$$
Hence, there exists a family of left-open, right-closed intervals $((a_n, b_n])_{n\in\mathbb{N}}$ such that $E \subseteq \bigcup_{n\in\mathbb{N}} (a_n, b_n]$ and $\sum_{n=1}^{\infty} \mu((a_n, b_n]) < \mu(E) + \frac{\varepsilon}{2}$. For each $n \in \mathbb{N}$, let $\delta_n > 0$ such that $\mu((a_n, b_n + \delta_n)) < \mu((a_n, b_n]) + 2^{-(n+1)}\varepsilon$. Such δ_n exists by continuity of μ from above. Then for the open set $U = \bigcup_{n\in\mathbb{N}} (a_n, b_n + \delta_n)$, we have
$$\mu(U) \leq \sum_{n=1}^{\infty} \mu((a_n, b_n + \delta_n)) < \sum_{n=1}^{\infty} (\mu((a_n, b_n]) + 2^{-(n+1)}\varepsilon) < \mu(E) + \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \mu(E) + \varepsilon.$$
But $\varepsilon > 0$ was arbitrary, so we are done.

(1) We will prove the chain of implications (i) $\Longrightarrow$ (ii) $\Longrightarrow$ (iii) $\Longrightarrow$ (i).

(i) $\Longrightarrow$ (ii). Suppose $E \in \mathcal{M}_\mu$, and let $\varepsilon > 0$. Let $E_n = E \cap (n, n+1]$ for $n \in \mathbb{Z}$. By (2), there exists an open set $G_n \supseteq E_n$ such that $\mu(G_n) < \mu(E_n) + 2^{-|n|}\frac{\varepsilon}{6}$. Let $G = \bigcup_{n\in\mathbb{Z}} G_n$. Then G is open, $E \subseteq G$, and
$$\mu(G \setminus E) \leq \sum_{n=-\infty}^{\infty} \mu(G_n \setminus E_n) < \frac{\varepsilon}{2}.$$

Applying the same argument to E^c, we find an open set $U \subseteq \mathbb{R}$ such that $E^c \subseteq U$ and $\mu(U \setminus E^c) < \frac{\varepsilon}{2}$. Let $F = U^c$. Then F is closed, $F \subseteq E$, and $\mu(E \setminus F) = \mu(U \setminus E^c) < \frac{\varepsilon}{2}$. Therefore, $\mu(G \setminus F) \leq \mu(G \setminus E) + \mu(E \setminus F) < \varepsilon$.

(ii) $\Longrightarrow$ (iii). For each $n \in \mathbb{N}$, choose $F_n \subseteq E \subseteq G_n$ such that $\mu(G_n \setminus F_n) < \frac{1}{n}$ by (ii). Let $A = \bigcup_{n\in\mathbb{N}} F_n$ and $B = \bigcap_{n\in\mathbb{N}} G_n$. Then A is an F_σ set, B is a G_δ set, $A \subseteq E \subseteq B$, and $\mu(B \setminus A) = 0$.

(iii) $\Longrightarrow$ (i). Suppose (iii) holds. Then we may write $E = A \cup N$, where A is an F_σ set and $N = E \setminus A \subseteq B \setminus A$. Since μ is complete and $\mathrm{Borel}(\mathbb{R}) \subseteq \mathcal{M}_\mu$, we have $E \in \mathcal{M}_\mu$.

(3) Let $E \in \mathcal{M}_\mu$, and let $\varepsilon > 0$. Then by (1), there exists a closed set $F \subseteq E$ such that $\mu(E \setminus F) < \varepsilon$. Let $F_n = F \cap [-n, n]$ for $n \in \mathbb{N}$. Then F_n is compact, $F_1 \subseteq F_2 \subseteq \dots$, and $F = \bigcup_{n\in\mathbb{N}} F_n$. Therefore, by continuity of μ from below, $\mu(F_n) \to \mu(F)$ as $n \to \infty$. Hence, $\sup\{\mu(K) : K \subseteq E \text{ compact}\} \geq \mu(F) \geq \mu(E) - \varepsilon$. □

Additional Reading

The construction of Lebesgue–Stieltjes measures presented in this chapter is along the same lines as in the book of Folland [**1**, Sections 1.4 and 1.5]. A similar approach is taken in [**6**, Section 3.3] and [**7**, Sections 1.7.1–1.7.3].

Exercises

5.1 Confirm that the examples given in Example 5.7 are π-systems or λ-systems as claimed.

5.2 Let X be a set. Suppose $\mathcal{P}$ is a π-system containing X. Let $\mathcal{F}$ be a family of functions from X to $\mathbb{R}$ with the following three properties

(1) for every $E \in \mathcal{P}$, one has $\mathbb{1}_E \in \mathcal{F}$;
(2) the space of functions $\mathcal{F}$ is a real vector space: if $f, g \in \mathcal{F}$ and $c \in \mathbb{R}$, then $cf + g \in \mathcal{F}$;
(3) if $0 \leq f_1 \leq f_2 \leq \dots$ is an increasing sequence in $\mathcal{F}$ and $f = \lim_{n\to\infty} f_n$ is bounded, then $f \in \mathcal{F}$.

Prove that $\mathcal{F}$ contains every bounded $\sigma(\mathcal{P})$-measurable function.

5.3 Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, and let $X : \Omega \to \mathbb{R}$ be a measurable function (called a *random variable* in probability theory). Prove that the cumulative distribution function $F_X : \mathbb{R} \to [0, 1]$ defined by $F(x) = \mathbb{P}(X \leq x)$ determines the measure induced by X, $\mathbb{P}_X : \mathrm{Borel}(\mathbb{R}) \to [0, 1]$ defined by $\mathbb{P}_X(A) = \mathbb{P}(X^{-1}(A))$.

5.4 Let $\lambda : \mathrm{Borel}(\mathbb{R}) \to [0, \infty]$ be the Lebesgue measure on the Borel subsets of $\mathbb{R}$. Let μ be another Borel measure on $\mathbb{R}$ such that

(1) for all $A \in \mathrm{Borel}(\mathbb{R})$ and $x \in \mathbb{R}$, $\mu(A + x) = \mu(A)$, and
(2) $0 < \mu((0, 1]) < \infty$.

Show that there is a number $\alpha > 0$ such that $\mu = \alpha\lambda$.

5.5 Let λ be the Lebesgue measure on $\mathbb{R}$, and let E be a Lebesgue measurable set. Prove: if $\lambda(E) > 0$, then E contains a non-measurable set.

5.6 Let μ be a locally finite Borel measure on $\mathbb{R}$ with distribution function F. For $x \in \mathbb{R}$, define $F(x^-) = \lim_{y\to x^-} F(y)$. The limit exists since F is an increasing function.

(a) Show that $\mu(\{x\}) = F(x) - F(x^-)$.
(b) Deduce that μ is a continuous measure if and only if F is a continuous function.

5.7 A family of sets $\mathcal{C} \subseteq \mathscr{P}(X)$ is called a *monotone class* if it is closed under countable monotone unions and intersections. That is,

- if $(A_n)_{n\in\mathbb{N}}$ is a sequence in $\mathcal{C}$ and $A_1 \subseteq A_2 \subseteq \dots$, then $\bigcup_{n\in\mathbb{N}} A_n \in \mathcal{C}$, and
- if $(B_n)_{n\in\mathbb{N}}$ is a sequence in $\mathcal{C}$ and $B_1 \supseteq B_2 \supseteq \dots$, then $\bigcap_{n\in\mathbb{N}} B_n \in \mathcal{C}$.

The goal of this problem is to prove the monotone class theorem.

(a) Show that a family of sets is a σ-algebra if and only if it is both an algebra and a monotone class.

(b) Show that the intersection of a family of monotone classes is a monotone class.

(c) Let $\mathcal{C}$ be a monotone class. Show that $\mathcal{C}' = \{E \subseteq X : X \setminus E \in \mathcal{C}\}$ is also a monotone class.

(d) Let $\mathcal{C}$ be a monotone class. Show that $\mathcal{C}_E = \{F \subseteq X : E \cup F \in \mathcal{C}\}$ is also a monotone class for every $E \subseteq X$.

Let $\mathcal{A} \subseteq \mathscr{P}(X)$ be an algebra on X, and let $\mathcal{C}(\mathcal{A})$ be the monotone class generated by $\mathcal{A}$. (This monotone class is well-defined by part (b).)

(e) Use part (c) to show that $\mathcal{C}(\mathcal{A})$ is closed under complementation.

(f) Use part (d) to show that $\mathcal{C}(\mathcal{A})$ is closed under finite unions.

(g) Deduce the monotone class theorem: $\mathcal{C}(\mathcal{A}) = \sigma(\mathcal{A})$.

CHAPTER 6

Borel Measures on Locally Compact Hausdorff Spaces

In the previous chapter, we constructed locally finite Borel measures on the real line. This is already sufficient for many purposes in probability theory, where the structure of the underlying measure space is often insignificant and the main object of study is (real-valued) random variables. However, in other contexts, the underlying structure of the measure space may play a prominent role (for example, if one is interested in measures on manifolds), and for this, we need additional tools to construct measures on more general topological spaces. The goal of this section is to construct Borel measures on locally compact Hausdorff spaces that are "compatible with the topology" in a sense that will be made precise below.

1. Locally Compact Hausdorff Spaces

Definition 6.1

A topological space X is

- *Hausdorff* if every pair of points can be separated by open sets: if $x, y \in X$ and $x \neq y$, then there are open set $U \ni x$ and $V \ni y$ such that $U \cap V = \emptyset$;
- *locally compact* if for every point has a compact neighborhood: for $x \in X$, there is an open set U and a compact set K such that $x \in U \subseteq K$.

If X is both locally compact and Hausdorff, we say X is a *locally compact Hausdorff space* or an *LCH space* for short.

Example 6.2

Examples of locally compact Hausdorff spaces include:

- the unit interval $[0, 1]$
- the middle-thirds Cantor set
- Euclidean space $\mathbb{R}^d$ for $d \in \mathbb{N}$
- topological manifolds
- discrete spaces

Non-examples include:

- the rational numbers $\mathbb{Q}$ (not locally compact)
- infinite-dimensional real or complex vector spaces (not locally compact)
- an infinite set with the co-finite topology (not Hausdorff)

Definition 6.3

Let X be an LCH space and $f : X \to \mathbb{C}$ a continuous function. The *support of f* is the set

$$\operatorname{supp}(f) = \overline{\{f \neq 0\}}.$$

We say that f is *compactly supported* if $\operatorname{supp}(f)$ is a compact subset of X.

NOTATION. We denote the space of compactly supported continuous functions on a topological space X by $C_c(X)$.

2. Radon Measures and the Riesz Representation Theorem

DEFINITION 6.4

Let X be an LCH space. A Borel measure μ on X is a *Radon measure* if μ is

- LOCALLY FINITE: $\mu(K) < \infty$ for every compact $K \subseteq X$;
- OUTER REGULAR: for every Borel set $E \subseteq X$,
$$\mu(E) = \inf\{\mu(U) : U \text{ is open and } E \subseteq U\};$$
and
- INNER REGULAR ON OPEN SETS: for every open set $G \subseteq X$,
$$\mu(G) = \sup\{\mu(K) : K \text{ is compact and } K \subseteq G\}.$$

Let X be an LCH space, and suppose μ is a Radon measure on X. Given $f \in C_c(X)$, we have
$$\int_X |f| \, d\mu \leq \sup_{x \in X} |f(x)| \cdot \mu(\text{supp}(f)).$$
The quantity $\sup_{x \in X} |f(x)|$ is actually a maximum and is finite by the extreme value theorem, while $\mu(\text{supp}(f)) < \infty$ since μ is locally finite. Hence, $C_c(X) \subseteq L^1(\mu)$. Integration against the measure μ thus induces a *positive linear functional* on $C_c(X)$.

DEFINITION 6.5

A *linear functional* on $C_c(X)$ is a linear map $\varphi : C_c(X) \to \mathbb{C}$. We say that a linear functional $\varphi : C_c(X) \to \mathbb{C}$ is *positive* if $\varphi(f) \geq 0$ for every $f \in C_c(X)$ with $f \geq 0$.

It turns out that all positive linear functionals on $C_c(X)$ arise via integration against a measure.

THEOREM 6.6: RIESZ REPRESENTATION THEOREM

Let X be a locally compact Hausdorff space. Given a positive linear functional $\varphi : C_c(X) \to \mathbb{C}$, there exists a unique Radon measure μ such that
$$\varphi(f) = \int_X f \, d\mu \tag{6.1}$$
for every $f \in C_c(X)$.

3. Topological Lemmas

LEMMA 6.7

Let X be an LCH space. Suppose $K \subseteq X$ is compact, $U \subseteq X$ is open, and $K \subseteq U$. Then there exists an open set $V \subseteq X$ such that $\overline{V}$ is compact and $K \subseteq V \subseteq \overline{V} \subseteq U$.

PROOF. We first handle the case $K = \{x\}$. Since X is locally compact, there is an open neighborhood $W \subseteq X$ such that $x \in W$ and $\overline{W}$ is compact. If $\overline{W} \subseteq U$, then we are done. Suppose $\overline{W} \not\subseteq U$. Then $L = \overline{W} \setminus U$ is a compact set, and $x \notin L$. For each point $y \in L$, let V_y and O_y be open sets such that $x \in V_y$, $y \in O_y$, and $V_y \cap O_y = \emptyset$. (The sets V_y and O_y

exist, since X is Hausdorff.) By compactness of L, there is a finite collection $y_1, \dots, y_n \in L$ such that $L \subseteq \bigcup_{j=1}^n O_{y_j}$. Let $V = W \cap \bigcap_{j=1}^n V_{y_j}$. Then V is open, $x \in V$, and $\overline{V} \subseteq \overline{W}$ is compact. Moreover, if $z \in \overline{V}$, then $z \in \overline{V}_{y_j}$ for every $j \in \{1, \dots, n\}$. Hence, $z \notin O_{y_j}$, so $z \notin L$. Therefore, $\overline{V} \subseteq \overline{W} \setminus L \subseteq U$.

Suppose now that K is an arbitrary compact set. By the above, we may find open sets V_x, $x \in K$, such that $x \in V_x \subseteq \overline{V}_x \subseteq U$. By compactness, there is a finite subcover $x_1, \dots, x_n \in K$ such that $K \subseteq \bigcup_{j=1}^n V_{x_j}$. We then take $V = \bigcup_{j=1}^n V_{x_j}$. □

Lemma 6.8: Urysohn's Lemma for LCH Spaces

Let X be an LCH space. Given a compact set $K \subseteq X$ and an open set $U \subseteq X$ with $K \subseteq U$, there exists a compactly supported continuous function $f : X \to [0,1]$ such that $f = 1$ on K and $\operatorname{supp}(f) \subseteq U$.

Proof. Let K be a compact subset of X and $U \subseteq X$ an open set with $K \subseteq U$. By Lemma 6.7, let V be an open set such that $\overline{V}$ is compact and $K \subseteq V \subseteq \overline{V} \subseteq U$. We construct a function f supported on $\overline{V}$ in terms of its sub-level sets.

Let $K(1) = K$, $V(0) = V$. Put $V(1) = \emptyset$ and $K(0) = \overline{V}$.

Claim 1. There are families of open sets $V(r)$ and compact sets $K(r)$ indexed by dyadic rationals $r \in [0,1]$ such that

- for every dyadic rational $r \in [0,1]$, $V(r) \subseteq K(r)$, and
- for dyadic rationals $r, s \in [0,1]$, if $r > s$, then $K(r) \subseteq V(s)$.

We will prove the claim by induction on the denominators of dyadic rationals. By Lemma 6.7, let $V(1/2)$ be an open set such that $K(1/2) = \overline{V(1/2)}$ is compact and $K = K(1) \subseteq V(1/2) \subseteq K(1/2) \subseteq V(0) = V$.

Suppose we have constructed sets $V(r)$ and $K(r)$ with the desired properties for dyadic rationals $r \in (0,1)$ with denominators 2^n for $n < N$. Let $r \in (0,1)$ by a dyadic rational with denominator 2^N, say $r = \frac{2j-1}{2^N}$, $j \in \{1, \dots, 2^{N-1}\}$. By the induction hypothesis, we have a compact set $K\left(\frac{j}{2^{N-1}}\right)$ and an open set $V\left(\frac{j-1}{2^{N-1}}\right)$ such that $K\left(\frac{j}{2^{N-1}}\right) \subseteq V\left(\frac{j-1}{2^{N-1}}\right)$. Applying Lemma 6.7, we then obtain an open set $V(r)$ such that $K(r) = \overline{U(r)}$ is compact, and $K\left(\frac{j}{2^{N-1}}\right) \subseteq V(r) \subseteq K(r) \subseteq V\left(\frac{j-1}{2^{N-1}}\right)$. The claim thus holds by induction.

Define $f(x) = 0$ if $x \notin V(0)$ and $f(x) = \sup\{r \geq 0 : x \in V(r)\}$ otherwise. By construction $K = K(1) \subseteq V(r)$ for every dyadic rational $r \in [0,1)$, so $f = 1$ on K. Moreover, $\operatorname{supp}(f) \subseteq K(0) \subseteq U$. It remains to show that f is continuous.

Claim 2. For $a \in \mathbb{R}$,

$$\{f > a\} = \begin{cases} \emptyset, & \text{if } a \geq 1; \\ \bigcup_{r>a} V(r), & \text{if } 0 \leq a < 1; \\ X, & \text{if } a < 0. \end{cases}$$

The function f takes values between 0 and 1, so the cases $a < 0$ and $a \geq 1$ are immediate. Let $0 \leq a < 1$. Suppose $x \in X$ and $f(x) > a$. Then by the definition of f, there exists $r > a$ such that $x \in V(r)$. Hence, $x \in \bigcup_{r>a} V(r)$. Conversely, if $x \in \bigcup_{r>a} V(r)$, then $f(x) \geq r > a$.

Claim 3. For $b \in \mathbb{R}$,

$$\{f < b\} = \begin{cases} X, & \text{if } b > 1; \\ \bigcup_{r<b}(X \setminus K(r)), & \text{if } 0 < b \leq 1; \\ \emptyset, & \text{if } b \leq 0. \end{cases}$$

As in the previous claim, since f takes values in the interval $[0, 1]$, the cases $b > 1$ and $b \leq 0$ are immediate. Let $0 < b \leq 1$. Suppose $f(x) < b$. Taking $s \in (f(x), b)$, we have $x \notin V(s)$. Let $r \in (s, b)$. Then since $K(r) \subseteq V(s)$, we conclude $x \notin K(r)$. Hence, $x \in \bigcup_{r<b}(X \setminus K(r))$. Conversely, if $x \notin K(r)$ for some $r < b$, then $x \notin V(s)$ for $s \geq r$, so $f(x) \leq r < b$.

Combining Claims 2 and 3, for any $a, b \in \mathbb{R}$, the set $\{a < f < b\}$ is an intersection of two open sets and therefore open. Thus, f is continuous. □

Notation. For a compact set $K \subseteq X$ and a function $f : X \to [0, 1]$, we write $K \prec f$ if $f = 1$ on K. Given an open set $U \subseteq X$ and a function $f : X \to [0, 1]$, we write $f \prec U$ if $\operatorname{supp}(f) \subseteq U$. With this notation, the conclusion of Urysohn's lemma reads $K \prec f \prec U$.

Corollary 6.9: Partition of Unity

Let X be an LCH space. Let $K \subseteq X$ be a compact set and $U_1, \ldots, U_N \subseteq X$ an open cover of K. Then there exist compactly supported continuous functions $h_n : X \to [0, 1]$ such that $h_n \prec U_n$ for each $n \in \{1, \ldots, N\}$ and $\sum_{n=1}^N h_n = 1$ on K.

Proof. First apply Lemma 6.7 to obtain open sets $V_1, \ldots, V_N$ such that $\overline{V}_n$ is compact and $K \subseteq V_n \subseteq \overline{V}_n \subseteq U_n$. Then by Urysohn's lemma, let $f_n \in C_c(X)$ with $\overline{V}_n \prec f_n \prec U_n$. Define $h_1 = f_1, h_2 = (1 - f_1)f_2, \ldots, h_N = (1 - f_1) \ldots (1 - f_{N-1})f_N$. For each $n \in \{1, \ldots, N\}$, $h_n \leq f_n$, so h_n has compact support, and $h_n \prec U_n$. Moreover,

$$\sum_{n=1}^N h_n = 1 - \prod_{n=1}^N (1 - f_n).$$

For $x \in K$, at least one of the functions f_n is equal to 1, so $\sum_{n=1}^N h_n = 1$. □

Lemma 6.10

Let X be a locally compact Hausdorff space, and suppose μ is a Radon measure on X. Then for any open set $U \subseteq X$,

$$\mu(U) = \sup \left\{ \int_X f \, d\mu : f \in C_c(X), 0 \leq f \prec U \right\}.$$

Proof. Clearly $\mu(U) \geq \int_X f \, d\mu$ for any $f \in C_c(x)$ with $0 \leq f \leq \mathbb{1}_U$. Let us prove the reverse inequality. Let $c < \mu(U)$ be arbitrary. Then by inner regularity of μ on open sets, there exists a compact set $K \subseteq U$ such that $\mu(K) > c$. Then by Urysohn's lemma, there is a continuous function $f \in C_c(X)$ such that $K \prec f \prec U$. By monotonicity of the integral, we then have

$$\int_X f \, d\mu \geq \mu(K) > c.$$

$\square$

Bibliography

[1] Gerald B. Folland. *Real Analysis*, 2nd edition (John Wiley & Sons, Inc., 1999).

[2] Alexander S. Kechris. *Classical Descriptive Set Theory*, Grad. Texts in Math., **156** (Springer, 1995).

[3] John C. Oxtoby. *Measure and Category*, 2nd edition, Grad. Texts in Math., **2** (Springer, 1980).

[4] Walter Rudin. *Real and Complex Analysis*, 3rd edition (McGraw-Hill Book Co., 1987).

[5] Robert M. Solovay. A model of set-theory in which every set of reals is Lebesgue measurable. *Ann. Math. (2)* **92** (1970) 1–56.

[6] Elias M. Stein and Rami Shakarchi. *Real Analysis* (Princeton University Press, 2005).

[7] Terence Tao. *An Introduction to Measure Theory*, Grad. Stud. Math., **126** (American Mathematical Society, 2011). Preliminary version available online at https://terrytao.wordpress.com/wp-content/uploads/2012/12/gsm-126-tao5-measure-book.pdf

[8] Anush Tserunyan. *Introduction to Descriptive Set Theory*. Online lecture notes, available at: https://www.math.mcgill.ca/atserunyan/Teaching_notes/dst_lectures.pdf